Contents

Introduction 5
1. What is a Hopf algebra? 6
   1.1. Algebras 7
   1.2. Coalgebras 8
   1.3. Morphisms, tensor products, and bialgebras 9
   1.4. Antipodes and Hopf algebras 13
   1.5. Commutativity, cocommutativity 20
   1.6. Duals 22
2. Review of symmetric functions Λ as Hopf algebra 28
   2.1. Definition of Λ 28
   2.2. Other Bases 30
   2.3. Comultiplications 34
   2.4. The antipode, the involution ω, and algebra generators 36
   2.5. Cauchy product, Hall inner product, self-duality 38
   2.6. Bialternants, Littlewood-Richardson: Stembridge’s concise proof 46
   2.7. The Pieri and Assaf-McNamara skew Pieri rule 50
   2.8. Skewing and Lam’s proof of the skew Pieri rule 54
   2.9. Assorted exercises on symmetric functions 56
3. Zelevinsky’s structure theory of positive self-dual Hopf algebras 70
   3.1. Self-duality implies polynomiality 70
   3.2. The decomposition theorem 73
   3.3. Λ is the unique indecomposable PSH 76
4. Complex representations for $S_n$, wreath products, $GL_n(F_q)$ 82
   4.1. Review of complex character theory 82
   4.2. Three towers of groups 90
   4.3. Bialgebra and double cosets 92
   4.4. Symmetric groups 99
   4.5. Wreath products 102
   4.6. General linear groups 104
   4.7. Steinberg’s unipotent characters 105
   4.8. Examples: $GL_2(F_2)$ and $GL_3(F_2)$ 106
   4.9. The Hall algebra 108
5. Quasisymmetric functions and $P$-partitions 114
   5.1. Definitions, and Hopf structure 114
   5.2. The fundamental basis and $P$-partitions 119
   5.3. The Hopf algebra NSym dual to QSym 126
6. Polynomial generators for QSym and Lyndon words 133
   6.1. Lyndon words 133
   6.2. Shuffles and Lyndon words 147
   6.3. Radford’s theorem on the shuffle algebra 159
   6.4. Polynomial freeness of QSym: statement and easy parts 162

Date: August 25, 2015.

Key words and phrases. Hopf algebra, combinatorics.
11.44. Solution to Exercise 2.65
11.45. Solution to Exercise 2.66
11.46. Solution to Exercise 2.67
11.47. Solution to Exercise 2.68
11.48. Solution to Exercise 2.69
11.49. Solution to Exercise 2.70
11.50. Solution to Exercise 2.71
11.51. Solution to Exercise 2.75
11.52. Solution to Exercise 2.76
11.53. Solution to Exercise 2.77
11.54. Solution to Exercise 2.79
11.55. Solution to Exercise 2.80
11.56. Solution to Exercise 2.82
11.57. Solution to Exercise 2.84
11.58. Solution to Exercise 2.85
11.59. Solution to Exercise 2.86
11.60. Solution to Exercise 2.87
11.61. Solution to Exercise 2.89
11.62. Solution to Exercise 2.90
11.63. Solution to Exercise 2.91
11.64. Solution to Exercise 2.92
11.65. Solution to Exercise 2.93
11.66. Solution to Exercise 2.94
11.67. Solution to Exercise 2.96
11.68. Solution to Exercise 2.97
11.69. Solution to Exercise 2.98
11.70. Solution to Exercise 3.7
11.71. Solution to Exercise 3.10
11.72. Solution to Exercise 4.1
11.73. Solution to Exercise 4.2
11.74. Solution to Exercise 4.3
11.75. Solution to Exercise 4.4
11.76. Solution to Exercise 4.6
11.77. Solution to Exercise 4.9
11.78. Solution to Exercise 4.10
11.79. Solution to Exercise 4.11
11.80. Solution to Exercise 4.12
11.81. Solution to Exercise 4.14
11.82. Solution to Exercise 4.15
11.83. Solution to Exercise 4.16
11.84. Solution to Exercise 4.20
11.85. Solution to Exercise 4.29
11.86. Solution to Exercise 4.31
11.87. Solution to Exercise 4.34
11.88. Solution to Exercise 4.35
11.89. Solution to Exercise 4.36
11.90. Solution to Exercise 4.38
11.91. Solution to Exercise 4.39
11.92. Solution to Exercise 4.48
11.93. Solution to Exercise 4.60
11.94. Solution to Exercise 5.21
11.95. Solution to Exercise 5.37
11.96. Solution to Exercise 5.38
11.97. Solution to Exercise 5.40
<table>
<thead>
<tr>
<th>Exercise Number</th>
<th>Solution to Exercise</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.44</td>
<td></td>
<td>614</td>
</tr>
<tr>
<td>5.45</td>
<td></td>
<td>621</td>
</tr>
<tr>
<td>5.46</td>
<td></td>
<td>623</td>
</tr>
<tr>
<td>6.3</td>
<td></td>
<td>632</td>
</tr>
<tr>
<td>6.7</td>
<td></td>
<td>634</td>
</tr>
<tr>
<td>6.9</td>
<td></td>
<td>636</td>
</tr>
<tr>
<td>6.10</td>
<td></td>
<td>636</td>
</tr>
<tr>
<td>6.11</td>
<td></td>
<td>636</td>
</tr>
<tr>
<td>6.12</td>
<td></td>
<td>637</td>
</tr>
<tr>
<td>6.21</td>
<td></td>
<td>638</td>
</tr>
<tr>
<td>6.22</td>
<td></td>
<td>642</td>
</tr>
<tr>
<td>6.23</td>
<td></td>
<td>642</td>
</tr>
<tr>
<td>6.24</td>
<td></td>
<td>643</td>
</tr>
<tr>
<td>6.29</td>
<td></td>
<td>645</td>
</tr>
<tr>
<td>6.31</td>
<td></td>
<td>647</td>
</tr>
<tr>
<td>6.32</td>
<td></td>
<td>649</td>
</tr>
<tr>
<td>6.33</td>
<td></td>
<td>658</td>
</tr>
<tr>
<td>6.34</td>
<td></td>
<td>675</td>
</tr>
<tr>
<td>6.35</td>
<td></td>
<td>685</td>
</tr>
<tr>
<td>6.36</td>
<td></td>
<td>686</td>
</tr>
<tr>
<td>6.39</td>
<td></td>
<td>686</td>
</tr>
<tr>
<td>6.40</td>
<td></td>
<td>687</td>
</tr>
<tr>
<td>6.41</td>
<td></td>
<td>695</td>
</tr>
<tr>
<td>6.49</td>
<td></td>
<td>707</td>
</tr>
<tr>
<td>6.51</td>
<td></td>
<td>710</td>
</tr>
<tr>
<td>6.53</td>
<td></td>
<td>712</td>
</tr>
<tr>
<td>6.57</td>
<td></td>
<td>716</td>
</tr>
<tr>
<td>6.59</td>
<td></td>
<td>717</td>
</tr>
<tr>
<td>6.61</td>
<td></td>
<td>720</td>
</tr>
<tr>
<td>6.63</td>
<td></td>
<td>721</td>
</tr>
<tr>
<td>6.66</td>
<td></td>
<td>722</td>
</tr>
<tr>
<td>6.67</td>
<td></td>
<td>725</td>
</tr>
<tr>
<td>6.70</td>
<td></td>
<td>726</td>
</tr>
<tr>
<td>6.74</td>
<td></td>
<td>727</td>
</tr>
<tr>
<td>6.77</td>
<td></td>
<td>729</td>
</tr>
<tr>
<td>6.78</td>
<td></td>
<td>731</td>
</tr>
<tr>
<td>6.79</td>
<td></td>
<td>734</td>
</tr>
<tr>
<td>6.81</td>
<td></td>
<td>736</td>
</tr>
<tr>
<td>6.85</td>
<td></td>
<td>749</td>
</tr>
<tr>
<td>6.87</td>
<td></td>
<td>753</td>
</tr>
<tr>
<td>6.91</td>
<td></td>
<td>753</td>
</tr>
<tr>
<td>6.92</td>
<td></td>
<td>757</td>
</tr>
<tr>
<td>6.94</td>
<td></td>
<td>760</td>
</tr>
<tr>
<td>6.98</td>
<td></td>
<td>761</td>
</tr>
<tr>
<td>6.99</td>
<td></td>
<td>769</td>
</tr>
<tr>
<td>6.101</td>
<td></td>
<td>769</td>
</tr>
<tr>
<td>6.103</td>
<td></td>
<td>771</td>
</tr>
<tr>
<td>6.106</td>
<td></td>
<td>771</td>
</tr>
<tr>
<td>6.109</td>
<td></td>
<td>772</td>
</tr>
<tr>
<td>6.111</td>
<td></td>
<td>774</td>
</tr>
<tr>
<td>6.114</td>
<td></td>
<td>784</td>
</tr>
<tr>
<td>6.116</td>
<td></td>
<td>784</td>
</tr>
<tr>
<td>6.118</td>
<td></td>
<td>785</td>
</tr>
<tr>
<td>6.120</td>
<td></td>
<td>786</td>
</tr>
</tbody>
</table>
Introduction

The concept of a Hopf algebra crystallized out of algebraic topology and the study of algebraic groups in the 1940s and 1950s (see [8] and [29] for its history). Being a fairly elementary algebraic notion itself, it subsequently found applications in other mathematical disciplines, and is now particularly commonplace in representation theory.\footnote{where it provides explanations for similarities between group representations and Lie algebra representations}

These notes concern themselves (after a brief introduction into the algebraic foundations of Hopf algebra theory in Chapter 1) with the Hopf algebras that appear in combinatorics. These Hopf algebras tend to have bases naturally parametrized by combinatorial objects (partitions, compositions, permutations, tableaux, graphs, trees, posets, polytopes, etc.), and their Hopf-algebraic operations often encode basic operations on these objects.\footnote{such as concatenating two compositions, or taking the disjoint union of two graphs – but, more often, operations which return a multiset of results, such as cutting a composition into two pieces at all possible places, or partitioning a poset into two subposets in every way that satisfies a certain axiom}. Combinatorial results can then be seen as particular cases of general algebraic properties of Hopf algebras (e.g., the multiplicativity of the Möbius function can be recovered from the fact that the antipode of a Hopf algebra is an algebra anti-endomorphism), and many interesting invariants of combinatorial objects turn out to be evaluations of Hopf morphisms. In some cases (particularly that of symmetric functions), the rigidity in the structure of a Hopf algebra can lead to enlightening proofs.

One of the most elementary interesting examples of a combinatorial Hopf algebra is that of the symmetric functions. We will devote all of Chapter 2 to studying it, deviating from the usual treatments (such as in Stanley [157, Ch. 7], Sagan [141] and Macdonald [109]) by introducing the Hopf-algebraic structure early on and using it to obtain combinatorial results. Chapter 3 will underpin the importance of this algebra by proving Zelevinsky’s main theorem of PSH theory, which (roughly) claims that a Hopf algebra over $\mathbb{Z}$ satisfying a certain set of axioms must be a tensor product of copies of the Hopf algebra of symmetric functions. These axioms are fairly restrictive, so this result is far from curtailing the diversity of combinatorial Hopf algebras; but they are natural enough that, as we will see in Chapter 4, they are satisfied for a Hopf algebra of representations of symmetric groups. As a consequence, this Hopf algebra will be revealed isomorphic to the symmetric functions – this is the famous Frobenius correspondence between symmetric functions and characters of symmetric groups, usually obtained through other ways ([54, §7.3], [141, §4.7]).

We will further elaborate on the representation theories of wreath products and general linear groups over finite fields; while Zelevinsky’s PSH theory does not fully explain the latter, it illuminates it significantly.

In the next chapters, we will study further examples of combinatorial Hopf algebras: the quasisymmetric functions and the noncommutative symmetric functions in Chapter 5, various other algebras (of graphs, posets, matroids, etc.) in Chapter 7, and the Malvenuto-Reutenauer Hopf algebra of permutations in Chapter 8.
The main prerequisite for reading these notes is a good understanding of graduate algebra\(^3\), in particular multilinear algebra (tensor products, symmetric powers and exterior powers)\(^4\) and basic categorical language. In Chapter 4, familiarity with representation theory of finite groups (over \(\mathbb{C}\)) is assumed, along with the theory of finite fields and (at some places) the rational canonical form of a matrix. Only basic knowledge of combinatorics is required (except for a few spots in Chapter 7), and familiarity with geometry and topology is needed only to understand some tangential remarks. The concepts of Hopf algebras and coalgebras and the basics of symmetric function theory will be introduced as needed. We will work over a commutative base ring most of the time, but no commutative algebra (besides, occasionally, properties of modules over a PID) will be used.

These notes began as an accompanying text for Fall 2012 Math 8680 Topics in Combinatorics, a graduate class taught by the second author at the University of Minnesota. The first author has since added many exercises (and solutions) as well as Chapter 6 on Lyndon words and the polynomiality of QSym. The notes might still grow, and any comments, corrections and complaints are welcome!

The course was an attempt to focus on examples that we find interesting, but which are hard to find fully explained currently in books or in one paper. Much of the subject of combinatorial Hopf algebras is fairly recent (1990s onwards) and still spread over research papers, although sets of lecture notes do exist, such as Foissy’s \([51]\). A reference which we discovered late, having a great deal of overlap with these notes is Hazewinkel, Gubareni, and Kirichenko \([69]\). References for the purely algebraic theory of Hopf algebras are much more frequent (see the beginning of Chapter 1 for a list).

Be warned that our notes are highly idiosyncratic in choice of topics, and they steal heavily from the sources in the bibliography.

### Warnings:

Unless otherwise specified ...

- \(\mathbf{k}\) here usually denotes a commutative ring\(^5\).
- all maps between \(\mathbf{k}\)-modules are \(\mathbf{k}\)-linear.
- every ring or \(\mathbf{k}\)-algebra is associative and has a 1, and every ring morphism or \(\mathbf{k}\)-algebra morphism preserves the 1’s.
- all \(\mathbf{k}\)-algebras \(A\) have the property that \((\lambda 1_A) a = a (\lambda 1_A) = \lambda a\) for all \(\lambda \in \mathbf{k}\) and \(a \in A\).
- all tensor products are over \(\mathbf{k}\).
- \(I\) will denote the multiplicative identity in some ring like \(\mathbf{k}\) or in some \(\mathbf{k}\)-algebra (sometimes also the identity of a group written multiplicatively).

### 1. What is a Hopf algebra?

The standard references for Hopf algebras are Abe [1] and Sweedler [163], and some other good ones are [28, 30, 38, 86, 120, 171, 69]. See also Foissy [51] and Manchon [116] for introductions to Hopf algebras tailored to combinatorial applications. Most texts only study Hopf algebras over fields (with exceptions such as Foissy’s [51] and Manchon [116] for introductions to Hopf algebras

---

\(^3\)William Schmitt’s expositions [146] are tailored to a reader interested in combinatorial Hopf algebras; his notes on modules and algebras cover a significant part of what we need from abstract algebra, whereas those on categories cover all category theory we will use and much more.

\(^4\)Keith Conrad’s expository notes [33] are useful, even if not comprehensive, sources for the latter.

\(^5\)As explained below, “ring” means “associative ring with 1”. The most important cases are when \(\mathbf{k}\) is a field or when \(\mathbf{k} = \mathbb{Z}\).
as \([30, 28, 171]\)). We will work over arbitrary commutative rings\(^6\), which requires some more care at certain points (but we will not go deep enough into the algebraic theory to witness the situation over commutative rings diverge seriously from that over fields).

Let’s build up the definition of Hopf algebra structure bit-by-bit, starting with the more familiar definition of algebras.

1.1. **Algebras.** The following definition of \(k\)-algebras may look unfamiliar, but it is merely a restatement of their classical definition using tensors\(^7\) and \(k\)-linear maps:

**Definition 1.1.** An **associative \(k\)-algebra** \(A\) is a \(k\)-module with a \(k\)-linear associative operation \(\otimes\) \(A \otimes A \rightarrow A\), and a \(k\)-linear **unit** \(\mathbf{1} \rightarrow A\) sending \(1\) in \(k\) to the two-sided multiplicative identity element \(1\) in \(A\). One can rephrase this by saying that these diagrams commute:

\[
\begin{array}{c}
\begin{array}{ccc}
A \otimes A & \otimes & A \\
\downarrow & & \downarrow \\
A & \otimes & A
\end{array}
\end{array}
\]

\[
\begin{array}{c}
\begin{array}{ccc}
A & \otimes & k \\
\downarrow & & \downarrow \\
k & \otimes & A
\end{array}
\end{array}
\]

where the maps \(A \rightarrow A \otimes k\) and \(A \rightarrow k \otimes A\) are the isomorphisms sending \(a \mapsto a \otimes 1\) and \(a \mapsto 1 \otimes a\).

We abbreviate “associative \(k\)-algebra” as “\(k\)-algebra” (associativity is assumed unless otherwise specified) or as “algebra” (when \(k\) is clear from the context).

Well-known examples of \(k\)-algebras are **tensor** and **symmetric algebras**, which we can think of as algebras of **words** and **multisets**, respectively.

**Example 1.2.** If \(V\) is a \(k\)-module and \(n \in \mathbb{N}\), then the **\(n\)-fold tensor power** \(V \otimes^\otimes V\) of \(V\) is the \(k\)-module \(V \otimes V \otimes \cdots \otimes V\). (For \(n = 0\), this is the \(k\)-module \(k\), spanned by the “empty tensor” \(1_k\).)

The **tensor algebra** \(T(V) = \bigoplus_{n \geq 0} V \otimes^\otimes V\) on a \(k\)-module \(V\) is an associative \(k\)-algebra spanned (as \(k\)-module) by decomposable tensors \(v_1 v_2 \cdots v_k := v_1 \otimes v_2 \otimes \cdots \otimes v_k\) with \(k \in \mathbb{N}\) and \(v_1, v_2, \ldots, v_k \in V\). Its multiplication is defined \(k\)-linearly by

\[
m(v_1 v_2 \cdots v_k \otimes w_1 w_2 \cdots w_\ell) := v_1 v_2 \cdots v_k w_1 w_2 \cdots w_\ell
\]

for all \(k, \ell \in \mathbb{N}\) and \(v_1, v_2, \ldots, v_k, w_1, w_2, \ldots, w_\ell \in V\). The unit map \(u : k \rightarrow T(V)\) sends \(1_k\) to the empty tensor \(1_{T(V)} = 1_k \in k = V \otimes^\otimes V\).

If \(V\) is a free \(k\)-module, say with \(k\)-basis \(\{x_i\}_{i \in I}\), then \(T(V)\) has a \(k\)-basis of decomposable tensors \(x_{i_1} \cdots x_{i_k} := x_{i_1} \otimes \cdots \otimes x_{i_k}\) indexed by **words** \((i_1, \ldots, i_k)\) in the alphabet \(I\), and the multiplication on this basis is given by concatenation of words:

\[
m(x_{i_1} \cdots x_{i_k} \otimes x_{j_1} \cdots x_{j_\ell}) = x_{i_1} \cdots x_{i_k} x_{j_1} \cdots x_{j_\ell}.
\]

\(^6\)and we will profit from this generality in Chapters 3 and 4, where we will be applying the theory of Hopf algebras to \(k = \mathbb{Z}\) in a way that would not be possible over \(k = \mathbb{Q}\).

\(^7\)i.e., the multiplication is encoded by a \(k\)-linear map \(A \otimes A \rightarrow A\) instead of a \(k\)-bilinear map \(A \times A \rightarrow A\).

\(^8\)Some remarks about our notation (which we are using here and throughout these notes) are in order.

Since we are working with tensor products of \(k\)-modules like \(T(V)\) – which themselves are made of tensors – here, we must specify what the \(\otimes\) sign means in expressions like \(a \otimes b\) where \(a\) and \(b\) are elements of \(T(V)\). Our convention is the following: When \(a\) and \(b\) are elements of a tensor algebra \(T(V)\), we always understand \(a \otimes b\) to mean the pure tensor \(a \otimes b \in T(V) \otimes T(V)\) rather than the product of \(a\) and \(b\) inside the tensor algebra \(T(V)\). The latter product will plainly be written \(ab\).

The operator precedence between \(\otimes\) and multiplication in \(T(V)\) is such that multiplication in \(T(V)\) binds more tightly than the \(\otimes\) sign; e.g., the term \(ab \otimes cd\) means \((ab) \otimes (cd)\).
Recall that in an algebra $A$, when one has a two-sided ideal $J \subset A$, meaning a $k$-submodule with $m(J \otimes A), m(A \otimes J) \subset J$, then one can form a quotient algebra $A/J$.

**Example 1.3.** The symmetric algebra $\text{Sym}(V) = \bigoplus_{n \geq 0} \text{Sym}^n(V)$ is the quotient of $T(V)$ by the two-sided ideal generated by all elements $xy - yx$ with $x, y \in V$. When $V$ is a free $k$-module with basis $\{x_i\}_{i \in I}$, this symmetric algebra $S(V)$ can be identified with a (commutative) polynomial algebra $k[x_i]_{i \in I}$, having a $k$-basis of (commutative) monomials $x_{i_1} \cdots x_{i_k}$ as $(i_1, \ldots, i_k)$ runs through all finite multisubsets of $I$, and with multiplication defined $k$-linearly via multiset union.

Note that the $k$-module $k$ itself canonically becomes a $k$-algebra. Its associative operation $m : k \otimes k \to k$ is the canonical isomorphism $k \otimes k \to k$, and its unit $u : k \to k$ is the identity map.

Topology and group theory give more examples.

**Example 1.4.** The cohomology algebra $H^*(X; k) = \bigoplus_{i \geq 0} H^i(X; k)$ with coefficients in $k$ for a topological space $X$ has an associative cup product. Its unit $k = H^*(\text{pt}; k) \xrightarrow{u} H^*(X; k)$ is induced from the unique (continuous) map $X \to \text{pt}$, where $\text{pt}$ is a one-point space.

**Example 1.5.** For a group $G$, the group algebra $kG$ has a $k$-basis $\{t_g\}_{g \in G}$ and multiplication defined $k$-linearly by $t_g t_h = t_{gh}$, and unit defined by $u(1) = t_e$, where $e$ is the identity element of $G$.

1.2. Coalgebras. If we are to think of the multiplication $A \otimes A \to A$ in an algebra as putting together two basis elements of $A$ to get a sum of basis elements of $A$, then coalgebra structure should be thought of as taking basis elements apart.

**Definition 1.6.** A co-associative $k$-coalgebra $C$ is a $k$-module $C$ with a comultiplication, that is, a $k$-linear map $C \xrightarrow{\Delta} C \otimes C$, and a $k$-linear counit $C \xrightarrow{\epsilon} k$ making commutative the diagrams as in (1.1), (1.2) but with all arrows reversed:

\[
\begin{array}{ccc}
C \otimes C & \xrightarrow{\Delta \otimes \text{id}} & C \otimes C \otimes C \\
\downarrow \Delta & & \downarrow \text{id} \otimes \Delta \\
C \otimes C & \xrightarrow{\Delta} & C \otimes C \\
\downarrow \Delta & & \downarrow \Delta \\
C & & C \\
\end{array}
\]

\[
\begin{array}{ccc}
C \otimes k & \xrightarrow{\text{id} \otimes \epsilon} & k \otimes C \\
\downarrow \text{id} \otimes \epsilon & & \downarrow \epsilon \otimes \text{id} \\
C \otimes C & \xrightarrow{\Delta} & C \otimes C \\
\end{array}
\]

Here the maps $C \otimes k \to C$ and $k \otimes C \to C$ are the isomorphisms sending $c \otimes 1 \mapsto c$ and $1 \otimes c \mapsto c$.

We abbreviate “co-associative $k$-coalgebra” as “$k$-coalgebra” (co-associativity, i.e., the commutativity of the diagram (1.3), is assumed unless otherwise specified) or as “coalgebra” (when $k$ is clear from the context). Sometimes, the word “coproduct” is used as a synonym for “comultiplication”\(^{10}\).

One often uses the Sweedler notation

\[\Delta(c) = \sum_{c_1 \otimes c_2} c_1 \otimes c_2 = \sum_{c_1 \otimes c_2} c_1 \otimes c_2\]

\(^9\)By a multisubset of a set $S$, we mean a multiset each of whose elements belongs to $S$ (but can appear arbitrarily often).\n
\(^{10}\)although the word “coproduct” already has a different meaning in algebra
Example 1.7. Let $k$ be a field. The homology $H_\ast(X;k)$ = $\bigoplus_{i \geq 0} H_i(X;k)$ for a topological space $X$ is naturally a coalgebra: the (continuous) diagonal embedding $X \to X \times X$ sending $x \mapsto (x,x)$ induces a coassociative map
\[ H_\ast(X;k) \to H_\ast(X \times X;k) \cong H_\ast(X;k) \otimes H_\ast(X;k) \]
in which the last isomorphism comes from the Künneth theorem with field coefficients $k$. As before, the unique (continuous) map $X \to pt$ induces the counit $H_\ast(X;k) \to H_\ast(pt;k) \cong k$.

Exercise 1.8. Given a $k$-module $C$ and a $k$-linear map $\Delta : C \to C \otimes C$. Prove that there exists at most one $k$-linear map $\epsilon : C \to k$ such that the diagram (1.4) commutes.

1.3. Morphisms, tensor products, and bialgebras.

Definition 1.9. A morphism of algebras $A \xrightarrow{\varphi} B$ makes these diagrams commute:

\[(1.5)\]
\[
\begin{array}{ccc}
A & \varphi \rightarrow & B \\
\downarrow m_A & & \downarrow m_B \\
A \otimes A & \varphi \otimes \varphi \rightarrow & B \otimes B \\
\end{array}
\]

Here the subscripts on $m_A, m_B, u_A, u_B$ indicate for which algebra they are part of the structure— we will occasionally use such conventions from now on.

Similarly a morphism of coalgebras is a $k$-linear map $C \xrightarrow{\varphi} D$ making the reverse diagrams commute:

\[(1.6)\]
\[
\begin{array}{ccc}
C & \varphi \rightarrow & D \\
\downarrow \Delta_C & & \downarrow \Delta_D \\
C \otimes C & \varphi \otimes \varphi \rightarrow & D \otimes D \\
\end{array}
\]

Example 1.10. Let $k$ be a field. Continuous maps $X \xrightarrow{f} Y$ of topological spaces induce algebra morphisms $H^\ast(Y;k) \to H^\ast(X;k)$, and coalgebra morphisms $H_\ast(X;k) \to H_\ast(Y;k)$.

Definition 1.11. Given two $k$-algebras $A, B$, their tensor product $A \otimes B$ also becomes a $k$-algebra defining the multiplication bilinearly via
\[ m((a \otimes b) \otimes (a' \otimes b')) := a a' \otimes b b' \]
or in other words $m_{A \otimes B}$ is the composite map
\[ A \otimes B \otimes A \otimes B \xrightarrow{id \otimes T \otimes id} A \otimes A \otimes B \otimes B \xrightarrow{m_A \otimes m_B} A \otimes B \]
where $T$ is the twist map $B \otimes A \to A \otimes B$ that sends $b \otimes a \mapsto a \otimes b$.

Here we are omitting the topologist’s sign in the twist map which should be present for graded algebras and coalgebras that come from cohomology and homology: for homogeneous elements $a$ and $b$ the topologist’s twist map sends
\[(1.7)\]
\[ b \otimes a \mapsto (-1)^{\deg(a) \cdot \deg(b)} a \otimes b. \]

This means that, if one is using the topologists’ conventions, most of our examples which we later call graded should actually be considered to live in only even degrees (which can be achieved, e.g., by artificially doubling their grading). We will, however, keep to our own definitions (so the twist map sends $b \otimes a \mapsto a \otimes b$) unless otherwise noted.

The unit element of $A \otimes B$ is $1_A \otimes 1_B$, meaning that the unit map $k \xrightarrow{u_A \otimes u_B} A \otimes B$ is the composite
\[ k \xrightarrow{u_A \otimes u_B} A \otimes B. \]

Similarly, given two coalgebras $C, D$, one can make $C \otimes D$ a coalgebra in which the comultiplication and counit maps are the composites of
\[ C \otimes D \xrightarrow{\Delta_C \otimes \Delta_D} C \otimes C \otimes D \otimes D \xrightarrow{id \otimes T \otimes id} C \otimes D \otimes C \otimes D \]
and

\[ C \otimes D \xrightarrow{\epsilon_C \otimes \epsilon_D} k \otimes k \twoheadrightarrow k \].

One of the first signs that these definitions interact nicely is the following straightforward proposition.

**Proposition 1.12.** When \( A \) is both a \( k \)-algebra and a \( k \)-coalgebra, the following are equivalent:

- \((\Delta, \epsilon)\) are morphisms for the algebra structure \((m, u)\).
- \((m, u)\) are morphisms for the coalgebra structure \((\Delta, \epsilon)\).
- These four diagrams commute:

\[ \begin{align*}
A \otimes A & \xrightarrow{\Delta \otimes \Delta} A \otimes A \\
\xrightarrow{id \otimes T \otimes id} A \otimes A \otimes A \otimes A & \xrightarrow{m} A \\
\xrightarrow{\epsilon \otimes \epsilon} k \otimes k & \xrightarrow{m} A \otimes A \\
\xrightarrow{\Delta} A & \xrightarrow{\epsilon} k
\end{align*} \]

**Exercise 1.13.** (a) If \( A, A', B \) and \( B' \) are four \( k \)-algebras, and \( f : A \to A' \) and \( g : B \to B' \) are two \( k \)-algebra homomorphisms, then show that \( f \otimes g : A \otimes B \to A' \otimes B' \) is a \( k \)-algebra homomorphism.

(b) If \( C, C', D \) and \( D' \) are four \( k \)-coalgebras, and \( f : C \to C' \) and \( g : D \to D' \) are two \( k \)-coalgebra homomorphisms, then show that \( f \otimes g : C \otimes D \to C' \otimes D' \) is a \( k \)-coalgebra homomorphism.

**Definition 1.14.** Call the \( k \)-module \( A \) a \( k \)-bialgebra if it is a \( k \)-algebra and \( k \)-coalgebra satisfying the three equivalent conditions in Proposition 1.12.

**Example 1.15.** For a group \( G \), one can make the group algebra \( kG \) a coalgebra with counit \( kG \xrightarrow{\epsilon} k \) mapping \( t_g \mapsto 1 \) for all \( g \) in \( G \), and with comultiplication \( kG \xrightarrow{\Delta} kG \otimes kG \) given by \( \Delta(t_g) := t_g \otimes t_g \). Checking the various diagrams in (1.8) commute is easy. For example, one can check the pentagonal diagram on each
Example 1.18. The tensor algebra $T(V) = \bigoplus_{n \geq 0} V \otimes^n$ is a coalgebra, with counit $\epsilon$ equal to the identity on $V \otimes^0 = \mathbb{k}$ and the zero map on $V \otimes^n$ for $n > 0$, and with comultiplication defined to make the elements $x$ in $V \otimes^1 = V$ all primitive:

$$\Delta(x) := 1 \otimes x + x \otimes 1 \quad \text{for } x \in V \otimes^1.$$ 

Since the elements of $V$ generate $T(V)$ as a $\mathbb{k}$-algebra, and since $T(V) \otimes T(V)$ is also an associative $\mathbb{k}$-algebra, the universal property of $T(V)$ as the free associative $\mathbb{k}$-algebra on the generators $V$ allows one to define $T(V) \xrightarrow{\Delta} T(V) \otimes T(V)$ arbitrarily on $V$, and extend it as an algebra morphism.

It may not be obvious that this $\Delta$ is coassociative, but one can note that
$$((\text{id} \otimes \Delta) \circ \Delta) (x) = x \otimes 1 \otimes 1 + 1 \otimes x \otimes 1 + 1 \otimes 1 \otimes x = ((\Delta \circ \text{id}) \circ \Delta) (x)$$

for every $x$ in $V$. Hence the two maps $(\text{id} \otimes \Delta) \circ \Delta$ and $(\Delta \circ \text{id}) \circ \Delta$, considered as algebra morphisms $T(V) \to T(V) \otimes T(V) \otimes T(V)$, must coincide on every element of $T(V)$ since they coincide on $V$. We leave it as an exercise to check the map $\epsilon$ defined as above satisfies the counit axioms (1.4).

Here is a sample calculation in $T(V)$ when $V$ has basis $\{x, y, z\}$:

$$\Delta(xyz) = \Delta(x) \Delta(y) \Delta(z) = (1 \otimes x + x \otimes 1)(1 \otimes y + y \otimes 1)(1 \otimes z + z \otimes 1)$$

$$= (1 \otimes xy + x \otimes y + y \otimes x + xy \otimes 1)(1 \otimes z + z \otimes 1)$$

$$= 1 \otimes xyz + x \otimes yz + y \otimes zx + z \otimes xy$$

$$+ xy \otimes z + xz \otimes y + yz \otimes x + xy \otimes z \otimes 1$$

This illustrates the idea that comultiplication “takes basis elements apart”. Here for any $v_1, v_2, \ldots, v_n$ in $V$ one has

$$\Delta(v_1 v_2 \cdots v_n) = \sum_{\langle j_1, j_2, \ldots, j_r \rangle, \langle k_1, k_2, \ldots, k_{n-r} \rangle \text{ of complementary subwords of the word } (1, 2, \ldots, n) 11\text{.}$$

Equivalently (and in a more familiar language),

$$(1.9) \quad \Delta(v_1 v_2 \cdots v_n) = \sum_{I \subset \{1, 2, \ldots, n\}} v_I \otimes v_{\{1, 2, \ldots, n\} \setminus I},$$

11More formally speaking, the sum is over all permutations $\langle j_1, j_2, \ldots, j_r, k_1, k_2, \ldots, k_{n-r} \rangle$ of $(1, 2, \ldots, n)$ satisfying $j_1 < j_2 < \cdots < j_r$ and $k_1 < k_2 < \cdots < k_{n-r}$. 

Remark 1.16. In fact, one can think of adding a bialgebra structure to a $\mathbb{k}$-algebra $A$ as a way of making $A$-modules $M, N$ have an $A$-module structure on their tensor product $M \otimes N$: the algebra $A \otimes A$ already acts naturally on $M \otimes N$, so one can let $a$ in $A$ act via $\Delta(a)$ in $A \otimes A$. In the theory of group representations over $\mathbb{k}$, that is, $\mathbb{k}G$-modules $M$, this is how one defines the diagonal action of $G$ on $M \otimes N$, namely $t_g$ acts as $t_g \otimes t_g$.

Definition 1.17. An element $x$ in a coalgebra for which $\Delta(x) = x \otimes x$ and $\epsilon(x) = 1$ is called group-like.

An element $x$ in a bialgebra for which $\Delta(x) = 1 \otimes x + x \otimes 1$ is called primitive.

basis element $t_g \otimes t_h$:
where \( v_J \) (for \( J \) a subset of \( \{1, 2, \ldots, n\} \)) denotes the product of all \( v_j \) with \( j \in J \) in the order of increasing \( j \).

Recall one can quotient a \( k \)-algebra \( A \) by a two-sided ideal \( J \) to obtain a quotient algebra \( A/J \).

**Definition 1.19.** In a coalgebra \( C \), a *two-sided coideal* is a \( k \)-submodule \( J \subset C \) for which
\[
\Delta(J) \subset J \otimes C + C \otimes J
\]
\[
\epsilon(J) = 0
\]
The quotient \( k \)-module \( C/J \) then inherits a coalgebra structure\(^\text{12}\). Similarly, in a bialgebra \( A \), a subset \( J \subset A \) which is both a two-sided ideal and two-sided coideal gives rise to a quotient bialgebra \( A/J \).

**Exercise 1.20.** Let \( A \) and \( C \) be two \( k \)-coalgebras, and \( f : A \to C \) a surjective coalgebra homomorphism.

(a) If \( f \) is surjective, then show that \( \ker f \) is a two-sided coideal of \( A \).

(b) If \( k \) is a field, then show that \( \ker f \) is a two-sided coideal of \( A \).

**Example 1.21.** The *symmetric algebra* \( \text{Sym}(V) \) was the quotient of \( T(V) \) by the two-sided ideal \( J \) generated by all *commutators* \( [x, y] = xy - yx \) for \( x, y \) in \( V \). Note that \( x, y \) are primitive elements in \( T(V) \), and the following very reusable calculation shows that the *commutator of two primitives is primitive*:
\[
\Delta[x, y] = \Delta(xy - yx) = (1 \otimes x + x \otimes 1)(1 \otimes y + y \otimes 1) - (1 \otimes y + y \otimes 1)(1 \otimes x + x \otimes 1)
\]
\[
= 1 \otimes xy - 1 \otimes yx + xy \otimes 1 - yx \otimes 1
\]
\[
+ x \otimes y + y \otimes x - x \otimes y - y \otimes x
\]
\[
= 1 \otimes (xy - yx) + (xy - yx) \otimes 1
\]
\[
= 1 \otimes [x, y] + [x, y] \otimes 1.
\]
In particular, the commutators \( [x, y] \) have \( \Delta[x, y] \in J \otimes T(V) + T(V) \otimes J \). They also satisfy \( \epsilon([x, y]) = 0 \). Since they are generators for \( J \) as a two-sided ideal, it is not hard to see this implies \( \Delta(J) \subset J \otimes T(V) + T(V) \otimes J \), and \( \epsilon(J) = 0 \). Thus \( J \) is also a two-sided coideal, and \( \text{Sym}(V) = T(V)/J \) inherits a bialgebra structure.

In fact we will see in Section 3.1 that symmetric algebras are the universal example of bialgebras which are *graded, connected, commutative, cocommutative*. But first we should define some of these concepts.

**Definition 1.22.** A *graded \( k \)-module* \( V \) is one with a \( k \)-module direct sum decomposition \( V = \bigoplus_{n \geq 0} V_n \). Elements \( x \) in \( V_n \) are called *homogeneous* of degree \( n \), or \( \deg(x) = n \).

One endows tensor products \( V \otimes W \) of graded \( k \)-modules \( V, W \) with graded module structure in which
\[
(V \otimes W)_n := \bigoplus_{i+j=n} V_i \otimes W_j
\]
A \( k \)-linear map \( V \xrightarrow{\varphi} W \) between two graded \( k \)-modules is called *graded* if \( \varphi(V_n) \subset W_n \) for all \( n \). Say that a \( k \)-algebra (coalgebra, bialgebra) is *graded* if it is a graded \( k \)-module and all of the relevant structure maps \((u, \epsilon, m, \Delta)\) are graded.

Say that a graded module \( V \) is *connected* if \( V_0 \cong k \).

**Example 1.23.** Let \( k \) be a field. A path-connected space \( X \) has its homology and cohomology
\[
H_i(X; k) = \bigoplus_{i \geq 0} H_i(X; k)
\]
\[
H^*(X; k) = \bigoplus_{i \geq 0} H^i(X; k)
\]
carrying the structure of graded connected coalgebras and algebras, respectively. If in addition, \( X \) is a topological group, or even less strongly, a homotopy-associative \( H \)-space (e.g. the loop space \( \Omega Y \) on some other space \( Y \)), the continuous multiplication map \( X \times X \to X \) induces an algebra structure on \( H_i(X; k) \) and a coalgebra structure on \( H^*(X; k) \), so that each become bialgebras in the topologist’s sense (i.e., with the twist as in (1.7)), and these bialgebras are dual to each other in a sense soon to be discussed. This was Hopf’s motivation: the (co-)homology of a compact Lie group carries bialgebra structure that explains why it takes a certain form; see Cartier [29, §2].

\(^{12}\)since \( J \otimes C + C \otimes J \) is contained in the kernel of the canonical map \( C \otimes C \to (C/J) \otimes (C/J) \).
Example 1.24. Tensor algebras $T(V)$ and symmetric algebras $\text{Sym}(V)$ are graded, once one picks a graded module structure for $V$; then
\[
\deg(x_{i_1} \cdots x_{i_k}) = \deg(x_{i_1}) + \cdots + \deg(x_{i_k})
\]
if $\{x_i\}_{i \in J}$ is a graded basis (and, more generally, $v_1 v_2 \cdots v_k$ is homogeneous of degree $i_1 + i_2 + \cdots + i_k$ if each $v_j$ is a homogeneous element of $V$ of degree $i_j$). Assuming that $V_0 = 0$, the algebras $T(V)$ and $\text{Sym}(V)$ are connected. For example, we will often say that all elements of $V$ are homogeneous of degree 1, but at other times, it will make sense to have $V$ live in different (positive) degrees.

Exercise 1.25. Let $A = \bigoplus_{n \geq 0} A_n$ be a graded $\mathbb{k}$-bialgebra. We denote by $p$ the set of all primitive elements of $A$.
(a) Show that $p$ is a graded $\mathbb{k}$-submodule of $A$ (that is, we have $p = \bigoplus_{n \geq 0} (p \cap A_n)$).
(b) Show that $p$ is a two-sided coideal of $A$.

Exercise 1.26. Check that for a graded connected $\mathbb{k}$-bialgebra $A$, we have that
(a) the $\mathbb{k}$-submodule $\mathbb{k} = \mathbb{k} \cdot 1_A$ of $A$ lies in $A_0$,
(b) $u$ is an isomorphism $\mathbb{k} \xrightarrow{u} A_0$, while
(c) the two-sided ideal $\ker \epsilon$ is the $\mathbb{k}$-module of positive degree elements $I = \bigoplus_{n > 0} A_n$.
(d) $\epsilon$ restricted to $A_0$ is the inverse isomorphism $A_0 \xrightarrow{\epsilon} \mathbb{k}$ to $u$.
(e) for every $x \in A$, we have
\[
\Delta(x) \in x \otimes 1 + A \otimes I;
\]
(f) every $x$ in $I$ has comultiplication of the form
\[
\Delta(x) = 1 \otimes x + x \otimes 1 + \Delta_+(x)
\]
where $\Delta_+(x)$ lies in $I \otimes I$.
(Use only the gradedness of the unit $u$ and counit $\epsilon$ maps, along with commutativity of diagrams (1.4), and (1.8) and the connectedness of $A$.)

1.4. Antipodes and Hopf algebras. There is one more piece of structure needed to make a bialgebra a Hopf algebra, although it will come for free in the graded connected case.

Definition 1.27. For any coalgebra $C$ and algebra $A$, one can endow the $\mathbb{k}$-linear maps $\text{Hom}(C, A)$ with an associative algebra structure called the convolution algebra: send $f, g$ in $\text{Hom}(C, A)$ to $f \ast g$ defined by $(f \ast g)(c) = \sum f(c_1)g(c_2)$, using the Sweedler notation $\Delta(c) = \sum c_1 \otimes c_2$. Equivalently, $f \ast g$ is the composite
\[
\begin{array}{c}
C \\
\xrightarrow{\Delta} \\
C \otimes C \\
\xrightarrow{f \otimes g} \\
A \otimes A \\
\xrightarrow{m} \\
A
\end{array}
\]
The associativity of this multiplication $\ast$ is easy to check (see Exercise 1.28 below).

The map $u \circ \epsilon$ is a two-sided identity element for $\ast$, meaning that every $f \in \text{Hom}(C, A)$ satisfies
\[
\sum f(c_1)\epsilon(c_2) = f(c) = \sum \epsilon(c_1)f(c_2)
\]
for all $c \in C$. One sees this by adding a top row to (1.4):

\[
\begin{array}{c}
A \otimes \mathbb{k} \\
\xrightarrow{f \otimes \text{id}} \\
A \\
\xleftarrow{\epsilon \otimes \text{id}} \\
\mathbb{k} \otimes A
\end{array}
\]
\[
\begin{array}{c}
C \otimes \mathbb{k} \\
\xrightarrow{\epsilon \otimes \text{id}} \\
C \\
\xrightarrow{\text{id} \otimes f} \\
\mathbb{k} \otimes C
\end{array}
\]
\[
\begin{array}{c}
C \otimes C \\
\xleftarrow{\Delta} \\
C \\
\xrightarrow{\Delta} \\
C \otimes C
\end{array}
\]

In particular, when one has a bialgebra $A$, the convolution product $\ast$ gives an associative algebra structure on $\text{End}(A) := \text{Hom}(A, A)$.

Exercise 1.28. Let $C$ be a $\mathbb{k}$-coalgebra and $A$ be a $\mathbb{k}$-algebra. Show that the binary operation $\ast$ on $\text{Hom}(C, A)$ is associative.
Exercise 1.29. Let $C$ and $D$ be two $k$-coalgebras, and let $A$ and $B$ be two $k$-algebras. Prove that:

(a) If $f : C \to A$, $f' : C \to A$, $g : D \to B$ and $g' : D \to B$ are four $k$-linear maps, then
\[
(f \otimes g) \ast (f' \otimes g') = (f \ast f') \otimes (g \ast g')
\]
in the convolution algebra $\text{Hom}(C \otimes D, A \otimes B)$.

(b) Let $R$ be the $k$-linear map $(\text{Hom}(C, A), \ast) \otimes (\text{Hom}(D, B), \ast) \to (\text{Hom}(C \otimes D, A \otimes B), \ast)$ which sends every tensor $f \otimes g \in (\text{Hom}(C, A), \ast) \otimes (\text{Hom}(D, B), \ast)$ to the map $f \otimes g : C \otimes D \to A \otimes B$. (Notice that the tensor $f \otimes g$ and the map $f \otimes g$ are different things which happen to be written in the same way.) Then, $R$ is a $k$-algebra homomorphism.

Definition 1.30. A bialgebra $A$ is called a Hopf algebra if there is an element $S$ (called an antipode for $A$) in $\text{End}(A)$ which is a 2-sided inverse under $\ast$ for the identity map $\text{id}_A$. In other words, this diagram commutes:

\[
\begin{array}{ccc}
A \otimes A & \xrightarrow{S \otimes \text{id}_A} & A \otimes A \\
\Delta & & m \\
A & \xrightarrow{\varepsilon} & k & \xrightarrow{u} & A \\
\Delta & & m \\
A \otimes A & \xrightarrow{\text{id}_A \otimes S} & A \otimes A
\end{array}
\]

Or equivalently, if $\Delta(a) = \sum a_1 \otimes a_2$, then
\[
\sum_{(a)} S(a_1) a_2 = u(\varepsilon(a)) = \sum_{(a)} a_1 S(a_2).
\]

Example 1.31. For a group algebra $kG$, one can define an antipode $k$-linearly via $S(t_g) = t_{g^{-1}}$. The top pentagon in the above diagram commutes because
\[
(S \ast \text{id})(t_g) = m((S \otimes \text{id})(t_g \otimes t_g)) = S(t_g) t_g = t_{g^{-1}} t_g = \varepsilon = (u \circ \varepsilon)(t_g).
\]

Note that when it exists, the antipode $S$ is unique, as with all 2-sided inverses in associative algebras: if $S, S'$ are both 2-sided $\ast$-inverses to $\text{id}_A$ then
\[
S' = (u \circ \varepsilon) \ast S' = (S \ast \text{id}_A) \ast S' = S \ast (\text{id}_A \ast S') = S \ast (u \circ \varepsilon) = S.
\]

On the other hand, the next property is not quite as obvious, but is useful when one wants to check that a certain map is the antipode in a particular Hopf algebra, by checking it on an algebra generating set.

Proposition 1.32. The antipode $S$ in a Hopf algebra $A$ is an algebra anti-endomorphism: $S(1) = 1$, and $S(ab) = S(b)S(a)$ for all $a, b$ in $A$.

Proof. (see [163, Chap. 4]) Since $\Delta$ is an algebra map, one has $\Delta(1) = 1 \otimes 1$, and therefore $1 = u(\varepsilon(1)) = S(1) \cdot 1 = S(1)$.

To show $S(ab) = S(b)S(a)$, consider $A \otimes A$ as a coalgebra and $A$ as an algebra. Then $\text{Hom}(A \otimes A, A)$ is an associative algebra with a convolution product $\otimes$ (to be distinguished from the convolution $\ast$ on $\text{Hom}(A)$), having two-sided identity element $u_{A \otimes A}$. We will show below that these three elements of $\text{Hom}(A \otimes A, A)$
\[
f(a \otimes b) = ab \\
g(a \otimes b) = S(b)S(a) \\
h(a \otimes b) = S(ab)
\]

have the property that
\[
h \otimes f = u_{A \otimes A} \ast \varepsilon_{A \otimes A} = f \otimes g
\]
which would then show the desired equality $h = g$ via associativity:
\[
h = h \otimes (u_{A \otimes A} \ast \varepsilon_{A \otimes A}) = h \otimes (f \otimes g) = (h \otimes f) \otimes g = (u_{A \otimes A} \ast \varepsilon_{A \otimes A}) \otimes g = g.
\]
So we evaluate the three elements in (1.14) on $a \otimes b$, assuming $\Delta(a) = \sum_{(a)} a_1 \otimes a_2$ and $\Delta(b) = \sum_{(b)} b_1 \otimes b_2$, and hence $\Delta(ab) = \sum_{(ab)} a_1 b_1 \otimes a_2 b_2$. One has
\[
(u_A \epsilon_A \otimes A)(a \otimes b) = u_A(\epsilon_A(a) \epsilon_A(b)) = u_A(\epsilon_A(ab)).
\]

\[
(h \otimes f)(a \otimes b) = \sum_{(a), (b)} h(a_1 \otimes b_1) f(a_2 \otimes b_2)
= \sum_{(a), (b)} S(a_1 b_1) a_2 b_2
= (S \star \text{id}_A)(ab) = u_A(\epsilon_A(ab)).
\]

\[
(f \otimes g)(a \otimes b) = \sum_{(a), (b)} f(a_1 \otimes b_1) g(a_2 \otimes b_2)
= \sum_{(a), (b)} a_1 b_1 S(b_2) S(a_2)
= \sum_{(a)} a_1 \cdot (\text{id}_A \ast S)(b) \cdot S(a_2)
= u_A(\epsilon_A(b)) \sum_{(a)} a_1 S(a_2) = u_A(\epsilon_A(b)) u_A(\epsilon_A(a)) = u_A(\epsilon_A(ab))
\]

Remark 1.33. Recall from Remark 1.16 that the comultiplication on a bialgebra $A$ allows one to define an $A$-module structure on the tensor product $M \otimes N$ of two $A$-modules $M, N$. Similarly, the anti-endomorphism $S$ in a Hopf algebra allows one to turn left $A$-modules into right $A$-modules, or vice-versa.\textsuperscript{13} E.g., left $A$-modules $M$ naturally have a right $A$-module structure on the dual $k$-module $M^* := \text{Hom}(M, k)$, defined via $(fa)(m) := f(am)$ for $f$ in $M^*$ and $a$ in $A$. The antipode $S$ can be used to turn this back into a left $A$-module $M^*$, via $(af)(m) = f(S(a)m)$.

For groups $G$ and left $kG$-modules (group representations) $M$, this is how one defines the contragredient action of $G$ on $M^*$, namely $t_g$ acts as $(t_g f)(m) = f(t_{g^{-1}} m)$.

More generally, if $A$ is a Hopf algebra and $M$ and $N$ are two left $A$-modules, then $\text{Hom}(M, N)$ (the Hom here means $\text{Hom}_A$, not $\text{Hom}_k$) canonically becomes a left $A$-module by setting
\[
(af)(m) = \sum_{(a)} a_1 f(S(a_2) m) \quad \text{for all } a \in A, \ f \in \text{Hom}(M, N) \text{ and } m \in M.
\]

\textsuperscript{14} When $A$ is the group algebra $kG$ of a group $G$, this leads to
\[
(t_g f)(m) = t_g f(t_{g^{-1}} m) \quad \text{for all } g \in G, \ f \in \text{Hom}(M, N) \text{ and } m \in M.
\]

This is precisely how one commonly makes $\text{Hom}(M, N)$ a representation of $G$ for two representations $M$ and $N$.

\textsuperscript{13}Be warned that these two transformations are not mutually inverse! Turning a left $A$-module into a right one and then again into a left one using the antipode might lead to a non-isomorphic $A$-module, unless the antipode $S$ satisfies $S^2 = \text{id}$.

\textsuperscript{14}In more abstract terms, this $A$-module structure is given by the composition
\[
\xymatrix{ A \ar[r]^-{\Delta} & A \otimes A \ar[r]^-{\text{id}_A \otimes S} & A \otimes A^\text{op} \ar[r] & \text{End}(\text{Hom}(M, N)) },
\]
where the last arrow is the morphism
\[
A \otimes A^\text{op} \rightarrow \text{End}(\text{Hom}(M, N)),
\]
\[
a \otimes b \mapsto (f \mapsto (M \rightarrow N, m \mapsto af(bm))).
\]

Here, $A^\text{op}$ denotes the opposite algebra of $A$, which is the $k$-algebra differing from $A$ only in the multiplication being twisted (the product of $a$ and $b$ in $A^\text{op}$ is defined to be the product of $b$ and $a$ in $A$). As $k$-modules, $A^\text{op} = A$, but we prefer to use $A^\text{op}$ instead of $A$ here to ensure that all morphisms in the above composition are algebra maps.
Along the same lines, whenever \( A \) is a \( k \)-bialgebra, we are supposed to think of the counit \( A \xrightarrow{\epsilon} k \) as giving a way to make \( k \) into a trivial \( A \)-module. This \( A \)-module \( k \) behaves as one would expect: the canonical isomorphisms \( k \otimes M \to M, M \otimes k \to M \) and (if \( A \) is a Hopf algebra) \( \text{Hom}(M,k) \to M^* \) are \( A \)-module isomorphisms for any \( A \)-module \( M \).

**Corollary 1.34.** Commutativity of \( A \) implies the antipode is an involution: \( S^2 = \text{id}_A \).

**Proof.** One checks that \( S^2 = S \circ S \) is a right \( \star \)-inverse to \( S \), and hence \( S^2 = \text{id}_A \):

\[
(S \star S^2)(a) = \sum_{(a)} S(a_1)S^2(a_2)
= S \left( \sum_{(a)} S(a_2)a_1 \right) \text{ by Proposition 1.32}
= S(u(\epsilon(a)))
= u(\epsilon(a)) \text{ since } S(1) = 1 \text{ by Proposition 1.32}.
\]

\( \square \)

**Remark 1.35.** We won’t need it, but it is easy to adapt the above proof to show that \( S^2 = \text{id}_A \) also holds for cocommutative Hopf algebras; see [120, Corollary 1.5.12] or [163, Chapter 4]. For a general Hopf algebra which is not finite-dimensional over a field \( k \), the antipode \( S \) may not even have finite order, even in the graded connected setting. E.g., Aguiar and Sottile [7] show that the Malvenuto-Reutenauer Hopf algebra of permutations has antipode of infinite order. In general, antipodes need not even be invertible [164].

In our frequent setting of graded connected bialgebras, antipodes come for free.

**Proposition 1.36.** A graded connected bialgebra \( A \) has a unique antipode \( S \), which is a graded map \( \xrightarrow{S} A \), endowing it with a Hopf structure.

**Proof.** Let us try to define a \((k\text{-linear})\) left \( \star \)-inverse \( S \) to \( \text{id}_A \) on each homogeneous component \( A_n \), via induction on \( n \).

In the base case \( n = 0 \), Proposition 1.32 and its proof show that one must define \( S(1) = 1 \) so \( S \) is the identity on \( A_0 = k \).

In the inductive step, recall from Exercise 1.26 that a homogeneous element \( a \) of degree \( n > 0 \) has \( \Delta(a) = a \otimes 1 + \sum a_1 \otimes a_2 \), with each \( \deg(a_1) < n \). Hence in order to have \( S \star \text{id}_A = \epsilon \), one must define \( S(a) \) in such a way that \( S(a) \cdot 1 + \sum S(a_1)a_2 = \epsilon(a) = 0 \) and hence \( S(a) := -\sum S(a_1)a_2 \), where \( S(a_1) \) have already been uniquely defined by induction. This does indeed define such a left \( \star \)-inverse \( S \) to \( \text{id}_A \), by induction. It is also a graded map by induction.

The same argument shows how to define a right \( \star \)-inverse \( S' \) to \( \text{id}_A \). Then \( S = S' \) is a two-sided \( \star \)-inverse to \( \text{id}_A \) by the associativity of \( \star \).

Here is another consequence of the fact that \( S(1) = 1 \).

**Proposition 1.37.** In bialgebras, primitive elements \( x \) have \( \epsilon(x) = 0 \), and in Hopf algebras, they have \( S(x) = -x \).

**Proof.** In a bialgebra, \( \epsilon(1) = 1 \). Hence \( \Delta(x) = 1 \otimes x + x \otimes 1 \) implies via (1.4) that \( 1 \cdot \epsilon(x) + \epsilon(1)x = x \), so \( \epsilon(x) = 0 \). It also implies via (1.12) that \( S(x)1 + S(1)x = \epsilon(x) = u(0) = 0 \), so \( S(x) = -x \).

Thus whenever \( A \) is a Hopf algebra generated as an algebra by its primitive elements, \( S \) is the unique anti-endomorphism that negates all primitive elements.

**Example 1.38.** The tensor and symmetric algebras \( T(V) \) and \( \text{Sym}(V) \) are each generated by \( V \), which contains only primitive elements in either case. Hence one has in \( T(V) \) that

\[
(1.15) \quad S(x_{i_1}x_{i_2} \cdots x_{i_k}) = (-x_{i_k}) \cdots (-x_{i_2})(-x_{i_1}) = (-1)^k x_{i_k} \cdots x_{i_2} x_{i_1}
\]

for each word \( (i_1, \ldots, i_k) \) in the alphabet \( I \) if \( V \) is a free \( k \)-module with basis \( \{x_i\}_{i \in I} \). The same holds in \( \text{Sym}(V) \) for each multiset \( (i_1, \ldots, i_k) \), recalling that the monomials are now commutative. In other words, for a commutative polynomial \( f(x) \) in \( \text{Sym}(V) \), the antipode \( S \) sends \( f \) to \( f(-x) \), negating all the variables.
The antipode for a graded connected Hopf algebra has an interesting formula due to Takeuchi [164], reminiscent of P. Hall's formula for the Möbius function of a poset. For the sake of stating this, consider (for every \( k \in \mathbb{N} \)) the \( k \)-fold tensor power \( A^{\otimes k} = A \otimes \cdots \otimes A \) (defined in Example 1.2) and define iterated multiplication and comultiplication maps

\[
A^{\otimes k} \xrightarrow{m^{(k-1)}} A \\
A \xrightarrow{\Delta^{(k-1)}} A^{\otimes k}
\]

by induction over \( k \), setting \( m^{(-1)} = u \), \( \Delta^{(-1)} = \epsilon \), \( m^{(0)} = \Delta^{(0)} = \text{id}_A \), and

\[
m^{(k)} = m \circ (\text{id}_A \otimes m^{(k-1)}) \quad \text{for every } k \geq 1; \\
\Delta^{(k)} = (\text{id}_A \otimes \Delta^{(k-1)}) \circ \Delta \quad \text{for every } k \geq 1.
\]

Using associativity and coassociativity, one can see that for \( k \geq 1 \) these maps also satisfy

\[
m^{(k)}(a^{(1)} \otimes \cdots \otimes a^{(k+1)}) = a^{(1)} \cdots a^{(k+1)} \\
\Delta^{(k)}(b) = \sum b_1 \otimes \cdots \otimes b_{k+1} \text{ in Sweedler notation.}
\]

**Exercise 1.39.** Let \( A \) be a \( k \)-algebra. Let us define, for every \( k \in \mathbb{N} \), a \( k \)-linear map \( m^{(k)} : A^{\otimes (k+1)} \to A \). Namely, we define these maps by induction over \( k \), with the induction base \( m^{(0)} = \text{id}_A \), and with the induction step \( m^{(k)} = m \circ (\text{id}_A \otimes m^{(k-1)}) \) for \( k \geq 1 \). (This generalizes our definition of \( m^{(k)} \) for Hopf algebras \( A \) given above, except for \( m^{(-1)} \) which we have omitted.)

(a) Show that \( m^{(k)} = m \circ (m^{(i)} \otimes m^{(k-1-i)}) \) for every \( k \geq 0 \) and \( 0 \leq i \leq k-1 \).

(b) Show that \( m^{(k)} = m \circ (m^{(k-1)} \otimes \text{id}_A) \) for every \( k \geq 1 \).

(c) Show that \( m^{(k)} = m^{(k-1)} \circ (\text{id}_A^{\otimes i} \otimes m \otimes \text{id}_A^{\otimes (k-1-i)}) \) for every \( k \geq 0 \) and \( 0 \leq i \leq k-1 \).

(d) Show that \( m^{(k)} = m^{(k-1)} \circ (\text{id}_A^{\otimes (i-1)} \otimes m) = m^{(k-1)} \circ (m \otimes \text{id}_A^{\otimes (k-1)}) \) for every \( k \geq 1 \).

**Exercise 1.40.** Let \( C \) be a \( k \)-coalgebra. Let us define, for every \( k \in \mathbb{N} \), a \( k \)-linear map \( \Delta^{(k)} : C \to C^{\otimes (k+1)} \). Namely, we define these maps by induction over \( k \), with the induction base \( \Delta^{(0)} = \text{id}_C \), and with the induction step \( \Delta^{(k)} = (\text{id}_C \otimes \Delta^{(k-1)}) \circ \Delta \) for every \( k \geq 1 \). (This generalizes our definition of \( \Delta^{(k)} \) for Hopf algebras \( A \) given above, except for \( \Delta^{(-1)} \) which we have omitted.)

(a) Show that \( \Delta^{(k)} = (\Delta^{(i)} \otimes \Delta^{(k-1-i)}) \circ \Delta \) for every \( k \geq 0 \) and \( 0 \leq i \leq k-1 \).

(b) Show that \( \Delta^{(k)} = (\Delta^{(k-1)} \otimes \text{id}_C) \circ \Delta \) for every \( k \geq 1 \).

(c) Show that \( \Delta^{(k)} = (\text{id}_C^{\otimes i} \otimes \Delta \otimes \text{id}_C^{\otimes (k-1-i)}) \circ \Delta^{(k-1)} \) for every \( k \geq 0 \) and \( 0 \leq i \leq k-1 \).

(d) Show that \( \Delta^{(k)} = (\text{id}_C^{\otimes (i-1)} \otimes \Delta) \circ \Delta^{(k-1)} = (\Delta \otimes \text{id}_C^{\otimes (k-1)}) \circ \Delta^{(k-1)} \) for every \( k \geq 1 \).

**Remark 1.41.** Exercise 1.39 holds more generally for nonunital associative algebras \( A \) (that is, \( k \)-modules \( A \) equipped with a \( k \)-linear map \( m : A \otimes A \to A \) such that the diagram (1.1) is commutative, but not necessarily admitting a unit map \( u \)). Similarly, Exercise 1.40 holds for non-counital coassociative coalgebras \( C \). The existence of a unit in \( A \), respectively a counit in \( C \), allows slightly extending these two exercises by additionally introducing maps \( m^{(-1)} = u : k \to A \) and \( \Delta^{(-1)} = \epsilon : C \to k \); however, not much is gained from this extension.\(^{16}\)

**Exercise 1.42.** For every \( k \in \mathbb{N} \) and every \( k \)-bialgebra \( H \), consider the map \( \Delta^{(k)}_H : H \to H^{\otimes (k+1)} \) (this is the map \( \Delta^{(k)} \) defined as in Exercise 1.40 for \( C = H \)), and the map \( m^{(k)}_H : H^{\otimes (k+1)} \to H \) (this is the map \( m^{(k)} \) defined as in Exercise 1.39 for \( A = H \)).

\(^{15}\)In fact, for incidence Hopf algebras, Takeuchi's formula generalizes Hall's formula—see Corollary 7.12.

\(^{16}\)The identity \( m^{(k)} = m \circ (\text{id}_A \otimes m^{(k-1)}) \) for a \( k \)-algebra \( A \) still holds when \( k = 0 \) if it is interpreted in the right way (viz., if \( A \) is identified with \( A \otimes k \) using the canonical homomorphism).
Let \( H \) be a \( k \)-bialgebra. Let \( k \in \mathbb{N} \). Show that:

(a) The map \( m_{H}^{(k)} : H \otimes (k+1) \to H \) is a \( k \)-coalgebra homomorphism.

(b) The map \( \Delta_{H}^{(k)} : H \to H \otimes (k+1) \) is a \( k \)-algebra homomorphism.

(c) We have \( m_{H}^{(\ell)} \otimes (k+1) \circ (\Delta_{H}^{(k)}) \otimes (\ell+1) = \Delta_{H}^{(k)} \circ m_{H}^{(\ell)} \) for every \( \ell \in \mathbb{N} \).

(d) We have \( (m_{H}^{(\ell)}) \otimes (k+1) \circ \Delta_{H}^{(k)} \circ m_{H}^{(\ell)} \) for every \( \ell \in \mathbb{N} \).

The iterated multiplication and comultiplication maps allow explicitly computing the convolution of multiple maps; the following formula will often be used without explicit mention:

**Exercise 1.43.** Let \( C \) be a \( k \)-coalgebra, and \( A \) be a \( k \)-algebra. Let \( k \in \mathbb{N} \). Let \( f_1, f_2, \ldots, f_k \) be \( k \) elements of Hom \((C, A)\). Show that

\[
f_1 \ast f_2 \ast \cdots \ast f_k = m_{A}^{(k-1)} \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_k) \circ \Delta_{C}^{(k-1)}.
\]

We are now ready to state Takeuchi’s formula for the antipode:

**Proposition 1.44.** In a graded connected Hopf algebra \( A \), the antipode has formula

\[
S = \sum_{k \geq 0} (-1)^k m_{A}^{(k-1)} f^{\otimes k} \Delta^{(k-1)}
\]

(1.16)

\[
= u_{\epsilon} - f + m \circ f^{\otimes 2} \circ \Delta - m^{(2)} \circ f^{\otimes 3} \circ \Delta^{(2)} + \cdots
\]

where \( f := \text{id}_A - u_{\epsilon} \) in End\((A)\).

**Proof.** We argue as in [164, proof of Lemma 14] or [7, §5]. For any \( f \) in End\((A)\) one has this explicit formula for its \( k \)-fold convolution power \( f^{*k} := f \ast \cdots \ast f \) in terms of its tensor powers \( f^{\otimes k} := f \otimes \cdots \otimes f \) (according to Exercise 1.43):

\[
f^{*k} = m_{A}^{(k-1)} \circ f^{\otimes k} \circ \Delta^{(k-1)}.
\]

Therefore any \( f \) annihilating \( A_0 \) will be locally \(*\)-nilpotent on \( A \), meaning that for each \( n \) one has that \( A_n \) is annihilated by \( f^{*m} \) for every \( m > n \): homogeneity forces that for \( a \in A_n \), every summand of \( \Delta^{(m-1)}(a) \) must contain among its \( m \) tensor factors at least one factor lying in \( A_0 \), so each summand is annihilated by \( f^{\otimes m} \), and \( f^{*m}(a) = 0 \).

In particular such \( f \) have the property that \( u_{\epsilon} + f \) has as two-sided \(*\)-inverse

\[
(u_{\epsilon} + f)^{*(-1)} = u_{\epsilon} - f + f \ast f - f \ast f \ast f + \cdots
\]

\[
= \sum_{k \geq 0} (-1)^k f^{*k} = \sum_{k \geq 0} (-1)^k m_{A}^{(k-1)} \circ f^{\otimes k} \circ \Delta^{(k-1)}.
\]

The proposition follows upon taking \( f := \text{id}_A - u_{\epsilon} \), which annihilates \( A_0 \). \( \square \)

**Remark 1.45.** In fact, one can see that Takeuchi’s formula applies more generally to define an antipode \( A \overset{S}{\rightarrow} A \) in any (not necessarily graded) bialgebra \( A \) where the map \( \text{id}_A - u_{\epsilon} \) is locally \(*\)-nilpotent.

It is also worth noting that the proof of Proposition 1.44 gives an alternate proof of Proposition 1.36.

To finish our discussion of antipodes, we mention some properties (taken from [163, Chap. 4]) relating antipodes to convolutional inverses. It also shows that a bialgebra morphism between Hopf algebras automatically respects the antipodes.

**Proposition 1.46.** Let \( H \) be a Hopf algebra with antipode \( S \).

(a) For any algebra \( A \) and algebra morphism \( H \overset{\alpha}{\rightarrow} A \), one has \( \alpha \circ S = \alpha^{*^{-1}} \), the convolutional inverse to \( \alpha \) in Hom\((H, A)\).

(b) For any coalgebra \( C \) and coalgebra morphism \( C \overset{\gamma}{\rightarrow} H \), one has \( S \circ \gamma = \gamma^{*^{-1}} \), the convolutional inverse to \( \gamma \) in Hom\((C, H)\).

(c) If \( H_1, H_2 \) are Hopf algebras with antipodes \( S_1, S_2 \), then any bialgebra morphism \( H_1 \overset{\beta}{\rightarrow} H_2 \) is a Hopf morphism, that is, it commutes with the antipodes, since \( \beta \circ S_1 = \beta^{*^{-1}} = S_2 \circ \beta \).

\( \text{The following statements are taken from [127]; specifically, part (c) is [127, Lem. 1.8].} \)
Proof. We prove (a); the proof of (b) is similar, and (c) follows immediately from (a),(b) as indicated in its statement. Begin with the following: given an algebra morphism \( \alpha : A \to A' \) and coalgebra morphism \( \phi : C \to C' \), the pre- and post-composition \( f \mapsto \alpha \circ f \circ \phi \) is a convolution algebra morphism \( \text{Hom}(C',A) \to \text{Hom}(C,A') \), as one can check that \( \varphi(u_A \circ \epsilon_{C'}) = \alpha \circ u_{A'} \circ \epsilon_{C'} \circ \gamma = u_{A'} \circ \epsilon_C \), and that every \( f \in \text{Hom}(C',A) \) and \( g \in \text{Hom}(C',A) \) satisfy
\[
\varphi(f \star g) = \alpha \circ (f \star g) \circ \gamma \\
= \alpha \circ m \circ (f \otimes g) \circ \Delta \circ \gamma \\
= m \circ (\alpha \otimes \alpha) \circ (f \otimes g) \circ (\gamma \otimes \gamma) \circ \Delta \\
= m \circ ((\alpha \circ f \circ \gamma) \otimes (\alpha \circ g \circ \gamma)) \circ \Delta \\
= \varphi(f) \star \varphi(g).
\]
(1.17)

For assertion (a), note that a special case of the above observation shows \( f \mapsto \alpha \circ f \) gives a convolution-algebra morphism \( \text{Hom}(H,H) \to \text{Hom}(H,A) \), and hence
\[
\alpha \circ S = \varphi(S) = (\varphi(id_H)^{-1}) = (\varphi(id_H)^*)^{-1} = (\alpha \circ id_H)^*^{-1} = \alpha^{-1}.
\]

\( \square \)

Exercise 1.47. Prove that the antipode \( S \) of a Hopf algebra \( A \) is a coalgebra anti-endomorphism, i.e., that it satisfies \( \epsilon \circ S = \epsilon \) and \( \Delta \circ S = T \circ (S \otimes S) \circ \Delta \), where \( T : A \otimes A \to A \otimes A \) is the twist map sending every \( a \otimes b \) to \( b \otimes a \).

Exercise 1.48. If \( C \) is a \( k \)-coalgebra and if \( A \) is a \( k \)-algebra, then a \( k \)-linear map \( f : C \to A \) is said to be \(*\)-invertible if it is invertible as an element of the \( k \)-algebra \( \text{Hom}(C,A) \), \( * \). In this case, the multiplicative inverse \( f^*(-1) \) of \( f \) in \( \text{Hom}(C,A) \), \( * \) is called the \( * \)-inverse of \( f \). For any two \( k \)-modules \( U \) and \( V \), let \( T_{U,V} : U \otimes V \to V \otimes U \) be the twist map (i.e., the \( k \)-linear map \( U \otimes V \to V \otimes U \) sending every \( u \otimes v \) to \( v \otimes u \)).

(a) If \( C \) is a \( k \)-bialgebra, if \( A \) is a \( k \)-algebra, and if \( r : C \to A \) is a \(*\)-invertible \( k \)-algebra homomorphism, then prove that the \(*\)-inverse \( r^*(-1) \) of \( r \) is a \( k \)-algebra anti-homomorphism\(^{18} \).

(b) If \( C \) is a \( k \)-bialgebra, if \( A \) is a \( k \)-coalgebra, and if \( r : A \to C \) is a \(*\)-invertible \( k \)-coalgebra homomorphism, then prove that the \(*\)-inverse \( r^*(-1) \) of \( r \) is a \( k \)-coalgebra anti-homomorphism\(^{19} \).

(c) Derive Proposition 1.32 from Exercise 1.48(a), and derive Exercise 1.47 from Exercise 1.48(b).

(d) Prove Corollary 1.34 again using Proposition 1.46.

(e) Prove that the antipode \( S \) of a cocommutative Hopf algebra \( A \) satisfies \( S^2 = \text{id}_A \). (This was a statement made in Remark 1.35.)

Exercise 1.49.  \( a \) Let \( A \) be a Hopf algebra. If \( P : A \to A \) is a \( k \)-linear map such that every \( a \in A \) satisfies
\[
\sum_{(a)} P(a_2) \cdot a_1 = u(\epsilon(a)),
\]
then prove that the antipode \( S \) of \( A \) is invertible and its inverse is \( P \).

(b) Let \( A \) be a Hopf algebra. If \( P : A \to A \) is a \( k \)-linear map such that every \( a \in A \) satisfies
\[
\sum_{(a)} a_2 \cdot P(a_1) = u(\epsilon(a)),
\]
then prove that the antipode \( S \) of \( A \) is invertible and its inverse is \( P \).

(c) Show that the antipode of a graded connected Hopf algebra is invertible.

(Compare this exercise to [120, Lemma 1.5.11].)

\(^{18}\)A \( k \)-algebra anti-homomorphism means a \( k \)-linear map \( \varphi : P \to Q \) between two \( k \)-algebras \( P \) and \( Q \) which satisfies
\[
\varphi \circ m_P = m_Q \circ (\varphi \otimes \varphi) \circ T_{P,P} \quad \text{and} \quad \varphi \circ u_P = u_Q.
\]

\(^{19}\)A \( k \)-coalgebra anti-homomorphism means a \( k \)-linear map \( \varphi : P \to Q \) between two \( k \)-coalgebras \( P \) and \( Q \) which satisfies
\[
\Delta_Q \circ \varphi = T_{Q,Q} \circ (\varphi \otimes \varphi) \circ \Delta_P \quad \text{and} \quad \epsilon_Q \circ \varphi = \epsilon_P.
\]
Definition 1.50. Let $C$ be a $k$-coalgebra. A subcoalgebra of $C$ means a $k$-coalgebra $D$ such that $D \subset C$ and such that the canonical inclusion map $D \rightarrow C$ is a $k$-coalgebra homomorphism\textsuperscript{20}. When $k$ is a field, we can equivalently define a subcoalgebra of $C$ as a $k$-submodule $D$ of $C$ such that $\Delta_C (D)$ is a subset of the $k$-submodule $D \otimes D$ of $C \otimes C$; however, this might no longer be equivalent when $k$ is not a field\textsuperscript{21}.

Similarly, a subbialgebra of a bialgebra $C$ is a $k$-bialgebra $D$ such that $D \subset C$ and such that the canonical inclusion map $D \rightarrow C$ is a $k$-bialgebra homomorphism. Also, a Hopf subalgebra of a Hopf algebra $C$ is a $k$-Hopf algebra $D$ such that $D \subset C$ and such that the canonical inclusion map $D \rightarrow C$ is a $k$-Hopf algebra homomorphism.\textsuperscript{22}

Exercise 1.51. Let $C$ be a $k$-coalgebra. Let $D$ be a $k$-submodule of $C$ such that $D$ is a direct summand of $C$ as a $k$-module (i.e., there exists a $k$-submodule $E$ of $C$ such that $C = D \oplus E$). (This is automatically satisfied if $k$ is a field.) Assume that $\Delta (D) \subset C \otimes D$ and $\Delta (D) \subset D \otimes C$. (Here, we are abusing the notation $C \otimes D$ to denote the $k$-submodule of $C \otimes C$ spanned by tensors of the form $c \otimes d$ with $c \in C$ and $d \in D$; similarly, $D \otimes C$ should be understood.) Show that there is a canonically defined $k$-coalgebra structure on $D$ which makes $D$ a subcoalgebra of $C$.

The next exercise is implicit in [4, §6.7]; other authors may use other definitions.

Exercise 1.52. Let $k$ be a field. Let $C$ be a $k$-coalgebra, and let $U$ be any $k$-module. Let $f : C \rightarrow U$ be a $k$-linear map. Recall the map $\Delta^{(2)} : C \rightarrow C^\otimes 3$ from Exercise 1.40. Let $K = \ker ((\text{id}_C \otimes f \otimes \text{id}_C) \circ \Delta^{(2)})$.

(a) Show that $K$ is a $k$-subcoalgebra of $C$.
(b) Show that every $k$-subcoalgebra of $C$ which is a subset of $\ker f$ must be a subset of $K$.

Exercise 1.53. (a) Let $C = \bigoplus_{n \geq 0} C_n$ be a graded $k$-coalgebra, and $A$ be any $k$-algebra. Notice that $C_0$ itself is a $k$-subcoalgebra of $C$. Let $h : C \rightarrow A$ be a $k$-linear map such that the restriction $h \mid_{C_0}$ is a $\ast$-invertible map in Hom ($C_0, A$). Prove that $h$ is a $\ast$-invertible map in Hom ($C, A$). (This is a weaker version of Takeuchi’s [164, Lemma 14].)
(b) Let $A = \bigoplus_{n \geq 0} A_n$ be a graded $k$-bialgebra. Notice that $A_0$ is a sub-bialgebra of $A$. Assume that $A_0$ is a Hopf algebra. Show that $A$ is a Hopf algebra.
(c) Obtain yet another proof of Proposition 1.36.

Exercise 1.54. Let $A = \bigoplus_{n \geq 0} A_n$ be a graded connected $k$-bialgebra. Let $p$ be the $k$-submodule of $A$ consisting of the primitive elements of $A$.

(a) If $I$ is a two-sided coideal of $A$ such that $I \cap p = 0$ and such that $I = \bigoplus_{n \geq 0} (I \cap A_n)$, then prove that $I = 0$.
(b) Let $f : A \rightarrow C$ be a graded surjective coalgebra homomorphism from $A$ to a graded $k$-coalgebra $C$. If $f \mid_p$ is injective, then prove that $f$ is injective.
(c) Assume that $k$ is a field. Show that the claim of Exercise 1.54(b) is valid even without requiring $f$ to be surjective.

Remark 1.55. Exercise 1.54 (b) and (c) are often used in order to prove that certain coalgebra homomorphisms are injective.

The word “bialgebra” can be replaced by “coalgebra” in Exercise 1.54, provided that the notion of a graded connected coalgebra is defined correctly (namely, as a graded coalgebra such that the restriction of $\epsilon$ to the 0-th graded component is an isomorphism), and the notion of the element 1 of a graded connected coalgebra is defined accordingly (namely, as the preimage of 1 in $k$ under the restriction of $\epsilon$ to the 0-th graded component).

1.5. Commutativity, cocommutativity.

\textsuperscript{20}In this definition, we follow [123, p. 55] and [171, §6.7]; other authors may use other definitions.

\textsuperscript{21}This is because the $k$-submodule $D \otimes D$ of $C \otimes C$ is generally not isomorphic to the $k$-module $D \otimes D$. See [123, p. 56] for specific counterexamples for the non-equivalence of the two notions of a subcoalgebra. Notice that the equivalence is salvaged if $D$ is a direct summand of $C$ as a $k$-module (see Exercise 1.51 for this).

\textsuperscript{22}By Proposition 1.46(c), we can also define it as a subbialgebra of $C$ which happens to be a Hopf algebra.
**Definition 1.56.** Say that the $k$-algebra $A$ is *commutative* if $ab = ba$, that is, this diagram commutes:

\begin{equation}
\begin{array}{ccc}
A \otimes A & \xrightarrow{T} & A \otimes A \\
\downarrow{m} & & \downarrow{m} \\
A & & A
\end{array}
\end{equation}

Say that the $k$-coalgebra $C$ is *cocommutative* if this diagram commutes:

\begin{equation}
\begin{array}{ccc}
C \otimes C & \xrightarrow{T} & C \otimes C \\
\downarrow{\Delta} & & \downarrow{\Delta} \\
C & & C
\end{array}
\end{equation}

**Example 1.57.** Group algebras $kG$ are always cocommutative, but commutative if and only if $G$ is abelian.

Tensor algebras $T(V)$ are always cocommutative, but not generally commutative\(^2\). Symmetric algebras $\text{Sym}(V)$ are always cocommutative and commutative.

Homology and cohomology of $H$-spaces are always cocommutative and commutative in the topologist’s sense where one reinterprets that twist map $A \otimes A \xrightarrow{T} A \otimes A$ to have the extra sign as in (1.7).

Note how the cocommutative Hopf algebras $T(V), \text{Sym}(V)$ have much of their structure controlled by their $k$-submodules $V$, which consist of primitive elements only (although, in general, not all of their primitive elements). This is not far from the truth in general, and closely related to Lie algebras.

**Exercise 1.58.** Recall that a *Lie algebra* over $k$ is a $k$-module $\mathfrak{g}$ with a $k$-bilinear map $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ that satisfies $[x, x] = 0$ for $x$ in $\mathfrak{g}$, and the *Jacobi identity*

\[
[x, [y, z]] = [[x, y], z] + [y, [x, z]], \text{ or equivalently }
[x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0.
\]

This $k$-bilinear map $[\cdot, \cdot]$ is called the *Lie bracket* of $\mathfrak{g}$.

(a) Check that any associative algebra $A$ gives rise to a Lie algebra by means of the commutator operation $[a, b] := ab - ba$.

(b) If $A$ is also a bialgebra, show that the $k$-submodule of primitive elements $p \subset A$ is closed under the Lie bracket, that is, $[p, p] \subset p$, and hence forms a Lie subalgebra.

Conversely, given a Lie algebra $p$, one constructs the *universal enveloping algebra* $U(p) := T(p)/J$ as the quotient of the tensor algebra $T(p)$ by the two-sided ideal $J$ generated by all elements $xy - yx - [x, y]$ for $x, y$ in $p$.

(c) Show that $J$ is also a two-sided coideal in $T(p)$ for its usual coalgebra structure, and hence the quotient $U(p)$ inherits the structure of a (cocommutative) Hopf algebra.

(d) Show that the antipode $S$ on $T(p)$ preserves $J$, meaning that $S(J) \subset J$, and hence $U(p)$ inherits the structure of a (cocommutative) Hopf algebra.

There are theorems, discussed in [29, §3.8], [120, Chap. 5], [45, §3.2] giving various mild hypotheses in addition to cocommutativity which imply that the inclusion of the $k$-module $p$ of primitives in a Hopf algebra $A$ extends to a Hopf isomorphism $U(p) \cong A$.

**Exercise 1.59.**

(a) Let $C$ be a $k$-coalgebra. Show that $C$ is cocommutative if and only if its comultiplication $\Delta_C : C \to C \otimes C$ is a $k$-coalgebra homomorphism.

(b) Let $A$ be a $k$-algebra. Show that $A$ is commutative if and only if its multiplication $m_A : A \otimes A \to A$ is a $k$-algebra homomorphism.

**Remark 1.60.** If $C$ is a $k$-coalgebra, then $\epsilon_C : C \to k$ is always a $k$-coalgebra homomorphism. Similarly, $u_A : k \to A$ is a $k$-algebra homomorphism whenever $A$ is a $k$-algebra.

\(^2\)If $k$ is a field, then $T(V)$ is commutative if and only if $\dim_k V \leq 1$.\]
Exercise 1.61. Let $A$ be a commutative $k$-algebra, and let $k \in \mathbb{N}$. The symmetric group $\mathfrak{S}_k$ acts on the $k$-fold tensor power $A^\otimes k$ by permuting the tensor factors: $\sigma (v_1 \otimes v_2 \otimes \cdots \otimes v_k) = v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \cdots \otimes v_{\sigma^{-1}(k)}$ for all $v_1, v_2, \ldots, v_k \in A$ and $\sigma \in \mathfrak{S}_k$. For every $\pi \in \mathfrak{S}_k$, denote by $\rho (\pi)$ the action of $\pi$ on $A^\otimes k$ (this is an endomorphism of $A^\otimes k$). Show that every $\pi \in \mathfrak{S}_k$ satisfies $\rho (\pi) = \rho (\pi^{-1})$.

Exercise 1.62. State and solve the analogue of Exercise 1.61 for cocommutative $k$-bialgebras.

Exercise 1.63. (a) If $H$ is a $k$-bialgebra and $A$ is a commutative $k$-algebra, and if $f$ and $g$ are two $k$-algebra homomorphisms $H \to A$, then prove that $f \circ g$ also is a $k$-algebra homomorphism $H \to A$.

(b) If $H$ is a $k$-bialgebra and $A$ is a commutative $k$-algebra, and if $f_1, f_2, \ldots, f_k$ are several $k$-algebra homomorphisms $H \to A$, then prove that $f_1 \star f_2 \star \cdots \star f_k$ also is a $k$-algebra homomorphism $H \to A$.

(c) If $H$ is a Hopf algebra and $A$ is a commutative $k$-algebra, and if $f : H \to A$ is a $k$-algebra homomorphism, then prove that $f \circ S : H \to A$ (where $S$ is the antipode of $H$) is again a $k$-algebra homomorphism, and is a $*$-inverse to $f$.

(d) If $A$ is a commutative $k$-algebra, then show that $m(k)$ is a $k$-algebra homomorphism for every $k \in \mathbb{N}$. (The map $m(k) : A^\otimes (k+1) \to A$ is defined as in Exercise 1.39.)

(e) If $C'$ and $C$ are two $k$-coalgebras, if $\gamma : C \to C'$ is a $k$-coalgebra homomorphism, if $A$ and $A'$ are two $k$-algebras, if $\alpha : A \to A'$ is a $k$-algebra homomorphism, and if $f_1, f_2, \ldots, f_k$ are several $k$-linear maps $C' \to A$, then prove that $\alpha \circ (f_1 \star f_2 \star \cdots \star f_k) \circ \gamma = (\alpha \circ f_1 \circ \gamma) \star (\alpha \circ f_2 \circ \gamma) \star \cdots \star (\alpha \circ f_k \circ \gamma)$.

(f) If $H$ is a commutative $k$-bialgebra, and $k$ and $\ell$ are two nonnegative integers, then prove that $\text{id}_{H}^{k} \circ \text{id}_{H}^{\ell} = \text{id}_{H}^{k+\ell}$.

(g) If $H$ is a commutative $k$-Hopf algebra, and $k$ and $\ell$ are two integers, then prove that $\text{id}_{H}^{k} \circ \text{id}_{H}^{\ell} = \text{id}_{H}^{k+\ell}$.

Remark 1.64. The maps $\text{id}_{H}^{k}$ for $k \in \mathbb{N}$ are known as the Adams operators of the bialgebra $H$; they are studied, inter alia, in [5]. Particular cases of Exercise 1.63(e) appear in [127, Corollaire II.9] and [59, Theorem 1]. Exercise 1.63(f) and its dual are [103, Prop. 1.6].

Exercise 1.65. Let $A$ be a cocommutative graded Hopf algebra with antipode $S$. Define a $k$-linear map $E : A \to A$ by having $E(a) = (\deg a) \cdot a$ for every homogeneous element $a \in A$.

(a) Prove that for every $a \in A$, the elements $(S \star E)(a)$ and $(E \star S)(a)$ (where $\star$ denotes convolution in $\text{Hom}(A, A)$) are primitive.

(b) Prove that for every primitive $p \in A$, we have $(S \star E)(p) = (E \star S)(p) = E(p)$.

(c) Prove that for every $a \in A$ and every primitive $p \in A$, we have $(S \star E)(ap) = [(S \star E)(a), p] + \epsilon(a)E(p)$, where $[u, v]$ denotes the commutator $uv - vu$ of $u$ and $v$.

(d) If $A$ is connected and $\mathbb{Q}$ is a subring of $k$, prove that the $k$-algebra $A$ is generated by the $k$-submodule $p$ consisting of the primitive elements of $A$.

(e) Assume that $A$ is the tensor algebra $T(V)$ of a $k$-module $V$, and that the $k$-submodule $V = V^{	ext{deg} 1}$ of $T(V)$ is the degree-1 homogeneous component of $A$. Show that $(S \star E)(x_1 x_2 \cdots x_n) = [\cdots [x_1, x_2], x_3, \ldots, x_n]$ for any $n \geq 1$ and any $x_1, x_2, \ldots, x_n \in V$.

Remark 1.66. Exercise 1.65 gives rise to a certain idempotent map $A \to A$ when $k$ is a commutative $\mathbb{Q}$-algebra and $A$ is a cocommutative connected graded $k$-Hopf algebra. Namely, the $k$-linear map $A \to A$ sending every homogeneous $a \in A$ to $\frac{1}{\deg a} (S \star E)(a)$ (or 0 if $\deg a = 0$) is idempotent and is a projection on the $k$-module of primitive elements of $A$. It is called the Dynkin idempotent; see [128] for more of its properties. Part (c) of the exercise is more or less Baker’s identity.

1.6. Duals. Recall that for $k$-modules $V$, taking the dual $k$-module $V^* := \text{Hom}(V, k)$ reverses $k$-linear maps. That is, every $k$-linear map $V \to W$ induces an adjoint map $W^* \to V^*$ defined uniquely by $(f, \varphi(v)) = (\varphi^*(f), v)$.

\[24\] We will see another such idempotent in Exercise 5.38.
in which \((f, v)\) is the bilinear pairing \(V^* \times V \to k\) sending \((f, v) \mapsto f(v)\). If \(V\) and \(W\) are finite free \(k\)-modules, more can be said: When \(\varphi\) is expressed in terms of a basis \(\{v_i\}_{i \in I}\) for \(V\) and a basis \(\{w_j\}_{j \in J}\) for \(W\) by some matrix, the map \(\varphi^*\) is expressed by the transpose matrix in terms of the dual bases of these two bases.

The correspondence \(\varphi \mapsto \varphi^*\) between \(k\)-linear maps \(V \rightarrow W\) and \(k\)-linear maps \(W^* \rightarrow V^*\) is one-to-one when \(W\) is finite free. However, this is not the case in many combinatorial situations (in which \(W\) is usually free but not finite free). Fortunately, many of the good properties of finite free modules carry over to a certain class of graded modules as long as the dual \(V^*\) is replaced by a smaller module \(V^o\) called the graded dual. Let us first introduce the latter:

When \(V = \bigoplus_{n \geq 0} V_n\) is a graded \(k\)-module, note that the dual \(V^* = \prod_{n \geq 0} (V_n)^*\) can contain functionals \(f\) supported on infinitely many \(V_n\). However, we can consider the \(k\)-submodule \(V^o := \bigoplus_{n \geq 0} (V_n)^* \subset \prod_{n \geq 0} (V_n)^* = V^*\), sometimes called the graded dual, consisting of the functions \(f\) that vanish on all but finitely many \(V_n\). Notice that \(V^o\) is graded, whereas \(V^*\) (in general) is not. If \(V \rightarrow W\) is a graded \(k\)-linear map, then the adjoint map \(W^* \rightarrow V^*\) restricts to a graded \(k\)-linear map \(W^o \rightarrow V^o\), which we (abusively) still denote by \(\varphi^*\).

When the graded \(k\)-module \(V = \bigoplus_{n \geq 0} V_n\) is of \textit{finite type}, meaning that each \(V_n\) is a finite free \(k\)-module, the graded \(k\)-module \(V^o\) is again of finite type and satisfies \((V^o)^o \cong V\). Many other properties of finite free modules are salvaged in this situation; most importantly: The correspondence \(\varphi \mapsto \varphi^*\) between graded \(k\)-linear maps \(V \rightarrow W\) and graded \(k\)-linear maps \(W^o \rightarrow V^o\) is one-to-one when \(W\) is of finite type.

Reversing the diagrams should then make it clear that, in the finite free or finite-type situation, duals of algebras are coalgebras, and vice-versa, and duals of bialgebras or Hopf algebras are bialgebras or Hopf algebras. For example, the product in a Hopf algebra \(A\) of finite type uniquely defines the coproduct of \(A^o\) via adjointness:

\[
(\Delta_{A^o}(f), a \otimes b)_{A \otimes A} = (f, ab)_A.
\]

Thus if \(A\) has a basis \(\{a_i\}_{i \in I}\) with \textit{product structure constants} \(\{c^i_{j,k}\}\), meaning

\[
a_j a_k = \sum_{i \in I} c^i_{j,k} a_i,
\]

then the dual basis \(\{f_i\}_{i \in I}\) has the same \(\{c^i_{j,k}\}\) as its \textit{coproduct structure constants}:

\[
\Delta_{A^o}(f_i) = \sum_{(j,k) \in I \times I} c^i_{j,k} f_j \otimes f_k.
\]

The assumption that \(A\) be of finite type was indispensable here; in general, the dual of a \(k\)-algebra does not become a \(k\)-coalgebra. However, the dual of a \(k\)-coalgebra still becomes a \(k\)-algebra, as shown in the following exercise:

**Exercise 1.67.** For any two \(k\)-modules \(U\) and \(V\), let \(\rho_{U,V} : U^* \otimes V^* \to (U \otimes V)^*\) be the \(k\)-linear map which sends every tensor \(f \otimes g \in U^* \otimes V^*\) to the composition \(U \otimes V \overset{\rho}{\twoheadrightarrow} k \otimes k \overset{m_k}{\rightarrow} k\) of the map with the canonical isomorphism \(k \otimes k \overset{m_k}{\rightarrow} k\). When \(k\) is a field and \(U\) is finite-dimensional, this map \(\rho_{U,V}\) is a
\( k \)-vector space isomorphism (and usually regarded as the identity); more generally, it is injective whenever \( k \) is a field\(^{32}\). Also, let \( s : k \to k^* \) be the canonical isomorphism. Prove that:

(a) If \( C \) is a \( k \)-coalgebra, then \( C^* \) becomes a \( k \)-algebra if we define its associative operation by \( m_{C^*} = \Delta_C \circ p_{C,C} : C^* \otimes C^* \to C^* \) and its unit map to be \( e_C \circ s : k \to C^* \).\(^{33}\)

(b) The \( k \)-algebra structure defined on \( C^* \) in part (a) is precisely the one defined on \( \text{Hom}(C,k) = C^* \) in Definition 1.27 applied to \( A = k \).

(c) If \( C \) is a graded \( k \)-coalgebra, then \( C^o \) is a \( k \)-subalgebra of the \( k \)-algebra \( C^* \) defined in part (a).

(d) If \( f : C \to D \) is a homomorphism of \( k \)-coalgebras, then \( f^* : D^* \to C^* \) is a homomorphism of \( k \)-algebras.

(e) Let \( U \) be a graded \( k \)-module (not necessarily of finite type), and let \( V \) be a graded \( k \)-module of finite type. Then, there is a 1-to-1 correspondence between graded \( k \)-linear maps \( U \to V \) and graded \( k \)-linear maps \( V^o \to U^o \) given by \( f \mapsto f^* \).

(f) Let \( C \) be a graded \( k \)-coalgebra (not necessarily of finite type), and let \( D \) be a graded \( k \)-coalgebra of finite type. Part (e) of this exercise shows that there is a 1-to-1 correspondence between graded \( k \)-linear maps \( C \to D \) and graded \( k \)-linear maps \( D^o \to C^o \) given by \( f \mapsto f^* \). This correspondence has the property that a given graded \( k \)-linear map \( f : C \to D \) is a \( k \)-coalgebra map if and only if \( f^* : D^o \to C^o \) is a \( k \)-algebra map.

Another example of a Hopf algebra is provided by the so-called shuffle algebra. Before we introduce it, let us define the **shuffles** of two words:

**Definition 1.68.** Given two words \( a = (a_1, a_2, \ldots, a_n) \) and \( b = (b_1, b_2, \ldots, b_m) \), the **multiset of shuffles of \( a \) and \( b \)** is defined as the multiset

\[
\{ (c_{w(1)}, c_{w(2)}, \ldots, c_{w(n+m)}) : w \in \text{Sh}_{n,m} \}_{\text{multiset}}
\]

where \( (c_1, c_2, \ldots, c_{n+m}) \) is the concatenation \( a \cdot b = (a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_m) \), and where \( \text{Sh}_{n,m} \) is the subset\(^{34}\)

\[
\{ w \in S_{n+m} : w^{-1}(1) < w^{-1}(2) < \cdots < w^{-1}(n); w^{-1}(n + 1) < w^{-1}(n + 2) < \cdots < w^{-1}(n + m) \}
\]

of the symmetric group \( S_{n+m} \). Informally speaking, the shuffles of the two words \( a \) and \( b \) are the words obtained by overlaying the words \( a \) and \( b \), after first moving their letters apart so that no letters get superimposed when the words are overlaid\(^{35}\). In particular, any shuffle of \( a \) and \( b \) contains \( a \) and \( b \) as subsequences. The multiset of shuffles of \( a \) and \( b \) has \( \binom{m+n}{n} \) elements (counted with multiplicity) and is denoted by \( a \shuffle b \). For instance, the shuffles of \((1,2,1)\) and \((3,2)\) are

\[
(1,2,1,3,2), (1,2,3,1,2), (1,3,2,1,2), (1,3,2,2,1),
(1,3,2,2,1), (3,1,2,2,1), (3,1,2,2,1), (3,2,1,2,1),
\]

listed here as often as they appear in the multiset \((1,2,1) \shuffle (3,2)\). Here we have underlined the letters taken from \( a \) – that is, the letters at positions \( w^{-1}(1), w^{-1}(2), \ldots, w^{-1}(n) \).

**Example 1.69.** When \( A = T(V) \) is the tensor algebra for a finite free \( k \)-module \( V \), having \( k \)-basis \( \{x_i\}_{i \in I} \), its graded dual \( A^o \) is another Hopf algebra whose basis \( \{y_{(i_1,\ldots,i_t)}\} \) (the dual basis of the basis \( \{x_{i_1}\cdots x_{i_t}\} \) of \( A = T(V) \)) is indexed by words in the alphabet \( I \). This Hopf algebra \( A^o \) could be called the **shuffle algebra** of \( V^* \). (To be more precise, it is isomorphic to the shuffle algebra of \( V^* \) introduced in Proposition 1.73

\(^{32}\)Over arbitrary rings it does not have to be even that!

\(^{33}\)If \( C \) is a finite free \( k \)-module, then this \( k \)-algebra structure is the same as the one defined above by adjointness. But the advantage of the new definition is that it works even if \( C \) is not a finite free \( k \)-module.

\(^{34}\)Warning: This definition of \( \text{Sh}_{n,m} \) is highly nonstandard, and many authors define \( \text{Sh}_{n,m} \) to be the set of the inverses of the permutations belonging to what we call \( \text{Sh}_{n,m} \).

\(^{35}\)For instance, if \( a = (1,3,2,1) \) and \( b = (2,4) \), then the shuffle \((1,3,2,2,4,1)\) of \( a \) and \( b \) can be obtained by moving the letters of \( a \) and \( b \) apart as follows:

\[
a = \begin{array}{cccc}
1 & 3 & 2 & 1 \\
b & 2 & 4 \\
\end{array}
\]

and then overlaying them to obtain \( 1 \ 2 \ 3 \ 2 \ 4 \ 1 \). Other ways of moving letters apart lead to further shuffles (not always distinct).
further below; we prefer not to call \( A^o \) itself the shuffle algebra of \( V^* \), since \( A^o \) has several disadvantages\(^\text{36}\).)

Duality shows that the cut coproduct in \( A^o \) is defined by

\[
\Delta y(i_1, \ldots, i_\ell) = \sum_{j=0}^{\ell} y(i_1, \ldots, i_j) \otimes y(i_{j+1}, i_{j+2}, \ldots, i_\ell).
\]

For example,

\[
\Delta y_{abcb} = y_{ab} \otimes y_{abcb} + y_b \otimes y_{abcb} + y_{abc} \otimes y_{abcb} + y_{abcb} \otimes y_b + y_{abcb} \otimes y_{ab}
\]

Duality also shows that the shuffle product in \( A^o \) will be given by

\[
y(i_1, \ldots, i_\ell) y(j_1, \ldots, j_m) = \sum_{k=(k_1, \ldots, k_{\ell+m}) \in i \shuffle j} y(k_1, \ldots, k_{\ell+m})
\]

where \( i \shuffle j \) (as in Definition 1.68) denotes the multiset of the \( \binom{\ell+m}{\ell} \) words obtained as shuffles of the two words \( i = (i_1, \ldots, i_\ell) \) and \( j = (j_1, \ldots, j_m) \). For example,

\[
y_{ab} y_{cb} = y_{abcb} + y_{acb} + y_{cab} + y_{acb} + y_{cabb} = y_{abcb} + 2y_{acb} + y_{cabb}
\]

Equivalently, one has

\[
y(i_1, i_2, \ldots, i_\ell) y(i_{\ell+1}, i_{\ell+2}, \ldots, i_{\ell+m}) = \sum_{w \in S_{\ell+m} \text{ for Sym}(V)} y(i_{w-1(1)} i_{w-1(2)} \cdots i_{w-1(\ell+m)})
\]

\[
= \sum_{\sigma \in S_{\ell+m}} y(i_{\sigma(1)} i_{\sigma(2)} \cdots i_{\sigma(\ell+m)})
\]

(1.23)

(using the notations of Definition 1.68 again). Lastly, the antipode \( S \) of \( A^o \) is the adjoint of the antipode of \( A = T(V) \) described in (1.15):

\[
Sy(i_1, i_2, \ldots, i_\ell) = (-1)^\ell y(i_\ell, \ldots, i_2, i_1).
\]

Since the coalgebra \( T(V) \) is cocommutative, its graded dual \( T(V)^o \) is commutative.

**Exercise 1.70.** Let \( V \) be a 1-dimensional free \( k \)-module with basis element \( x \), so \( \text{Sym}(V) \cong k[x] \), with \( k \)-basis \( \{1 = x^0, x^1, x^2, \ldots\} \).

(a) Check that the powers \( x^i \) satisfy

\[
x^i \cdot x^j = x^{i+j}
\]

\[
\Delta(x^n) = \sum_{i+j=n} \binom{n}{i} x^i \otimes x^j
\]

\[
S(x^n) = (-1)^n x^n
\]

(b) Check that the dual basis elements \( \{f(0), f(1), f(2), \ldots\} \) for \( \text{Sym}(V)^o \), defined by \( f(i)(x^j) = \delta_{i,j} \), satisfy

\[
f(i) f(j) = \binom{i+j}{i} f(i+j)
\]

\[
\Delta(f(n)) = \sum_{i+j=n} f(i) \otimes f(j)
\]

\[
S(f(n)) = (-1)^n f(n)
\]

(c) Show that if \( \mathbb{Q} \) is a subring of \( k \), then the \( k \)-linear map \( \text{Sym}(V)^o \to \text{Sym}(V) \) sending \( f(n) \mapsto \frac{x^n}{n!} \) is a graded Hopf isomorphism.

For this reason, the Hopf structure on \( \text{Sym}(V)^o \) is called a divided power algebra.

(d) Show that when \( k \) is a field of characteristic \( p > 0 \), one has \( f(1)^p = 0 \), and hence why there can be no Hopf isomorphism \( \text{Sym}(V)^o \to \text{Sym}(V) \).

\(^{36}\text{Specifically, } A^o \text{ has the disadvantages of being defined only when } V^* \text{ is the dual of a finite free } k \text{-module } V, \text{ and depending on a choice of basis, whereas Proposition 1.73 will define shuffle algebras in full generality and canonically.} \)
Exercise 1.71. Let $V$ have $k$-basis $\{x_1, \ldots, x_n\}$, and let $V \oplus V$ have $k$-basis $\{x_1, \ldots, x_n, y_1, \ldots, y_n\}$, so that one has isomorphisms

$$\text{Sym}(V \oplus V) \cong k[x, y] \cong k[x] \otimes k[y] \cong \text{Sym}(V) \otimes \text{Sym}(V).$$

(a) Show that our usual coproduct on $\text{Sym}(V)$ can be re-expressed as follows:

\[
\begin{align*}
\text{Sym}(V) & \quad \Delta \quad \text{Sym}(V) \\
\begin{array}{c|c}
   k[x] & \Delta \\
   f(x_1, \ldots, x_n) & \mapsto f(x_1 + y_1, \ldots, x_n + y_n)
\end{array}
\end{align*}
\]

In other words, it is induced from the diagonal map

\[
V \quad \longrightarrow \quad V \oplus V \\
x_i \quad \longmapsto \quad x_i + y_i
\]

(b) One can similarly define a coproduct on the exterior algebra $\wedge V$, which is the quotient $T(V)/J$ where $J$ is the two-sided ideal generated by the elements $\{x^2 = x \otimes x\}_{x \in V}$ in $T^2(V)$. This becomes a graded commutative algebra

\[
\wedge V = \bigoplus_{d=0}^{\infty} \wedge^d V = \bigoplus_{d=0}^{\infty} \wedge^d V,
\]

if one views the elements of $V = \wedge^1 V$ as having odd degree, and uses the topologist’s sign conventions (as in (1.7)). One again has $\wedge(V \oplus V) = \wedge V \otimes \wedge V$ as graded algebras. Show that one can again let the diagonal map (1.24) induce a map

\[
\begin{align*}
\wedge(V) & \quad \Delta \quad \wedge V \otimes \wedge V \\
\begin{array}{c|c}
   \sum c_{i_1, \ldots, i_d} \cdot x_{i_1} \wedge \cdots \wedge x_{i_d} & \sum c_{i_1, \ldots, i_d} \cdot (x_{i_1} + y_{i_1}) \wedge \cdots \wedge (x_{i_d} + y_{i_d})
\end{array}
\end{align*}
\]

which makes $\wedge V$ into a graded connected Hopf algebra.

(c) Show that in the tensor algebra $T(V)$, if one views the elements of $V = V^{\otimes 1}$ as having odd degree, and uses the convention (1.7) in the twist map when defining $T(V)$, then for any $x$ in $V$ one has $\Delta(x^2) = 1 \otimes x^2 + x^2 \otimes 1$.

(d) Use part (c) to show that the two-sided ideal $J \subset T(V)$ generated by $\{x^2\}_{x \in V}$ is also a two-sided coideal, and hence the quotient $\wedge V = T(V)/J$ inherits the structure of a bialgebra. Check that the coproduct on $\wedge V$ inherited from $T(V)$ is the same as the one defined in part (b).

[Hint: The ideal $J$ in part (b) is a graded $k$-submodule of $T(V)$, but this is not completely obvious (not all elements of $V$ have to be homogeneous!).]

Exercise 1.72. Let $C$ be a $k$-coalgebra. As we know from Exercise 1.67(a), this makes $C^*$ into a $k$-algebra. Let $A$ be a $k$-algebra which is finite free as $k$-module. This makes $A^*$ into a $k$-coalgebra.

Let $f : C \rightarrow A$ and $g : C \rightarrow A$ be two $k$-linear maps. Show that $f^* \ast g^* = (f \ast g)^*$.

The above arguments might have created the impression that duals of bialgebras have good properties only under certain restrictive conditions (e.g., the dual of a bialgebra $H$ does not generally become a bialgebra unless $H$ is of finite type), and so they cannot be used in proofs and constructions unless one is willing to sacrifice some generality (e.g., we had to require $V$ to be finite free in Example 1.69). While the first part of this impression is true, the second is not always true; often there is a way to gain back the generality lost from using duals. As an example of this, let us define the shuffle algebra of an arbitrary $k$-module (not just of a dual of a finite free $k$-module as in Example 1.69):

**Proposition 1.73.** Let $V$ be a $k$-module. Define a $k$-linear map $\Delta_{\text{shuffle}} : T(V) \rightarrow T(V) \otimes T(V)$ by setting

\[
\Delta_{\text{shuffle}}(v_1 v_2 \cdots v_n) = \sum_{k=0}^{n} (v_1 v_2 \cdots v_k) \otimes (v_{k+1} v_{k+2} \cdots v_n) \quad \text{for all } n \in \mathbb{N} \text{ and } v_1, v_2, \ldots, v_n \in V.
\]
Define a \( k \)-bilinear map \( \shuffle : T(V) \times T(V) \to T(V) \), which will be written in infix notation (that is, we will write \( a \shuffle b \) instead of \( \shuffle(a,b) \)), by setting\(^{38}\)

\[
(v_1 v_2 \cdots v_\ell) \shuffle (v_{\ell+1} v_{\ell+2} \cdots v_{\ell+m}) = \sum_{\sigma \in \text{Sh}_{\ell,m}} v_{\sigma(1)} v_{\sigma(2)} \cdots v_{\sigma(\ell+m)}
\]

for all \( \ell, m \in \mathbb{N} \) and \( v_1, v_2, \ldots, v_{\ell+m} \in V \).

Consider also the comultiplication \( \epsilon \) of the Hopf algebra \( T(V) \).

Then, the \( k \)-module \( T(V) \), endowed with the multiplication \( \shuffle \), the unit \( 1_{T(V)} \in V^\otimes 0 \subset T(V) \), the comultiplication \( \Delta_\shuffle \) and the counit \( \epsilon \), becomes a commutative Hopf algebra. This Hopf algebra is called the shuffle algebra of \( V \), and denoted by \( \text{Sh}(V) \). The antipode of the Hopf algebra \( \text{Sh}(V) \) is precisely the antipode \( S \) of \( T(V) \).

**Exercise 1.74.** Prove Proposition 1.73.

**Hint:** When \( V \) is a finite free \( k \)-module, Proposition 1.73 follows from Example 1.69. The trick is to derive the general case from this specific one. Every \( k \)-linear map \( f : W \to V \) between two \( k \)-modules \( W \) and \( V \) induces a map \( T(f) : T(W) \to T(V) \) which preserves \( \Delta_\shuffle \), \( \epsilon \), \( 1_{T(V)} \), \( S \) (in the appropriate meanings – e.g., preserving \( \Delta_\shuffle \) means \( \Delta_\shuffle \circ T(f) = (T(f) \otimes T(f)) \circ \Delta_\shuffle \)). Show that each of the equalities that need to be proven in order to verify Proposition 1.73 can be “transported” along such a map \( T(f) \) from a \( T(W) \) for a suitably chosen finite free \( k \)-module \( W \).

It is also possible to prove Proposition 1.73 “by foot”, as long as one is ready to make combinatorial arguments about cutting shuffles.

**Remark 1.75.** (a) Let \( V \) be a finite free \( k \)-module. The Hopf algebra \( T(V)^* \) (studied in Example 1.69) is naturally isomorphic to the shuffle algebra \( \text{Sh}(V^*) \) (defined as in Proposition 1.73 but for \( V^* \) instead of \( V \)) as Hopf algebras, by the obvious isomorphism (namely, the direct sum of the isomorphisms \( (V^\otimes n)^* \to (V^*)^\otimes n \) over all \( n \in \mathbb{N} \)).

(b) The same statement applies to the case when \( V \) is a graded \( k \)-module of finite type satisfying \( V_0 = 0 \) rather than a finite free \( k \)-module, provided that \( V^* \) and \( (V^\otimes n)^* \) are replaced by \( V^\circ \) and \( (V^\otimes n)^\circ \).

We shall return to shuffle algebras in Section 6.3, where we will show that under certain conditions (\( \mathbb{Q} \) being a subring of \( k \), and \( V \) being a free \( k \)-module) the algebra structure on a shuffle algebra \( \text{Sh}(V) \) is a polynomial algebra in an appropriately chosen set of generators\(^{41}\).

---

\(^{37}\)This is well-defined, because the right hand side is \( n \)-multilinear in \( v_1, v_2, \ldots, v_n \), and because any \( n \)-multilinear map \( V^\times n \to M \) into a \( k \)-module \( M \) gives rise to a unique \( k \)-linear map \( V^\otimes n \to M \).

\(^{38}\)Many authors use the symbol \( \circledast \) instead of \( \shuffle \) here, but we prefer to reserve the former notation for the shuffle product of words.

\(^{39}\)Again, this is well-defined by the \( \ell + m \)-multilinearity of the right hand side.

\(^{40}\)This can be verified by comparing (1.20) with the definition of \( \Delta_\shuffle \), and comparing (1.23) with the definition of \( \shuffle \).

\(^{41}\)This says nothing about the coalgebra structure on \( \text{Sh}(V) \) – which is much more complicated in these generators.
2. Review of symmetric functions \( \Lambda \) as Hopf algebra

Here we review the ring of symmetric functions, borrowing heavily from standard treatments, such as Macdonald [109, Chap. 1], Sagan [141, Chap. 4], and Stanley [157, Chap. 7], but emphasizing the Hopf structure early on.

2.1. Definition of \( \Lambda \). As before, \( k \) here is a commutative ring (hence could be a field or the integers \( \mathbb{Z} \); these are the usual choices). Given an infinite variable set \( x = (x_1, x_2, \ldots) \), a monomial \( x^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \cdots \) is indexed by an element \( \alpha = (\alpha_1, \alpha_2, \ldots) \in \mathbb{N}^\infty \) having finite support\(^{42} \); such \( \alpha \) are called weak compositions. The nonzero ones among the integers \( \alpha_1, \alpha_2, \ldots \) are called the parts of the weak composition \( \alpha \). We will consider the ring \( R(x) \) of formal power series \( f(x) = \sum c_\alpha x^\alpha \) with \( c_\alpha \in k \) of bounded degree, that is, where there exists some bound \( d = d(f) \) for which \( \deg(x^\alpha) := \sum \alpha_i > d \) implies \( c_\alpha = 0 \). It is easy to see that the product of two such power series is well-defined, and also has bounded degree.

The symmetric group \( S_n \) permuting the first \( n \) variables \( x_1, \ldots, x_n \) acts as a group of automorphisms on \( R(x) \), as does the union \( S(\infty) = \bigcup_{n \geq 0} S_n \) of the infinite ascending chain \( S_0 \subset S_1 \subset S_2 \subset \cdots \) of symmetric groups\(^{43} \). This group \( S(\infty) \) can also be described as the group of all permutations of the set \( \{1, 2, 3, \ldots\} \) which leave all but finitely many elements invariant.

**Definition 2.1.** The ring of symmetric functions in \( x \) with coefficients in \( k \), denoted \( \Lambda = \Lambda_k = \Lambda(x) = \Lambda_k(x) \), is the \( S(\infty) \)-invariant subalgebra \( R(x)^{S(\infty)} \) of \( R(x) \):

\[
\Lambda := \left\{ f = \sum_\alpha c_\alpha x^\alpha \in R(x) : c_\alpha = c_\beta \text{ if } \alpha, \beta \text{ lie in the same } S(\infty)\text{-orbit} \right\}.
\]

We refer to the elements of \( \Lambda \) as symmetric functions (over \( k \)); however, despite this terminology, they are not functions in the usual sense.\(^{44} \)

Note that \( \Lambda \) is a graded \( k \)-algebra, since \( \Lambda = \bigoplus_{n \geq 0} \Lambda_n \) where \( \Lambda_n \) are the symmetric functions \( f = \sum_\alpha c_\alpha x^\alpha \) which are homogeneous of degree \( n \), meaning \( \deg(x^\alpha) = n \) for all \( c_\alpha \neq 0 \).

**Exercise 2.2.** Let \( f \in R(x) \). Let \( A \) be a commutative \( k \)-algebra, and \( a_1, a_2, \ldots, a_k \) be finitely many elements of \( A \). Show that substituting \( a_1, a_2, \ldots, a_k, 0, 0, \ldots \) for \( x_1, x_2, x_3, \ldots \) in \( f \) yields an infinite sum in which all but finitely many addends are zero. Hence, this sum has a value in \( A \), which is commonly denoted by \( f(a_1, a_2, \ldots, a_k) \).

**Definition 2.3.** A partition \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell, 0, 0, \ldots) \) is a weak composition whose entries weakly decrease: \( \lambda_1 \geq \cdots \geq \lambda_\ell > 0 \). The (uniquely defined) \( \ell \) is said to be the length of the partition \( \lambda \) and denoted by \( \ell(\lambda) \). One sometimes omits trailing zeroes from a partition; e.g., one can write the partition \( (3, 1, 0, 0, 0, \ldots) \) as \( (3, 1) \). We will often (but not always) write \( \lambda_i \) for the \( i \)-th entry of the partition \( \lambda \) (for instance, \( \lambda_2 = 3 \) if \( \lambda = (5, 3, 1, 1) \)) and call it the \( i \)-th part of \( \lambda \). The sum \( \lambda_1 + \lambda_2 + \cdots + \lambda_\ell = \lambda_1 + \lambda_2 + \cdots + \lambda_\ell \) (where \( \ell = \ell(\lambda) \)) of all parts of \( \lambda \) is called the size of \( \lambda \) and denoted by \( |\lambda| \); for a given integer \( n \), the partitions of size \( n \) are referred to as the partitions of \( n \). The empty partition \( \varnothing \) is denoted by \( \varnothing \).

Every weak composition \( \alpha \) lies in the \( S(\infty) \)-orbit of a unique partition \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell, 0, 0, \ldots) \) with \( \lambda_1 \geq \cdots \geq \lambda_\ell > 0 \). For any partition \( \lambda \), define the monomial symmetric function

\[(2.1) \quad m_\lambda := \sum_{\alpha \in S(\infty) \lambda} x^\alpha.\]

Letting \( \lambda \) run through the set \( \text{Par} \) of all partitions, this gives the monomial \( k \)-basis \( \{m_\lambda \} \) of \( \Lambda \). Letting \( \lambda \) run only through the set \( \text{Par}_n \) of partitions of \( n \) gives the monomial \( k \)-basis for \( \Lambda_n \).

\(^{42}\)The support of a sequence \( \alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots) \in \mathbb{N}^\infty \) is defined to be the set of all positive integers \( i \) for which \( \alpha_i \neq 0 \).

\(^{43}\)This ascending chain is constructed as follows: For every \( n \in \mathbb{N} \), there is an injective group homomorphism \( \iota_n : S_n \rightarrow S_{n+1} \) which sends every permutation \( \sigma \in S_n \) to the permutation \( \iota_n(\sigma) = \tau \in S_{n+1} \) defined by

\[
\tau(i) = \begin{cases} 
\sigma(i), & \text{if } i \leq n; \\
 i, & \text{if } i = n + 1
\end{cases}
\]

for all \( i \in \{1, 2, \ldots, n + 1\} \).

These homomorphisms \( \iota_n \) for all \( n \) form a chain \( S_0 \xrightarrow{\iota_0} S_1 \xrightarrow{\iota_1} S_2 \xrightarrow{\iota_2} \cdots \), which is often regarded as a chain of inclusions.

\(^{44}\)Being power series, they can be evaluated at appropriate families of variables. But this does not make them functions (no more than polynomials are functions). The terminology “symmetric function” is thus not well-chosen; but it is standard.
Example 2.4. For \( n = 3 \), one has
\[
m_{(3)} = x_1^3 + x_2^3 + x_3^3 + \ldots
\]
\[
m_{(2,1)} = x_1^2 x_2 + x_1 x_2^2 + x_1^2 x_3 + x_1 x_3^2 + \ldots
\]
\[
m_{(1,1,1)} = x_1 x_2 x_3 + x_1 x_2 x_4 + x_1 x_3 x_4 + x_2 x_3 x_4 + x_1 x_2 x_5 + \ldots
\]

Remark 2.5. We have defined the symmetric functions as the elements of \( R(\mathbf{x}) \) invariant under the group \( \mathfrak{S}_\infty \). However, they also are the elements of \( R(\mathbf{x}) \) invariant under the group \( \mathfrak{S}_\infty \) of all permutations of the set \( \{1, 2, 3, \ldots\} \).

Remark 2.6. It is sometimes convenient to work with finite variable sets \( x_1, \ldots, x_n \), which one justifies as follows. Note that the algebra homomorphism
\[
R(\mathbf{x}) \rightarrow R(x_1, \ldots, x_n) = k[x_1, \ldots, x_n]
\]
which sends \( x_{n+1}, x_{n+2}, \ldots \) to 0 restricts to an algebra homomorphism
\[
\Lambda_k(\mathbf{x}) \rightarrow \Lambda_k(x_1, \ldots, x_n) = k[x_1, \ldots, x_n]^\mathfrak{S}_n.
\]
Furthermore, this last homomorphism is a \( k \)-linear isomorphism when restricted to \( \Lambda_i \) for \( 0 \leq i \leq n \), since it sends the monomial basis elements \( m_\lambda(\mathbf{x}) \) to the monomial basis elements \( m_\lambda(x_1, \ldots, x_n) \). Thus when one proves identities in \( \Lambda_k(x_1, \ldots, x_n) \) for all \( n \), they are valid in \( \Lambda \), that is, \( \Lambda \) is the inverse limit of the \( \Lambda(x_1, \ldots, x_n) \) in the category of graded \( k \)-algebras.

One can also define a comultiplication on \( \Lambda \) as follows. Note that when one decomposes the variables into two sets \( (x, y) = (x_1, x_2, \ldots, y_1, y_2, \ldots) \), one has a ring homomorphism
\[
R(x) \otimes R(y) \rightarrow R(x, y)
\]
\[
f(x) \otimes g(y) \rightarrow f(x)g(y)
\]
This restricts to an isomorphism
\[
2.2 \quad \Lambda \otimes \Lambda = R(\mathbf{x})^\mathfrak{S}_\infty \otimes R(\mathbf{y})^\mathfrak{S}_\infty \rightarrow R(\mathbf{x}, \mathbf{y})^\mathfrak{S}_\infty \times \mathfrak{S}_\infty
\]
where \( \mathfrak{S}_\infty \times \mathfrak{S}_\infty \) denotes permutations of \( (\text{finite subsets of}) \) the \( \mathbf{x} \) and separate permutations of \( (\text{finite subsets of}) \) the \( \mathbf{y} \), because \( R(\mathbf{x}, \mathbf{y})^\mathfrak{S}_\infty \times \mathfrak{S}_\infty \) has \( k \)-basis \( \{m_\lambda(\mathbf{x})m_\mu(\mathbf{y})\}_{\lambda, \mu \in \text{Par}} \). As \( \mathfrak{S}_\infty \times \mathfrak{S}_\infty \) is a subgroup of the group \( \mathfrak{S}_\infty \times \mathfrak{S}_\infty \) (the group of all permutations of \( \{x_1, x_2, \ldots, y_1, y_2, \ldots\} \) leaving all but finitely many variables invariant) acting on all \( (x, y) \), one gets an inclusion of rings
\[
\Lambda(\mathbf{x}, \mathbf{y}) = R(\mathbf{x}, \mathbf{y})^\mathfrak{S}_\infty \rightarrow R(\mathbf{x}, \mathbf{y})^\mathfrak{S}_\infty \times \mathfrak{S}_\infty \equiv \Lambda \otimes \Lambda
\]
where the last isomorphism is the inverse of the one in \( (2.2) \). This gives a comultiplication
\[
\Delta(\mathbf{x}, \mathbf{y}) = f(x_1, x_2, \ldots) \rightarrow f(x_1, x_2, \ldots, y_1, y_2, \ldots).
\]
Here, \( f(x_1, x_2, \ldots, y_1, y_2, \ldots) \) means the result of choosing some bijection \( \phi : \{x_1, x_2, x_3, \ldots\} \rightarrow \{x_1, x_2, \ldots, y_1, y_2, \ldots\} \) and substituting \( \phi(x_i) \) for every \( x_i \) in \( f \). (The choice of \( \phi \) is irrelevant since \( f \) is symmetric.)

\( \text{Proof.} \) We need to show that \( \Lambda = R(\mathbf{x})^\mathfrak{S}_\infty \). Since
\[
\Lambda = \left\{ f = \sum_\alpha c_\alpha x^\alpha \in R(\mathbf{x}) : c_\alpha = c_\beta \text{ if } \alpha, \beta \text{ lie in the same } \mathfrak{S}_\infty \text{-orbit} \right\}
\]
and
\[
R(\mathbf{x})^\mathfrak{S}_\infty = \left\{ f = \sum_\alpha c_\alpha x^\alpha \in R(\mathbf{x}) : c_\alpha = c_\beta \text{ if } \alpha, \beta \text{ lie in the same } \mathfrak{S}_\infty \text{-orbit} \right\},
\]
this will follow immediately if we can show that two weak compositions \( \alpha \) and \( \beta \) lie in the same \( \mathfrak{S}_\infty \)-orbit if and only if they lie in the same \( \mathfrak{S}_\infty \)-orbit. But this is straightforward to check (in fact, two weak compositions \( \alpha \) and \( \beta \) lie in the same orbit under either group if and only if they have the same multiset of nonzero entries).

\( \text{To be more precise, the choice of } \phi \text{ is irrelevant because } f \text{ is } \mathfrak{S}_\infty \text{-invariant, with the notations of Remark 2.5.} \)
Example 2.7. One has
\[ \Delta m_{(2,1)} = m_{(2,1)}(x_1, x_2, \ldots, y_1, y_2, \ldots) \]
\[ = x_1^2 x_2 + x_1 x_2^2 + \cdots \]
\[ + x_1^2 y_1 + x_1 y_2^2 + \cdots \]
\[ + x_1 y_1^2 + x_1 y_2^2 + \cdots \]
\[ + y_1^2 y_2 + y_1 y_2^2 + \cdots \]
\[ = m_{(2,1)}(x) + m_{(2)}(x)m_{(1)}(y) + m_{(1)}(x)m_{(2)}(y) + m_{(2,1)}(y) \]
\[ = m_{(2,1)} \otimes 1 + m_{(2)} \otimes m_{(1)} + m_{(1)} \otimes m_{(2)} + 1 \otimes m_{(2,1)}. \]

This example generalizes easily to the following formula
\[ \Delta m_\lambda = \sum_{\mu \sqcup \nu = \lambda} m_\mu \otimes m_\nu, \]
in which $\mu \sqcup \nu$ is the partition obtained by taking the multiset union of the parts of $\mu$ and $\nu$, and then reordering them to make them weakly decreasing.

Checking that $\Delta$ is coassociative amounts to checking that
\[ (\Delta \otimes \text{id}) \circ \Delta f = f(\text{id} \otimes \Delta) \circ \Delta f \]
inside $\Lambda(x, y, z)$ as a subring of $\Lambda \otimes \Lambda \otimes \Lambda$. The counit $\Lambda \rightarrow k$ is defined in the usual fashion for graded connected coalgebras, namely $\epsilon$ annihilates $I = \bigoplus_{n>0} \Lambda_n$, and $\epsilon$ is the identity on $\Lambda_0 = k$; alternatively $\epsilon$ sends a symmetric function $f(x)$ to its constant term $f(0,0,\ldots)$.

Note that $\Delta$ is an algebra morphism $\Lambda \rightarrow \Lambda \otimes \Lambda$ because it is a composition of maps which are all algebra morphisms. As the unit and counit axioms are easily checked, $\Lambda$ becomes a graded connected $k$-bialgebra of finite type, and hence also a Hopf algebra by Proposition 1.36. We will identify its antipode more explicitly in Section 2.4 below.

2.2. Other Bases. We introduce the usual other bases of $\Lambda$, and explain their significance later.

Definition 2.8. Define the families of power sum symmetric functions $p_n$, elementary symmetric functions $e_n$, and complete homogeneous symmetric functions $h_n$, for $n = 1, 2, 3, \ldots$ by
\[ p_n := x_1^n + x_2^n + \cdots = m_{(n)} \]
\[ e_n := \sum_{i_1 < \cdots < i_n} x_{i_1} \cdots x_{i_n} = m_{(1^n)} \]
\[ h_n := \sum_{i_1 \leq \cdots \leq i_n} x_{i_1} \cdots x_{i_n} = \sum_{\lambda \in \text{Par}_n} m_\lambda \]

Here, we are using the multiplicative notation for partitions: whenever $(m_1, m_2, m_3, \cdots)$ is a weak composition, $(1^{m_1}2^{m_2}3^{m_3} \cdots)$ denotes the partition $\lambda$ such that for every $i$, the multiplicity of the part $i$ in $\lambda$ is $m_i$. The $i^{m_i}$ satisfying $m_i = 0$ are often omitted from this notation, and so the $(1^n)$ in (2.5) means $\begin{pmatrix} 1,1,\ldots,1 \end{pmatrix}_{n \text{ ones}}$. (For another example, $(123^34^3) = (1^{10}2^31^34^35^06^07^0 \cdots)$ means the partition $(4,4,4,3,1,1,\ldots)$.)

By convention, also define $h_0 = e_0 = 1$, and $h_{n} = e_{n} = 0$ if $n < 0$. Extend these multiplicatively to partitions $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ with $\lambda_1 \geq \cdots \geq \lambda_\ell > 0$:
\[ p_\lambda := p_{\lambda_1}p_{\lambda_2} \cdots p_{\lambda_\ell} \]
\[ e_\lambda := e_{\lambda_1}e_{\lambda_2} \cdots e_{\lambda_\ell} \]
\[ h_\lambda := h_{\lambda_1}h_{\lambda_2} \cdots h_{\lambda_\ell} \]

Also define the Schur function
\[ s_\lambda := \sum_T x^{\text{cont}(T)} \]


where $T$ runs through all column-strict tableaux of shape $\lambda$, that is, $T$ is an assignment of entries in $\{1, 2, 3, \ldots\}$ to the cells of the Ferrers diagram\(^{47}\) for $\lambda$, weakly increasing left-to-right in rows, and strictly increasing top-to-bottom in columns. Here $\text{cont}(T)$ denotes the weak composition $\{\lvert T^{-1}(1)\rvert, \lvert T^{-1}(2)\rvert, \lvert T^{-1}(3)\rvert, \ldots\}$, so that $x^{\text{cont}(T)} = \prod_i x_i^{\lvert T^{-1}(i)\rvert}$. For example,\(^8\) 

\[
 T = \begin{bmatrix} 1 & 1 & 1 & 4 & 7 \\
 2 & 3 & 3 \\
 4 & 4 & 6 \\
 6 & 7 
\end{bmatrix}
\]

is a column-strict tableau of shape $\lambda = (5, 3, 3, 2)$ with $x^{\text{cont}(T)} = x_1^3x_2^1x_3^2x_4^3x_5^0x_6^2x_7^4$.

**Example 2.9.** One has 

\[
 m(1) = p(1) = e(1) = h(1) = s(1) = x_1 + x_2 + x_3 + \cdots \\
 s(n) = h_n \\
 s(1^n) = e_n
\]

**Example 2.10.** One has for $\lambda = (2, 1)$ that 

\[
 p(2,1) = p_2p_1 = (x_1^2 + x_2^2 + \cdots)(x_1 + x_2 + \cdots) \\
 = m_{(2,1)} + m_{(3)} \\
 e(2,1) = e_2e_1 = (x_1x_2 + x_1x_3 + \cdots)(x_1 + x_2 + \cdots) \\
 = m_{(2,1)} + 3m_{(1,1,1)} \\
 h(2,1) = h_2h_1 = (x_1^2 + x_2^2 + \cdots + x_1x_2 + x_1x_3 + \cdots)(x_1 + x_2 + \cdots) \\
 = m_{(3)} + 2m_{(2,1)} + 3m_{(1,1,1)}
\]

and 

\[
 s_{(2,1)} = x_1^2x_2 + x_1x_2x_3 + x_1x_2^2 + x_1x_2x_3 + x_1x_2x_3 + x_1x_2x_3 + x_1x_2x_4 + \cdots \\
 = m_{(2,1)} + 2m_{(1,1,1)}
\]

\(^{47}\)The Ferrers diagram of a partition $\lambda$ is defined as the set of all pairs $(i, j) \in \{1, 2, 3, \ldots\}^2$ satisfying $j \leq \lambda_i$. This is a set of cardinality $\lvert \lambda \rvert$. Usually, one visually represents a Ferrers diagram by drawing its elements $(i, j)$ as points on the plane, although (unlike the standard convention for drawing points on the plane) one lets the $x$-axis go top-to-bottom (i.e., the point $(i, j + 1)$ is one step below the point $(i, j)$), and the $y$-axis go left-to-right (i.e., the point $(i, j + 1)$ is one step to the right of the point $(i, j)$). (This is the so-called English notation, also known as the matrix notation because it is precisely the way one labels the entries of a matrix. Other notations appear in literature, such as the French notation used, e.g., in Malvenuto’s\(^{112}\), and the Russian notation used, e.g., in parts of Kerov’s\(^{80}\).) These points are drawn either as dots or as square boxes; in the latter case, the boxes are centered at the points they represent, and they have sidelength 1 so that the boxes centered around $(i, j)$ and $(i, j + 1)$ touch each other along a sideline. For example, the Ferrers diagram of the partition $(3, 2, 2)$ is represented as

\[
\begin{array}{cccc}
\cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot \\
\end{array}
\]

or as

\[
\begin{array}{cccc}
\Box & \Box & \Box & \Box \\
\Box & \Box & \Box & \Box \\
\end{array}
\]

The Ferrers diagram of a partition $\lambda$ uniquely determines $\lambda$. One refers to the elements of the Ferrers diagram of $\lambda$ as the cells (or boxes) of this diagram (which is particularly natural when one represents them by boxes) or, briefly, as the cells of $\lambda$. Notation like “west”, “north”, “left”, “right”, “row” and “column” concerning cells of Ferrers diagrams normally refers to their visual representation.

Ferrers diagrams are also known as Young diagrams.

One can characterize the Ferrers diagrams of partitions as follows: A finite subset $S$ of $\{1, 2, 3, \ldots\}^2$ is the Ferrers diagram of some partition if and only if for every $(i, j) \in S$ and every $(i', j') \in \{1, 2, 3, \ldots\}^2$ satisfying $i' \leq i$ and $j' \leq j$, we have $(i', j') \in S$. In other words, a finite subset $S$ of $\{1, 2, 3, \ldots\}^2$ is the Ferrers diagram of some partition if and only if it is a lower set of the poset $\{1, 2, 3, \ldots\}^2$ with respect to the componentwise order.

\(^{48}\)To visually represent a column-strict tableau $T$ of shape $\lambda$, we draw the same picture as when representing the Ferrers diagram of $\lambda$, but with a little difference: a cell $(i, j)$ is no longer represented by a dot or box, but instead is represented by the entry of $T$ assigned to this cell. Accordingly, the entry of $T$ assigned to a given cell $c$ is often referred to as the entry of $T$ in $c$.\)
In fact, one has these transition matrices for \( n = 3 \) expressing elements in terms of the monomial basis \( m_\lambda \):

\[
\begin{aligned}
&\begin{pmatrix}
1 & 1 & 1 \\
0 & 1 & 3 \\
0 & 0 & 6
\end{pmatrix} & & \begin{pmatrix}
1 & 1 & 1 \\
0 & 1 & 3 \\
1 & 3 & 6
\end{pmatrix} & & \begin{pmatrix}
0 & 0 & 1 \\
0 & 1 & 3 \\
1 & 2 & 3
\end{pmatrix} & & \begin{pmatrix}
0 & 0 & 1 \\
1 & 1 & 0 \\
1 & 3 & 6
\end{pmatrix}
\end{aligned}
\]

Our next goal is to show that \( e_\lambda, s_\lambda, h_\lambda \) (and, under some conditions, the \( p_\lambda \) as well) all give bases for \( \Lambda \).

However at the moment it is not yet even clear that \( s_\lambda \) are symmetric!

**Proposition 2.11.** Schur functions \( s_\lambda \) are symmetric, that is, they lie in \( \Lambda \).

**Proof.** It suffices to show \( s_\lambda \) is symmetric under swapping the variables \( x_i, x_{i+1} \), by providing an involution \( \iota \) on the set of all column-strict tableaux \( T \) of shape \( \lambda \) which switches the \( \text{cont}(T) \) for \( (i, i + 1) \text{cont}(T) \). Restrict attention to the entries \( i, i + 1 \) in \( T \), which must look something like this:

\[
\begin{array}{cccccccc}
& i & i & i & i & i & i & \ \\
i + 1 & i + 1 & i & i & i + 1 & i + 1 & i + 1 & i + 1
\end{array}
\]

One finds several vertically aligned pairs \( i, i + 1 \). If one were to remove all such pairs, the remaining entries would be a sequence of rows, each looking like this:

\[
\begin{array}{ccccccc}
& i, \ldots, i, & i + 1, \ldots, i + 1 \\
s \text{occurrences} & r \text{occurrences}
\end{array}
\]

An involution due to Bender and Knuth tells us to leave fixed all the vertically aligned pairs \( i, i + 1 \) but change each sequence as in (2.8) to this:

\[
\begin{array}{ccccccc}
& i, \ldots, i, & i + 1, \ldots, i + 1 \\
s \text{occurrences} & r \text{occurrences}
\end{array}
\]

For example, the above configuration in \( T \) would change to

\[
\begin{array}{ccccccc}
i & i & i & i + 1 & i + 1 & i + 1 & i + 1 \\
i & i + 1 & i + 1
\end{array}
\]

It is easily checked that this map is an involution, and that it has the effect of swapping \( (i, i + 1) \) in \( \text{cont}(T) \).

**Remark 2.12.** The symmetry of Schur functions allows one to reformulate them via column-strict tableaux defined with respect to *any* total ordering \( \mathcal{L} \) on the positive integers, rather than the usual \( 1 < 2 < 3 < \cdots \).

For example, one can use the reverse order\(^{49}\) \( 49 < \cdots < 3 < 2 < 1 \), or even more exotic orders, such as

\[
1 < 3 < 5 < 7 < \cdots < 2 < 4 < 6 < 8 < \cdots
\]

Say that an assignment \( T \) of entries in \( \{1, 2, 3, \ldots\} \) to the cells of the Ferrers diagram of \( \lambda \) is an \( \mathcal{L} \)-column-strict tableau if it is weakly \( \mathcal{L} \)-increasing left-to-right in rows, and strictly \( \mathcal{L} \)-increasing top-to-bottom in columns.

\(^{49}\)This reverse order is what one uses when one defines a Schur function as a generating function for reverse semistandard tableaux or column-strict plane partitions; see Stanley [157, Proposition 7.10.4].
Proposition 2.13. For any total order $\mathcal{L}$ on the positive integers,
\begin{equation}
    s_\lambda = \sum_T x^{\text{cont}(T)}
\end{equation}
as $T$ runs through all $\mathcal{L}$-column-strict tableaux of shape $\lambda$.

Proof. Given a weak composition $\alpha = (\alpha_1, \alpha_2, \ldots)$ with $\alpha_{n+1} = \alpha_{n+2} = \cdots = 0$, assume that the integers $1, 2, \ldots, n$ are totally ordered by $\mathcal{L}$ as $w(1) <_{\mathcal{L}} \cdots <_{\mathcal{L}} w(n)$ for some $w \in \mathcal{S}_n$. Then the coefficient of $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ on the right side of (2.9) is the same as the coefficient of $x^{w^{-1}(\alpha)}$ on the right side of (2.7) defining $s_\lambda$, which by symmetry of $s_\lambda$ is the same as the coefficient of $x^\alpha$ on the right side of (2.7). \hfill $\Box$

It is now not hard to show that $p_\lambda, e_\lambda, s_\lambda$ give bases by a triangularity argument. For this purpose, let us introduce a useful partial order on partitions.

Definition 2.14. The dominance or majorization partial order on $\text{Par}_n$, written $\lambda \triangleright \mu$, is defined by
\[ \lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k \quad \text{for } k = 1, 2, \ldots, n. \]

(The definition of dominance would not change if we would replace “for $k = 1, 2, \ldots, n$” by “for every positive integer $k$” or by “for every $k \in \mathbb{N}$.”)

Definition 2.15. For a partition $\lambda$, its conjugate or transpose partition $\lambda'$ is the one whose Ferrers diagram is obtained from that of $\lambda$ by exchanging rows for columns. Alternatively, one has this formula for its $i^{th}$ part:
\begin{equation}
    (\lambda')_i := |\{ j : \lambda_j \geq i \}|.
\end{equation}

Exercise 2.16. Let $\lambda, \mu \in \text{Par}_n$. Show that $\lambda \triangleright \mu$ if and only if $\lambda' \triangleright \mu'$.

Proposition 2.17. The sets $\{e_\lambda\}, \{s_\lambda\}$ as $\lambda$ runs through all partitions give $k$-bases for $\Lambda_k$ for any commutative ring $k$. The same holds for $\{p_\lambda\}$ when $Q$ is a subring of $k$.

Proof. One can restrict attention to each homogeneous component $\Lambda_n$ and partitions $\lambda$ of $n$. We check that in each case, the proposed basis expands triangularly in the $\{m_\lambda\}$ with some choice of orderings on $\text{Par}_n$ indexing the rows and columns, as illustrated in Example 2.10. Moreover, we check that the diagonal entries of this triangular expansion are invertible, which then yields that the proposed basis is indeed a basis.

One has $s_\lambda = \sum_\mu K_{\lambda, \mu} m_\mu$ where the coefficient $K_{\lambda, \mu}$ is the Koszul number counting the column-strict tableaux $T$ of shape $\lambda$ having $\text{cont}(T) = \mu$; this follows because both sides are symmetric functions, and $K_{\lambda, \mu}$ is the coefficient of $x^\mu$ on both sides. Since for each positive integer $k$, the entries $1, 2, \ldots, k$ in $T$ must all lie within the first $k$ rows of $\lambda$, one has that $\lambda_1 + \lambda_2 + \cdots + \lambda_k \geq \mu_1 + \mu_2 + \cdots + \mu_k$, that is, $\lambda \triangleright \mu$ (if $K_{\lambda, \mu} \neq 0$). Hence, the expansion $s_\lambda = \sum_\mu K_{\lambda, \mu} m_\mu$ becomes $s_\lambda = \sum_{\mu, \lambda, \mu} K_{\lambda, \mu} m_\mu$. One can also check that $K_{\lambda, \lambda} = 1$, so this expansion is untriangular with appropriate ordering of rows and columns.

Before we show $\{e_\lambda\}$ is a basis, we define a few notations regarding integer matrices. A $\{0, 1\}$-matrix means a matrix whose entries belong to the set $\{0, 1\}$. If $A \in \mathbb{N}^{m \times m}$ is a matrix, then the row sums of $A$ means the $\ell$-tuple $(r_1, r_2, \ldots, r_\ell)$, where each $r_i$ is the sum of all entries in the $i$-th row of $A$; similarly, the column sums of $A$ means the $\ell$-tuple $(c_1, c_2, \ldots, c_\ell)$, where each $c_j$ is the sum of all entries in the $j$-th column of $A$.

One has $e_\lambda = \sum_{\mu} a_{\lambda, \mu} m_\mu$ where $a_{\lambda, \mu}$ counts $\{0, 1\}$-matrices having row sums $\lambda$ and column sums $\mu$: when one expands $e_\lambda e_{\lambda_2} \cdots$, choosing the monomial $x_{j_1} \cdots x_{j_\lambda}$ in the $e_\lambda$ factor corresponds to putting $1$’s in the $i^{th}$ row and columns $j_1, \ldots, j_\lambda$ of the $\{0, 1\}$-matrix. It is not hard to check $^{50}$ that $a_{\lambda, \mu}$ vanishes unless $\lambda' \triangleright \mu$. One can also check that $a_{\lambda, \lambda'} = 1$, so this expansion is again untriangular with appropriate ordering of rows and columns.

Assume now that $Q$ is a subring of $k$. One has $p_\lambda = \sum_\mu b_{\lambda, \mu} m_\mu$ where $b_{\lambda, \mu}$ counts the ways to partition the nonzero parts $\lambda_1, \ldots, \lambda_\ell$ into blocks such that the sums of the blocks give $\mu$; more formally, $b_{\lambda, \mu}$ is the number of maps $\varphi : \{1, 2, \ldots, \ell\} \rightarrow \{1, 2, 3, \ldots\}$ having $\mu_j = \sum_{i : \varphi(i) = j} \lambda_i$ for $j = 1, 2, \ldots$. Again it is not hard to check that $b_{\lambda, \mu}$ vanishes unless $\lambda \leq \mu$, and hence this expansion is triangular, for appropriate ordering of rows and columns (but not untriangular, as $b_{\lambda, \lambda'} \neq 1$ in general). The diagonal entries $b_{\lambda, \lambda}$ are positive integers and thus invertible in $k$, so $\{p_\lambda\}$ is a basis.

\footnote{This is the easy implication in the Gale-Ryser Theorem.}
Remark 2.18. When \( Q \) is not a subring of \( k \), the family \( \{ p_\lambda \} \) is not (in general) a basis of \( \Lambda_k \); for instance, \( e_2 = \frac{1}{2} (p_{(1,1)} - p_2) \in A_Q \) is not in the \( Z \)-span of this family. However, if we define \( b_{\lambda,\mu} \) as in the above proof, then the \( Z \)-linear span of all \( p_\lambda \) equals the \( Z \)-linear span of all \( b_{\lambda,\mu} m_\lambda \). Indeed, if \( \mu = (\mu_1, \mu_2, \ldots, \mu_k) \) with \( k = \ell(\mu) \), then \( b_{\mu,\mu} \) is the size of the subgroup of \( S_k \) consisting of all permutations \( \sigma \in S_k \) having each \( i \) satisfy \( \mu_{\sigma(i)} = \mu_i \). As a consequence, \( b_{\mu,\mu} \) divides \( b_{\lambda,\mu} \) for every partition \( \mu \) of the same size as \( \lambda \) (because this group acts \(^51\) freely on the set which is enumerated by \( b_{\lambda,\mu} \)), so that by rescaling \( m_\lambda \) with the factor \( b_{\lambda,\lambda} \) we obtain a unitriangular integer change-of-basis matrix from the family \( \{ p_\lambda \}_{\lambda \in \text{Par}_n} \) to the family \( \{ b_{\lambda,\lambda} m_\lambda \}_{\lambda \in \text{Par}_n} \).

2.3. Comultiplications. Thinking about comultiplication \( \Lambda \rightarrow \Lambda \otimes \Lambda \) on Schur functions forces us to immediately confront the following.

Definition 2.19. For partitions \( \mu, \lambda \) say that \( \mu \subseteq \lambda \) if \( \mu_i \leq \lambda_i \) for \( i = 1, 2, \ldots \), so the Ferrers diagram for \( \mu \) is a subset of the cells for the Ferrers diagram of \( \lambda \). In this case, define the skew (Ferrers) diagram \( \lambda/\mu \) to be their set difference.\(^52\)

Then define the skew Schur function \( s_{\lambda/\mu}(x) \) to be the sum \( s_{\lambda/\mu} := \sum_T x^{\text{cont}(T)} \), where the sum ranges over all column-strict tableaux \( T \) of shape \( \lambda/\mu \); that is, assignments of a value in \( \{1, 2, 3, \ldots \} \) to each cell of \( \lambda/\mu \), weakly increasing left-to-right in rows, and strictly increasing top-to-bottom in columns.

Example 2.20.

\[
\begin{array}{ccc}
\cdot & \cdot & 2 & 5 \\
2 & 2 & 4 \\
4 & 5 \\
\end{array}
\]

is a column-strict tableau of shape \( \lambda/\mu = (5, 3, 2)/(3, 1, 0, 0) \) and it has \( x^{\text{cont}(T)} = x_1^2 x_2^4 x_3^2 x_5^2 \).

Remark 2.21. If \( \mu \) and \( \lambda \) are partitions such that \( \mu \subseteq \lambda \), then \( s_{\lambda/\mu} \in \Lambda \). (This is proven similarly as Proposition 2.11.) Actually, if \( \mu \subseteq \lambda \), then \( s_{\lambda/\mu} \in \Lambda\{\lambda/\mu\} \), where \( |\lambda/\mu| \) denotes the number of cells of the skew shape \( \lambda/\mu \) (so \( |\lambda/\mu| = |\lambda| - |\mu| \)).

It is customary to define \( s_{\lambda/\mu} \) to be 0 if we don’t have \( \mu \subseteq \lambda \). This can also be seen by a literal reading of the definition \( s_{\lambda/\mu} := \sum_T x^{\text{cont}(T)} \), as long as we understand that there are no column-strict tableaux of shape \( \lambda/\mu \) when \( \lambda/\mu \) is not defined.

Clearly, every partition \( \lambda \) satisfies \( s_\lambda = s_{\lambda/\emptyset} \).

Exercise 2.22. (a) State and prove an analogue of Proposition 2.13 for skew Schur functions.

(b) Let \( \lambda, \mu, \lambda' \) and \( \mu' \) be partitions such that \( \mu \subseteq \lambda \) and \( \mu' \subseteq \lambda' \). Assume that the skew Ferrers diagram \( \lambda'/\mu' \) can be obtained from the skew Ferrers diagram \( \lambda/\mu \) by a 180° rotation.\(^53\) Prove that \( s_{\lambda/\mu} = s_{\lambda'/\mu'} \).

Exercise 2.23. Let \( \lambda \) and \( \mu \) be two partitions, and let \( k \in \mathbb{N} \) be such that\(^54\) \( \mu_k \geq \lambda_{k+1} \). Let \( F \) be the skew Ferrers diagram \( \lambda/\mu \). Let \( F_{\text{rows} \leq k} \) denote the subset of \( F \) consisting of all \( (i, j) \in F \) satisfying \( i \leq k \). Let \( F_{\text{rows} > k} \) denote the subset of \( F \) consisting of all \( (i, j) \in F \) satisfying \( i > k \). Let \( \alpha \) and \( \beta \) be two partitions such that \( \beta \subseteq \alpha \) and such that the skew Ferrers diagram \( \alpha/\beta \) can be obtained from \( F_{\text{rows} \leq k} \) by parallel...
translates. Let $\gamma$ and $\delta$ be two partitions such that $\delta \subseteq \gamma$ and such that the skew Ferrers diagram $\gamma/\delta$ can be obtained from $F_{\text{row}}^{-k}$ by parallel translation.\footnote{This is an example of the situation: $\lambda = (6,5,5,2,2), \mu = (4,4,3,1), k = 3$ (satisfying $\mu_k = \mu_3 = 3 \geq 2 = \lambda_4 = \lambda_{k+1}$), $\alpha = (3,2,2), \beta = (1,1), \gamma = (2,2)$, and $\delta = (1)$.}

\textbf{Proposition 2.24.} The comultiplication $\Lambda \xrightarrow{\Delta} \Lambda \otimes \Lambda$ has the following effect on the symmetric functions discussed so far:\footnote{The abbreviated summation indexing $\sum_{i+j=n}t_{i,j}$ used here is intended to mean $\sum_{0 \leq i,j \leq n}$, $i+j=n$.}

(i) $\Delta p_n = 1 \otimes p_n + p_n \otimes 1$ for every $n \geq 1$, that is, the power sums $p_n$ are primitive.
(ii) $\Delta e_n = \sum_{i+j=n} e_i \otimes e_j$ for every $n \in \mathbb{N}$.
(iii) $\Delta h_n = \sum_{i+j=n} h_i \otimes h_j$ for every $n \in \mathbb{N}$.
(iv) $\Delta s_\lambda = \sum_{\mu \subseteq \lambda} s_\mu \otimes s_\lambda/\mu$ for any partition $\lambda$.
(v) $\Delta s_\lambda/\nu = \sum_{\nu \subseteq \mu \subseteq \lambda} s_\mu/\nu \otimes s_\lambda/\mu$ for any partitions $\lambda$ and $\nu$.

\textbf{Proof.} Recall that $\Delta$ sends $f(x) \mapsto f(x,y)$, and one can easily check that

(i) $p_n(x,y) = \sum_i x_i^n + \sum_i y_i^n = p_n(x) \cdot 1 + 1 \cdot p_n(y)$
(ii) $e_n(x,y) = \sum_{i+j=n} e_i(x) e_j(y)$
(iii) $h_n(x,y) = \sum_{i+j=n} h_i(x) h_j(y)$

For assertion (iv), note that by (2.9), one has

$$s_\lambda(x,y) = \sum_T (x,y)^{\text{cont}(T)},$$

where the sum is over column-strict tableaux $T$ of shape $\lambda$ having entries in the linearly ordered alphabet $x_1 < x_2 < \cdots < y_1 < y_2 < \cdots$.

\textbf{Example:}

$$T = \begin{array}{cccc}
x_1 & x_1 & x_2 & y_5 \\
y_2 & y_1 & y_1 & y_1 \\
y_4 & y_2 & y_2 & y_4 \\
\end{array}$$

is such a tableau of shape $\lambda = (5,3,3,2)$. Note that the restriction of $T$ to the alphabet $x$ gives a column-strict tableau $T_x$ of some shape $\mu \subseteq \lambda$, and the restriction of $T$ to the alphabet $y$ gives a column-strict tableau $T_y$ of shape $\lambda/\mu$ (e.g. for $T$ in the example above, the tableau $T_y$ appeared in Example 2.20). Consequently, one has

$$s_\lambda(x,y) = \sum_T x^{\text{cont}(T_x)} \cdot y^{\text{cont}(T_y)} = \sum_{\mu \subseteq \lambda} \left( \sum_{T_x} x^{\text{cont}(T_x)} \right) \left( \sum_{T_y} y^{\text{cont}(T_y)} \right) = \sum_{\mu \subseteq \lambda} s_\mu(x) s_{\lambda/\mu}(y).$$

Assertion (v) is obvious in the case when we don’t have $\nu \subseteq \lambda$ (in fact, in this case, both $s_{\lambda/\nu}$ and $\sum_{\mu \subseteq \lambda} s_\mu/\nu \otimes s_{\lambda/\mu}$ are clearly zero). In the remaining case, the proof of assertion (v) is similar to that of
Exercise 2.25. (a) Show that the Hopf algebra $\Lambda$ is cocommutative.
(b) Show that $\Delta s_{\lambda/\mu} = \sum_{\nu \subseteq \mu \subseteq \lambda} s_{\nu} \otimes s_{\lambda/\nu}$ for any partitions $\lambda$ and $\mu$.

Exercise 2.26. Let $n \in \mathbb{N}$. Consider the finite variable set $(x_1, x_2, \ldots, x_n)$ as a subset of $\mathbf{x} = (x_1, x_2, x_3, \ldots)$. Recall that $f(x_1, x_2, \ldots, x_n)$ is a well-defined element of $k[x_1, x_2, \ldots, x_n]$ for every $f \in R(\mathbf{x})$ (and therefore also for every $f \in \Lambda$, since $\Lambda \subseteq R(\mathbf{x})$), according to Exercise 2.2.
(a) Show that any two partitions $\lambda$ and $\mu$ satisfy

$$s_{\lambda/\mu}(x_1, x_2, \ldots, x_n) = \sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu; \text{ all entries of } T \text{ belong to } \{1, 2, \ldots, n\}} x^{\text{cont}(T)}.$$  

(b) If $\lambda$ is a partition having more than $n$ parts (where the word “parts” means “nonzero parts”), then show that $s_{\lambda}(x_1, x_2, \ldots, x_n) = 0$.  

Remark 2.27. An analogue of Proposition 2.17 holds for symmetric polynomials in finitely many variables: Let $N \in \mathbb{N}$. Then, we have
(a) The set $\{m_\lambda(x_1, x_2, \ldots, x_N)\}$, as $\lambda$ runs through all partitions having length $\leq N$, is a basis of the $k$-module $\Lambda(x_1, x_2, \ldots, x_N) = k[x_1, x_2, \ldots, x_N]^S_N$.
(b) For any partition $\lambda$ having length $> N$, we have $m_\lambda(x_1, x_2, \ldots, x_N) = 0$.
(c) The set $\{e_\lambda(x_1, x_2, \ldots, x_N)\}$, as $\lambda$ runs through all partitions whose parts are all $\leq N$, is a basis of the $k$-module $\Lambda(x_1, x_2, \ldots, x_N)$.
(d) The set $\{s_\lambda(x_1, x_2, \ldots, x_N)\}$, as $\lambda$ runs through all partitions having length $\leq N$, is a basis of the $k$-module $\Lambda(x_1, x_2, \ldots, x_N)$.
(e) If $\mathbb{Q}$ is a subring of $k$, then the set $\{p_\lambda(x_1, x_2, \ldots, x_N)\}$, as $\lambda$ runs through all partitions having length $\leq N$, is a basis of the $k$-module $\Lambda(x_1, x_2, \ldots, x_N)$.
(f) If $\mathbb{Q}$ is a subring of $k$, then the set $\{p_\lambda(x_1, x_2, \ldots, x_N)\}$, as $\lambda$ runs through all partitions whose parts are all $\leq N$, is a basis of the $k$-module $\Lambda(x_1, x_2, \ldots, x_N)$.

Indeed, the claims (a) and (b) are obvious, while the claims (c), (d) and (e) are proven similarly to our proof of Proposition 2.17. We leave the proof of (f) to the reader.

Claim (c) can be rewritten as follows: The elementary symmetric polynomials $e_i(x_1, x_2, \ldots, x_N)$, for $i \in \{1, 2, \ldots, N\}$, form an algebraically independent generating set of $\Lambda(x_1, x_2, \ldots, x_N)$. This is precisely the well-known theorem (due to Gauss)\(^{58}\) that every symmetric polynomial in $N$ variables $x_1, x_2, \ldots, x_N$ can be written uniquely as a polynomial in the $N$ elementary symmetric polynomials.

2.4. The antipode, the involution $\omega$, and algebra generators. Since $\Lambda$ is a graded connected $k$-bialgebra, it will have an antipode $\Sigma^S \Lambda \rightarrow \Lambda$ making it a Hopf algebra by Proposition 1.36. However, several issues will be resolved by identifying $S$ more explicitly now.

Proposition 2.28. Each of $\{e_n\}_{n=1,2,\ldots}$, $\{h_n\}_{n=1,2,\ldots}$ are algebraically independent, and generate $\Lambda_k$ as a polynomial algebra for any commutative ring $k$. The same holds for $\{p_n\}_{n=1,2,\ldots}$ when $\mathbb{Q}$ is a subring of $k$.

Furthermore, the antipode $S$ acts as follows:
(i) $S(p_n) = -p_n$
(ii) $S(e_n) = (-1)^n h_n$
(iii) $S(h_n) = (-1)^n e_n$

Proof. The assertions that $\{e_n\}, \{p_n\}$ are algebraically independent and generate $\Lambda$ are equivalent to Proposition 2.17 asserting $\{e_\lambda\}, \{p_\lambda\}$ give bases for $\Lambda$. The assertion $S(p_n) = -p_n$ follows from Proposition 1.37 since $p_n$ is primitive by Proposition 2.24(i).

\(^{58}\)See, e.g., [33, Symmetric Polynomials, Theorem 5 and Remark 17].
For the remaining assertions, start with the easy generating function identities

\[ H(t) := \prod_{i=1}^{\infty} (1 - x_i t)^{-1} = 1 + h_1(x) t + h_2(x) t^2 + \cdots = \sum_{n \geq 0} h_n(x) t^n \]

\[ E(t) := \prod_{i=1}^{\infty} (1 + x_i t) = 1 + e_1(x) t + e_2(x) t^2 + \cdots = \sum_{n \geq 0} e_n(x) t^n \]

which shows that

\[ 1 = E(-t) H(t) = \left( \sum_{n \geq 0} e_n(x)(-t)^n \right) \left( \sum_{n \geq 0} h_n(x) t^n \right) \]

and hence, equating coefficients of powers of \( t \), that for \( n = 0, 1, 2, \ldots \) one has

\[ \sum_{i+j=n} (-1)^i e_i h_j = \delta_{0,n}. \]

This lets one recursively express the \( e_n \) in terms of \( h_n \) and vice-versa:

\begin{align*}
e_0 & := 1 =: h_0 \\
e_n & = e_{n-1} h_1 - e_{n-2} h_2 + e_{n-3} h_3 - \cdots \\
h_n & = h_{n-1} e_1 - h_{n-2} e_2 + h_{n-3} e_3 - \cdots
\end{align*}

for \( n = 1, 2, \ldots \) Thus if one uses the algebraic independence of the generators \( \{e_n\} \) for \( \Lambda \) to define an algebra endomorphism as follows

\[ \Lambda \xrightarrow{\omega} \Lambda \]

\[ e_n \xrightarrow{\omega} h_n, \]

then the identical form of the two recursions in (2.18) shows that \( \omega \) also sends \( h_n \mapsto e_n \). Therefore \( \omega \) is an involutive automorphism of \( \Lambda \), and the \( \{h_n\} \) are another algebraically independent generating set for \( \Lambda \).

For the assertion about the antipode \( S \) applied to \( e_n \) or \( h_n \), note that the coproduct formulas for \( e_n, h_n \) in Proposition 2.24(ii),(iii) show that the defining relations for their antipodes (1.13) will in this case be

\[ \sum_{i+j=n} S(e_i)e_j = \delta_{0,n} = \sum_{i+j=n} e_i S(e_j), \]

\[ \sum_{i+j=n} S(h_i)h_j = \delta_{0,n} = \sum_{i+j=n} h_i S(h_j) \]

because \( u\epsilon(e_n) = u\epsilon(h_n) = \delta_{0,n} \). Comparing these to (2.17), one concludes via induction on \( n \) that \( S(e_n) = (-1)^n h_n \) and \( S(h_n) = (-1)^n e_n \).

The \( k \)-algebra endomorphism \( \omega \) of \( \Lambda \) defined in the proof of Proposition 2.28 is known as the fundamental involution on \( \Lambda \). We record for future use two results that were shown in the above proof: that \( \omega \) is an involution, and that it sends \( e_n \) to \( h_n \) and vice versa for every positive integer \( n \).

Proposition 2.28 shows that the antipode \( S \) on \( \Lambda \) is, up to sign, the same as the fundamental involution \( \omega \): one has

\[ S(f) = (-1)^n \omega(f) \text{ for } f \in \Lambda_n \]

since this formula holds for all elements of the generating set \( \{e_n\} \) (or \( \{h_n\} \)).

**Remark 2.29.** Up to now we have not yet derived how the involution \( \omega \) and the antipode \( S \) act on (skew) Schur functions, which is quite beautiful: If \( \lambda \) and \( \mu \) are partitions satisfying \( \mu \subseteq \lambda \), then

\[ \omega(s_{\lambda/\mu}) = s_{\lambda'}/\mu' \]

\[ S(s_{\lambda/\mu}) = (-1)^{|\lambda/\mu|} s_{\lambda'}/\mu' \]

where recall that \( \lambda' \) is the transpose or conjugate partition to \( \lambda \), and \( |\lambda/\mu| \) is the number of squares in the skew diagram \( \lambda/\mu \), that is, \( |\lambda/\mu| = n-k \) if \( \lambda, \mu \) lie in \( \text{Par}_n, \text{Par}_k \) respectively.

We will deduce this later in three ways (once as an exercise using the Pieri rules in Exercise 2.69, once again using skewing operators in Exercise 2.76, and for the third time from the action of the antipode in
QSym on $P$-partition enumerators in Corollary 5.29). However, one could also deduce it immediately from our knowledge of the action of $\omega$ and $S$ on $e_n, h_n$, if we were to prove the following famous Jacobi-Trudi and dual Jacobi-Trudi formulas:\footnote{The second of the following identities is also known as the von N"agelsbach-Kostka identity.}

**Theorem 2.30.** Skew Schur functions are the following polynomials in $\{h_n, \{e_n\}$:

\begin{align}
\label{2.22}
s_{\lambda/\mu} &= \det(h_{\lambda_i - \mu_j - i + j})_{i,j=1,2,\ldots,\ell} \\
\label{2.23}s_{\lambda'/\mu'} &= \det(e_{\lambda_i - \mu_j - i + j})_{i,j=1,2,\ldots,\ell}
\end{align}

for any two partitions $\lambda$ and $\mu$ and any $\ell \in \mathbb{N}$ satisfying $\ell(\lambda) \leq \ell$ and $\ell(\mu) \leq \ell$.

Since we appear not to need these formulas in the sequel, we will not prove them right away. However, a proof is sketched in the solution to Exercise 2.71, and various proofs are well-explained in [95, (39) and (41)], [109, §I.5], [139, Thm. 7.1], [141, §4.5], [157, §7.16], [169, Thms. 3.5 and 3.5∗]; also, a simultaneous generalization of both formulas is shown in [63, Theorem 11], and two others in [136, 1.9] and [64, Thm. 3.1]. An elegant treatment of Schur polynomials taking the Jacobi-Trudi formula (2.22) as the definition of $s_\lambda$ is given by Tamvakis [165].

2.5. **Cauchy product, Hall inner product, self-duality.** The Schur functions, although a bit unmotivated right now, have special properties with regard to the Hopf structure. One property is intimately connected with the following Cauchy identity.

**Theorem 2.31.** In the power series ring $k[[x, y]] := k[[x_1, x_2, \ldots, y_1, y_2, \ldots]]$, one has the following expansion:

\begin{equation}
\label{2.24}
\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \mathbb{P}} s_\lambda(x)s_\lambda(y).
\end{equation}

**Remark 2.32.** The left hand side of (2.24) is known as the Cauchy product, or Cauchy kernel.

An equivalent version of the equality (2.24) is obtained by replacing each $x_i$ by $x_i t$, and writing the resulting identity in the power series ring $R(x, y)[[t]]$:

\begin{equation}
\label{2.25}
\prod_{i,j=1}^{\infty} (1 - tx_i y_j)^{-1} = \sum_{\lambda \in \mathbb{P}} t^{\lambda} s_\lambda(x)s_\lambda(y).
\end{equation}

(Recall that $|\lambda| = \lambda_1 + \lambda_2 + \cdots + \lambda_\ell$ for any partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$.)

**Proof of Theorem 2.31.** We follow the standard combinatorial proof (see [141, §4.8], [157, §7.11-7.12]), which rewrites the left and right sides of (2.25), and then compares them with the Robinson-Schensted-Knuth (RSK) bijection.\footnote{The RSK bijection has been introduced by Knuth [82], where we call “biletters” is referred to as “two-line arrays”. The most important ingredient of this algorithm – the RS-insertion operation – however goes back to Schensted. The special case of the RSK algorithm where the biletter has to be a permutation (written in two-line notation) and the two tableaux have to be standard (i.e., each of them has content (1$^n$), where $n$ is the size of their shape) is the famous Robinson-Schensted correspondence [99]. More about these algorithms can be found in [141, Chapter 3], [157, §7.11-7.12], [54, Chapters 1 and A], [23, §3, §6] and various other places.} On the left side, expanding out each geometric series

\begin{equation}
(1 - tx_i y_j)^{-1} = 1 + tx_i y_j + (tx_i y_j)^2 + (tx_i y_j)^3 + \cdots
\end{equation}

and thinking of $(x_i y_j)^m$ as $m$ occurrences of a biletter\footnote{A biletter here simply means a pair of letters, written as a column vector. A letter means a positive integer.} ($i_j$), we see that the left hand side can be rewritten as the sum of $t^{\ell}(x_{i_1} y_{j_1}) (x_{i_2} y_{j_2}) \cdots (x_{i_\ell} y_{j_\ell})$ over all multisets $\{i_1^{j_1}, \ldots, i_\ell^{j_\ell}\}$ of biletters. Order the biletters in such a multiset in the lexicographic order $\leq_{lex}$, which is the total order on the set of all biletters defined by

\begin{equation}
\left\{\begin{array}{l}
(i_1) \leq_{lex} (i_2) \\
(j_1) \leq_{lex} (j_2)
\end{array}\right\} \iff \left\{\begin{array}{l}
(i_1) \leq_{lex} (i_2) \\
(j_1) \leq_{lex} (j_2)
\end{array}\right\}
\end{equation}

(we have $i_1 \leq i_2$, and if $i_1 = i_2$, then $j_1 \leq j_2$).

Defining a biword to be an array $^{(i_1)}_{(j_1)}$ in which the biletters are ordered $^{(i_1)}_{(j_1)} \leq_{lex} \cdots \leq_{lex} ^{(i_\ell)}_{(j_\ell)}$, then the left side of (2.25) is the sum $\sum t^{\ell} x^{\text{cont}(i)} y^{\text{cont}(j)}$ over all biwords $^{(i_1)}_{(j_1)}$, where $\ell$ stands for the number
of bileters in the biword. On the right side, expanding out the Schur functions as sums of tableaux gives \( \sum_{(P,Q)} t^i x^\operatorname{cont}(Q) y^\operatorname{cont}(P) \) in which the sum is over all ordered pairs \((P,Q)\) of column-strict tableaux having the same shape\(^{62} \), with \( \ell \) cells. (We shall refer to such pairs as tableau pairs from now on.)

The Robinson-Schensted-Knuth algorithm gives us a bijection between the biwords \( (1^i) \) and the tableau pairs \((P,Q)\), which has the property that
\[
\operatorname{cont}(i) = \operatorname{cont}(Q),
\]
\[
\operatorname{cont}(j) = \operatorname{cont}(P)
\]
(and that the length \( \ell \) of the biword \( (1^i) \) equals the size \(|\lambda|\) of the common shape of \( P \) and \( Q \); but this follows automatically from \( \operatorname{cont}(i) = \operatorname{cont}(Q) \)). Clearly, once such a bijection is constructed, the equality (2.25) will follow.

Before we define this algorithm, we introduce a simpler operation known as \textit{RS-insertion} (short for Robinson-Schensted insertion). RS-insertion takes as input a column-strict tableau \( P \) and a letter \( j \), and returns a new column-strict tableau \( P' \) along with a corner cell\(^{63} \) \( c \) of \( P' \), which is constructed as follows:

- Start out by setting \( P' = P \). The letter \( j \) tries to insert itself into the first row of \( P' \) by either bumping out the leftmost letter in the first row strictly larger than \( j \), or else placing itself at the right end of the row if no such larger letter exists. If a letter was bumped from the first row, this letter follows the same rules to insert itself into the second row, and so on\(^{64} \). This series of bumps must eventually come to an end\(^{65} \). At the end of the bumping, the tableau \( P' \) created has an extra corner cell not present in \( P \). If we call this corner cell \( c \), then \( P' \) (in its final form) and \( c \) are what the RS-insertion operation returns. One says that \( P' \) is the result of \textit{inserting} \( c \) into the tableau \( P \). It is straightforward to see that this resulting filling \( P' \) is a column-strict tableau\(^{66} \).

**Example 2.33.** To give an example of this operation, let us insert the letter \( j = 3 \) into the column-strict tableau

\[
\begin{array}{cccc}
1 & 1 & 3 & 3 \\
2 & 2 & 4 & 6 \\
3 & 4 & 7 & \text{done}
\end{array}
\]

(we are showing all intermediate states of \( P' \); the underlined letter is always the one that is going to be bumped out at the next step):

- \( \text{insert 3;} \)
- \( \text{bump out 4;} \)

\[
\begin{array}{cccc}
1 & 1 & 3 & 3 \\
2 & 2 & 4 & 6 \\
3 & 4 & 7 & \\
5 &
\end{array}
\]

- \( \text{insert 4;} \)
- \( \text{bump out 6;} \)

\[
\begin{array}{cccc}
1 & 1 & 3 & 3 \\
2 & 2 & 4 & 4 \\
3 & 4 & 7 & \\
5 &
\end{array}
\]

- \( \text{insert 7;} \)
- \( \text{done;} \)

\[
\begin{array}{cccc}
1 & 1 & 3 & 3 \\
2 & 2 & 4 & 4 \\
3 & 4 & 6 & \\
5 & 7 &
\end{array}
\]

The last tableau in this sequence is the column-strict tableau that is returned. The corner cell that is returned is the second cell of the fourth row (the one containing 7).

---

\(^{62}\)And this shape should be the Ferrers diagram of a partition (not just a skew partition).

\(^{63}\)A \textit{corner cell} of a tableau or a Ferrers diagram is defined to be a cell \( c \) which belongs to the tableau (resp. diagram) but whose immediate neighbors to the east and to the south don’t.

\(^{64}\)Here, rows are allowed to be empty – so it is possible that a letter is bumped from the last nonempty row of \( P' \) and settles in the next, initially empty, row.

\(^{65}\)since we can only bump out entries from nonempty rows

\(^{66}\)This terminology is reminiscent of insertion into binary search trees, a basic operation in theoretical computer science. This is more than superficial similarity; there are, in fact, various analogies between Ferrers diagrams (and their fillings) and unlabelled plane binary trees (resp. their labellings), and one of them is the analogy between RS-insertion and binary search tree insertion. See [72, §4.1].

\(^{67}\)Indeed, the reader can check that \( P' \) remains a column-strict tableau throughout the algorithm that defines RS-insertion. (The only part of this that isn’t obvious is showing that when a letter \( t \) bumped out of some row \( k \) is inserted into row \( k + 1 \), the property that the letters increase strictly down columns is preserved. Argue that the bumping-out of \( t \) from row \( k \) was caused by the insertion of another letter \( u < t \), and that the cell of row \( k + 1 \) into which \( t \) is then being inserted is in the same column as this \( u \), or in a column further left than it.)
RS-insertion will be used as a step in the RSK algorithm; the construction will rely on a simple fact known as the row bumping lemma. Let us first define the notion of a bumping path (or bumping route): If \( P \) is a column-strict tableau, and \( j \) is a letter, then some letters are inserted into some cells when RS-insertion is applied to \( P \) and \( j \). The sequence of these cells (in the order in which they see letters inserted into them) is called the bumping path for \( P \) and \( j \). This bumping path always ends with the corner cell \( c \) which is returned by RS-insertion. As an example, when \( j = 1 \) is inserted into the tableau \( P \) shown below, the result \( P' \) is shown with all entries on the bumping path underlined:

\[
\begin{array}{cccc}
1 & 1 & 2 & 3 \\
2 & 2 & 4 & 4 \\
3 & 4 & 5 & \\
4 & 6 & 6 & \\
\end{array}
\quad \xrightarrow{\text{insert}} \quad
\begin{array}{cccc}
\text{insert} & \\
1 & 1 & 1 & 2 & 3 \\
2 & 2 & 2 & 4 \\
3 & 4 & 4 & \\
4 & 5 & 6 & \\
6 & & & \\
\end{array}
\]

A first simple observation about bumping paths is that bumping paths trend weakly left – that is, if the bumping path of \( P \) and \( j \) is \((c_1, c_2, \ldots, c_k)\), then, for each \( 1 \leq i < k \), the cell \( c_{i+1} \) lies in the same column as \( c_i \) or in a column further left.\(^{68}\) A subtler property of bumping paths is the following row bumping lemma ([54, p. 9]):

**Row bumping lemma:** Let \( P \) be a column-strict tableau, and let \( j \) and \( j' \) be two letters. Applying RS-insertion to the tableau \( P \) and the letter \( j \) yields a new column-strict tableau \( P' \) and a corner cell \( c \). Applying RS-insertion to the tableau \( P' \) and the letter \( j' \) yields a new column-strict tableau \( P'' \) and a corner cell \( c' \).

(a) Assume that \( j \leq j' \). Then, the bumping path for \( P' \) and \( j' \) stays strictly to the right, within each row, of the bumping path for \( P \) and \( j \). The cell \( c' \) (in which the bumping path for \( P' \) and \( j' \) ends) is in the same row as the cell \( c \) (in which the bumping path for \( P \) and \( j \) ends) or in a row further up; it is also in a column further right than \( c \).

(b) Assume instead that \( j > j' \). Then, the bumping path for \( P' \) and \( j' \) stays weakly to the left, within each row, of the bumping path for \( P \) and \( j \). The cell \( c' \) (in which the bumping path for \( P' \) and \( j' \) ends) is in a row further down than the cell \( c \) (in which the bumping path for \( P \) and \( j \) ends); it is also in the same column as \( c \) or in a column further left.

This lemma can be easily proven by induction over the row.\(^{69}\)

We can now define the actual RSK algorithm. Let \((i,j)\) be a biword. Starting with the pair \((P_0, Q_0) = (\emptyset, \emptyset)\) and \(m = 0\), the algorithm applies the following steps (see Example 2.34 below):

- If \( i_{m+1} \) does not exist (that is, \( m \) is the length of \( i \)), stop.
- Apply RS-insertion to the column-strict tableau \( P_m \) and the letter \( j_{m+1} \) (the bottom letter of \( \binom{m+1}{j_{m+1}} \)).
  - Let \( P_{m+1} \) be the resulting column-strict tableau, and let \( c_{m+1} \) be the resulting corner cell.
- Create \( Q_{m+1} \) from \( Q_m \) by adding the top letter \( i_{m+1} \) of \( \binom{m+1}{i_{m+1}} \) to \( Q_m \) in the cell \( c_{m+1} \) (which, as we recall, is the extra corner cell of \( P_{m+1} \) not present in \( P_m \)).
- Set \( m \) to \( m + 1 \).

After all of the biletters have been thus processed, the result of the RSK algorithm is \((P_r, Q_\ell) =: (P, Q)\).

\(^{68}\)This follows easily from the preservation of column-strictness during RS-insertion.

\(^{69}\)We leave the details to the reader, only giving the main idea for (a) (the proof of (b) is similar). To prove the first claim of (a), it is enough to show that for every \( i \), if any letter is inserted into row \( i \) during RS-insertion for \( P' \) and \( j' \), then some letter is also inserted into row \( i \) during RS-insertion for \( P \) and \( j \), and the former insertion happens in a cell strictly to the right of the cell where the latter insertion happens. This follows by induction over \( i \). In the induction step, we need to show that if, for a positive integer \( i \), we try to consecutively insert two letters \( k \) and \( k' \), in this order, into the \( i \)-th row of a column-strict tableau, possibly bumping out existing letters in the process, and if we have \( k \leq k' \), then the cell into which \( k \) is inserted is strictly to the left of the cell into which \( k' \) is inserted, and the letter bumped out by the insertion of \( k \) is \( \leq \) to the letter bumped out by the insertion of \( k' \) (or else the insertion of \( k' \) bumps out no letter at all – but it cannot happen that \( k' \) bumps out a letter but \( k \) does not). This statement is completely straightforward to check (by only studying the \( i \)-th row). This way, the first claim of (a) is proven, and this entails that the cell \( c' \) (being the last cell of the bumping path for \( P' \) and \( j' \)) is in the same row as the cell \( c \) or in a row further up. It only remains to show that \( c' \) is in a column further right than \( c \). This follows by noticing that, if \( k \) is the row in which the cell \( c' \) lies, then \( c' \) is in a column further right than the entry of the bumping path for \( P \) and \( j \) in row \( k \) (by the first claim of (a)), and this latter entry is further right than or in the same column as the ultimate entry \( c \) of this bumping path (since bumping paths trend weakly left).
Example 2.34. The term in the expansion of the left side of (2.24) corresponding to

\[(x_1 y_2)^1 (x_1 y_1)^1 (x_2 y_1)^1 (x_3 y_1)^1 (x_4 y_3)^2 (x_5 y_2)^1\]

is the biword \( (_1^1) = (124445)_{(2411332)} \), whose RSK algorithm goes as follows:

\[
\begin{align*}
P_0 &= \emptyset & Q_0 &= \emptyset \\
P_1 &= 2 & Q_1 &= 1 \\
P_2 &= 2 4 & Q_2 &= 1 1 \\
P_3 &= 1 4 & Q_3 &= 1 1 \\
P_4 &= 1 1 4 & Q_4 &= 1 1 \\
P_5 &= 1 1 3 2 4 & Q_5 &= 1 1 4 2 4 \\
P_6 &= 1 1 3 3 2 4 & Q_6 &= 1 1 4 4 2 4 \\
P := P_7 &= 2 3 4 1 1 2 3 & Q := Q_7 &= 2 4 1 1 4 4 5
\end{align*}
\]

The bumping rule obviously maintains the property that \( P_m \) is a column-strict tableau of some Ferrers shape throughout. It should be clear that \( (P_m, Q_m) \) have the same shape at each stage. Also, the construction of \( Q_m \) shows that it is at least weakly increasing in rows and weakly increasing in columns throughout. What is perhaps least clear is that \( Q_m \) remains strictly increasing down columns. That is, when one has a string of equal letters on top \( i_m = i_{m+1} = \cdots = i_{m+r} \), so that on bottom one bumps in \( j_m \leq j_{m+1} \leq \cdots \leq j_{m+r} \), one needs to know that the new cells form a horizontal strip, that is, no two of them lie in the same column. This follows from (the last claim of) part (a) of the row bumping lemma. Hence, the result \( (P, Q) \) of the RSK algorithm is a tableau pair.

To see that the RSK map is a bijection, we show how to recover \( (_1^1) \) from \( (P, Q) \). This is done by reverse bumping from \( (P_{m+1}, Q_{m+1}) \) to recover both the biletter \( (_{m+1}^{m+1}) \) and the tableaux \( (P_m, Q_m) \), as follows. Firstly, \( i_{m+1} \) is the maximum entry of \( Q_{m+1} \), and \( Q_m \) is obtained by removing the rightmost occurrence of this letter \( i_{m+1} \) from \( Q_{m+1} \). To produce \( P_m \) and \( j_{m+1} \), find the position of the rightmost occurrence of \( i_{m+1} \) in \( Q_{m+1} \), and start reverse bumping in \( P_{m+1} \) from the entry in this same position, where reverse bumping an entry means inserting it into one row higher by having it bump out the rightmost entry which is strictly smaller. The entry bumped out of the first row is \( j_{m+1} \), and the resulting tableau is \( P_m \).

---

70Actually, each of these new cells (except for the first one) is in a column further right than the previous one. We will use this stronger fact further below.

71It necessarily has to be the rightmost occurrence, since (according to the previous footnote) the cell into which \( i_{m+1} \) was filled at the step from \( Q_m \) to \( Q_{m+1} \) lies further right than any existing cell of \( Q_m \) containing the letter \( i_{m+1} \).

72Let us give a few more details on this “reverse bumping” procedure. Reverse bumping (also known as RS-deletion or reverse RS-insertion) is an operation which takes a column-strict tableau \( P' \) and a corner cell \( c \) of \( P' \), and constructs a column-strict tableau \( P \) and a letter \( j \) such that RS-insertion for \( P \) and \( j \) yields \( P' \) and \( c \). It starts by setting \( P = P' \), and removing the entry in the cell \( c \) from \( P \). This removed entry is then denoted by \( k \), and is inserted into the row of \( P \) above \( c \), bumping out the rightmost entry which is smaller than \( k \). The letter which is bumped out -- say, \( \ell \) --, in turn, is inserted into the row above it, bumping out the rightmost entry which is smaller than \( \ell \). This procedure continues in the same way until an entry is bumped out of the first row (which will eventually happen). The reverse bumping operation returns the resulting tableau \( P \) and the entry which is bumped out of the first row.

It is straightforward to check that the reverse bumping operation is well-defined (i.e., \( P \) does stay a column-strict tableau throughout the procedure) and is the inverse of the RS-insertion operation. (In fact, these two operations undo each other step by step.)
Finally, to see that the RSK map is surjective, one needs to show that the reverse bumping procedure can be applied to any pair \((P,Q)\) of column-strict tableaux of the same shape, and will result in a (lexicographically ordered) biword \((\updownarrow)\). We leave this verification to the reader.\(^{73}\)

This is by far not the only known proof of Theorem 2.31. Two further proofs will be sketched in Exercise 2.68 and Exercise 2.66.

**Corollary 2.35.** In the Schur function basis \(\{s_\lambda\}\) for \(\Lambda\), the structure constants for multiplication and comultiplication are the same, that is, if one defines \(c_{\mu,\nu}^\lambda, \tilde{c}_{\mu,\nu}^\lambda\) via the unique expansions

\[
\begin{align*}
    s_\mu s_\nu &= \sum_\lambda c_{\mu,\nu}^\lambda s_\lambda \\
    \Delta(s_\lambda) &= \sum_{\mu,\nu} \tilde{c}_{\mu,\nu}^\lambda s_\mu \otimes s_\nu
\end{align*}
\]

(2.26)

then \(c_{\mu,\nu}^\lambda = \tilde{c}_{\mu,\nu}^\lambda\).

**Proof.** Work in the ring \(k[[x,y,z]]\), where \(y = (y_1, y_2, y_3, \ldots)\) and \(z = (z_1, z_2, z_3, \ldots)\) are two new sets of variables. The identity (2.24) lets one interpret both \(c_{\mu,\nu}^\lambda, \tilde{c}_{\mu,\nu}^\lambda\) as the coefficient\(^{74}\) of \(s_\mu(x) s_\nu(y) s_\lambda(z)\) in the product

\[
\prod_{i,j=1}^\infty (1 - x_i z_j)^{-1} \prod_{i,j=1}^\infty (1 - y_i z_j)^{-1} \sum_\lambda s_\lambda(x,y) s_\lambda(z) = \sum_{\mu,\nu} c_{\mu,\nu}^\lambda s_\mu(x) s_\nu(y) s_\lambda(z)
\]

since, regarding \(x_1, x_2, \ldots, y_1, y_2, \ldots\) as lying in a single variable set \((x,y)\), separate from the variables \(z\), the Cauchy identity (2.24) expands the same product as

\[
\prod_{i,j=1}^\infty (1 - x_i z_j)^{-1} \prod_{i,j=1}^\infty (1 - y_i z_j)^{-1} = \sum_\lambda \left(\sum_{\mu,\nu} \tilde{c}_{\mu,\nu}^\lambda s_\mu(x) s_\nu(y)\right) s_\lambda(z).
\]

\[\square\]

**Definition 2.36.** The coefficients \(c_{\mu,\nu}^\lambda = \tilde{c}_{\mu,\nu}^\lambda\) appearing in the expansions (2.26) are called Littlewood-Richardson coefficients.

**Remark 2.37.** We will interpret \(c_{\mu,\nu}^\lambda\) combinatorially in Section 2.6. By now, however, we can already prove some properties of these coefficients:

We have

\[
\begin{align*}
    c_{\mu,\nu}^\lambda &= c_{\nu,\mu}^\lambda \quad \text{for all } \lambda, \mu, \nu \in \text{Par}
\end{align*}
\]

(by comparing coefficients in \(\sum_\lambda c_{\mu,\nu}^\lambda s_\lambda = s_\mu s_\nu = s_\nu s_\mu = \sum_\lambda \tilde{c}_{\mu,\nu}^\lambda s_\lambda\)). Also, noting on one hand the expansion

\[
\begin{align*}
    s_\lambda(x,y) &= \Delta(s_\lambda) = \sum_{\mu,\nu} c_{\mu,\nu}^\lambda s_\mu(x) s_\nu(y)
\end{align*}
\]

\[\square\]

\(^{73}\)It is easy to see that repeatedly applying reverse bumping to \((P,Q)\) will result in a sequence \((i_1, i_2, \ldots, i_{\ell-1}, i_{\ell})\) of biletters such that applying the RSK algorithm to \((i_1, i_2, \ldots, i_{\ell})\) gives back \((P,Q)\). The question is why we have \(i_1 \leq i_2 \leq \cdots \leq i_{\ell}\). Since the chain of inequalities \(i_1 \leq i_2 \leq \cdots \leq i_{\ell}\) is clear from the choice of entry to reverse-bump, it only remains to show that for every string \(i_m = i_{m+1} = \cdots = i_{m+r}\) of equal top letters, the corresponding bottom letters weakly increase (that is, \(j_m \leq j_{m+1} \leq \cdots \leq j_{m+r}\)). One way to see it is to apply part (b) of the row bumping lemma.

\(^{74}\)Speaking of coefficients makes sense here since \(\Lambda(x) \otimes \Lambda(y) \otimes \Lambda(z)\) embeds into \(k[[x,y,z]]\).
and on the other hand

\[ s_\lambda(x, y) = \sum_{\mu \subseteq \lambda} s_\mu(x)s_{\lambda/\mu}(y) = \sum_{\mu \in \text{Par}} s_\mu(x)s_{\lambda/\mu}(y) \]

one concludes another standard interpretation for \( c_{\mu,\nu}^\lambda \):

\[ s_{\lambda/\mu} = \sum_{\nu} c_{\mu,\nu}^\lambda s_\nu \]

whenever \( \lambda \) and \( \mu \) are two partitions (not necessarily satisfying \( \mu \subseteq \lambda \)). In particular, \( c_{\mu,\nu}^\lambda \) vanishes unless \( \mu \subseteq \lambda \). Consequently, \( c_{\mu,\nu}^\lambda \) vanishes unless \( \nu \subseteq \lambda \) as well (since \( c_{\mu,\nu}^\lambda = c_{\nu,\mu}^\lambda \)) and furthermore vanishes unless the equality \( |\mu| + |\nu| = |\lambda| \) holds. Altogether, we conclude that \( c_{\mu,\nu}^\lambda \) vanishes unless \( \mu, \nu \subseteq \lambda \) and \( |\mu| + |\nu| = |\lambda| \).

**Exercise 2.38.** Show that any four partitions \( \kappa, \lambda, \varphi \) and \( \psi \) satisfy

\[ \sum_{\rho \in \text{Par}} c_{\kappa,\lambda}^{\rho} c_{\varphi,\psi}^{\rho} = \sum_{(\alpha,\beta,\gamma,\delta) \in \text{Par}^4} c_{\alpha,\beta}^{\gamma} c_{\alpha,\gamma}^{\delta} c_{\beta,\delta}^{\psi} c_{\gamma,\delta}^{\psi}. \]

**Exercise 2.39.**

(a) For any partition \( \mu \), prove that

\[ \sum_{\lambda \in \text{Par}} s_\lambda(x) s_{\lambda/\mu}(y) = s_\mu(x) \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} \]

in the power series ring \( \mathbf{k}[[x,y]] = \mathbf{k}[[x_1, x_2, x_3, \ldots, y_1, y_2, y_3, \ldots]] \).

(b) Let \( \alpha \) and \( \beta \) be two partitions. Show that

\[ \sum_{\lambda \in \text{Par}} s_{\lambda/\alpha}(x) s_{\lambda/\beta}(y) = \left( \sum_{\rho \in \text{Par}} s_{\rho/\beta}(y) s_{\alpha/\rho}(x) \right) \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} \]

in the power series ring \( \mathbf{k}[[x,y]] = \mathbf{k}[[x_1, x_2, x_3, \ldots, y_1, y_2, y_3, \ldots]] \).

[Hint: For (b), expand the product

\[ \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} \prod_{i,j=1}^{\infty} (1 - x_i w_j)^{-1} \prod_{i,j=1}^{\infty} (1 - z_i y_j)^{-1} \prod_{i,j=1}^{\infty} (1 - z_i w_j)^{-1} \]

in the power series ring \( \mathbf{k}[[x_1, x_2, x_3, \ldots, y_1, y_2, y_3, \ldots, z_1, z_2, z_3, \ldots, w_1, w_2, w_3, \ldots]] \) in two ways: once by applying Theorem 2.31 to the two variable sets \( (z, x) \) and \( (w, y) \) and then using (2.13); once again by applying (2.24) to the two variable sets \( z \) and \( w \) and then applying Exercise 2.39(a) twice.]

The statement of Exercise 2.39(b) is known as the *skew Cauchy identity*, and appears in Sagan-Stanley [142, Cor. 6.12], Stanley [157, exercise 7.27(c)] and Macdonald [109, §I.5, example 26]; it seems to be due to Zelevinsky. It generalizes the statement of Exercise 2.39(a), which in turn is a generalization of Theorem 2.31.

**Definition 2.40.** Define the *Hall inner product* on \( \Lambda \) to be the \( \mathbf{k} \)-bilinear form \( (\cdot, \cdot) \) which makes \( \{s_\lambda\} \) an orthonormal basis, that is, \( (s_\lambda, s_\nu) = \delta_{\lambda,\nu} \).

**Exercise 2.41.**

(a) If \( n \) and \( m \) are two distinct nonnegative integers, and if \( f \in \Lambda_n \) and \( g \in \Lambda_m \), then show that \( (f, g) = 0 \).

(b) If \( n \in \mathbb{N} \) and \( f \in \Lambda_n \), then prove that \( (h_n, f) = f(1) \) (where \( f(1) \) is defined as in Exercise 2.2).

The Hall inner product induces a \( \mathbf{k} \)-module homomorphism \( \Lambda \to \Lambda^\circ \) (sending every \( f \in \Lambda \) to the \( \mathbf{k} \)-linear map \( \Lambda \to \mathbf{k}, \ g \mapsto (f, g) \)). This homomorphism is invertible (since the Hall inner product has an orthonormal basis), so that \( \Lambda^\circ \cong \Lambda \) as \( \mathbf{k} \)-modules. But in fact, more can be said:

**Corollary 2.42.** The isomorphism \( \Lambda^\circ \cong \Lambda \) induced by the Hall inner product is an isomorphism of Hopf algebras.

---

75In the last equality, we removed the condition \( \mu \subseteq \lambda \) on the addends of the sum; this does not change the value of the sum (because we have \( s_{\lambda/\mu} = 0 \) whenever we don’t have \( \mu \subseteq \lambda \)).

76In fact, this is clear when we don’t have \( \mu \subseteq \lambda \). When we do have \( \mu \subseteq \lambda \), this follows from observing that \( s_{\lambda/\mu} \in \Lambda_{|\lambda/\mu|} \) has zero coefficient before \( s_{\mu} \) whenever \( |\mu| + |\nu| \neq |\lambda| \).
Proof. We have seen that the orthonormal basis \( \{ s_\lambda \} \) of Schur functions is self-dual, in the sense that its multiplication and comultiplication structure constants are the same. Thus the isomorphism \( \Lambda^* \cong \Lambda \) induced by the Hall inner product is an isomorphism of bialgebras, and hence also a Hopf algebra isomorphism by Proposition 1.46(c). \( \square \)

We next identify two other dual pairs of bases, by expanding the Cauchy product in two other ways.

Proposition 2.43. One can also expand

\[
\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \Par} h_\lambda(x) m_\lambda(y) = \sum_{\lambda \in \Par} z_\lambda^{-1} p_\lambda(x) p_\lambda(y)
\]

where \( z_\lambda := m_1! \cdot 1^{m_1} \cdot m_2! \cdot 2^{m_2} \cdots \) if \( \lambda = (1^{m_1}, 2^{m_2}, \ldots) \) with multiplicity \( m_i \) for the part \( i \). (Here, we assume that \( \mathbb{Q} \) is a subring of \( \mathbf{k} \) for the last equality.)

Remark 2.44. It is relevant later (and explains the notation) that \( z_\lambda \) is the size of the \( \mathfrak{S}_n \)-centralizer subgroup for a permutation having cycle type\(^{77}\) \( \lambda \) with \( |\lambda| = n \). This is a classical (and fairly easy) result (see, e.g., [141, Prop. 1.1.1] or [157, Prop. 7.7.3] for a proof).

Proof of Proposition 2.43. For the first expansion, note that (2.14) shows

\[
\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \prod_{j=1}^{\infty} \sum_{n \geq 0} h_n(x) y_j^n
\]

\[
= \sum_{\text{weak compositions } (n_1, n_2, \ldots)} (h_{n_1}(x) h_{n_2}(x) \cdots)(y_1^{n_1} y_2^{n_2} \cdots)
\]

\[
= \sum_{\lambda \in \Par} \sum_{\text{weak compositions } (n_1, n_2, \ldots)} (h_{n_1}(x) h_{n_2}(x) \cdots) = h_\lambda(x) \text{ (since } (n_1, n_2, \ldots) \in \mathfrak{S}_\lambda(\infty)\text{)}
\]

\[
= \sum_{\lambda \in \Par} h_\lambda(x) m_\lambda(y).
\]

For the second expansion (and for later use in the proof of Theorem 4.59) note that

\[
\log H(t) = \log \prod_{i=1}^{\infty} (1 - x_i t)^{-1} = \sum_{i=1}^{\infty} -\log(1 - x_i t) = \sum_{i=1}^{\infty} \sum_{m=1}^{\infty} \frac{(x_i t)^m}{m} = \sum_{m=1}^{\infty} \frac{1}{m} p_m(x)t^m
\]

so that taking \( \frac{d}{dt} \) then shows that

\[
P(t) := \sum_{m \geq 0} \frac{1}{m} p_{m+1} t^m = \frac{H'(t)}{H(t)} = H'(t) E(-t).
\]

\(^{77}\)If \( \sigma \) is a permutation of a finite set \( X \), then the cycle type of \( \sigma \) is defined as the list of the lengths of all cycles of \( \sigma \) (that is, of all orbits of \( \sigma \) acting on \( X \)) written in decreasing order. This is clearly a partition of \( |X| \). (Some other authors write it in increasing order instead, or treat it as a multiset.)

For instance, the permutation of the set \( \{0, 3, 6, 9, 12\} \) which sends 0 to 3, 3 to 9, 6 to 6, 9 to 0, and 12 to 12 is \( (3, 1, 1) \), since the cycles of this permutation have lengths 3, 1 and 1.

It is known that two permutations in \( \mathfrak{S}_n \) have the same cycle type if and only if they are conjugate, thus, for a given partition \( \lambda \) with \( |\lambda| = n \), any two permutations in \( \mathfrak{S}_n \) having cycle type \( \lambda \) are conjugate and therefore their \( \mathfrak{S}_n \)-centralizer subgroups have the same size.
A similar calculation shows that
\[
(2.31) \quad \log \prod_{i,j=1}^{\infty} (1 - x_iy_j)^{-1} = \sum_{m=1}^{\infty} \frac{1}{m} \frac{p_m(x)p_m(y)}{
}
and hence
\[
\prod_{i,j=1}^{\infty} (1 - x_iy_j)^{-1} = \exp \left( \sum_{m=1}^{\infty} \frac{1}{m} \frac{p_m(x)p_m(y)}{\sqrt{m}} \right) = \prod_{m=1}^{\infty} \exp \left( \frac{1}{m} \frac{p_m(x)p_m(y)}{\sqrt{m}} \right) = \prod_{m=1}^{\infty} \left( 1 + \frac{1}{m} \frac{p_m(x)p_m(y)}{\sqrt{m}} \right)^k \]
(by the product rule)
\[
= \sum_{\text{weak compositions } (k_1,k_2,k_3,\ldots)} \prod_{m=1}^{\infty} \frac{(p_m(x)p_m(y))^{k_m}}{k_m!^{m^{k_m}}} = \sum_{\text{weak compositions } (k_1,k_2,k_3,\ldots)} \prod_{m=1}^{\infty} \frac{p_m(x)^{k_m}p_m(y)^{k_m}}{k_m!^{m^{k_m}}}
\]
due to the fact that every partition can be uniquely written in the form \((1^{k_1}2^{k_2}3^{k_3}\ldots)\) with \((k_1,k_2,k_3,\ldots)\) a weak composition.

Corollary 2.45. \(\Box\)

(a) With respect to the Hall inner product on \(\Lambda\), one also has dual bases \(\{h_\lambda\}\) and \(\{m_\lambda\}\).

(b) If \(\mathbb{Q}\) is a subring of \(\mathbb{K}\), then \(\{p_\lambda\}\) and \(\{z_\lambda^{-1}p_\lambda\}\) are also dual bases with respect to the Hall inner product on \(\Lambda\).

(c) If \(\mathbb{R}\) is a subring of \(\mathbb{K}\), then \(\left\{ \frac{p_\lambda}{\sqrt{z_\lambda}} \right\}\) is an orthonormal basis of \(\Lambda\) with respect to the Hall inner product.

Proof. Since (2.24) and (2.28) showed
\[
\prod_{i,j=1}^{\infty} (1 - x_iy_j)^{-1} = \sum_{\lambda \in \text{Par}} s_\lambda(x)s_\lambda(y) = \sum_{\lambda \in \text{Par}} h_\lambda(x)m_\lambda(y) = \sum_{\lambda \in \text{Par}} p_\lambda(x)z_\lambda^{-1}p_\lambda(y) = \sum_{\lambda \in \text{Par}} \frac{p_\lambda(x)}{\sqrt{z_\lambda}} \frac{p_\lambda(y)}{\sqrt{z_\lambda}}
\]
it suffices to show that any pair of graded bases \(\{u_\lambda\}\), \(\{v_\lambda\}\) having
\[
\sum_{\lambda \in \text{Par}} s_\lambda(x)s_\lambda(y) = \sum_{\lambda \in \text{Par}} u_\lambda(x)v_\lambda(y)
\]
will be dual with respect to \((\cdot,\cdot)\). Write transition matrices \(A = (a_{\nu,\lambda}), B = (b_{\nu,\lambda})\) uniquely expressing
\[
u_\lambda = \sum_{\nu} a_{\nu,\lambda}s_\nu,
\]
\[
u_\lambda = \sum_{\nu} b_{\nu,\lambda}s_\nu.
\]
Then orthonormality of the \(s_\lambda\) gives \((u_\alpha, v_\beta) = \sum_{\nu} a_{\nu,\alpha}b_{\nu,\beta}\), and hence we want \(A'B = I\), that is, \(B^{-1} = A'\).

On the other hand, one has
\[
\sum_{\lambda} s_\lambda(x)s_\lambda(y) = \sum_{\lambda} u_\lambda(x)v_\lambda(y) = \sum_{\lambda} \sum_{\nu} a_{\nu,\lambda}s_\nu(x) \sum_{\rho} b_{\nu,\lambda}s_\rho(y).
\]
\textsuperscript{78} A basis \(\{w_\lambda\}\) of \(\Lambda\) indexed by the partitions \(\lambda\) is said to be graded if every \(n \in \mathbb{N}\) and \(\lambda \in \text{Par}_n\) satisfy \(w_\lambda \in \Lambda_n\).
Comparing coefficients\textsuperscript{79} of $s_{\nu}(x)s_{\rho}(y)$ forces $\sum_{\lambda} a_{\nu,\lambda}b_{\mu,\lambda} = \delta_{\nu,\rho}$, or in other words, $AB' = I$. Since $A$ is invertible (being a transition matrix between two bases\textsuperscript{80}), this yields $B'A = I$, and hence $A'B = I$, as desired.

Corollary 2.45 is a known and fundamental fact\textsuperscript{81}. However, our definition of the Hall inner product is unusual; most authors (e.g., Macdonald in [109, §I.4, (4.5)], Hazewinkel/Gubaren/Gi/Kirichenko in [69, Def. 4.121], and Stanley in [157, (7.30)]) define the Hall inner product as the bilinear form satisfying $(h_{\lambda}, m_{\mu}) = \delta_{\lambda,\mu}$ (or, alternatively, $(m_{\lambda}, h_{\mu}) = \delta_{\lambda,\mu}$), and only later prove that the basis \{$s_{\lambda}$\} is orthonormal with respect to this scalar product. (Of course, the fact that this definition is equivalent to our Definition 2.40 follows either from this orthonormality, or from our Corollary 2.45(a).)

The tactic applied in the proof of Corollary 2.45 can not only be used to show that certain bases of $A$ are dual, but also, with a little help from linear algebra over rings (Exercise 2.46), it can be strengthened to show that certain families of symmetric functions are bases to begin with, as we will see in Exercise 2.47.

**Exercise 2.46.** (a) Prove that if an endomorphism of a finitely generated $k$-module is surjective, then this endomorphism is a $k$-module isomorphism.

(b) Let $A$ be a free finite $k$-module with finite basis $(\gamma_{i})_{i \in I}$. Let $(\beta_{i})_{i \in I}$ be a family of elements of $A$ which spans the $k$-module $A$. Prove that $(\beta_{i})_{i \in I}$ is a $k$-basis of $A$.

**Exercise 2.47.** (a) Assume that for every partition $\lambda$, two homogeneous elements $u_{\lambda}$ and $v_{\lambda}$ of $\Lambda$, both having degree $|\lambda|$, are given. Assume further that

\[
\sum_{\lambda \in \Par} s_{\lambda}(x)s_{\lambda}(y) = \sum_{\lambda \in \Par} u_{\lambda}(x)v_{\lambda}(y)
\]

in $k[[x,y]] = k[[x_{1}, x_{2}, \ldots, x_{r}, y_{1}, y_{2}, \ldots, y_{s}]]$. Show that $(u_{\lambda})_{\lambda \in \Par}$ and $(v_{\lambda})_{\lambda \in \Par}$ are $k$-bases of $\Lambda$, and actually are dual bases with respect to the Hall inner product on $\Lambda$.

(b) Use this to give a new proof of the fact that $(h_{\lambda})_{\lambda \in \Par}$ is a $k$-basis of $\Lambda$.

**Exercise 2.48.** Prove that $\sum_{m \geq 0} p_{m+1} t^{m} = \frac{H'(t)}{H(t)}$. (This was proven in (2.30) in the case when $Q$ is a subring of $k$, but here we make no requirements on $k$.)

2.6. Alternants, Littlewood-Richardson: Stembridge’s concise proof. There is a more natural way in which Schur functions arise as a $k$-basis for $\Lambda$, coming from consideration of polynomials in a finite variable set, and the relation between those which are symmetric and those which are alternating.

For the remainder of this section, fix a nonnegative integer $n$, and let $x = (x_{1}, \ldots, x_{n})$ be a finite variable set. This means that $s_{\lambda/\mu} = s_{\lambda/\mu}(x) = \sum T^{x_{\text{cont}}(T)}$ is a generating function for column-strict tableaux $T$ as in Definition 2.19, but with the extra condition that $T$ have entries in $\{1, 2, \ldots, n\}$.\textsuperscript{82} As a consequence, $s_{\lambda/\mu}$ is a polynomial in $k[x_{1}, x_{2}, \ldots, x_{n}]$ (not just a power series), since there are only finitely many column-strict tableaux $T$ of shape $\lambda/\mu$ having all their entries in $\{1, 2, \ldots, n\}$. We will assume without further mention that all partitions appearing in the section have at most $n$ parts.

**Definition 2.49.** Let $k$ be the ring $Z$ or a field of characteristic not equal to 2. (We require this to avoid certain annoyances in the discussion of alternating polynomials in characteristic 2.)

Say that a polynomial $f(x) = f(x_{1}, \ldots, x_{n})$ is alternating if for every permutation $w$ in $S_{n}$ one has that

\[
(wf)(x) = f(x_{w(1)}, \ldots, x_{w(n)}) = \text{sgn}(w)f(x).
\]

Let $\Lambda^{\text{sgn}} \subset k[x_{1}, \ldots, x_{n}]$ denote the subset of alternating polynomials.$^{83}$

---

\textsuperscript{79}Comparing coefficients is legitimate since $\Lambda(x) \otimes \Lambda(y)$ embeds into $k[[x,y]]$.

\textsuperscript{80}Alternatively, we can use the fact that $A$ and $B'$ are block-diagonal matrices with each block being a square matrix of finite size (as $\Lambda$ is graded), and such matrices are always invertible when they are right-invertible.

\textsuperscript{81}For example, Corollary 2.45(a) appears in [95, Corollary 3.3] (though the definition of Schur functions in [95] is different from ours; we will meet this alternative definition later on), and parts (b) and (c) of Corollary 2.45 are equivalent to [109, §I.4, (4.7)] (though Macdonald defines the Hall inner product using Corollary 2.45(a)).

\textsuperscript{82}See Exercise 2.26(a) for this.

\textsuperscript{83}When $k$ has characteristic 2 (or, more generally, is an arbitrary commutative ring), it is probably best to define the alternating polynomials $\Lambda^{\text{sgn}}_{k}$ as the $k$-submodule $\Lambda^{\text{sgn}} \otimes_{Z} k$ of $Z[x_{1}, \ldots, x_{n}] \otimes_{Z} k \cong k[x_{1}, \ldots, x_{n}]$. 
As with $\Lambda$ and its monomial basis $\{m_\lambda\}$, there is an obvious $k$-basis for $\Lambda^{\sgn}$, coming from the fact that a polynomial $f = \sum a_u x^u$ is alternating if and only if $c_{w(\alpha)} = \sgn(w)c_\alpha$ for every $w \in S_n$ and every $\alpha \in \mathbb{N}^n$. This means that every alternating $f$ is a $k$-linear combination of the following elements.

**Definition 2.50.** For $\alpha = (\alpha_1, \ldots, \alpha_n)$ in $\mathbb{N}^n$, define the *alternant*

$$a_\alpha := \sum_{w \in S_n} \sgn(w)w(x^\alpha) = \det \begin{bmatrix} x_1^{\alpha_1} & \cdots & x_n^{\alpha_n} \\ x_1^{\alpha_1} & \cdots & x_n^{\alpha_n} \\ \vdots & & \vdots \\ x_1^{\alpha_1} & \cdots & x_n^{\alpha_n} \end{bmatrix}.$$  

**Example 2.51.** One has

$$a_{(1,5,0)} = x_1^1 x_2^5 x_3^0 - x_1^5 x_2^0 x_3^0 - x_1^1 x_2^0 x_3^5 - x_1^0 x_2^5 x_3 + x_1^0 x_2^0 x_3 + x_1^5 x_2^0 x_3 = -a_{(5,1,0)}.$$  

Similarly, $a_{w(\alpha)} = \sgn(w)a_\alpha$ for every $w \in S_n$ and every $\alpha \in \mathbb{N}^n$.

Meanwhile, $a_{(5,2,2)} = 0$ since the transposition $t = (\frac{123}{132})$ fixes $(5,2,2)$ and hence $a_{(5,2,2)} = t(a_{(5,2,2)}) = \sgn(t)a_{(5,2,2)} = -a_{(5,2,2)}$.

**Proposition 2.52.** Let $k$ be the ring $\mathbb{Z}$ or a field of characteristic not equal to 2.

The alternants $\{a_\lambda\}$ as $\lambda$ runs through the partitions with at most $n$ parts form a $k$-basis for $\Lambda^{\sgn}$. In addition, the *bialternants* $\left\{\frac{a_{\lambda+\rho}}{a_\rho}\right\}$ as $\lambda$ runs through the same set form a $k$-basis for $\Lambda(x_1, \ldots, x_n) = k[x_1, \ldots, x_n]^{\mathfrak{S}_n}$.

**Proof.** The first assertion should be clear from our previous discussion: the alternants $\{a_\lambda\}$ span $\Lambda^{\sgn}$ by definition, and they are $k$-linearly independent because they are supported on disjoint sets of monomials $x^\alpha$.

The second assertion follows from the first, after proving the following **Claim**: $f(x)$ lies in $\Lambda^{\sgn}$ if and only if $f(x) = a_\rho \cdot g(x)$ where $g(x)$ lies in $k[x]^{\mathfrak{S}_n}$ and where

$$a_\rho = \det(x_i^{n-j})_{i,j=1,2,\ldots,n} = \prod_{1 \leq i < j \leq n} (x_i - x_j)$$

is the *Vandermonde determinant/product*. In other words

$$\Lambda^{\sgn} = a_\rho \cdot k[x]^{\mathfrak{S}_n}$$

is a free $k[x]^{\mathfrak{S}_n}$-module of rank one, with $a_\rho$ as its $k[x]^{\mathfrak{S}_n}$-basis element.

To see the Claim, first note the inclusion

$$\Lambda^{\sgn} \supset a_\rho \cdot k[x]^{\mathfrak{S}_n}$$

since the product of a symmetric polynomial and an alternating polynomial is an alternating polynomial. For the reverse inclusion, note that since an alternating polynomial $f(x)$ changes sign whenever one exchanges two distinct variables $x_i, x_j$, it must vanish upon setting $x_i = x_j$, and therefore be divisible by $x_i - x_j$, so

---

84 One subtlety should be addressed: We want to prove that $a_{(5,2,2)} = 0$ in $k[x_1, \ldots, x_n]$ for every commutative ring $k$. It is clearly enough to prove that $a_{(5,2,2)} = 0$ in $\mathbb{Z}[x_1, \ldots, x_n]$. Since 2 is not a zero-divisor in $\mathbb{Z}[x_1, \ldots, x_n]$, we can achieve this by showing that $a_{(5,2,2)} = -a_{(5,2,2)}$. We would not be able to make this argument directly over an arbitrary commutative ring $k$.

85 The name is owed to its Ferrers shape. For instance, if $n = 5$, then the Ferrers diagram of $\rho$ (represented using dots) has the form

```
• • • •
• • •
• •
•
•
```
divisible by the entire product $\prod_{1 \leq i < j \leq n} (x_i - x_j) = a_\rho$. But then the quotient $g(x) = \frac{f(x)}{a_\rho}$ is symmetric, as it is a quotient of two alternating polynomials.

Let us now return to the general setting, where $k$ is an arbitrary commutative ring. We are not requiring that the assumptions of Proposition 2.52 be valid; we can still study the $a_{\alpha}$ of Definition 2.50, but we cannot use Proposition 2.52 anymore. We will show that the fraction $\frac{a_\lambda + x}{a_\rho}$ is nevertheless a well-defined polynomial in $\Lambda(x_1, \ldots, x_n)$ whenever $\lambda$ is a partition, and in fact equals the Schur function $s_\lambda(x)$. As a consequence, the mysterious bialternant basis $\{\frac{a_\lambda + x}{a_\rho}\}$ of $\Lambda(x_1, \ldots, x_n)$ defined in Proposition 2.52 still exists in the general setting, and is plainly the Schur functions $\{s_\lambda(x)\}$. Stembridge [160] noted that one could give a remarkably concise proof of an even stronger assertion, which simultaneously gives one of the standard combinatorial interpretations for the Littlewood-Richardson coefficients $c_{\mu,\nu}^\lambda$. For the purposes of stating it, we introduce for a tableau $T$ the notation $T|_{\text{cols} \geq j}$ (resp. $T|_{\text{cols} \leq j}$) to indicate the subtableau which is the restriction of $T$ to the union of its columns $j, j+1, j+2, \ldots$ (resp. columns $1, 2, \ldots, j$).

**Theorem 2.53.** For partitions $\lambda, \mu, \nu$ with $\mu \subseteq \lambda$, one has

$$a_{\nu} s_{\lambda/\mu} = \sum_T a_{\nu + \text{cont}(T)} + \rho$$

where $T$ runs through all column-strict tableaux with entries in $\{1, 2, \ldots, n\}$ of shape $\lambda/\mu$ with the property that for all $j = 1, 2, \ldots$ one has $\nu + \text{cont}(T|_{\text{cols} \geq j})$ a partition.

Before proving Theorem 2.53, let us see some of its consequences.

**Corollary 2.54.** For any partition $\lambda$, we have

$$s_\lambda(x) = \frac{a_{\lambda + \rho}}{a_\rho}.$$  

**Proof.** Take $\nu = \mu = \emptyset$ in Theorem 2.53. Note that for any $\lambda$, there is only one column-strict tableau $T$ of shape $\lambda$ having each cont$(T|_{\text{cols} \geq j})$ a partition, namely the one having every entry in row $i$ equal to $i$:

\[
\begin{array}{cccccc}
1 & 1 & 1 & 1 & 1 \\
2 & 2 & 2 \\
3 & 3 & 3 \\
4 & 4 \\
\end{array}
\]

Furthermore, this $T$ has cont$(T) = \lambda$, so the theorem says $a_\rho s_\lambda = a_{\lambda + \rho}$. \hfill $\Box$

**Example 2.55.** For $n = 2$, so that $\rho = (1,0)$, if we take $\lambda = (4,2)$, then one has

$$a_{\lambda + \rho} = a_{(4,2) + (1,0)} = \frac{a_{(5,2)}}{a_{(1,0)}} = \frac{x_1^5 x_2^5 - x_1^2 x_2^8}{x_1^2 - x_2} = x_1^2 x_2^2 + x_1^3 x_2^3 + x_1^2 x_2^4$$

$$= x \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 \\ \end{pmatrix} + x \begin{pmatrix} 1 & 1 \\ 2 \\ \end{pmatrix} + x \begin{pmatrix} 2 \\ 2 \\ \end{pmatrix} = s_{(4,2)} = s_\lambda.$$

---

86 This can also be deduced by base change from the $k = \mathbb{Z}$ case of Proposition 2.52.

87 Again, we can drop the requirement that $\mu \subseteq \lambda$, provided that we understand that there are no column-strict tableaux of shape $\lambda/\mu$ unless $\mu \subseteq \lambda$.

88 Notice that division by $a_\rho$ is unambiguous in the ring $k[x_1, \ldots, x_n]$, since $a_\rho$ is not a zero-divisor (in fact, $a_\rho = \prod_{1 \leq i < j \leq n} (x_i - x_j)$ is the product of the binomials $x_i - x_j$, none of which is a zero-divisor).
Some authors use the equality in Corollary 2.54 to define the Schur polynomial $s_\lambda(x_1, x_2, \ldots, x_n)$ in $n$ variables; this definition, however, has the drawback of not generalizing easily to infinitely many variables or to skew Schur functions.  

Next divide through by $a_\rho$ on both sides of Theorem 2.53 (and use Corollary 2.54) to give the following.

**Corollary 2.56.** For partitions $\lambda, \mu, \nu$ having at most $n$ parts, one has

\[
\sum_T S_{\nu|\lambda/\mu} = \sum_{T \ni \nu} s_{\nu+\cont(T)}
\]

where $T$ runs through the same set as in Theorem 2.53. In particular, taking $\nu = \emptyset$, we obtain

\[
\sum_T S_{\lambda/\mu} = \sum_{T \ni \lambda/\mu} s_{\cont(T)}
\]

where in the sum $T$ runs through all column-strict tableaux of shape $\lambda/\mu$ for which each $\cont(T)_{\cols \geq j}$ is a partition.

*Proof of Theorem 2.53.* Start by rewriting the left side of the theorem, and using the fact that $w(s_{\lambda/\mu}) = s_{\lambda/\mu}$ for any $w$ in $S_n$:  

\[
\begin{align*}
\sum_{T \ni \lambda/\mu} s_{\nu+\cont(T)} & = \sum_{T \ni \lambda/\mu} \sum_{w \in S_n} \sgn(w) x^{w(\nu+\rho)} w(s_{\lambda/\mu}) \\
& = \sum_{w \in S_n} \sgn(w) x^{w(\nu+\rho)} \sum_{T \ni \lambda/\mu} \sum_{\text{column-strict } T} x^{w(\cont(T))} \\
& = \sum_{\text{column-strict } T \ni \lambda/\mu} \sum_{w \in S_n} \sgn(w) x^{w(\nu+\cont(T)+\rho)} \\
& = \sum_{\text{column-strict } T \ni \lambda/\mu} \sum_{\text{column-strict } T} a_{\nu+\cont(T)+\rho}
\end{align*}
\]

We wish to cancel out all the summands indexed by column-strict tableaux $T$ which fail any of the conditions that $\nu + \cont(T)_{\cols \geq j}$ be a partition. Given such a $T$, find the maximal $j$ for which it fails this condition, and then find the minimal $k$ for which

\[
\nu_k + \cont_k(T)_{\cols \geq j} < \nu_{k+1} + \cont_{k+1}(T)_{\cols \geq j}.
\]

Maximality of $j$ forces

\[
\nu_k + \cont_k(T)_{\cols \geq j+1} \geq \nu_{k+1} + \cont_{k+1}(T)_{\cols \geq j+1}.
\]

Since column-strictness implies that column $j$ of $T$ can contain at most one occurrence of $k$ or of $k + 1$ (or neither or both), the previous two inequalities imply that column $j$ must contain an occurrence of $k + 1$ and no occurrence of $k$, so that

\[
\nu_k + \cont_k(T)_{\cols \geq j} + 1 = \nu_{k+1} + \cont_{k+1}(T)_{\cols \geq j}.
\]

This implies that the adjacent transposition $t_{k,k+1}$ swapping $k$ and $k+1$ fixes the vector $\nu + \cont(T)_{\cols \geq j} + \rho$.

Now create a new tableau $T^*$ from $T$ by applying the Bender-Knuth involution (from the proof of Proposition 2.11) on letters $k, k + 1$, but *only to columns $1, 2, \ldots, j - 1$ of $T$, leaving columns $j, j + 1, j + 2, \ldots$ unchanged*. One should check that $T^*$ is still column-strict, but this holds because column $j$ of $T$ has no occurrences of letter $k$. Note that

\[
t_{k,k+1} \cont(T)_{\cols \leq j-1} = \cont(T^*)_{\cols \leq j-1}
\]

and hence

\[
t_{k,k+1}(\nu + \cont(T) + \rho) = \nu + \cont(T^*) + \rho
\]

so that $a_{\nu+\cont(T)+\rho} = -a_{\nu+\cont(T^*)+\rho}$.

---

89. With some effort, it is possible to use Corollary 2.54 in order to define the Schur function $s_\lambda$ in infinitely many variables. Indeed, one can define this Schur function as the unique element of $\Lambda$ whose evaluation at $(x_1, x_2, \ldots, x_n)$ equals $s_{\lambda(x_1, x_2, \ldots, x_n)}$ for every $n \in \mathbb{N}$. If one wants to use such a definition, however, one needs to check that such an element exists. This is the approach to defining $s_\lambda$ taken in [95, Definition 1.4.2] and in [109, §3].

90. Such a $j$ exists because $\nu + \cont(T)_{\cols \geq j}$ is a partition for all sufficiently high $j$ (in fact, $\nu$ itself is a partition).
Because $T, T^*$ have exactly the same columns $j, j+1, j+2, \ldots$, the tableau $T^*$ is also a violator of at least one of the conditions that $\nu + \text{cont}(T^*|_{\text{cols} \geq j})$ be a partition, and has the same choice of maximal $j$ and minimal $k$ as did $T$. Hence the map $T \mapsto T^*$ is an involution on the violators that lets one cancel their summands $a_{\nu + \text{cont}(T) + \rho}$ and $a_{\nu + \text{cont}(T^*) + \rho}$ in pairs. □

So far (in this section) we have worked with a finite set of variables $x_1, x_2, \ldots, x_n$ (where $n$ is a fixed nonnegative integer) and with partitions having at most $n$ parts. We now drop these conventions and restrictions; thus, partitions again mean arbitrary partitions, and $\mathbf{x}$ again means the infinite family $(x_1, x_2, x_3, \ldots)$ of variables. In this setting, we have the following analogue of Corollary 2.56:

**Corollary 2.57.** For partitions $\lambda, \mu, \nu$ (of any lengths), one has

$$s_\nu s_{\lambda/\mu} = \sum_{T} s_{\nu + \text{cont}(T)}$$

(2.34)

where $T$ runs through all column-strict tableaux of shape $\lambda/\mu$ with the property that for all $j = 1, 2, \ldots$ one has $\nu + \text{cont}(T|_{\text{cols} \geq j})$ a partition. In particular, taking $\nu = \emptyset$, we obtain

$$s_{\lambda/\mu} = \sum_{T} s_{\text{cont}(T)}$$

(2.35)

where in the sum $T$ runs through all column-strict tableaux of shape $\lambda/\mu$ for which each $\text{cont}(T|_{\text{cols} \geq j})$ is a partition.

**Proof of Corollary 2.57.** Essentially, Corollary 2.57 is obtained from Corollary 2.56 by “letting $n$ (that is, the number of variables) tend to $\infty$”. This can be formalized in different ways: One way is to endow the ring of power series $k[[\mathbf{x}]]$ with the coefficientwise topology\(^{\text{91}}\), and to show that the left hand side of (2.32) tends to the left hand side of (2.34) when $n \to \infty$, and the same holds for the right hand sides. A different approach proceeds by regarding $\Lambda$ as the inverse limit of the $\Lambda$ ($n \in \mathbb{N}$), all sufficiently high $n$ satisfy

\[ (\text{the coefficient of } \mathbf{m} \text{ in } a_n) = (\text{the coefficient of } \mathbf{m} \text{ in } a) . \]

Comparing coefficients of a given Schur function $s_\nu$ in (2.35), we obtain the following version of the Littlewood-Richardson rule.

**Corollary 2.58.** For partitions $\lambda, \mu, \nu$ (of any lengths), the Littlewood-Richardson coefficient $c_{\mu, \nu}^\lambda$ counts column-strict tableaux $T$ of shape $\lambda/\mu$ with $\text{cont}(T) = \nu$ having the property that each $\text{cont}(T|_{\text{cols} \geq j})$ is a partition.

2.7. **The Pieri and Assaf-McNamara skew Pieri rule.** The classical Pieri rule refers to two special cases of the Littlewood-Richardson rule. To state them, recall that a skew shape is called a horizontal (resp. vertical) strip if no two of its cells lie in the same column (resp. row). A horizontal (resp. vertical) $n$-strip (for $n \in \mathbb{N}$) shall mean a horizontal (resp. vertical) strip of size $n$ (that is, having exactly $n$ cells).

**Theorem 2.59.** For every partition $\lambda$ and any $n \in \mathbb{N}$, we have

$$s_{\lambda} h_n = \sum_{\lambda^+ : \lambda^+ / \lambda \text{ is a horizontal } n\text{-strip}} s_{\lambda^+}$$

(2.36)

$$s_{\lambda} e_n = \sum_{\lambda^+ : \lambda^+ / \lambda \text{ is a vertical } n\text{-strip}} s_{\lambda^+}$$

(2.37)
Example 2.60.

\[
\begin{array}{c}
\begin{array}{c}
\begin{array}{ccc}
\square & \square & \square \\
\square & \square & \bullet \\
\square & \square & \square
\end{array}
+ \begin{array}{ccc}
\square & \square & \square
\end{array}
\end{array}
\end{array}
\]

\[
\begin{array}{c}
\begin{array}{c}
\begin{array}{ccc}
\square & \square & \square \\
\square & \square & \square
\end{array}
\end{array}
\end{array}
\]

\[
\begin{array}{c}
\begin{array}{c}
\begin{array}{ccc}
\square & \square & \square \\
\square & \square & \square
\end{array}
\end{array}
\end{array}
\]

\[
\begin{array}{c}
\begin{array}{c}
\begin{array}{ccc}
\square & \square & \square \\
\square & \square & \square
\end{array}
\end{array}
\end{array}
\]

\[
\begin{array}{c}
\begin{array}{c}
\begin{array}{ccc}
\square & \square & \square \\
\square & \square & \square
\end{array}
\end{array}
\end{array}
\]

Proof of Theorem 2.59. For the first Pieri formula involving \( h_n \), as \( h_n = s(n) \) one has

\[
\sum_{\lambda} c_{\lambda, (n) \lambda} \lambda (n) = s_{\lambda h_n} = \sum_{\lambda} c_{\lambda, n}^{\lambda^+} s_{\lambda^+}.
\]

Corollary 2.58 says \( c_{\lambda, (n) \lambda}^{\lambda^+} \) counts column-strict tableaux \( T \) of shape \( \lambda^+ / \lambda \) having \( \text{cont}(T) = (n) \) (i.e. all entries of \( T \) are 1’s), with an extra condition. Since its entries are all equal, such a \( T \) must certainly have shape being a horizontal strip, and more precisely a horizontal \( n \)-strip (since it has \( n \) cells). Conversely, for any horizontal \( n \)-strip, there is a unique such filling, and it will trivially satisfy the extra condition that \( \text{cont}(T_{\text{cols} \geq j}) \) is a partition for each \( j \). Hence \( c_{\lambda, (n) \lambda}^{\lambda^+} \) is 1 if \( \lambda^+ / \lambda \) is a horizontal \( n \)-strip, and 0 else.

For the second Pieri formula involving \( e_n \), using \( e_n = s(n) \) one has

\[
\sum_{\lambda} c_{\lambda, (1^n) \lambda}^{\lambda^+} s_{\lambda^+} = s_{\lambda e_n} = \sum_{\lambda} c_{\lambda, (1^n) \lambda}^{\lambda^+} s_{\lambda^+}.
\]

Corollary 2.58 says \( c_{\lambda, (1^n) \lambda}^{\lambda^+} \) counts column-strict tableaux \( T \) of shape \( \lambda^+ / \lambda \) having \( \text{cont}(T) = (1^n) \), so its entries are \( 1, 2, \ldots, n \) each occurring once, with the extra condition that \( 1, 2, \ldots, n \) appear from right to left. Together with the tableau condition, this forces at most one entry in each row, that is \( \lambda^+ / \lambda \) is a vertical strip, and then there is a unique way to fill it (maintaining column-strictness and the extra condition that \( 1, 2, \ldots, n \) appear from right to left). Thus \( c_{\lambda, (1^n) \lambda}^{\lambda^+} \) is 1 if \( \lambda^+ / \lambda \) is a vertical \( n \)-strip, and 0 else. \( \square \)

Assaf and McNamara [9] recently proved an elegant generalization.

Theorem 2.61. For any partitions \( \lambda \) and \( \mu \) and any \( n \in \mathbb{N} \), we have\(^{92}\)

\[
s_{\lambda / \mu} h_n = \sum_{\lambda^+, \mu^-: \lambda^+ / \lambda \text{ a horizontal strip}} (-1)^{|\mu^-|} s_{\lambda^+ / \mu^-}.
\]

\[
s_{\lambda / \mu} e_n = \sum_{\lambda^+, \mu^-: \lambda^+ / \lambda \text{ a vertical strip}} (-1)^{|\mu^-|} s_{\lambda^+ / \mu^-}.
\]

\(^{92}\)Note that \( \mu \subseteq \lambda \) is not required. (The left hand sides are 0 otherwise, but this does not trivialize the equalities.)
Example 2.62.

\[
\begin{align*}
\begin{array}{c}
\begin{array}{c}
\includegraphics[width=0.2\textwidth]{example_2.62}
\end{array}
\end{array}
\end{align*}
\]

Theorem 2.61 is proven in the next section, using an important Hopf algebra tool.

Exercise 2.63. Let \( \lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots) \) and \( \mu = (\mu_1, \mu_2, \mu_3, \ldots) \) be two partitions such that \( \mu \subseteq \lambda \).

(a) Show that \( \lambda/\mu \) is a horizontal strip if and only if every \( i \in \{1, 2, 3, \ldots\} \) satisfies \( \mu_i \geq \lambda_i + 1 \).

(b) Show that \( \lambda/\mu \) is a vertical strip if and only if every \( i \in \{1, 2, 3, \ldots\} \) satisfies \( \lambda_i \leq \mu_i + 1 \).

Exercise 2.64. (a) Let \( \lambda \) and \( \mu \) be two partitions such that \( \mu \subseteq \lambda \). Let \( n \in \mathbb{N} \). Show that \( (h_n, s_{\lambda/\mu}) \) equals 1 if \( \lambda/\mu \) is a horizontal \( n \)-strip, and equals 0 otherwise.

(b) Use part (a) to give a new proof of (2.36).

Exercise 2.65. Prove Theorem 2.59 again using the ideas of the proof of Theorem 2.31.

Exercise 2.66. Let \( A \) be a commutative ring, and \( n \in \mathbb{N} \).

(a) Let \( a_1, a_2, \ldots, a_n \) be \( n \) elements of \( A \). Let \( b_1, b_2, \ldots, b_n \) be \( n \) further elements of \( A \). If \( a_i - b_j \) is an invertible element of \( A \) for every \( i \in \{1, 2, \ldots, n\} \) and \( j \in \{1, 2, \ldots, n\} \), then prove that

\[
\det \left( \frac{1}{a_i - b_j} \right)_{i,j=1,2,\ldots,n} = \frac{\prod_{1 \leq j < i \leq n} ((a_i - a_j)(b_j - b_i))}{\prod_{(i,j) \in \{1,2,\ldots,n\}^2} (a_i - b_j)}. 
\]

(b) Let \( a_1, a_2, \ldots, a_n \) be \( n \) elements of \( A \). Let \( b_1, b_2, \ldots, b_n \) be \( n \) further elements of \( A \). If \( 1 - a_i b_j \) is an invertible element of \( A \) for every \( i \in \{1, 2, \ldots, n\} \) and \( j \in \{1, 2, \ldots, n\} \), then prove that

\[
\det \left( \frac{1}{1 - a_i b_j} \right)_{i,j=1,2,\ldots,n} = \frac{\prod_{1 \leq j < i \leq n} ((a_i - a_j)(b_i - b_j))}{\prod_{(i,j) \in \{1,2,\ldots,n\}^2} (1 - a_i b_j)}. 
\]

\footnote{In other words, \( \lambda/\mu \) is a horizontal strip if and only if \( (\lambda_2, \lambda_3, \lambda_4, \ldots) \subseteq \mu \). This simple observation has been used by Pak and Postnikov [125, §10] for a new approach to RSK-type algorithms.}
(c) Use the result of part (b) to give a new proof for Theorem 2.31.\footnote{This approach to Theorem 2.31 is taken in \cite{37, \S 4} (except that \cite{37} only works with finitely many variables).}

The determinant on the left hand side of Exercise 2.66(a) is known as the Cauchy determinant.

**Exercise 2.67.** Prove that $s_{(a,b)} = h_a h_b - h_a h_{b-1}$ for any two integers $a \geq b \geq 0$ (where we set $h_{-1} = 0$ as usual).

(Note that this is precisely the Jacobi-Trudi formula (2.22) in the case when $\lambda = (a, b)$ is a partition with at most two entries and $\mu = \emptyset$.)

**Exercise 2.68.** If $\lambda$ is a partition and $\mu$ is a weak composition, let $K_{\lambda, \mu}$ denote the number of column-strict tableaux $T$ of shape $\lambda$ having $\text{cont} (T) = \mu$. (This $K_{\lambda, \mu}$ is called the $(\lambda, \mu)$-Kostka number.)

(a) Use Theorem 2.59 to show that every partition $\mu$ satisfies $h_\mu = \sum_\lambda K_{\lambda, \mu} s_\lambda$, where the sum ranges over all partitions $\lambda$.

(b) Use this to give a new proof for Theorem 2.31.\footnote{Of course, this gives a new proof of Theorem 2.31 only when coupled with a proof of Theorem 2.59 which does not rely on Theorem 2.31. The proof of Theorem 2.59 we gave in the text above did not rely on Theorem 2.31, whereas the proof of (2.30) given in Exercise 2.64(b) did.}

(c) Give a new proof of the fact that $K_{\lambda, \mu}$ is called the $(\lambda,\mu)$-Kostka number.

**Exercise 2.69.** (a) Define a $k$-linear map $\mathfrak{Z} : \Lambda \rightarrow \Lambda$ by having it send $s_\lambda$ to $s_{\lambda'}$ for every partition $\lambda$. (This is clearly well-defined, since $(s_\lambda)_{\lambda \in \text{Par}}$ is a $k$-basis of $\Lambda$.) Show that

$$\mathfrak{Z} (f h_\lambda) = \mathfrak{Z} (f) \cdot \mathfrak{Z} (h_\lambda) \quad \text{for every } f \in \Lambda \text{ and every } n \in \mathbb{N}. $$

(b) Show that $\mathfrak{Z} = \omega$.

(c) Show that $c^{\lambda'}_{\mu', \nu} = c^{\lambda'}_{\mu, \nu'}$ for any three partitions $\lambda, \mu$ and $\nu$.

(d) Use this to prove (2.21).\footnote{The first author learned this approach to (2.21) from Alexander Postnikov.}

**Exercise 2.70.** (a) Show that

$$\prod_{i,j=1}^{\infty} (1 + x_i y_j) = \sum_{\lambda \in \text{Par}} s_\lambda (x) s_{\lambda'} (y) = \sum_{\lambda \in \text{Par}} e_\lambda (x) m_\lambda (y)$$

in the power series ring $k [[x, y]] = k [[x_1, x_2, x_3, \ldots , y_1, y_2, y_3, \ldots ]]$.

(b) Assume that $\mathbb{Q}$ is a subring of $k$. Show that

$$\prod_{i,j=1}^{\infty} (1 + x_i y_j) = \sum_{\lambda \in \text{Par}} (-1)^{|\lambda| - \ell (\lambda)} z_\lambda^{-1} p_\lambda (x) p_\lambda (y)$$

in the power series ring $k [[x, y]] = k [[x_1, x_2, x_3, \ldots , y_1, y_2, y_3, \ldots ]]$, where $z_\lambda$ is defined as in Proposition 2.43.

The first equality of Exercise 2.70(a) appears in \cite{157, Thm. 7.14.3}, \cite{141, Thm. 4.8.6} and several other references under the name of the dual Cauchy identity, and is commonly proven using a “dual” analogue of the Robinson-Schensted-Knuth algorithm.

**Exercise 2.71.** Prove Theorem 2.30.

[**Hint:**\footnote{This is the proof given in Stanley \cite[\S 7.16, Second Proof of Thm. 7.16.1]{157} and Macdonald \cite[proof of (5.4)]{109}.} Switch $x$ and $y$ in the formula of Exercise 2.39(a), and specialize the resulting equality by replacing $y$ by a finite set of variables $(y_1, y_2, \ldots , y_\ell)$; then, set $n = \ell$ and $\rho = (n-1, n-2, \ldots , 0)$, and multiply with the alternant $a_\rho (y_1, y_2, \ldots , y_\ell)$, using Corollary 2.54 to simplify the result; finally, extract the coefficient of $y^{\lambda + \rho}$.]
2.8. Skewing and Lam’s proof of the skew Pieri rule. We codify here the operation $s^\perp_\mu$ of skewing by $s_\mu$, acting on Schur functions via

$$s^\perp_\mu(s_\lambda) = s_{\lambda/\mu}$$

(where, as before, one defines $s_\lambda/\mu = 0$ if $\mu \not\subseteq \lambda$). These operations play a crucial role

- in Lam’s proof of the skew Pieri rule,
- in Lam, Lauve, and Sottile’s proof [90] of a more general skew Littlewood-Richardson rule that had been conjectured by Assaf and McNamara, and
- in Zelevinsky’s structure theory of PSH’s to be developed in the next chapter.

**Definition 2.72.** Given a Hopf algebra $A$ of finite type, and its (graded) dual $A^\circ$, let $(\cdot, \cdot)_A$ be the pairing $(f, a) := f(a)$ for $f$ in $A^\circ$ and $a$ in $A$. Then define for each $f$ in $A^\circ$ an operator $A \xrightarrow{f^\perp} A$ as follows\(^98\): for $a$ in $A$ with $\Delta(a) = \sum a_1 \otimes a_2$, let

$$f^\perp(a) = \sum (f, a_1)a_2.$$  

Note that when one takes $A = \Lambda = A^\circ$ \(^99\), the element $a = s_\lambda$ has $\Delta s_\lambda = \sum s_\mu \otimes s_{\lambda/\mu}$. Hence if $f = s_\mu$, then one has $f^\perp(a) = s_{\lambda/\mu} = s^\perp_\mu(s_\lambda)$ as desired.

**Proposition 2.73.** The $f^\perp$ operators $A \to A$ have the following properties.

(i) For every $f \in A^\circ$, the map $f^\perp$ is adjoint to left multiplication $A^\circ \xrightarrow{f} A^\circ$ in the sense that

$$(g, f^\perp(a)) = (fg, a).$$

(ii) For every $f, g \in A^\circ$, we have $(fg)^\perp(a) = g^\perp(f^\perp(a))$, that is, $A$ becomes a right $A^\circ$-module via the $f^\perp$ action.

(iii) If an $f \in A^\circ$ satisfies $\Delta(f) = \sum f_1 \otimes f_2$, then

$$f^\perp(ab) = \sum f_1^\perp(a)f_2^\perp(b).$$

In particular, if $f$ is primitive in $A^\circ$, so that $\Delta(f) = f \otimes 1 + 1 \otimes f$, then $f^\perp$ is a derivation:

$$f^\perp(ab) = f^\perp(a) \cdot b + a \cdot f^\perp(b).$$

**Proof.** For (i), note that

$$(g, f^\perp(a)) = \sum (f, a_1)(g, a_2) = (f \otimes g, \Delta_A(a)) = (m_{A^\circ}(f \otimes g), a) = (fg, a).$$

For (ii), using (i) and considering any $h$ in $A^\circ$, one has that

$$(h, (fg)^\perp(a)) = (fg, a) = (gh, f^\perp(a)) = (h, g^\perp(f^\perp(a))).$$

For (iii), noting that

$$\Delta(ab) = \Delta(a)\Delta(b) = \left(\sum_{(a)} a_1 \otimes a_2\right)\left(\sum_{(b)} b_1 \otimes b_2\right) = \sum_{(a),(b)} a_1b_1 \otimes a_2b_2,$$

one has that

$$f^\perp(ab) = \sum_{(a),(b)} (f, a_1b_1)A a_2b_2 = \sum_{(a),(b)} (\Delta(f), a_1 \otimes b_1)_{A \otimes A} a_2b_2$$

$$= \sum_{(f),(a),(b)} (f_1, a_1)A(f_2, b_1)A a_2b_2$$

$$= \sum_{(f)} \left(\sum_{(a)} (f_1, a_1)A a_2\right) \left(\sum_{(b)} (f_2, b_1)A b_2\right) = \sum_{(f)} f_1^\perp(a)f_2^\perp(b).$$

\(\square\)

\(^98\)This $f^\perp(a)$ is called $a \leftarrow f$ in Montgomery [120, Example 1.6.5].

\(^99\)Here, we are using the isomorphism $\Lambda^\circ \cong \Lambda$ induced by the Hall inner product (see Corollary 2.42 for why it is an isomorphism) to identify $\Lambda$ with $\Lambda^\circ$.\n
The following interaction between multiplication and $h^+$ is the key to deducing the skew Pieri formula from the usual Pieri formulas.

**Lemma 2.74.** For any $f, g \in \Lambda$ and any $n \in \mathbb{N}$, one has

$$f \cdot h_n^+(g) = \sum_{k=0}^{n} (-1)^k h_{n-k}^+(e_k^+(f) \cdot g).$$

**Proof.** Starting with the right side, first apply Proposition 2.73(iii):

$$\sum_{k=0}^{n} (-1)^k h_{n-k}^+(e_k^+(f) \cdot g)$$

$$= \sum_{k=0}^{n} (-1)^k \sum_{j=0}^{n-k} h_j^+(e_j^+(f)) \cdot h_{n-k-j}(g)$$

$$= \sum_{i=0}^{n} (1)^{n-i} \left( \sum_{j=0}^{n-i} (-1)^j h_j^+(e_{n-i-j}^+(f)) \right) \cdot h_i^+(g) \quad \text{(reindexing } i := n - k - j \text{)}$$

$$= \sum_{i=0}^{n} (1)^{n-i} \left( \sum_{j=0}^{n-i} (-1)^j e_{n-i-j}^+(f) h_j^+(g) \right) \quad \text{(by Proposition 2.73(ii))}$$

$$= 1^+(f) \cdot h_n^+(g) = f \cdot h_n^+(g)$$

where the second-to-last equality used (2.17). □

**Proof of Theorem 2.61.** We prove the first skew Pieri rule; the second is analogous, swapping $h_i \leftrightarrow e_i$ and swapping the words “vertical” $\leftrightarrow$ “horizontal”. Symmetry of $(\cdot, \cdot)_\Lambda$, commutativity of $\Lambda$, and Proposition 2.73(i) imply for $f$ in $\Lambda$,

$$(s_{\lambda/\mu}, f) = (s_{\mu}^+(s_{\lambda}), f) = (s_{\lambda}, s_{\mu}f)$$

Hence for any $g$ in $\Lambda$, one can compute that

$$(h_n s_{\lambda/\mu}, g) \overset{Prop. 2.73(i)}{=} (s_{\lambda/\mu}, h_n^+ g) \overset{(2.38)}{=} (s_{\lambda}, s_{\mu}^+ h_n^+ g)$$

(2.39)

The first Pieri rule in Theorem 2.59 lets one rewrite $h_{n-k}s_{\lambda} = \sum_{\lambda^+, \mu^+} s_{\lambda^+}$, with the sum running through $\lambda^+$ for which $\lambda^+/\lambda$ is a horizontal $(n-k)$-strip. The second Pieri rule in Theorem 2.59 lets one rewrite $e^+_k s_{\mu} = \sum_{\mu^-} s_{\mu^-}$, with the sum running through $\mu^-$ for which $\mu/\mu^-$ is a vertical $k$-strip, since $(s_{\mu^-}, e^+_k s_{\mu}) = (e^+_k s_{\mu^-}, s_{\mu})$. Thus the last line of (2.39) becomes

$$\sum_{k=0}^{n} (1)^k \left( \sum_{\lambda^+} s_{\lambda^+} \cdot s_{\mu^-} \cdot g \right) \overset{(2.38)}{=} \left( \sum_{k=0}^{n} (1)^k \sum_{\lambda^+, \mu^-} s_{\lambda^+/\mu^-} \cdot g \right)$$

where the sum is over the pairs $(\lambda^+, \mu^-)$ for which $\lambda^+/\lambda$ is a horizontal $(n-k)$-strip and $\mu/\mu^-$ is a vertical $k$-strip. □

**Exercise 2.75.** Let $n \in \mathbb{N}$.

(a) For every $k \in \mathbb{N}$, let $p(n, k)$ denote the number of partitions of $n$ of length $k$. Let $c(n)$ denote the number of *self-conjugate* partitions of $n$ (that is, partitions $\lambda$ of $n$ satisfying $\lambda^t = \lambda$). Show that

$$(-1)^n c(n) = \sum_{k=0}^{n} (-1)^k p(n, k).$$
Exercise 2.76. The goal of this exercise is to prove (2.21) using the skewing operators that we have developed. Recall the involution \( \omega : \Lambda \to \Lambda \) defined in (2.19).

(a) Show that \( \omega(p_\lambda) = (-1)^{|\lambda| - \ell(\lambda)} p_\lambda \) for any \( \lambda \in \text{Par} \), where \( \ell(\lambda) \) denotes the length of the partition \( \lambda \).

(b) Show that \( \omega \) is an isometry.

(c) Show that this same map \( \omega : \Lambda \to \Lambda \) is a Hopf automorphism.

(d) Prove that \( \omega(a^\perp b) = (\omega(a))^\perp (\omega(b)) \) for every \( a \in \Lambda \) and \( b \in \Lambda \).

(e) For any partition \( \lambda = (\lambda_1, \ldots, \lambda_\ell) \) with length \( \ell(\lambda) = \ell \), prove that
\[
{e_\ell^\perp} s_\lambda = s(\lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_\ell - 1).
\]

(f) For any partition \( \lambda = (\lambda_1, \lambda_2, \ldots) \), prove that
\[
h_{\lambda_1}^\perp s_\lambda = s(\lambda_2, \lambda_3, \lambda_4, \ldots).
\]

(g) Prove (2.21).

2.9. Assorted exercises on symmetric functions. Over a hundred exercises on symmetric functions are collected in Stanley’s [157, chapter 7], and even more (but without any hints or references) on his website.

Further sources for results related to symmetric functions are Macdonald’s work, including his monograph [109] and his expository [110]. In this section, we gather a few exercises that are not too difficult to handle with the material given above.

Exercise 2.77. (a) Let \( m \in \mathbb{Z} \). Prove that, for every \( f \in \Lambda \), the infinite sum \( \sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e_i^\perp f \) is convergent in the discrete topology (i.e., all but finitely many addends of this sum are zero). Hence, we can define a map \( B_m : \Lambda \to \Lambda \) by setting
\[
B_m(f) = \sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e_i^\perp f \quad \text{for all } f \in \Lambda.
\]

Show that this map \( B_m \) is \( \mathbb{k} \)-linear.

(b) Let \( \lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots) \) be a partition, and let \( m, n \in \mathbb{Z} \) be such that \( m \geq \lambda_1 \). Show that
\[
\sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e_i^\perp s_\lambda = s(m, \lambda_1, \lambda_2, \lambda_3, \ldots).
\]

(c) Let \( n \in \mathbb{N} \). For every \( n \)-tuple \( (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n \), we define an element \( \overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} \in \Lambda \) by
\[
\overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = \det \left( (h_{\alpha_i - i+j})_{i,j=1,2,\ldots,n} \right).
\]

Show that
\begin{align}
(2.40) \\
\overline{s}_\lambda = \overline{s}_{(\lambda_1, \lambda_2, \ldots, \lambda_n)}
\end{align}

for every partition \( \lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots) \) having at most \( n \) parts (where “part” means “nonzero part”).

Furthermore, show that for every \( n \)-tuple \( (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n \), the symmetric function \( \overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} \) either is 0 or equals \( \pm s_\nu \) for some partition \( \nu \) having at most \( n \) parts.

Finally, show that for any \( n \)-tuples \( (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n \) and \( (\beta_1, \beta_2, \ldots, \beta_n) \in \mathbb{N}^n \), we have
\begin{align}
(2.41) \\
\overline{s}_{(\beta_1, \beta_2, \ldots, \beta_n)} \overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = \det \left( (h_{\alpha_i - \beta_j + i+j})_{i,j=1,2,\ldots,n} \right).
\end{align}

\textsuperscript{100} \( \text{http://math.mit.edu/~rstan/ec/ch7supp.pdf} \)
(d) For every \( n \in \mathbb{N} \), every \( m \in \mathbb{Z} \) and every \( n \)-tuple \((\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n\), prove that

\[
\sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e_i s_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = s_{(m, \alpha_1, \alpha_2, \ldots, \alpha_n)},
\]

where we are using the notations of Exercise 2.77(c).

(e) For every \( n \in \mathbb{N} \) and every \( n \)-tuple \((\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n\), prove that

\[
s_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = (B_{\alpha_1} \circ B_{\alpha_2} \circ \cdots \circ B_{\alpha_n})(1),
\]

where we are using the notations of Exercise 2.77(c) and Exercise 2.77(a).

Remark 2.78. The map \( B_n \) defined in Exercise 2.77(a) is the so-called \( m \)-th Bernstein creation operator; it appears in Zelevinsky [173, §4.20(a)] and has been introduced by J.N. Bernstein, who found the result of Exercise 2.77(b). It is called a “Schur row adder” in [55]. Exercise 2.77(e) appears in Berg/Bergeron/Saliola/Serrano/Zabrocki [16, Theorem 2.3], where it is used as a prototype for defining noncommutative analogues of Schur functions, the so-called immaculate functions. The particular case of Exercise 2.77(e) for \((\alpha_1, \alpha_2, \ldots, \alpha_n)\) a partition of length \( n \) (a restatement of Exercise 2.77(b)) is proven in [109, §3.3, example 29].

Exercise 2.79. (a) Prove that there exists a unique family \((x_n)_{n \geq 1}\) of elements of \( \Lambda \) such that

\[
H(t) = \prod_{n=1}^{\infty} (1 - x_n t^n)^{-1}.
\]

Denote this family \((x_n)_{n \geq 1}\) by \((w_n)_{n \geq 1}\). For instance,

\[
w_1 = s_{(1)}, \quad w_2 = -s_{(1,1)}, \quad w_3 = -s_{(2,1)},
\]

\[
w_4 = -s_{(1,1,1,1)} - s_{(2,1,1)} - s_{(2,2)} - s_{(3,1)}, \quad w_5 = -s_{(2,1,1,1)} - s_{(2,2,1)} - s_{(3,1,1)} - s_{(3,2)} - s_{(4,1)}.
\]

(b) Show that \(w_n\) is homogeneous of degree \( n\) for every positive integer \( n\).

(c) For every partition \( \lambda \), define \( w_\lambda \in \Lambda \) by \( w_\lambda = w_{\lambda_1} w_{\lambda_2} \cdots w_{\lambda_\ell} \) (where \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell) \) with \( \ell = \ell(\lambda) \)). Notice that \(w_\lambda\) is homogeneous of degree \( |\lambda|\). Prove that \( \sum_{\lambda \in \text{Par}_n} w_\lambda = h_n \) for every \( n \in \mathbb{N}\).

(d) Show that \( \{w_\lambda\}_{\lambda \in \text{Par}} \) is a \( \mathbb{k}\)-basis of \( \Lambda \). (This basis is called the Witt basis\(^{101}\); it is studied in [66, §9-§10].\(^{102}\)

(e) Prove that \( p_n = \sum_{d|n} d w_n^d \) for every positive integer \( n\). (Here, the summation sign \( \sum_{d|n} \) means a sum over all positive divisors \( d \) of \( n\).)

(f) We are going to show that \(-w_n\) is a sum of Schur functions (possibly with repetitions, but without signs!) for every \( n \geq 2\). (For \( n = 1\), the opposite is true: \( w_1\) is a single Schur function.) This proof goes back to Doran [42].\(^{103}\)

For any positive integers \( n \) and \( k \), define \( f_{n,k} \in \Lambda \) by \( f_{n,k} = \sum_{\lambda \in \text{Par}_n, \min \lambda \geq k} w_\lambda \), where \( \min \lambda \) denotes the smallest nonzero part of \( \lambda \). Show that

\[
f_{n,k} = \sum_{i=2}^{k-1} f_{1,i} f_{n-i,i} \quad \text{for every } n \geq k \geq 2.
\]

Conclude that \(-f_{n,k}\) is a sum of Schur functions for every \( n \in \mathbb{N} \) and \( k \geq 2\). Conclude that \(-w_n\) is a sum of Schur functions for every \( n \geq 2\).

(g) For every partition \( \lambda \), define \( r_\lambda \in \Lambda \) by \( r_\lambda = \prod_{i \geq 1} h_{v_i} (x_1^i, x_2^i, x_3^i, \ldots) \), where \( v_i \) is the number of occurrences of \( i \) in \( \lambda \). Show that \( \sum_{\lambda \in \text{Par}} w_\lambda(x) r_\lambda(y) = \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} \).

(h) Show that \( \{r_\lambda\}_{\lambda \in \text{Par}} \) and \( \{w_\lambda\}_{\lambda \in \text{Par}} \) are dual bases of \( \Lambda \).

\(^{101}\)This is due to its relation with Witt vectors in the appropriate sense. Most of the work on this basis has been done by Reutenauer and Hazewinkel.

\(^{102}\)It also implicitly appears in [11, §5]. Indeed, the \( q_n \) of [11] are our \( w_n \) (for \( k = R \)).

\(^{103}\)See also Stanley [157, Exercise 7.46].
Exercise 2.80. For this exercise, set \( k = \mathbb{Z} \), and consider \( \Lambda = \mathbb{Z} \) as a subring of \( \Lambda_Q \). Also, consider \( \Lambda \otimes \mathbb{Z} \) as a subring of \( \Lambda_Q \otimes \Lambda_Q \).\(^{104}\) Recall that the family \( (p_n)_{n \geq 1} \) generates the \( \mathbb{Q} \)-algebra \( \Lambda_Q \), but does not generate the \( \mathbb{Z} \)-algebra \( \Lambda \).

(a) Define a \( \mathbb{Q} \)-linear map \( Z : \Lambda_Q \to \Lambda_Q \) by setting
\[
Z(p_\lambda) = z_\lambda p_\lambda
\]
where \( z_\lambda \) is defined as in Proposition 2.43.\(^{105}\) Show that \( Z(\Lambda) \subset \Lambda \).

(b) Define a \( \mathbb{Q} \)-algebra homomorphism \( \Delta_x : \Lambda_Q \to \Lambda_Q \otimes \Lambda_Q \) by setting
\[
\Delta_x(p_n) = p_n \otimes p_n
\]
for every positive integer \( n \).\(^{106}\) Show that \( \Delta_x(\Lambda) \subset \Lambda \otimes \mathbb{Z} \).

(c) Let \( r \in \mathbb{Z} \). Define a \( \mathbb{Q} \)-algebra homomorphism \( \epsilon_r : \Lambda_Q \to \mathbb{Q} \) by setting
\[
\epsilon_r(p_n) = r
\]
for every positive integer \( n \).\(^{107\)} Show that \( \epsilon_r(\Lambda) \subset \mathbb{Z} \).

(d) Let \( r \in \mathbb{Z} \). Define a \( \mathbb{Q} \)-algebra homomorphism \( i_r : \Lambda_Q \to \Lambda_Q \) by setting
\[
i_r(p_n) = rp_n
\]
for every positive integer \( n \).\(^{108\)} Show that \( i_r(\Lambda) \subset \Lambda \).

(e) Define a \( \mathbb{Q} \)-linear map \( Sq : \Lambda_Q \to \Lambda_Q \) by setting
\[
Sq(p_\lambda) = p_\lambda^2
\]
for every partition \( \lambda \).\(^{109\)} Show that \( Sq(\Lambda) \subset \Lambda \).

(f) Let \( r \in \mathbb{Z} \). Define a \( \mathbb{Q} \)-algebra homomorphism \( \Delta_r : \Lambda_Q \to \Lambda_Q \otimes \Lambda_Q \) by setting
\[
\Delta_r(p_n) = \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + r \otimes p_n + p_n \otimes r
\]
for every positive integer \( n \).\(^{110\)} Show that \( \Delta_r(\Lambda) \subset \Lambda \otimes \mathbb{Z} \).

(g) Consider the map \( \Delta_x \) introduced in Exercise 2.80(b) and the map \( \epsilon_1 \) introduced in Exercise 2.80(c). Show that the \( \mathbb{Q} \)-algebra \( \Lambda_Q \), endowed with the comultiplication \( \Delta_x \) and the counit \( \epsilon_1 \), becomes a cocommutative \( \mathbb{Q} \)-bialgebra.\(^{111\)}

(h) Define a \( \mathbb{Q} \)-bilinear map \( * : \Lambda_Q \times \Lambda_Q \to \Lambda_Q \), which will be written in infix notation (that is, we will write \( a * b \) instead of \( * (a, b) \)), by setting
\[
p_\lambda * p_\mu = \delta_{\lambda,\mu} z_\lambda p_\lambda
\]
for any partitions \( \lambda \) and \( \mu \) (where \( z_\lambda \) is defined as in Proposition 2.43).\(^{112\)} Show that \( f * g \in \Lambda \) for any \( f \in \Lambda \) and \( g \in \Lambda \).

(i) Show that \( \epsilon_1(f) = f(1) \) for every \( f \in \Lambda_Q \) (where we are using the notation \( \epsilon_r \) defined in Exercise 2.80(c)).

**Hint:**

\(^{104}\) Here is how this works: We have \( \Lambda_Q \cong \mathbb{Q} \otimes \Lambda \). But fundamental properties of tensor products yield
\[
(2.43) \quad \mathbb{Q} \otimes (\Lambda \otimes \Lambda) \cong (\mathbb{Q} \otimes \Lambda) \otimes (\mathbb{Q} \otimes \Lambda) \cong \Lambda_Q \otimes \Lambda_Q
\]
as \( \mathbb{Q} \)-algebras. But \( \Lambda \otimes \mathbb{Z} \) is a free \( \mathbb{Z} \)-module (since \( \Lambda \) is a free \( \mathbb{Z} \)-module), and so the canonical ring homomorphism \( \Lambda \otimes \mathbb{Z} \to \mathbb{Q} \otimes (\Lambda \otimes \Lambda) \) sending every \( u \) to \( 1_\mathbb{Q} \otimes u \) is injective. Composing this ring homomorphism with the \( \mathbb{Q} \)-algebra isomorphism of \( (2.43) \) gives an injective ring homomorphism \( \Lambda \otimes \mathbb{Z} \to \Lambda_Q \otimes \Lambda_Q \). We use this latter homomorphism to identify \( \Lambda \otimes \mathbb{Z} \) with a subring of \( \Lambda_Q \otimes \Lambda_Q \).

\(^{105}\) This is well-defined, since \( (p_\lambda)_{\lambda \in \text{Par}} \) is a \( \mathbb{Q} \)-module basis of \( \Lambda_Q \).

\(^{106}\) This is well-defined, since the family \( (p_n)_{n \geq 1} \) generates the \( \mathbb{Q} \)-algebra \( \Lambda_Q \) and is algebraically independent.

\(^{107}\) This is well-defined, since the family \( (p_n)_{n \geq 1} \) generates the \( \mathbb{Q} \)-algebra \( \Lambda_Q \) and is algebraically independent.

\(^{108}\) This is well-defined, since the family \( (p_n)_{n \geq 1} \) generates the \( \mathbb{Q} \)-algebra \( \Lambda_Q \) and is algebraically independent.

\(^{109}\) This is well-defined, since \( (p_\lambda)_{\lambda \in \text{Par}} \) is a \( \mathbb{Q} \)-module basis of \( \Lambda_Q \).

\(^{110}\) This is well-defined, since the family \( (p_n)_{n \geq 1} \) generates the \( \mathbb{Q} \)-algebra \( \Lambda_Q \) and is algebraically independent.

\(^{111}\) But unlike \( \Lambda_Q \) with the usual coalgebra structure, it is neither graded nor a Hopf algebra.

\(^{112}\) This is well-defined, since \( (p_\lambda)_{\lambda \in \text{Par}} \) is a \( \mathbb{Q} \)-module basis of \( \Lambda_Q \).
\[ \Delta (x_1 y_1) = \Delta x \otimes \Delta y \]

Remark 2.71. The map \( \Delta \) defined in Exercise 2.80(b) is known as the \textit{internal comultiplication} (or \textit{Kronecker comultiplication}) on \( \Lambda Q \). Unlike the standard comultiplication \( \Delta_{\Lambda Q} \), it is not a graded map, but rather sends every homogeneous component \( \Lambda Q_n \) into \( \Lambda Q_n \otimes \Lambda Q_n \). The bilinear map \( \ast \) from Exercise 2.80(h) is the so-called \textit{internal multiplication} (or \textit{Kronecker multiplication}), and is similarly not graded but rather takes \( \Lambda Q_m \times \Lambda Q_n \) to \( \Lambda Q_{m+n} \) if \( n = m \) and to 0 otherwise.

The analogy between the two internal structures is not perfect: While we saw in Exercise 2.80(g) how the internal comultiplication yields another bialgebra structure on \( \Lambda Q \), it is not true that the internal multiplication (combined with the usual coalgebra structure of \( \Lambda Q \)) forms a bialgebra structure as well. What is missing is a multiplicative unity; if we would take the closure of \( \Lambda Q \) with respect to the grading, then \( 1 + h_1 + h_2 + h_3 + \cdots \) would be such a unity.

The structure constants of the internal comultiplication on the Schur basis \( (s_\lambda)_{\lambda \in \text{Par}} \) are equal to the structure constants of the internal multiplication on the Schur basis \textsuperscript{114}, and are commonly referred to as the \textit{Kronecker coefficients}. They are known to be nonnegative integers (this follows from Exercise 4.38(c) \textsuperscript{115}), but no combinatorial proof is known for their nonnegativity. Combinatorial interpretations for these coefficients akin to the Littlewood-Richardson rule have been found only in special cases (cf., e.g., [138]).

The map \( \Delta_r \) of Exercise 2.80(f) also has some classical theory behind it, relating to Chern classes of tensor products ([117], [109, §1.4, example 5]).

Parts (b), (c), (d), (e) and (f) of Exercise 2.80 are instances of a general phenomenon: Many \( \mathbb{Z} \)-algebra homomorphisms \( \Lambda \to A \) (with \( A \) a commutative ring, usually torsionfree) are easiest to define by first defining a \( \mathbb{Q} \)-algebra homomorphism \( \Lambda Q \to A \otimes \mathbb{Q} \) and then showing that this homomorphism restricts to a \( \mathbb{Z} \)-algebra homomorphism \( \Lambda \to A \). One might ask for general criteria when this is possible; specifically, for what choices

\textsuperscript{113} e.g., it involves summing infinitely many \( x_i \)'s if \( f = e_1 \)

\textsuperscript{114} This can be obtained, e.g., from (2.44).

\textsuperscript{115} Their integrality can also be easily deduced from Exercise 2.80(b).
of \((b_n)_{n \geq 1} \in A^{(1, 2, 3, \ldots)}\) does there exist a \(Z\)-algebra homomorphism \(A \to A\) sending the \(p_n\) to \(b_n\)? Such choices are called \textit{ghost-Witt vectors} in Hazewinkel [66], and we can give various equivalent conditions for a family \((b_n)_{n \geq 1}\) to be a ghost-Witt vector:

\textbf{Exercise 2.82.} Let \(A\) be a commutative ring.

For every \(n \in \{1, 2, 3, \ldots\}\), let \(\varphi_n : A \to A\) be a ring endomorphism of \(A\). Assume that the following properties hold:

- We have \(\varphi_n \circ \varphi_m = \varphi_{nm}\) for any two positive integers \(n\) and \(m\).
- We have \(\varphi_1 = \text{id}\).
- We have \(\varphi_p (a) \equiv a^p \mod pA\) for every \(a \in A\) and every prime number \(p\).

(For example, when \(A = \mathbb{Z}\), one can set \(\varphi_n = \text{id}\) for all \(n\); this simplifies the exercise somewhat. More generally, setting \(\varphi_n = \text{id}\) works whenever \(A\) is a binomial ring.\(^{116}\) However, the results of this exercise are at their most useful when \(A\) is a multivariate polynomial ring \(\mathbb{Z}[x_1, x_2, \ldots] \) over \(\mathbb{Z}\) and the homomorphism \(\varphi_n\) sends every \(P \in A\) to \(P (x_1^n, x_2^n, \ldots)\).

Let \(\mu\) denote the \textit{number-theoretic M"obius function}; this is the function \(\{1, 2, 3, \ldots\} \to \mathbb{Z}\) defined by

\[
\mu (m) = \begin{cases} 
0, & \text{if } m \text{ is not squarefree;} \\
(1)^{(\text{number of prime factors of } m)}, & \text{if } m \text{ is squarefree} 
\end{cases}
\]

for every positive integer \(m\).

Let \(\phi\) denote the \textit{Euler totient function}; this is the function \(\{1, 2, 3, \ldots\} \to \mathbb{N}\) which sends every positive integer \(m\) to the number of elements of \(\{1, 2, \ldots, m\}\) coprime to \(m\).

Let \((b_n)_{n \geq 1} \in A^{(1, 2, 3, \ldots)}\) be a family of elements of \(A\). Prove that the following seven assertions are equivalent:

- \textbf{Assertion C:} For every positive integer \(n\) and every prime factor \(p\) of \(n\), we have

\[
\varphi_p \left( \frac{b_n}{p} \right) \equiv b_n \mod p^{\nu_p (n)} A.
\]

Here, \(\nu_p (n)\) denotes the exponent of \(p\) in the prime factorization of \(n\).

- \textbf{Assertion D:} There exists a family \((\alpha_n)_{n \geq 1} \in A^{(1, 2, 3, \ldots)}\) of elements of \(A\) such that every positive integer \(n\) satisfies

\[
b_n = \sum_{d \mid n} d \alpha_d^{n/d}.
\]

\(^{117}\)

- \textbf{Assertion E:} There exists a family \((\beta_n)_{n \geq 1} \in A^{(1, 2, 3, \ldots)}\) of elements of \(A\) such that every positive integer \(n\) satisfies

\[
b_n = \sum_{d \mid n} d \varphi_n (\beta_d).
\]

- \textbf{Assertion F:} Every positive integer \(n\) satisfies

\[
\sum_{d \mid n} \mu (d) \varphi_d \left( \frac{b_n}{d} \right) \in nA.
\]

- \textbf{Assertion G:} Every positive integer \(n\) satisfies

\[
\sum_{d \mid n} \phi (d) \varphi_d \left( \frac{b_n}{d} \right) \in nA.
\]

\(^{116}\)A \textit{binomial ring} is defined to be a torsionfree (as an additive group) commutative ring \(A\) which has one of the following equivalent properties:

- For every \(n \in \mathbb{N}\) and \(a \in A\), we have \(a (a - 1) \cdots (a - n + 1) \in n! \cdot A\). (That is, binomial coefficients \(\binom{a}{n}\) with \(a \in A\) and \(n \in \mathbb{N}\) are defined in \(A\).)

- We have \(a^p \equiv a \mod pA\) for every \(a \in A\) and every prime number \(p\).

See [172] and the references therein for studies of these rings. It is not hard to check that \(\mathbb{Z}\) and every localization of \(\mathbb{Z}\) are binomial rings, and so is any commutative \(\mathbb{Q}\)-algebra as well as the ring

\[
\{ P \in \mathbb{Q}[X] \mid P (n) \in \mathbb{Z} \text{ for every } n \in \mathbb{Z} \}
\]

(but not the ring \(\mathbb{Z}[X]\) itself).

\(^{117}\)Here and in the following, summations of the form \(\sum_{d \mid n}\) range over all \textbf{positive} divisors of \(n\).
• **Assertion \( \mathcal{H} \):** Every positive integer \( n \) satisfies

\[
\sum_{i=1}^{n} \varphi_{n / \gcd(i,n)}(b_{\gcd(i,n)}) \in nA.
\]

• **Assertion \( \mathcal{J} \):** There exists a ring homomorphism \( A_{\mathbb{Z}} \to A \) which, for every positive integer \( n \), sends \( p_{n} \) to \( b_{n} \).

**[Hint]:** The following identities hold for every positive integer \( n \):

\[
\begin{align*}
(2.45) & \quad \sum_{d|n} \phi(d) = n; \\
(2.46) & \quad \sum_{d|n} \mu(d) = \delta_{n,1}; \\
(2.47) & \quad \sum_{d|n} \mu(d) \frac{n}{d} = \phi(n); \\
(2.48) & \quad \sum_{d|n} \mu(d) \phi\left(\frac{n}{d}\right) = \mu(n).
\end{align*}
\]

Furthermore, the following simple lemma is useful: If \( k \) is a positive integer, and if \( p \in \mathbb{Z} \), \( a \in A \) and \( b \in A \) are such that \( a \equiv b \mod p^{k}A \), then \( ap^{\ell} \equiv bp^{\ell} \mod p^{k+\ell}A \) for every \( \ell \in \mathbb{N} \).]

**Remark 2.83.** Much of Exercise 2.82 is folklore, but it is hard to pinpoint concrete appearances in literature. The equivalence \( \mathcal{C} \iff \mathcal{D} \) appears in Hesselholt [71, Lemma 1] (in slightly greater generality), where it is referred to as Dwork’s lemma and used in the construction of the Witt vector functor. This equivalence is also [66, Lemma 9.93]. The equivalence \( \mathcal{D} \iff \mathcal{F} \iff \mathcal{G} \iff \mathcal{H} \) in the case \( A = \mathbb{Z} \) is [43, Corollary on p. 10], where it is put into the context of Burnside rings and necklace counting. The equivalence \( \mathcal{C} \iff \mathcal{F} \) for finite families \( (b_{n})_{n \in \{1,2,\ldots,m\}} \) in lieu of \( (b_{n})_{n \geq 1} \) is [157, Exercise 5.2 a]. One of the likely oldest relevant sources is Schur’s [148], which proves the equivalence \( \mathcal{C} \iff \mathcal{D} \iff \mathcal{F} \) for finite families \( (b_{n})_{n \in \{1,2,\ldots,m\}} \), as well as a “finite version” of \( \mathcal{C} \iff \mathcal{J} \) (Schur did not have \( \Lambda \), but was working with actual power sums of roots of polynomials).

**Exercise 2.84.** Let \( A \) denote the ring \( \mathbb{Z} \). For every \( n \in \{1,2,3,\ldots\} \), let \( \varphi_{n} \) denote the identity endomorphism \( \text{id} \) of \( A \). Prove that the seven equivalent assertions \( \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}, \mathcal{H} \) and \( \mathcal{J} \) of Exercise 2.82 are satisfied for each of the following families \( (b_{n})_{n \geq 1} \in \mathbb{Z}^{1,2,3,\ldots} \):

- the family \( (b_{n})_{n \geq 1} = (q^n)_{n \geq 1} \), where \( q \) is a given integer.
- the family \( (b_{n})_{n \geq 1} = (\binom{q^n}{r^n})_{n \geq 1} \), where \( r \in \mathbb{Q} \) and \( q \in \mathbb{Z} \) are given. (Here, a binomial coefficient \( \binom{a}{b} \) has to be interpreted as 0 when \( b \not\in \mathbb{N} \).)
- the family \( (b_{n})_{n \geq 1} = (\binom{q^n - 1}{r^n - 1})_{n \geq 1} \), where \( r \in \mathbb{Z} \) and \( q \in \mathbb{Z} \) are given.

**Exercise 2.85.** For every \( n \in \{1,2,3,\ldots\} \), define a map \( f_{n} : \Lambda \to \Lambda \) by setting

\[
f_{n}(a) = a(x_{1}^{n}, x_{2}^{n}, x_{3}^{n}, \ldots) \quad \text{for every } a \in \Lambda.
\]

(So what \( f_{n} \) does to a symmetric function is replacing all variables \( x_{1}, x_{2}, x_{3}, \ldots \) by their \( n \)-th powers.)

(a) Show that \( f_{n} : \Lambda \to \Lambda \) is a \( k \)-algebra homomorphism for every \( n \in \{1,2,3,\ldots\} \).
(b) Show that \( f_{n} \circ f_{m} = f_{nm} \) for any two positive integers \( n \) and \( m \).
(c) Show that \( f_{1} = \text{id} \).
(d) Prove that \( f_{n} : \Lambda \to \Lambda \) is a Hopf algebra homomorphism for every \( n \in \{1,2,3,\ldots\} \).
(e) Prove that \( f_{2}(h_{m}) = \sum_{i=0}^{2m} (-1)^{i} h_{i}h_{2m-i} \) for every \( m \in \mathbb{N} \).
(f) Assume that \( k = \mathbb{Z} \). Prove that \( f_{p}(a) = a^{p} \mod pA \) for every \( a \in \Lambda \) and every prime number \( p \).
(g) Use Exercise 2.82 to obtain new solutions to parts (b), (c), (d), (e) and (f) of Exercise 2.80.
The maps $f_n$ constructed in Exercise 2.85 are known as the Frobenius endomorphisms of $Λ$. They are a (deceptively) simple particular case of the notion of plethysms ([157, Chapter 7, Appendix 2]), and are often used as intermediate steps in computing more complicated plethysms\(^\text{118}\).

**Exercise 2.86.** For every $n \in \{1, 2, 3, \ldots\}$, define a $k$-algebra homomorphism $v_n : Λ \to Λ$ by

$$v_n(h_m) = \begin{cases} h_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases}$$

for every positive integer $m$

\(^{119}\).

(a) Show that any positive integers $n$ and $m$ satisfy

$$v_n(p_m) = \begin{cases} np_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases}.$$

(b) Show that any positive integers $n$ and $m$ satisfy

$$v_n(e_m) = \begin{cases} (-1)^{m-m/n} ne_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases}.$$

(c) Prove that $v_n \circ v_m = v_{nm}$ for any two positive integers $n$ and $m$.

(d) Prove that $v_1 = \text{id}$.

(e) Prove that $v_n : Λ \to Λ$ is a Hopf algebra homomorphism for every $n \in \{1, 2, 3, \ldots\}$.

Now, consider also the maps $f_n : Λ \to Λ$ defined in Exercise 2.85. Fix a positive integer $n$.

(f) Prove that the maps $f_n : Λ \to Λ$ and $v_n : Λ \to Λ$ are adjoint with respect to the Hall inner product on $Λ$.

(g) Show that $v_n \circ f_n = \text{id}^*n$.

(h) Prove that $f_n \circ v_m = v_m \circ f_n$ whenever $m$ is a positive integer coprime to $n$.

Finally, recall the $w_m \in Λ$ defined in Exercise 2.79.

(i) Show that any positive integer $m$ satisfies

$$v_n(w_m) = \begin{cases} w_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases}.$$

The homomorphisms $v_n : Λ \to Λ$ defined in Exercise 2.86 are called the Verschiebung endomorphisms of $Λ$; this name comes from German, where “Verschiebung” means “shift”. This terminology, as well as that of Frobenius endomorphisms, originates in the theory of Witt vectors, and the connection between the Frobenius and Verschiebung endomorphisms of $Λ$ and the identically named operators on Witt vectors is elucidated in [66, Chapter 13]\(^\text{120}\).

**Exercise 2.87.** Fix $n \in \mathbb{N}$. For any $n$-tuple $w = (w_1, w_2, \ldots, w_n)$ of integers, define the descent set $\text{Des}(w)$ of $w$ to be the set $\{i \in \{1, 2, \ldots, n-1\} : w_i > w_{i+1}\}$.

(a) We say that an $n$-tuple $w = (w_1, w_2, \ldots, w_n)$ is Smirnov if every $i \in \{1, 2, \ldots, n-1\}$ satisfies $w_i \neq w_{i+1}$.

Fix $k \in \mathbb{N}$, and let $X_{n,k} \in \mathbb{k}[[x]]$ denote the sum of the monomials $x_{w_1}x_{w_2}\cdots x_{w_n}$ over all Smirnov $n$-tuples $w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n$ satisfying $|\text{Des}(w)| = k$. Prove that $X_{n,k} \in Λ$.

(b) For any $n$-tuple $w = (w_1, w_2, \ldots, w_n)$, define the stagnation set $\text{Stag}(w)$ of $w$ to be the set $\{i \in \{1, 2, \ldots, n-1\} : w_i = w_{i+1}\}$. (Thus, an $n$-tuple is Smirnov if and only if its stagnation set is empty.)

For any $d \in \mathbb{N}$ and $s \in \mathbb{N}$, define a power series $X_{n,d,s} \in \mathbb{k}[[x]]$ as the sum of the monomials $x_{w_1}x_{w_2}\cdots x_{w_n}$ over all $n$-tuples $w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n$ satisfying $|\text{Des}(w)| = d$ and $|\text{Stag}(w)| = s$. Prove that $X_{n,d,s} \in Λ$ for any nonnegative integers $d$ and $s$.

\(^{118}\)In the notations of [157, (A2.160)], the value $f_n(a)$ for an $a \in Λ$ can be written as $a[p_n]$ or (when $k = \mathbb{Z}$) as $p_n[a]$.

\(^{119}\)This is well-defined, since the family $(h_m)_{m \geq 1}$ generates the $k$-algebra $Λ$ and is algebraically independent.

\(^{120}\)which is also where most of the statements of Exercises 2.85 and 2.86 come from.
(c) Assume that $n$ is positive. For any $d \in \mathbb{N}$ and $s \in \mathbb{N}$, define three further power series $U_{n,d,s}$, $V_{n,d,s}$ and $W_{n,d,s}$ in $k[[x]]$ by the following formulas:

\begin{align}
U_{n,d,s} &= \sum_{w=(w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n; \quad |\text{Des}(w)| = d; \quad |\text{Stag}(w)| = s; \quad w_1 < w_n} x_{w_1} x_{w_2} \cdots x_{w_n}; \quad (2.49) \\
V_{n,d,s} &= \sum_{w=(w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n; \quad |\text{Des}(w)| = d; \quad |\text{Stag}(w)| = s; \quad w_1 = w_n} x_{w_1} x_{w_2} \cdots x_{w_n}; \quad (2.50) \\
W_{n,d,s} &= \sum_{w=(w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n; \quad |\text{Des}(w)| = d; \quad |\text{Stag}(w)| = s; \quad w_1 > w_n} x_{w_1} x_{w_2} \cdots x_{w_n}. \quad (2.51)
\end{align}

Prove that these three power series $U_{n,d,s}$, $V_{n,d,s}$ and $W_{n,d,s}$ belong to $\Lambda$.

Remark 2.88. The function $X_{n,k}$ in Exercise 2.87(a) is a simple example ([151, Example 2.5, Theorem C.3]) of a chromatic quasisymmetric function that happens to be symmetric. See Shareshian/Wachs [151] for more general criteria for such functions to be symmetric, as well as deeper results. For example, [151, Theorem 6.3] gives an expansion for a wide class of chromatic quasisymmetric functions in the Schur basis of $\Lambda$, which, in particular, shows that our $X_{n,k}$ satisfies

\[ X_{n,k} = \sum_{\lambda \in \Par_n} a_{\lambda,k} s_\lambda, \]

where $a_{\lambda,k}$ is the number of all assignments $T$ of entries in $\{1,2,\ldots,n\}$ to the cells of the Ferrers diagram of $\lambda$ such that the following four conditions are satisfied:

- Every element of $\{1,2,\ldots,n\}$ is used precisely once in the assignment (i.e., we have $\cont(T) = (1^n)$).
- Whenever a cell $y$ of the Ferrers diagram lies immediately to the right of a cell $x$, we have $T(y) \leq T(x)$.
- Whenever a cell $y$ of the Ferrers diagram lies immediately below a cell $x$, we have $T(y) \geq T(x)$.
- There exist precisely $k$ elements $i \in \{1,2,\ldots,n-1\}$ such that the cell $T^{-1}(i)$ lies in a row below $T^{-1}(i+1)$.

Are there any such rules for the $X_{n,d,s}$ of part (b)?

Smirnov $n$-tuples are more usually called Smirnov words, or (occasionally) Carlitz words.

Exercise 2.89. (a) Let $n \in \mathbb{N}$. Define a matrix $A_n = (a_{i,j})_{i,j=1,2,\ldots,n}$ by

\[ a_{i,j} = \begin{cases} p_{i-j+1}, & \text{if } i \geq j; \\
i, & \text{if } i = j - 1; \\
0, & \text{if } i < j - 1 \end{cases} \quad \text{for all } (i,j) \in \{1,2,\ldots,n\}^2. \]

This matrix $A_n$ looks as follows:

\[ A_n = \begin{pmatrix}
p_1 & 1 & 0 & \cdots & 0 & 0 \\
p_2 & p_1 & 2 & \cdots & 0 & 0 \\
p_3 & p_2 & p_1 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
p_{n-1} & p_{n-2} & p_{n-3} & \cdots & p_1 & n-1 \\
p_n & p_{n-1} & p_{n-2} & \cdots & p_2 & p_1 \\
\end{pmatrix}. \]

Show that $\det(A_n) = n!e_n$.

(b) Let $n$ be a positive integer. Define a matrix $B_n = (b_{i,j})_{i,j=1,2,\ldots,n}$ by

\[ b_{i,j} = \begin{cases} ie_i, & \text{if } j = 1; \\
e_{i-j+1}, & \text{if } j > 1 \end{cases} \quad \text{for all } (i,j) \in \{1,2,\ldots,n\}^2. \]
The matrix $B_n$ looks as follows:

$$B_n = \begin{pmatrix}
e_1 & e_0 & e_{-1} & \cdots & e_{-n+3} & e_{-n+2} \\
2e_2 & e_1 & e_0 & \cdots & e_{-n+4} & e_{-n+3} \\
3e_3 & e_2 & e_1 & \cdots & e_{-n+5} & e_{-n+4} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
(n-1)e_{n-1} & e_{n-2} & e_{n-3} & \cdots & e_1 & e_0 \\
n^2e_n & e_{n-1} & e_{n-2} & \cdots & e_2 & e_1
\end{pmatrix}$$

Show that $\det(B_n) = p_n$.

The formulas of Exercise 2.89, for finitely many variables, appear in Prasolov’s [130, §4.1][121]. In [130, §4.2], Prasolov gives four more formulas, which express $e_n$ as a polynomial in the $h_1, h_2, h_3, \ldots, h_n$ as a polynomial in the $e_1, e_2, e_3, \ldots, p_n$ as a polynomial in the $h_1, h_2, h_3, \ldots, n!h_n$ as a polynomial in the $p_1, p_2, p_3, \ldots$. These are not novel for us, since the first two of them are particular cases of Theorem 2.30, whereas the latter two can be derived from Exercise 2.89 by applying $\omega$. (Note that $\omega$ is only well-defined on symmetric functions in infinitely many indeterminates, so we need to apply $\omega$ before evaluating at finitely many indeterminates; this explains why Prasolov has to prove the latter two identities separately.)

**Exercise 2.90.** In the following, if $k \in \mathbb{N}$, we shall use the notation $1^k$ for $1, 1, \ldots, 1$ (in contexts such as $(n, 1^m)$). So, for example, $(3, 1^4)$ is the partition $(3, 1, 1, 1, 1)$.

(a) Show that $e_n h_m = s_{(m+1, 1^{n-1})} + s_{(m, 1^n)}$ for any two positive integers $n$ and $m$.

(b) Show that

$$\sum_{i=0}^{b} (-1)^i h_{a+i-1} e_{b-i} = s_{(a+1, 1^b)}$$

for any $a \in \mathbb{N}$ and $b \in \mathbb{N}$.

(c) Show that

$$\sum_{i=0}^{b} (-1)^i h_{a+i+1} e_{b-i} = (-1)^b \delta_{a+b, -1}$$

for any negative integer $a$ and every $b \in \mathbb{N}$. (As usual, we set $h_j = 0$ for $j < 0$ here.)

(d) Show that

$$\Delta s_{(a+1, 1^b)} = 1 \otimes s_{(a+1, 1^b)} + s_{(a+1, 1^b)} \otimes 1 + \sum_{(c,d,e) \in \mathbb{N}^3} s_{(c+1, 1^d)} \otimes s_{(e+1, 1^f)} + \sum_{(c,d,e) \in \mathbb{N}^3} s_{(c+1, 1^d)} \otimes s_{(e+1, 1^f)}$$

for any $a \in \mathbb{N}$ and $b \in \mathbb{N}$.

Our next few exercises survey some results on Littlewood-Richardson coefficients.

**Exercise 2.91.** Let $m \in \mathbb{N}$ and $k \in \mathbb{N}$. Let $\lambda$ and $\mu$ be two partitions such that $\ell(\lambda) \leq k$ and $\ell(\mu) \leq k$.

Assume that all parts of $\lambda$ and all parts of $\mu$ are $\leq m$ (that is, $\lambda_i \leq m$ and $\mu_i \leq m$ for every positive integer $i$).

Let $\lambda'$ and $\mu'$ denote the $k$-tuples $(m - \lambda_k, m - \lambda_{k-1}, \ldots, m - \lambda_1)$ and $(m - \mu_k, m - \mu_{k-1}, \ldots, m - \mu_1)$, respectively.

\[\text{where our symmetric functions } e_k, h_k, p_k, \text{ evaluated in finitely many indeterminates, are denoted } s_k, p_k, s_k, \text{ respectively.}\]

\[\text{As usual, we are denoting by } \nu_i \text{ the } i\text{-th part of a partition } \nu \text{ here.}\]
(a) Show that $\lambda^\nu$ and $\mu^\nu$ are partitions, and that $s_{\lambda/\mu} = s_{\mu^\nu/\lambda^\nu}$.

(b) Show that $c^\lambda_{\mu,\nu} = c^\mu_{\lambda^\nu,\nu}$ for any partition $\nu$.

(c) Let $\nu$ be a partition such that $\ell(\nu) \leq k$, and such that all parts of $\nu$ are $\leq m$. Let $\nu^\nu$ denote the $k$-tuple $(m - \nu_k, m - \nu_{k-1}, \ldots, m - \nu_1)$. Show that $\nu^\nu$ is a partition, and satisfies

$$c^\lambda_{\mu,\nu} = c^\nu_{\lambda^\nu,\mu} = c^\nu_{\mu^\nu,\nu^\nu} = c^\nu_{\nu^\nu,\mu^\nu}.$$

(d) Show that

$$s_{\lambda^\nu}(x_1, x_2, \ldots, x_k) = (x_1 x_2 \cdots x_k)^m \cdot s_\lambda(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1})$$

in the Laurent polynomial ring $k[x_1, x_2, \ldots, x_k]$.

(e) Let $r$ be a nonnegative integer. Show that $(r + \lambda_1, r + \lambda_2, \ldots, r + \lambda_k)$ is a partition and satisfies

$$s_{r + \lambda_1, r + \lambda_2, \ldots, r + \lambda_k}(x_1, x_2, \ldots, x_k) = (x_1 x_2 \cdots x_k)^r \cdot s_\lambda(x_1, x_2, \ldots, x_k)$$

in the polynomial ring $k[x_1, x_2, \ldots, x_k]$.

**Exercise 2.92.** Let $m \in \mathbb{N}$, $n \in \mathbb{N}$ and $k \in \mathbb{N}$. Let $\mu$ and $\nu$ be two partitions such that $\ell(\mu) \leq k$ and $\ell(\nu) \leq k$. Assume that all parts of $\mu$ are $\leq m$ (that is, $\mu_i \leq m$ for every positive integer $i$) and that all parts of $\nu$ are $\leq n$ (that is, $\nu_i \leq n$ for every positive integer $i$). Let $\mu^\nu(m)$ denote the $k$-tuple $(m - \mu_k, m - \mu_{k-1}, \ldots, m - \mu_1)$, and let $\nu^\nu(n)$ denote the $k$-tuple $(n - \nu_k, n - \nu_{k-1}, \ldots, n - \nu_1)$.

(a) Show that $\nu^\nu(n)$ are partitions.

Now, let $\lambda$ be a further partition such that $\ell(\lambda) \leq k$.

(b) If not all parts of $\lambda$ are $\leq m + n$, then show that $c^\lambda_{\mu,\nu} = 0$.

(c) If all parts of $\lambda$ are $\leq m + n$, then show that $c^\lambda_{\mu,\nu} = c^\nu_{\lambda^\nu,\nu^\nu}$, where $\lambda^\nu(m+n)$ denotes the $k$-tuple $(m + n - \lambda_k, m + n - \lambda_{k-1}, \ldots, m + n - \lambda_1)$.

The results of Exercise 2.69(c) and Exercise 2.91(c) are two symmetries of Littlewood-Richardson coefficients; combining them yields further such symmetries. While these symmetries were relatively easy consequences of our algebraic definition of the Littlewood-Richardson coefficients, it is a much more challenging task to derive them bijectively from a combinatorial definition of these coefficients (such as the one given in Corollary 2.58). Some such derivations appear in [168], in [15, Example 3.6, Proposition 5.11 and references therein], [54, §5.1, §A.1, §A.4 and 87, (2.12)] (though a different combinatorial interpretation of $c^\lambda_{\mu,\nu}$ is used in the latter three).

**Exercise 2.93.** Recall our usual notations: For every partition $\lambda$ and every positive integer $i$, the $i$-th entry of $\lambda$ is denoted by $\lambda_i$. The sign $\triangleright$ stands for dominance order. We let $\lambda^t$ denote the conjugate partition of a partition $\lambda$.

For any two partitions $\mu$ and $\nu$, we define two new partitions $\mu + \nu$ and $\mu \triangleright \nu$ of $|\mu| + |\nu|$ as follows:

- The partition $\mu + \nu$ is defined as $(\mu_1 + \nu_1, \mu_2 + \nu_2, \mu_3 + \nu_3, \ldots)$.
- The partition $\mu \triangleright \nu$ is defined as the result of sorting the list $(\mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}, \nu_1, \nu_2, \ldots, \nu_{\ell(\nu)})$ in decreasing order.

(a) Show that any two partitions $\mu$ and $\nu$ satisfy $(\mu + \nu)^t = \mu^t + \nu^t$ and $(\mu \triangleright \nu)^t = \mu^t + \nu^t$.

(b) Show that any two partitions $\mu$ and $\nu$ satisfy $c^\mu_{\mu^t,\nu^t} = 1$ and $c^\mu_{\mu^t,\nu} = 1$.

(c) If $k \in \mathbb{N}$ and $n \in \mathbb{N}$ satisfy $k \leq n$, and if $\mu \in \text{Par}_k$, $\nu \in \text{Par}_{n-k}$ and $\lambda \in \text{Par}_n$ are such that $c^\lambda_{\mu,\nu} \neq 0$, then prove that $\mu + \nu \triangleright \lambda \triangleright \mu \triangleright \nu$.

(d) If $n \in \mathbb{N}$ and $m \in \mathbb{N}$ and $\alpha, \beta \in \text{Par}_n$ and $\gamma, \delta \in \text{Par}_m$ are such that $\alpha \triangleright \beta$ and $\gamma \triangleright \delta$, then show that $\alpha + \gamma \triangleright \beta + \delta$ and $\alpha \triangleright \gamma \triangleright \beta \triangleright \delta$.

(e) Let $m \in \mathbb{N}$ and $k \in \mathbb{N}$, and let $\lambda$ be the partition $(m^k) = \left(\begin{array}{c} m, m, \ldots, m \\ k \text{ times} \end{array}\right)$. Show that any two partitions $\mu$ and $\nu$ satisfy $c^\lambda_{\mu,\nu} \in \{0, 1\}$.

(f) Let $a \in \mathbb{N}$ and $b \in \mathbb{N}$, and let $\lambda$ be the partition $(a + 1, b^k)$ (using the notation of Exercise 2.90). Show that any two partitions $\mu$ and $\nu$ satisfy $c^\lambda_{\mu,\nu} \in \{0, 1\}$.

---

123 As usual, we are denoting by $\nu_i$ the $i$-th part of a partition $\nu$ here.

124 The result of Exercise 2.92(c) can also be regarded as a symmetry of Littlewood-Richardson coefficients; see [10, §3.3].
Exercise 2.93(g) is part of Stembridge’s [161, Thm. 2.1]; we refer to that article for further results of its kind.

The Littlewood-Richardson rule comes in many different forms, whose equivalence is not always immediate. Our version (Corollary 2.58) has the advantage of being the simplest to prove and one of the simplest to state. Other versions can be found in [157, appendix to Ch. 7], Fulton’s [54, Ch. 5] and van Leeuwen’s [98]. We restrict ourselves to proving some very basic equivalences that allow us to restate parts of Corollary 2.58:

Exercise 2.94. We shall use the following notations:

- If $T$ is a column-strict tableau and $j$ is a positive integer, then we use the notation $T|_{\text{cols} \geq j}$ for the restriction of $T$ to the union of its columns $j, j + 1, j + 2, \ldots$ (This notation has already been used in Section 2.6.)
- If $T$ is a column-strict tableau and $S$ is a set of cells of $T$, then we write $T|_S$ for the restriction of $T$ to the set $S$ of cells.
- If $T$ is a column-strict tableau, then an NE-set of $T$ means a set $S$ of cells of $T$ such that whenever $s \in S$, every cell of $T$ which lies northeast of $s$ must also belong to $S$.
- The Semitic reading word of a column-strict tableau $T$ is the concatenation $r_1 r_2 r_3 \cdots$, where $r_i$ is the word obtained by reading the $i$-th row of $T$ from right to left.
- If $w = (w_1, w_2, \ldots, w_n)$ is a word, then a prefix of $w$ means a word of the form $(w_1, w_2, \ldots, w_i)$ for some $i \in \{0, 1, \ldots, n\}$. (In particular, both $w$ and the empty word are prefixes of $w$.)

A word $w$ over the set of positive integers is said to be Yamanouchi if for any prefix $v$ of $w$ and any positive integer $i$, there are at least as many $i$’s among the letters of $v$ as there are $(i + 1)$’s among them.

Prove the following two statements:

(a) Let $\mu$ be a partition. Let $b_{i,j}$ be a nonnegative integer for every two positive integers $i$ and $j$. Assume that $b_{i,j} = 0$ for all but finitely many pairs $(i,j)$.

The following two assertions are equivalent:
- **Assertion A**: There exist a partition $\lambda$ and a column-strict tableau $T$ of shape $\lambda/\mu$ such that all $(i,j) \in \{1, 2, 3, \ldots\}^2$ satisfy

\[
(2.52) \quad b_{i,j} = \text{(the number of all entries } i \text{ in the } j\text{-th row of } T). 
\]

- **Assertion B**: The inequality

\[
(2.53) \quad \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) \leq \mu_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j})
\]

holds for all $(i,j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}$.

---

125 A partition is called rectangular if it has the form \((m^k) = \left(\begin{array}{c} m, m, \ldots, m \\ k \text{ times} \end{array}\right)\) for some $m \in \mathbb{N}$ and $k \in \mathbb{N}$.
126 This restriction $T|_S$ is not necessarily a tableau of skew shape; it is just a map from $S$ to $\{1, 2, 3, \ldots\}$. The content $\text{cont}(T|_S)$ is nevertheless well-defined (in the usual way: $\text{cont}(T|_S)_i = \left[\begin{array}{c} T|_S \end{array}\right]^{-1}(i)$).
127 A cell $(r,c)$ is said to lie northeast of a cell $(r', c')$ if and only if we have $r \leq r'$ and $c \geq c'$.
128 The notation comes from [98] and is a reference to the Arabic and Hebrew way of writing.
129 If $s_1, s_2, s_3, \ldots$ are several words (finitely or infinitely many), then the concatenation $s_1 s_2 s_3 \cdots$ is defined as the word which is obtained by starting with the empty word, then appending $s_1$ to its end, then appending $s_2$ to the end of the result, then appending $s_3$ to the end of the result, etc.
130 For example, the Semitic reading word of the tableau

\[
\begin{array}{cccc}
3 & 4 & 4 & 5 \\
1 & 4 & 6 \\
3 & 5
\end{array}
\]

is 544364153.

The Semitic reading word of a tableau $T$ is what is called the reverse reading word of $T$ in [157, §A1.3].
131 For instance, the words 11213231232 and 1121323123111 are not. The Dyck words (written using 1’s and 2’s) are precisely the Yamanouchi words whose letters are 1’s and 2’s.

Yamanouchi words are often called lattice permutations.
(b) Let λ and µ be two partitions, and let T be a column-strict tableau of shape λ/µ. Then, the following
five assertions are equivalent:
- Assertion C: For every positive integer j, the weak composition \(\text{cont}(T|_{\text{cols} \geq j})\) is a partition.
- Assertion D: For every positive integers j and i, the number of entries \(i+1\) in the first j rows\(^{132}\) of T is \(\leq\) to the number of entries i in the first \(j-1\) rows of T.
- Assertion E: For every NE-set S of T, the weak composition \(\text{cont}(T|_S)\) is a partition.
- Assertion F: The Semitic reading word of T is Yamanouchi.
- Assertion G: There exists a column-strict tableau S whose shape is a partition and which satisfies the following property: For any positive integers i and j, the number of entries i in the j-th row of T equals the number of entries j in the i-th row of S.

Remark 2.95. The equivalence of Assertions C and F in Exercise 2.94(b) is the “not-too-difficult exercise” mentioned in [160]. It yields the equivalence between our version of the Littlewood-Richardson rule (Corollary 2.58) and that in [157, A1.3.3].

In the next exercises, we shall restate Corollary 2.57 in a different form. While Corollary 2.57 provided a decomposition of the product of a skew Schur function with a Schur function into a sum of Schur functions, the different form that we will encounter in Exercise 2.97(b) will give a combinatorial interpretation for the Hall inner product between two skew Schur functions. Let us first generalize Exercise 2.94(b):

**Exercise 2.96.** Let us use the notations of Exercise 2.94. Let \(\kappa, \lambda, \mu\) be three partitions, and let T be a column-strict tableau of shape \(\lambda/\mu\).

(a) Prove that the following five assertions are equivalent:
- Assertion \(\mathcal{C}(\kappa)\): For every positive integer j, the weak composition \(\kappa + \text{cont}(T|_{\text{cols} \geq j})\) is a partition.
- Assertion \(\mathcal{D}(\kappa)\): For every positive integers j and i, we have
  \[\kappa_{i+1} + (\text{the number of entries } i+1 \text{ in the first } j \text{ rows of } T)\]
  \[\leq \kappa_i + (\text{the number of entries } i \text{ in the first } j - 1 \text{ rows of } T).\]
- Assertion \(\mathcal{E}(\kappa)\): For every NE-set S of T, the weak composition \(\kappa + \text{cont}(T|_S)\) is a partition.
- Assertion \(\mathcal{F}(\kappa)\): For every prefix v of the Semitic reading word of T, and for every positive integer i, we have
  \[\kappa_i + (\text{the number of } i^{\prime}s \text{ among the letters of } v)\]
  \[\geq \kappa_{i+1} + (\text{the number of } (i+1)^{\prime}s \text{ among the letters of } v).\]
- Assertion \(\mathcal{G}(\kappa)\): There exist a partition \(\zeta\) and a column-strict tableau S of shape \(\zeta/\kappa\) which satisfies the following property: For any positive integers i and j, the number of entries i in the j-th row of T equals the number of entries j in the i-th row of S.

(b) Let \(\tau\) be a partition such that \(\tau = \kappa + \text{cont } T\). Consider the five assertions \(\mathcal{C}(\kappa), \mathcal{D}(\kappa), \mathcal{E}(\kappa), \mathcal{F}(\kappa)\) and \(\mathcal{G}(\kappa)\) introduced in Exercise 2.96(a). Let us also consider the following assertion:
- Assertion \(\mathcal{H}(\kappa)\): There exists a column-strict tableau S of shape \(\tau/\kappa\) which satisfies the following property: For any positive integers i and j, the number of entries i in the j-th row of T equals the number of entries j in the i-th row of S.

Prove that the six assertions \(\mathcal{C}(\kappa), \mathcal{D}(\kappa), \mathcal{E}(\kappa), \mathcal{F}(\kappa), \mathcal{G}(\kappa)\) and \(\mathcal{H}(\kappa)\) are equivalent.

Clearly, Exercise 2.94(b) is the particular case of Exercise 2.96 when \(\kappa = \emptyset\).

Using Exercise 2.96, we can restate Corollary 2.57 in several ways:

**Exercise 2.97.** Let \(\lambda, \mu, \kappa\) be three partitions.

(a) Show that
\[s_{\kappa}s_{\lambda/\mu} = \sum_T s_{\kappa+\text{cont } T},\]
where the sum ranges over all column-strict tableaux T of shape \(\lambda/\mu\) satisfying the five equivalent assertions \(\mathcal{C}(\kappa), \mathcal{D}(\kappa), \mathcal{E}(\kappa), \mathcal{F}(\kappa)\) and \(\mathcal{G}(\kappa)\) introduced in Exercise 2.96(a).

\(^{132}\)The “first j rows” mean the 1-st row, the 2-nd row, etc., the j-th row (even if some of these rows are empty).
(b) Let $\tau$ be a partition. Show that $(s_{\lambda/\mu}, s_{\tau/\kappa})_A$ is the number of all column-strict tableaux $T$ of shape $\lambda/\mu$ satisfying $\tau = \kappa + \text{cont } T$ and also satisfying the six equivalent assertions $C^{(\kappa)}, D^{(\kappa)}, E^{(\kappa)}, F^{(\kappa)}, G^{(\kappa)}$ and $H^{(\kappa)}$ introduced in Exercise 2.96.

Exercise 2.97(a) is merely Corollary 2.57, rewritten in light of Exercise 2.96. Various parts of it appear in the literature. For instance, [95, (53)] easily reveals to be a restatement of the fact that $s_e s_{\lambda/\mu} = \sum_T s_{\mu+\text{cont } T}$, where the sum ranges over all column-strict tableaux $T$ of shape $\lambda/\mu$ satisfying Assertion $D^{(\kappa)}$.

Exercise 2.97(b) is one version of a “skew Littlewood-Richardson rule” that goes back to Zelevinsky [174] (although Zelevinsky’s version uses both a different language and a combinatorial interpretation which is not obviously equivalent to ours). It appears in various sources; for instance, [95, Theorem 5.2, second formula] says that $(s_{\lambda/\mu}, s_{\tau/\kappa})_A$ is the number of all column-strict tableaux $T$ of shape $\lambda/\mu$ satisfying $\tau = \kappa + \text{cont } T$ and the assertion $H^{(\kappa)}$, whereas [56, Theorem 1.2] says that $(s_{\lambda/\mu}, s_{\tau/\kappa})_A$ is the number of all column-strict tableaux $T$ of shape $\lambda/\mu$ satisfying $\tau = \kappa + \text{cont } T$ and the assertion $F^{(\kappa)}$. (Notice that Gasharov’s proof of [56, Theorem 1.2] uses the same involutions as Stembridge’s proof of Theorem 2.53; it can thus be regarded as a close precursor to Stembridge’s proof. However, it uses the Jacobi-Trudi identities, while Stembridge’s does not.)

**Exercise 2.98.** Let $K$ be a field.\(^{133}\) If $N \in K^{n \times n}$ is a nilpotent matrix, then the Jordan type of $N$ is defined to be the list of the sizes of the Jordan blocks in the Jordan normal form of $N$, sorted in decreasing order.\(^{134}\) This Jordan type is a partition of $n$, and uniquely determines $N$ up to similarity (i.e., two nilpotent $n \times n$-matrices $N$ and $N'$ are similar if and only if the Jordan types of $N$ and $N'$ are equal). If $f$ is a nilpotent endomorphism of a finite-dimensional $K$-vector space $V$, then we define the Jordan type of $f$ as the Jordan type of any matrix representing $f$ (the choice of the matrix does not matter, since the Jordan type of a matrix remains unchanged under conjugation).

(a) Let $n \in \mathbb{N}$. Let $N \in K^{n \times n}$ be a nilpotent matrix. Let $\lambda \in \text{Par}_n$. Show that the matrix $N$ has Jordan type $\lambda$ if and only if for every $k \in \mathbb{N}$ satisfies
\[
\dim \left( \text{ker} \left( N^k \right) \right) = (\lambda^t)_1 + (\lambda^t)_2 + \ldots + (\lambda^t)_k.
\]
(Here, we are using the notation $\lambda^t$ for the transpose of a partition $\lambda$, and the notation $\nu_i$ for the $i$-th entry of a partition $\nu$.)

(b) Let $f$ be a nilpotent endomorphism of a finite-dimensional $K$-vector space $V$. Let $U$ be an $f$-stable $K$-vector subspace of $V$ (that is, a $K$-vector subspace of $V$ satisfying $f(U) \subset U$). Then, restricting $f$ to $U$ gives a nilpotent endomorphism $f|U$ of $U$, and the endomorphism $f$ also induces a nilpotent endomorphism $\overline{f}$ of the quotient space $V/U$. Let $\lambda, \mu$ and $\nu$ be the Jordan types of $f, f|U$ and $\overline{f}$, respectively. Show that $c^{\lambda,\mu,\nu}_{\lambda',\mu',\nu'} \neq 0$ if $Z$ is a subring of $K$.

[Hint: For (b), Exercise 2.69(c) shows that it is enough to prove that $c^{\lambda',\mu',\nu'}_{\lambda,\mu,\nu} \neq 0$. Due to Corollary 2.58, this only requires constructing a column-strict tableau $T$ of shape $\lambda'/\mu'$ with $\text{cont } T = \nu'$ which has the property that each $\text{cont } (T|_{\text{cols} \geq j})$ is a partition. Construct this tableau by defining $\alpha_{i,j} = \dim \left( \left( f^t \right)^{-1} (U) \cap \ker \left( f^j \right) \right)$ for all $(i,j) \in \mathbb{N}^2$, and requiring that the number of entries $i$ in the $j$-th row of $T$ be $a_{i,j} - a_{i-1,j-1} + a_{i-1,j-1}$ for all $(i,j) \in \{1,2,3,\ldots\}$. Use Exercise 2.94(a) to prove that this indeed defines a column-strict tableau, and Exercise 2.94(b) to verify that it satisfies the condition on $\text{cont } (T|_{\text{cols} \geq j})$.]

**Remark 2.99.** Exercise 2.98 is a taste of the connections between the combinatorics of partitions and the Jordan normal form. Much more can, and has, been said. Marc van Leeuwen’s [96] is dedicated to some of these connections; in particular, our Exercise 2.98(a) is [96, Proposition 1.1], and a far stronger version of Exercise 2.98(b) appears in [96, Theorem 4.3 (2)], albeit only for the case of an infinite $K$. One can prove a converse to Exercise 2.98(b) as well: If $c^{\lambda,\mu,\nu}_{\lambda',\mu',\nu'} \neq 0$, then there exist $V, f$ and $U$ satisfying the premises of Exercise 2.98(b). When $K$ is a finite field, we can ask enumerative questions, such as how many $U$’s are there for given $V, f, \lambda, \mu$ and $\nu$; we will see a few answers in Section 4.9 (specifically, Proposition 4.58), and a more detailed treatment is given in [109, Ch. 2].

---

\(^{133}\)This field has no relation to the ring $k$, over which our symmetric functions are defined.

\(^{134}\)The Jordan normal form of $N$ is well-defined even if $K$ is not algebraically closed, because $N$ is nilpotent (so the characteristic polynomial of $N$ is $X^n$).
The relationship between partitions and Jordan normal forms can be exploited to provide linear-algebraic proofs of purely combinatorial facts. See [23, Sections 6 and 9] for some examples. Note that [23, Lemma 9.10] is the statement that, under the conditions of Exercise 2.98(b), we have $\nu \subseteq \lambda$. This is a direct consequence of Exercise 2.98(b) (since $c_{\mu,\nu}^\lambda \neq 0$ can happen only if $\nu \subseteq \lambda$).
3. Zelevinsky’s structure theory of positive self-dual Hopf algebras

Chapter 2 showed that, as a Z-basis for the Hopf algebra \( \Lambda = \Lambda_Z \), the Schur functions \( \{ s_\lambda \} \) have two special properties: they have the same structure constants \( c_{\lambda,\mu,\nu}^\lambda \) for their multiplication as for their comultiplication (Corollary 2.35), and these structure constants are all nonnegative integers (Corollary 2.58). Zelevinsky [173, §2.3] isolated these two properties as crucial.

**Definition 3.1.** Say that a graded connected Hopf algebra \( A \) over \( k = \mathbb{Z} \) with a distinguished \( \mathbb{Z} \)-basis \( \{ \sigma_\lambda \} \) consisting of homogeneous elements is a **positive self-dual Hopf algebra** (or PSH) if it satisfies the two further axioms

- **(self-duality)** The same structure constants \( a^\lambda_{\mu,\nu} \) appear for the product \( \sigma_\mu \sigma_\nu = \sum_\lambda a^\lambda_{\mu,\nu} \sigma_\lambda \) and the coproduct \( \Delta \sigma_\lambda = \sum_{\mu,\nu} a^\lambda_{\mu,\nu} \sigma_\mu \otimes \sigma_\nu \).

- **(positivity)** The \( a^\lambda_{\mu,\nu} \) are all nonnegative (integers).

Call \( \{ \sigma_\lambda \} \) the **PSH-basis** of \( A \).

He then developed a beautiful structure theory for PSH’s, explaining how they can be uniquely expressed as tensor products of copies of PSH’s each isomorphic to \( \Lambda \) after rescaling their grading. The next few sections explain this, following his exposition closely.

3.1. **Self-duality implies polynomiality.** We begin with a property that forces a Hopf algebra to have algebra structure which is a polynomial algebra, specifically the symmetric algebra \( \text{Sym}(\mathfrak{p}) \), where \( \mathfrak{p} \) is the \( k \)-submodule of primitive elements.

Recall from Exercise 1.26 that for a graded connected Hopf algebra \( A = \bigoplus_{n=0}^\infty A_n \), every \( x \) in the two-sided ideal \( I := \ker \epsilon = \bigoplus_{n>0} A_n \) has the property that its comultiplication takes the form

\[ \Delta(x) = 1 \otimes x + x \otimes 1 + \Delta_+(x) \]

where \( \Delta_+(x) \) lies in \( I \otimes I \). Recall also that the elements \( x \) for which \( \Delta_+(x) = 0 \) are called the **primitives**. Denote by \( p \) the \( k \)-submodule of primitive elements inside \( A \).

Given a PSH \( A \) (over \( k = \mathbb{Z} \)) with a PSH-basis \( \{ \sigma_\lambda \} \), we consider the bilinear form \( (\cdot, \cdot)_A : A \times A \to \mathbb{Z} \) on \( A \) that makes this basis orthonormal. Similarly, the elements \( \{ \sigma_\lambda \otimes \sigma_\mu \} \) give an orthonormal basis for a form \( (\cdot, \cdot)_{A \otimes A} \) on \( A \otimes A \). The bilinear form \( (\cdot, \cdot)_A \) on the PSH \( A \) gives rise to a \( \mathbb{Z} \)-linear map \( A \to A^* \), which is easily seen to be injective and a \( \mathbb{Z} \)-algebra homomorphism. We thus identify \( A \) with a subalgebra of \( A^* \). When \( A \) is of finite type, this map is a Hopf algebra isomorphism, thus allowing us to identify \( A \) with \( A^o \).

This is an instance of the following notion of self-duality.

**Definition 3.2.**

(a) If \( (\cdot, \cdot) : V \times W \to k \) is a bilinear form on the product \( V \times W \) of two \( k \)-modules \( V = \bigoplus_{n \geq 0} V_n \) and \( W = \bigoplus_{n \geq 0} W_n \), then we say that this form \( (\cdot, \cdot) \) is **graded** if every two distinct nonnegative integers \( n \) and \( m \) satisfy \( (V_n, W_m) = 0 \) (that is, if every two homogeneous elements \( v \in V \) and \( w \in W \) having distinct degrees satisfy \( (v, w) = 0 \)).

(b) If \( (\cdot, \cdot)_V : V \times V \to k \) and \( (\cdot, \cdot)_W : W \times W \to k \) are two symmetric bilinear forms on some \( k \)-modules \( V \) and \( W \), then we can canonically define a symmetric bilinear form \( (\cdot, \cdot)_{V \otimes W} \) on the \( k \)-module \( V \otimes W \) by letting

\[ (v \otimes w, v' \otimes w')_{V \otimes W} = (v, v')_V (w, w')_W \quad \text{for all } v, v' \in V \text{ and } w, w' \in W. \]

This new bilinear form is graded if the original two forms \( (\cdot, \cdot)_V \) and \( (\cdot, \cdot)_W \) were graded (presuming that \( V \) and \( W \) are graded).

(c) Say that a bialgebra \( A \) is **self-dual** with respect to a given symmetric bilinear form \( (\cdot, \cdot) : A \times A \to k \) if one has \( (a, m(b \otimes c))_A = (\Delta(a), b \otimes c)_{A \otimes A} \) and \( (1_A, a) = \epsilon(a) \) for \( a, b, c \) in \( A \). If \( A \) is a graded Hopf algebra of finite type, and this form \( (\cdot, \cdot) \) is graded, then this is equivalent to the \( k \)-module map \( A \to A^o \) induced by \( (\cdot, \cdot)_A \) giving a Hopf algebra homomorphism.

Thus, any PSH \( A \) is self-dual with respect to the bilinear form \( (\cdot, \cdot)_A \) that makes its PSH-basis orthonormal.

**Proposition 3.3.** Let \( A \) be a Hopf algebra over \( k = Z \) or \( k = Q \) which is graded, connected, and self-dual with respect to a positive definite graded\(^{36} \) bilinear form. Then:

\(^{35}\)not necessarily indexed by partitions

\(^{36}\)That is, \( (A_i, A_j) = 0 \) for \( i \neq j \).
Proof. (a) Note that \( Q_3.4 \). Remark \( \square \) case we will ever encounter in practice). See Exercise 3.7 for the general proof.

perpendicular to \( y \) \( \text{where the second equality uses self-duality, while the third equality uses the fact that} \ A \) can be used.

\[ \otimes \] \( \text{so that by induction they will commute with} \ 1 \]\( \text{implies that} \ \Delta + \text{of} \ \Lambda \text{which is commutative.} \]

\[ \text{for which one shows that} \ \Delta \text{is the index} \ 2 \text{lattice of} \ \Lambda \text{which is} \ \mathbb{Z} \text{-spanned by} \ \{p_2,e^1\}, \text{containing} \ 2e_2, \text{but not containing} \ e_2 \text{itself.} \]

Already the fact that \( p \cap I^2 = 0 \) has a strong implication.

Lemma 3.5. A graded connected Hopf algebra \( A \) over any ring \( k \) having \( p \cap I^2 = 0 \) has algebra structure which is commutative.

Proof. The component \( A_0 = k \) commutes with all of \( A \). This forms the base case for an induction on \( i + j \) in which one shows that any elements \( x \in A_i \) and \( y \in A_j \) with \( i,j > 0 \) will have \( [x,y] := xy - yx = 0 \). Since \( [x,y] \) lies in \( I^2 \), it suffices to show that \( [x,y] \) also lies in \( p \):

\[ \Delta[x,y] = [\Delta(x),\Delta(y)] \]

\[ = [1 \otimes x + x \otimes 1, 1 \otimes y + y \otimes 1 + \Delta_+(y)] \]

\[ = [1 \otimes x + x \otimes 1, 1 \otimes y + y \otimes 1] \]

\[ + [1 \otimes x + x \otimes 1, \Delta_+(y)] + [\Delta_+(x), 1 \otimes y + y \otimes 1] + [\Delta_+(x), \Delta_+(y)] \]

\[ = [1 \otimes x + x \otimes 1, 1 \otimes y + y \otimes 1] \]

\[ = 1 \otimes [x,y] + [x,y] \otimes 1 \]

showing that \( [x,y] \) lies in \( p \). Here the second-to-last equality used the inductive hypotheses: homogeneity implies that \( \Delta_+(x) \) is a sum of homogeneous tensors of the form \( z_1 \otimes z_2 \) satisfying \( \deg(z_1),\deg(z_2) < i \), so that by induction they will commute with \( 1 \otimes y, y \otimes 1 \), thus proving that \( [\Delta_+(x), 1 \otimes y + y \otimes 1] = 0 \); a symmetric argument shows \( [1 \otimes x + x \otimes 1, \Delta_+(y)] = 0 \), and, a similar argument shows \( [\Delta_+(x), \Delta_+(y)] = 0 \). The last equality is an easy calculation, and was done already in (1.10).

Remark 3.6. Zelevinsky actually shows [173, Proof of A.1.3, p. 150] that the assumption of \( p \cap I^2 = 0 \) (along with hypotheses of unit, connit, graded, connected, and \( \Delta \) being a morphism for multiplication) already implies the associativity of the multiplication in \( A \) ! One shows by induction on \( i + j + k \) that any \( x,y,z \) in \( A_i, A_j, A_k \) with \( i,j,k > 0 \) have vanishing associator \( \text{assoc}(x,y,z) := x(yz) - (xy)z \). In the inductive step, one first notes that \( \text{assoc}(x,y,z) \) lies in \( I^2 \), and then checks that \( \text{assoc}(x,y,z) \) also lies in \( p \), by a calculation very similar to the one above, repeatedly using the fact that \( \text{assoc}(x,y,z) \) is multilinear in its three arguments.

Exercise 3.7. Prove Proposition 3.3(c) in the general case.

This leads to a general structure theorem.

Theorem 3.8. If a graded, connected Hopf algebra \( A \) over a field \( k \) of characteristic zero has \( I = p \oplus I^2 \), then the inclusion \( p \hookrightarrow A \) extends to a Hopf algebra isomorphism from the symmetric algebra \( \text{Sym}_k(p) \to A \). In particular, \( A \) is both commutative and cocommutative.

\[ ^{137} \text{Specifically, either the existence of an orthogonal projection on a subspace of a finite-dimensional inner-product space over} \ \mathbb{Q}, \text{or the fact that} \ \dim(W^+) = \dim V - \dim W \text{for a subspace} \ W \text{of a finite-dimensional inner-product space} \ V \text{over} \ \mathbb{Q} \text{can be used.} \]
Note that these hypotheses are valid, using Proposition 3.3(c), whenever $A$ is obtained from a PSH (over $\mathbb{Z}$) by tensoring with $\mathbb{Q}$.

Proof of Theorem 3.8. Since Lemma 3.5 implies that $A$ is commutative, the universal property of $\text{Sym}_k(p)$ as a free commutative algebra on generators $p$ shows that the inclusion $p \hookrightarrow A$ at least extends to an algebra morphism $\text{Sym}_k(p) \rightarrow A$. Since the Hopf structure on $\text{Sym}_k(p)$ makes the elements of $p$ primitive (see Example 1.21), this $\varphi$ is actually a coalgebra morphism (since $\Delta \circ \varphi = (\varphi \otimes \varphi) \circ \Delta$ and $\epsilon \circ \varphi = \epsilon$ need only to be checked on algebra generators), hence a bialgebra morphism, hence a Hopf algebra morphism (by Proposition 1.46(c)). It remains to show that $\varphi$ is surjective, and injective.

For the surjectivity of $\varphi$, note that the hypothesis $I = p \otimes I^2$ implies that the composite $p \hookrightarrow I \rightarrow I/I^2$ gives a $k$-vector space isomorphism. What follows is a standard argument to deduce that $p$ generates $A$ as a commutative graded $k$-algebra. One shows by induction on $n$ that any homogeneous element $a$ in $A_n$ lies in the $k$-subalgebra generated by $p$. The base case $n = 0$ is trivial as $a$ lies in $k$. In the inductive step where $a$ lies in $I$, write $a \equiv p \mod I^2$ for some $p$ in $p$. Thus $a = p + \sum b_i c_i$, where $b_i, c_i$ lie in $I$ but have strictly smaller degree, so that by induction they lie in the subalgebra generated by $p$, and hence so does $a$.

Note that the surjectivity argument did not use the assumption that $k$ has characteristic zero, but we will now use it in the injectivity argument for $\varphi$, to establish the following

\[ \text{(3.1) Claim: Every primitive element of } \text{Sym}(p) \text{ lies in } p = \text{Sym}^1(p). \]

Note that this claim fails in positive characteristic, e.g. if $k$ has characteristic 2 then $x^2$ lies in $\text{Sym}^2(p)$, however

\[ \Delta(x^2) = 1 \otimes x^2 + 2x \otimes x + x^2 \otimes 1 = 1 \otimes x^2 + x^2 \otimes 1. \]

To see the claim, assume not, so that by gradedness, there must exist some primitive element $y \neq 0$ lying in some $\text{Sym}^n(p)$ with $n \geq 2$. This would mean that the composite map $f$ that follows the coproduct with a component projection

\[ \text{Sym}^n(p) \xrightarrow{\Delta} \bigoplus_{i+j=n} \text{Sym}^i(p) \otimes \text{Sym}^j(p) \rightarrow \text{Sym}^1(p) \otimes \text{Sym}^{n-1}(p) \]

has $f(y) = 0$. However, one can check on a basis that the multiplication backward $\text{Sym}^1(p) \otimes \text{Sym}^{n-1}(p) \rightarrow \text{Sym}^n(p)$ has the property that $m \circ f = n \cdot \text{id}_{\text{Sym}^n(p)}$;

\[ (m \circ f)(x_1 \cdots x_n) = m \left( \sum_{j=1}^n x_j \otimes x_1 \cdots \widehat{x_j} \cdots x_n \right) = n \cdot x_1 \cdots x_n \]

for $x_1, \ldots, x_n$ in $p$. Then $n \cdot y = m(f(y)) = m(0) = 0$ leads to the contradiction that $y = 0$, since $k$ has characteristic zero.

Now one can argue the injectivity of the (graded) map\(^{138}\) $\varphi$ by assuming that one has a nonzero homogeneous element $u$ in $\ker(\varphi)$ of minimum degree. In particular, $\deg(u) \geq 1$. Also since $p \hookrightarrow A$, one has that $u$ is not in $\text{Sym}^1(p) = p$, and hence $u$ is not primitive by the previous Claim. Consequently $\Delta_+(u) \neq 0$, and one can find a nonzero component $u^{(i,j)}$ of $\Delta_+(u)$ lying in $\text{Sym}(p)_i \otimes \text{Sym}(p)_j$ for some $i, j > 0$. Since this forces $i, j < \deg(u)$, one has that $\varphi$ maps both $\text{Sym}(p)_i, \text{Sym}(p)_j$ injectively into $A_i, A_j$. Hence the tensor product map

\[ \text{Sym}(p)_i \otimes \text{Sym}(p)_j \xrightarrow{\varphi \otimes \varphi} A_i \otimes A_j \]

\(^{138}\)The grading on $\text{Sym}(p)$ is induced from the grading on $p$, a homogeneous subspace of $I \subset A$ as it is the kernel of the graded map $I \rightarrow A \otimes A$. 


\[ 0 = \Delta^A_+(0) = \Delta^A_+(\varphi(u)) = (\varphi \otimes \varphi)(\Delta^\text{Sym}(p)(u)) \]

contains the nonzero \( A_i \otimes A_j \)-component \((\varphi \otimes \varphi)(u^{(i,j)})\).

An alternative proof of the injectivity of \( \varphi \) proceeds as follows: By (3.1), the subspace of primitive elements of \( \text{Sym}(p) \) is \( p \), and clearly \( \varphi \mid _p \) is injective. Hence, Exercise 1.54(b) (applied to the homomorphism \( \varphi \)) shows that \( \varphi \) is injective.

Before closing this section, we mention one nonobvious corollary of the Claim (3.1), when applied to the ring of symmetric functions \( \Lambda_Q \) with \( \mathbb{Q} \)-coefficients, since Proposition 2.28 says that \( \Lambda_Q = \mathbb{Q}[p_1, p_2, \ldots] = \text{Sym}(V) \) where \( V = \mathbb{Q}[p_1, p_2, \ldots] \).

**Corollary 3.9.** The subspace \( p \) of primitives in \( \Lambda_Q \) is one-dimensional in each degree \( n = 1, 2, \ldots, \) and spanned by \( \{p_1, p_2, \ldots\} \).

We note in passing that this corollary can also be obtained in a simpler fashion and a greater generality:

**Exercise 3.10.** Let \( k \) be any commutative ring. Show that the primitive elements of \( \Lambda \) are precisely the elements of the \( k \)-linear span of \( p_1, p_2, p_3, \ldots \).

3.2. **The decomposition theorem.** Our goal here is Zelevinsky’s theorem [173, Theorem 2.2] giving a canonical decomposition of any PSH as a tensor product into PSH’s that each have only one primitive element in their PSH-basis. For the sake of stating it, we introduce some notation.

**Definition 3.11.** Given \( A \) a PSH with PSH-basis \( \Sigma \), let \( C := \Sigma \cap p \) be the primitive elements in \( \Sigma \). For each \( \rho \) in \( C \), let \( A(\rho) \subset A \) be the \( Z \)-span of

\[ \Sigma(\rho) := \{ \sigma \in \Sigma : \text{ there exists } n \geq 0 \text{ with } (\sigma, \rho^n) \neq 0 \} \]

**Theorem 3.12.** Any PSH \( A \) has a canonical tensor product decomposition\(^{140}\)

\[ A = \bigotimes_{\rho \in C} A(\rho) \]

with \( A(\rho) \) a PSH, and \( \rho \) the only primitive element in its PSH-basis \( \Sigma(\rho) \).

Although in all the applications, \( C \) will be finite, when \( C \) is infinite one should interpret the tensor product in the theorem as the inductive limit of tensor products over finite subsets of \( C \), that is, linear combinations of basic tensors \( \otimes_{\rho} a_{\rho} \) in which there are only finitely many factors \( a_{\rho} \neq 1 \).

The first step toward the theorem uses a certain unique factorization property.

**Lemma 3.13.** Given \( P \) a set of pairwise orthogonal primitives in a PSH \( A \),

\[ (\rho_1 \cdots \rho_r, \pi_1 \cdots \pi_s) = 0 \]

for \( \rho_i, \pi_j \) in \( P \) unless \( r = s \) and one can reindex so that \( \rho_i = \pi_i \).

**Proof.** Induct on \( r \). For \( r > 0 \), one has

\[ (\rho_1 \cdots \rho_r, \pi_1 \cdots \pi_s) = (\rho_2 \cdots \rho_r, \rho_1^+ (\pi_1 \cdots \pi_s)) \]

\[ = (\rho_2 \cdots \rho_r, \sum_{j=1}^s (\pi_1 \cdots \pi_{j-1} \cdot \rho_1^+ (\pi_j) \cdot \pi_{j+1} \cdots \pi_s)) \]

\(^{139}\)One needs to know that for two injective maps \( V_1 \overset{\varphi_1}{\rightarrow} W_1 \) of \( k \)-vector spaces \( V_i, W_i \) with \( i = 1, 2 \), the tensor product \( \varphi_1 \otimes \varphi_2 \) is also injective. Factoring it as \( \varphi_1 \otimes \varphi_2 = (\text{id} \otimes \varphi_2) \circ (\varphi_1 \otimes \text{id}) \), one sees that it suffices to show that for an injective map \( V \overset{\varphi_1}{\rightarrow} W \) of free \( k \)-modules, and any free \( k \)-module \( U \), the map \( V \otimes U \overset{\varphi_1 \otimes \text{id}}{\rightarrow} W \otimes U \) is also injective. Since tensor products commute with direct sums, and \( U \) is (isomorphic to) a direct sum of copies of \( k \), this reduces to the easy-to-check case where \( U = k \).

Note that some kind of freeness or flatness hypothesis on \( U \) is needed here since, e.g. the injective \( \mathbb{Z} \)-module maps \( \mathbb{Z} \overset{\varphi_1 = (\times 2)}{\rightarrow} \mathbb{Z} \) and \( \mathbb{Z}/2\mathbb{Z} \overset{\varphi_1 \otimes \varphi_2}{\rightarrow} \mathbb{Z}/2\mathbb{Z} \) have \( \varphi_1 \otimes \varphi_2 = 0 \) on \( \mathbb{Z} \otimes \mathbb{Z} /2\mathbb{Z} \) but \( \mathbb{Z}/2\mathbb{Z} \).

\(^{140}\)The tensor product of two PSHs \( A_1 \) and \( A_2 \) with PSH-bases \( \Sigma_1 \) and \( \Sigma_2 \) is defined as the graded Hopf algebra \( A_1 \otimes A_2 \) with PSH-basis \( \{\sigma_1 \otimes \sigma_2\}_{(\sigma_1, \sigma_2) \in \Sigma_1 \times \Sigma_2} \). The tensor product of any family of PSHs is defined similarly (for the empty family, it is the graded connected Hopf algebra \( \mathbb{Z} \) with PSH-basis \( \{1\} \)).
from Proposition 2.73(iii) because $\rho_1$ is primitive\textsuperscript{141}. On the other hand, since each $\pi_j$ is primitive, one has $\rho_1^\perp(\pi_j) = (\rho_1, 1) \cdot \pi_j + (\rho_1, \pi_j) \cdot 1 = (\rho_1, \pi_j)$ which vanishes unless $\rho_1 = \pi_j$. Hence $(\rho_1 \cdots \rho_r, \pi_1 \cdots \pi_s) = 0$ unless $\rho_1 \in \{\pi_1, \ldots, \pi_s\}$, in which case after reindexing so that $\pi_1 = \rho_1$, it equals

$$n \cdot (\rho_1, \rho_1) \cdot (\rho_2 \cdots \rho_r, \pi_2 \cdots \pi_s)$$

if there are exactly $n$ occurrences of $\rho_1$ among $\pi_1, \ldots, \pi_s$. Now apply induction. \hfill $\square$

So far the positivity hypothesis for a PSH has played little role. Now we use it to introduce a certain partial order on the PSH $A$, and then a semigroup grading.

**Definition 3.14.** For a subset $S$ of an abelian group, let $\mathbb{Z}S$ (resp. $\mathbb{N}S$) denote the subgroup of $\mathbb{Z}$-linear combinations (resp. submonoid of $\mathbb{N}$-linear combinations\textsuperscript{142}) of the elements of $S$.

In a PSH $A$ with PSH-basis $\Sigma$, the subset $\mathbb{N}\Sigma$ forms a submonoid, and lets one define a partial order on $A$ via $a \leq b$ if $b - a$ lies in $\mathbb{N}\Sigma$.

We note a few trivial properties of this partial order:

- The positivity hypothesis implies that $\mathbb{N}\Sigma : \mathbb{N}\Sigma \subset \mathbb{N}\Sigma$.
- Hence multiplication by an element $c \geq 0$ (meaning $c$ lies in $\mathbb{N}\Sigma$) preserves the order: $a \leq b$ implies $ac \leq bc$ since $(b - a)c$ lies in $\mathbb{N}\Sigma$.
- Thus $0 \leq c \leq d$ and $0 \leq a \leq b$ implies $ac \leq bc \leq bd$.

This allows one to introduce a semigroup grading on $A$.

**Definition 3.15.** Let $\mathbb{N}\Sigma_{\text{fin}}^C$ denote the additive submonoid of $\mathbb{N}\Sigma$ consisting of those $\alpha = (\alpha_\rho)_{\rho \in C}$ with finite support.

Note that for any $\alpha$ in $\mathbb{N}\Sigma_{\text{fin}}^C$, one has that the product $\prod_{\rho \in C} \rho^{\alpha_\rho} \geq 0$. Define

$$\Sigma(\alpha) := \{\sigma \in \Sigma : \sigma \leq \prod_{\rho \in C} \rho^{\alpha_\rho}\},$$

that is, the subset of $\Sigma$ on which $\prod_{\rho \in C} \rho^{\alpha_\rho}$ has support. Also define

$$A(\alpha) := \mathbb{Z}\Sigma(\alpha) \subset A.$$

**Proposition 3.16.** The PSH $A$ has an $\mathbb{N}\Sigma_{\text{fin}}^C$-semigroup-grading: one has an orthogonal direct sum decomposition

$$A = \bigoplus_{\alpha \in \mathbb{N}\Sigma_{\text{fin}}^C} A(\alpha)$$

for which

\begin{align*}
A(\alpha)A(\beta) & \subset A(\alpha + \beta) \\
\Delta A(\alpha) & \subset \bigoplus_{\gamma \in \mathbb{N}\Sigma} A(\beta) \otimes A(\gamma),
\end{align*}

**Proof.** We will make free use of the fact that a PSH $A$ is commutative, since it embeds in $A \otimes_{\mathbb{Z}} \mathbb{Q}$, which is commutative by Theorem 3.8.

Note that the orthogonality $(A(\alpha), A(\beta)) = 0$ for $\alpha \neq \beta$ is equivalent to the assertion that

$$\left(\prod_{\rho \in C} \rho^{\alpha_\rho}, \prod_{\rho \in C} \rho^{\beta_\rho}\right) = 0,$$

which follows from Lemma 3.13.

\textsuperscript{141}Strictly speaking, this argument needs further justification since $A$ might not be of finite type (and if it is not, Proposition 2.73(iii) cannot be applied). It is more adequate to refer to the proof of Proposition 2.73(iii), which indeed goes through with $\rho_1$ taking the role of $f$.

\textsuperscript{142}Recall that $\mathbb{N} := \{0, 1, 2, \ldots\}$. 
Next let us deal with the assertion (3.2). It suffices to check that when \( \tau, \omega \) in \( \Sigma \) lie in \( A_{(\alpha)}, A_{(\beta)} \), respectively, then \( \tau \omega \) lies in \( A_{(\alpha+\beta)} \). But note that any \( \sigma \) in \( \Sigma \) having \( \sigma \leq \tau \omega \) will then have
\[
\sigma \leq \tau \omega \leq \prod_{\rho \in C} \rho^{\alpha_{\rho}} \cdot \prod_{\rho \in C} \rho^{\beta_{\rho}} = \prod_{\rho \in C} \rho^{\alpha_{\rho}+\beta_{\rho}},
\]
so that \( \sigma \) lies in \( A_{(\alpha+\beta)} \). This means that \( \tau \omega \) lies in \( A_{(\alpha+\beta)} \).

This lets us check that \( \bigoplus_{\alpha \in \mathbb{N}_{\infty}^{C}} A_{(\alpha)} \) exhaust \( A \). It suffices to check that any \( \sigma \) in \( \Sigma \) lies in some \( A_{(\alpha)} \). Proceed by induction on \( \deg(\sigma) \), with the case \( \sigma = 1 \) being trivial; the element 1 always lies in \( \Sigma \), and hence lies in \( A_{(\alpha)} \) for \( \alpha = 0 \). For \( \sigma \) lying in \( I \), one either has \( (\sigma, a) \neq 0 \) for some \( a \) in \( I^2 \), or else \( \sigma \) lies in \( (I^2)^\perp = p \) (by Proposition 3.3(a)), so that \( \sigma \) is in \( C \) and we are done. If \( (\sigma, a) \neq 0 \) with \( a \) in \( I^2 \), then \( \sigma \) appears in the support of some \( \mathbb{Z} \)-linear combination of elements \( \tau, \omega \) where \( \tau, \omega \) lie in \( \Sigma \) and have strictly smaller degree than \( \sigma \) has. There exists at least one such pair \( \tau, \omega \) for which \( (\sigma, \tau \omega) \neq 0 \), and therefore \( \sigma \leq \tau \omega \). Then by induction \( \tau, \omega \) lie in some \( A_{(\alpha)}, A_{(\beta)} \), respectively, so \( \tau \omega \) lies in \( A_{(\alpha+\beta)} \), and hence \( \sigma \) lies in \( A_{(\alpha+\beta)} \) also.

Self-duality shows that (3.2) implies (3.3): if \( a, b, c \) lie in \( A_{(\alpha)}, A_{(\beta)}, A_{(\gamma)} \), respectively, then \( (\Delta a, b \otimes c)_{A \otimes A} = (a, bc)_A = 0 \) unless \( \alpha = \beta + \gamma \).

**Proposition 3.17.** For \( \alpha, \beta \) in \( \mathbb{N}_{\infty}^{C} \) with disjoint support, one has a bijection
\[
\Sigma(\alpha) \times \Sigma(\beta) \longrightarrow \Sigma(\alpha + \beta)
\]
\[
(\sigma, \tau) \longmapsto \sigma \tau.
\]

Thus, the multiplication map \( A_{(\alpha)} \otimes A_{(\beta)} \rightarrow A_{(\alpha+\beta)} \) is an isomorphism.

**Proof.** We first check that for \( \sigma_1, \sigma_2 \) in \( \Sigma(\alpha) \) and \( \tau_1, \tau_2 \) in \( \Sigma(\beta) \), one has
\[
(\sigma_1 \tau_1, \sigma_2 \tau_2) = \delta(\sigma_1, \tau_1)(\sigma_2, \tau_2).
\]
Note that this is equivalent to showing both
- that \( \sigma \tau \) lie in \( \Sigma(\alpha + \beta) \) so that the map is well-defined, since it shows \( (\sigma \tau, \sigma \tau) = 1 \), and
- that the map is injective.

One calculates
\[
(\sigma_1 \tau_1, \sigma_2 \tau_2)_A = (\sigma_1 \tau_1, m(\sigma_2 \otimes \tau_2))_A
\]
\[
= (\Delta(\sigma_1 \tau_1), \sigma_2 \otimes \tau_2)_{A \otimes A}
\]
\[
= (\Delta(\sigma_1) \Delta(\tau_1), \sigma_2 \otimes \tau_2)_{A \otimes A}.
\]

Note that due to (3.3), \( \Delta(\sigma_1) \Delta(\tau_1) \) lies in \( \sum A_{(\alpha' + \beta')} \otimes A_{(\alpha'' + \beta'')} \) where
\[
\alpha' + \alpha'' = \alpha
\]
\[
\beta' + \beta'' = \beta.
\]

Since \( \sigma_2 \otimes \tau_2 \) lies in \( A_{(\alpha)} \otimes A_{(\beta)} \), the only nonvanishing terms in the inner product come from those with
\[
\alpha' + \beta' = \alpha
\]
\[
\alpha'' + \beta'' = \beta.
\]
As \( \alpha, \beta \) have disjoint support, this can only happen if
\[
\alpha' = \alpha, \; \alpha'' = 0, \; \beta' = 0, \; \beta'' = \beta,
\]
that is, the only nonvanishing term comes from \( (\sigma_1 \otimes 1)(1 \otimes \tau_1) = \sigma_1 \otimes \tau_1 \). Hence
\[
(\sigma_1 \tau_1, \sigma_2 \tau_2)_A = (\sigma_1 \otimes \tau_1, \sigma_2 \otimes \tau_2)_{A \otimes A} = \delta(\sigma_1, \tau_1)(\sigma_2, \tau_2).
\]

To see that the map is surjective, express
\[
\prod_{\rho \in C} \rho^{\alpha_{\rho}} = \sum_{i} \sigma_i
\]
\[
\prod_{\rho \in C} \rho^{\beta_{\rho}} = \sum_{j} \tau_j
\]
with \( \sigma_i \in \Sigma(\alpha) \) and \( \tau_j \in \Sigma(\beta) \). Then each product \( \sigma_i \tau_j \) is in \( \Sigma(\alpha + \beta) \) by (3.4), and
\[
\prod_{\rho \in \mathcal{C}} \rho^\alpha \rho^\beta = \sum_{i,j} \sigma_i \tau_j
\]
shows that \( \{\sigma_i \tau_j\} \) exhausts \( \Sigma(\alpha + \beta) \). This gives surjectivity.

**Proof of Theorem 3.12.** Recall from Definition 3.11 that for each \( \rho \in \mathcal{C} \), one defines \( A(\rho) \subset A \) to be the \( \mathbb{Z} \)-span of
\[
\Sigma(\rho) := \{ \sigma \in \Sigma : \text{ there exists } n \geq 0 \text{ with } (\sigma, \rho^n) \neq 0 \}.
\]
In other words, \( A(\rho) := \bigoplus_{n \geq 0} A(n, e_\rho) \) where \( e_\rho \) in \( \mathbb{N}_0^{\mathbb{N}} \) is the standard basis element indexed by \( \rho \). Proposition 3.16 then shows that \( A(\rho) \) is a Hopf subalgebra of \( A \). Since every \( \alpha \) in \( \mathbb{N}_0^{\mathbb{N}} \) can be expressed as the (finite) sum \( \sum \alpha_\rho e_\rho \), and the \( e_\rho \) have disjoint support, iterating Proposition 3.17 shows that \( A = \bigotimes_{\rho \in \mathcal{C}} A(\rho) \). Lastly, \( \Sigma(\rho) \) is clearly a PSH-basis for \( A(\rho) \), and if \( \sigma \) is any primitive element in \( \Sigma(\rho) \) then \( (\sigma, \rho^n) \neq 0 \) lets one conclude via Lemma 3.13 that \( \sigma = \rho \) (and \( n = 1 \)). \( \square \)

3.3. \( A \) is the unique indecomposable PSH. The goal here is to prove the rest of Zelevinsky’s structure theory for PSH’s. Namely, if \( A \) has only one primitive \( \rho \) in its PSH-basis \( \Sigma \), then \( A \) must be isomorphic as a PSH to the ring of symmetric functions \( \Lambda \), after one rescales the grading of \( A \). Note that every \( \sigma \) in \( \Sigma \) has \( \sigma \leq \rho^n \) for some \( n \), and hence has degree divisible by the degree of \( \rho \). Thus one can divide all degrees by that of \( \rho \) and assume \( \rho \) has degree 1.

The idea is to find within \( A \) and \( \Sigma \) a set of elements that play the role of
\[
\{h_n = s(n)\}_{n=0,1,2,...}, \{e_n = s(1^n)\}_{n=0,1,2,...}
\]
within \( A = \Lambda \) and its PSH-basis of Schur functions \( \Sigma = \{s_\lambda\} \). Zelevinsky’s argument does this by isolating some properties that turn out to characterize these elements:

(a) \( h_0 = e_0 = 1 \), and \( h_1 = e_1 =: \rho \) has \( \rho^2 \) a sum of two elements of \( \Sigma \), namely
\[
\rho^2 = h_2 + e_2.
\]

(b) For all \( n = 0,1,2,... \), there exist unique elements \( h_n, e_n \) in \( A_n \cap \Sigma \) that satisfy
\[
\begin{align*}
\delta^h h_n &= 0, \\
\delta^e h_n &= 0
\end{align*}
\]
with \( h_2, e_2 \) being the two elements of \( \Sigma \) introduced in (a).

(c) For \( k = 0,1,2,..., n \) one has
\[
\begin{align*}
\delta^h h_n &= h_{n-k} \quad \text{and} \quad \delta^\tau h_n = 0 \quad \text{for } \sigma \in \Sigma \setminus \{h_0, h_1, \ldots, h_n\} \\
\delta^\tau e_n &= e_{n-k} \quad \text{and} \quad \delta^\tau e_n = 0 \quad \text{for } \sigma \in \Sigma \setminus \{e_0, e_1, \ldots, e_n\}.
\end{align*}
\]
In particular, \( \delta^h e_n = 0 = \delta^e h_n \) for \( k \geq 2 \).

(d) Their coproducts are
\[
\begin{align*}
\Delta(h_n) &= \sum_{i+j=n} h_i \otimes h_j, \\
\Delta(e_n) &= \sum_{i+j=n} e_i \otimes e_j.
\end{align*}
\]

We will prove Zelevinsky’s result [173, Theorem 3.1] as a combination of the following two theorems.

**Theorem 3.18.** Let \( A \) be a PSH with PSH-basis \( \Sigma \) containing only one primitive \( \rho \), and assume that the grading has been rescaled so that \( \rho \) has degree 1. Then, after renaming \( \rho = e_1 = h_1 \), one can find unique sequences \( \{h_n\}_{n=0,1,2,...}, \{e_n\}_{n=0,1,2,...} \) of elements of \( \Sigma \) having properties (a),(b),(c),(d) listed above.

The second theorem uses the following notion.
Definition 3.19. A PSH-morphism $A \xrightarrow{\varphi} A'$ between two PSH’s $A, A'$ having PSH-bases $\Sigma, \Sigma'$ is a graded Hopf algebra morphism for which $\varphi(\mathbb{N}\Sigma) \subseteq \mathbb{N}\Sigma'$. If $A = A'$ and $\Sigma = \Sigma'$ it will be called a PSH-endomorphism. If $\varphi$ is an isomorphism and restricts to a bijection $\Sigma \to \Sigma'$, it will be called a PSH-isomorphism; if it is both a PSH-isomorphism and an endomorphism, it is a PSH-automorphism.\[^{143}\]

**Theorem 3.20.** The elements $\{h_n\}_{n=0,1,2,\ldots}, \{e_n\}_{n=0,1,2,\ldots}$ in Theorem 3.18 also satisfy the following.

(e) The elements $h_n, e_n$ in $A$ satisfy the same relation (2.17)

$$\sum_{i+j=n} (-1)^i e_i h_j = \delta_{0,n}$$

as their counterparts in $\Lambda$, along with the property that

$A = \mathbb{Z}[h_1, h_2, \ldots] = \mathbb{Z}[e_1, e_2, \ldots]$.

(f) There is exactly one nontrivial automorphism $A \xrightarrow{\varphi} A$ as a PSH, swapping $h_n \leftrightarrow e_n$.

(g) There are exactly two PSH-isomorphisms $A \to \Lambda$,

- one sending $h_n$ to the complete homogeneous symmetric functions $h_n(x)$, while sending $e_n$ to the elementary symmetric functions $e_n(x)$,
- the second one (obtained by composing the first with $\omega$) sending $h_n \mapsto e_n(x)$ and $e_n \mapsto h_n(x)$.

Before embarking on the proof, we mention one more bit of convenient terminology: say that an element $\sigma$ in $\Sigma$ is a constituent of $a$ in $\mathbb{N}\Sigma$ when $\sigma \leq a$, that is, $\sigma$ appears with nonzero coefficient $c_\sigma$ in the unique expansion $a = \sum_{\tau \in \Sigma} c_\tau \tau$.

**Proof of Theorem 3.18.** One fact that occurs frequently is this:

$$\sum_{\tau \in \Sigma} c_\tau \tau = 0.$$\(^{(3.5)}\)

Every $\sigma$ in $\Sigma \cap A_n$ is a constituent of $\rho^n$.

This follows from Theorem 3.12, since $\rho$ is the only primitive element of $\Sigma$: one has $A = A(\rho)$ and $\Sigma = \Sigma(\rho)$, so that $\sigma$ is a constituent of some $\rho^m$, and homogeneity considerations force $m = n$.

Notice that $A$ is of finite type (due to (3.5)). Thus, $A^o$ is a graded Hopf algebra isomorphic to $A$.

**Assertion (a).** Note that

$$(\rho^2, \rho^2) = (\rho^2, \rho^2) = (2\rho, \rho) = 2$$

using the fact that $\rho^2$ is a derivation since $\rho$ is primitive (Proposition 2.73(iii)). On the other hand, expressing $\rho^2 = \sum_{\sigma \in \Sigma} c_\sigma \sigma$ with $c_\sigma$ in $\mathbb{N}$, one has $(\rho^2, \rho^2) = \sum c_\sigma^2$. Hence exactly two of the $c_\sigma = 1$, so $\rho^2$ has exactly two distinct constituents. Denote them by $h_2$ and $e_2$. One concludes that $\Sigma \cap A_2 = \{h_2, e_2\}$ from (3.5).

Note also that the same argument shows $\Sigma \cap A_1 = \{\rho\}$, so that $A_1 = \mathbb{Z}\rho$. Since $\rho^2 h_2$ lies in $A_1 = \mathbb{Z}\rho$ and $(\rho^2 h_2, \rho) = (h_2, \rho^2) = 1$, we have $\rho^2 h_2 = \rho$. Similarly $\rho^2 e_2 = \rho$.

**Assertion (b).** We will show via induction on $n$ the following three assertions for $n \geq 1$:

- There exists an element $h_n$ in $\Sigma \cap A_n$ with $c_2^2 h_n = 0$.
- This element $h_n$ is unique.
- Furthermore $\rho^2 h_n = h_{n-1}$.

In the base cases $n = 1, 2$, it is not hard to check that our previously labelled elements, $h_1, h_2$ (namely $h_1 := \rho$, and $h_2$ as named in part (a)) really are the unique elements satisfying these hypotheses.

\[^{143}\]The reader should be warned that not every invertible PSH-endomorphism is necessarily a PSH-automorphism. For instance, it is an easy exercise to check that $\Lambda \otimes \Lambda \to \Lambda \otimes \Lambda$, $f \otimes g \mapsto \sum f_i \otimes g_i$ is a well-defined invertible PSH-endomorphism of the PSH $\Lambda \otimes \Lambda$ with PSH-basis $(s\lambda \otimes s\mu)(\lambda, \mu) \in \text{Par} \times \text{Par}$, but not a PSH-automorphism.
In the inductive step, it turns out that we will find $h_n$ as a constituent of $\rho h_{n-1}$. Thus we again use the derivation property of $\rho^\perp$ to compute that $\rho h_{n-1}$ has exactly two constituents:

\[
(rh_{n-1}, \rho h_{n-1}) = (r^\perp (rh_{n-1}), h_{n-1}) = (h_{n-1} + \rho \cdot r^\perp h_{n-1}, h_{n-1}) = (h_{n-1} + \rho h_{n-2}, h_{n-1}) = 1 + (h_{n-2}, \rho \cdot r^\perp h_{n-1}) = 1 + (h_{n-2}, h_{n-2}) = 1 + 1 = 2
\]

where the inductive hypothesis $\rho^\perp h_{n-1} = h_{n-2}$ was used twice. We next show that exactly one of the two constituents of $\rho h_{n-1}$ is annihilated by $e_2^\perp$. Note that since $e_2$ lies in $A_2$, and $A_1$ has $\mathbb{Z}$-basis element $\rho$, there is a constant $c$ in $\mathbb{Z}$ such that

\[
(3.7) \quad \Delta(e_2) = e_2 \otimes 1 + c \rho \otimes \rho + 1 \otimes e_2.
\]

On the other hand, (a) showed

\[
1 = (e_2, \rho^2)_A = (\Delta(e_2), \rho \otimes \rho)_{A \otimes A}
\]

so one must have $c = 1$. Therefore by Proposition 2.73(iii) again,

\[
(3.8) \quad e_2^\perp (\rho h_{n-1}) = e_2^\perp (\rho) h_{n-1} + \rho^\perp (\rho) \rho^\perp (h_{n-1}) + \rho e_2^\perp (h_{n-1}) = 0 + h_{n-2} + 0 = h_{n-2}.
\]

where the first term vanished due to degree considerations and the last term vanished by the inductive hypothesis. Bearing in mind that $\rho h_{n-1}$ lies in $\mathbb{N}_\Sigma$, and in a PSH with PSH-basis $\Sigma$, any skewing operator $\sigma^\perp$ for $\sigma$ in $\Sigma$ will preserve $\mathbb{N}_\Sigma$, one concludes from (3.8) that

- one of the two distinct constituents of the element $\rho h_{n-1}$ must be sent by $e_2^\perp$ to $h_{n-2}$, and
- the other constituent of $\rho h_{n-1}$ must be annihilated by $e_2^\perp$; call this second constituent $h_n$.

Lastly, to see that this $h_n$ is unique, it suffices to show that any element $\sigma$ of $\Sigma \cap A_n$ which is killed by $e_2^\perp$ must be a constituent of $\rho h_{n-1}$. This holds for the following reason. We know $\sigma \leq \rho^2$ by (3.5), and hence $0 \neq (\rho^2, \sigma) = (\rho^{n-1}, \rho^\perp \sigma)$, implying that $\rho^\perp \sigma \neq 0$. On the other hand, since $0 = \rho^\perp e_2^\perp \sigma = e_2^\perp \rho^\perp \sigma$, one has that $\rho^\perp \sigma$ is annihilated by $e_2^\perp$, and hence $\rho^\perp \sigma$ must be a (positive) multiple of $h_{n-1}$ by part of our inductive hypothesis. Therefore $(\sigma, \rho h_{n-1}) = (\rho^\perp \sigma, h_{n-1})$ is positive, that is, $\sigma$ is a constituent of $\rho h_{n-1}$.

The preceding argument, applied to $\sigma = h_n$, shows that $\rho^\perp h_n = \rho^\perp h_{n-1}$ for some $c$ in $\{1, 2, \ldots\}$. Since $(\rho^\perp h_n, h_{n-1}) = (h_n, h_{n-1}) = 1$, this $c$ must be 1, so that $\rho^\perp h_n = h_{n-1}$. This completes the induction step in the proof of (3.6).

One can then argue, swapping the roles of $\epsilon_n, h_n$ in the above argument, the existence and uniqueness of a sequence $\{\epsilon_n\}_{n=0}^\infty$ in $\Sigma$ satisfying the properties analogous to (3.6), with $\epsilon_0 := 1, \epsilon_1 := \rho$.

**Assertion (c).** Iterating the property from (b) that $\rho^\perp h_n = h_{n-1}$ shows that $(\rho^k)^\perp h_n = h_{n-k}$ for $0 \leq k \leq n$. However one also has an expansion

\[
\rho^k = ch_k + \sum_{\sigma \in \Sigma \cap A_k \setminus \{e\}} c_{\sigma} \sigma
\]

for some integers $c, c_\sigma > 0$, since every $\sigma$ in $\Sigma \cap A_k$ is a constituent of $\rho^k$. Hence

\[
1 = (h_{n-k}, h_{n-k}) = ((\rho^k)^\perp h_n, (\rho^k)^\perp h_n) \geq c^2 (h_k^\perp h_n, h_k^\perp h_n)
\]

using Proposition 2.73(ii). Hence if we knew that $h_k^\perp h_n \neq 0$ this would force

\[
h_k^\perp h_n = (\rho^k)^\perp h_n = h_{n-k}
\]

as well as $\sigma^\perp h_n = 0$ for all $\sigma \notin \{h_0, h_1, \ldots, h_n\}$. But

\[
(p^{n-k})^\perp h_k^\perp h_n = h_k^\perp (p^{n-k})^\perp h_n = h_k^\perp h_k = 1 \neq 0
\]

so $h_k^\perp h_n \neq 0$, as desired. The argument for $e_2^\perp \epsilon_n = \epsilon_{n-k}$ is symmetric.

The last assertion in (c) follows if one checks that $\epsilon_n \neq h_n$ for each $n \geq 2$, but this holds since $e_2^\perp (h_n) = 0$ but $e_2^\perp (\epsilon_n) = e_{n-2}$. 
Assertion (d). Part (c) implies that
\[(\Delta h_n, \sigma \otimes \tau)_{A \otimes A} = (h_n, \sigma \tau)_A = (\sigma^\perp h_n, \tau)_A = 0\]
unless \(\sigma = h_k\) for some \(k = 0, 1, 2, \ldots, n\) and \(\tau = h_{n-k}\). Also one can compute
\[(\Delta h_n, h_k \otimes h_{n-k}) = (h_n, h_k h_{n-k}) = (h_k^\perp h_n, h_{n-k}) \overset{(c)}{=} (h_{n-k}, h_{n-k}) = 1.\]
This is equivalent to the assertion for \(\Delta h_n\) in (d). The argument for \(\Delta e_n\) is symmetric. \(\square\)

Before proving Theorem 3.20, we note some consequences of Theorem 3.18. Define for each partition \(\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell)\) the elements of \(A\)
\[h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots,\]
\[e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots.\]
Also, define the lexicographic order on \(\text{Par}_n\) by saying \(\lambda <_{\text{lex}} \mu\) if \(\lambda \neq \mu\) and the smallest index \(i\) for which \(\lambda_i \neq \mu_i\) has \(\lambda_i < \mu_i\). Recall also that \(\lambda^t\) denotes the conjugate or transpose partition to \(\lambda\), obtained by swapping rows and columns in the Ferrers diagram.

The following unitriangularity lemma will play a role in the proof of Theorem 3.20(e).

**Lemma 3.21.** Under the hypotheses of Theorem 3.18, for \(\lambda, \mu\) in \(\text{Par}_n\), one has
\[e_\mu^\perp h_\lambda = \begin{cases} 1 & \text{if } \mu = \lambda^t \\ 0 & \text{if } \mu >_{\text{lex}} \lambda^t. \end{cases}\]
Consequently
\[\det [(e_\mu^t, h_\lambda)_{\lambda, \mu \in \text{Par}_n}] = \pm 1.\]

**Proof.** Notice that \(A\) is of finite type (as shown in the proof of Theorem 3.18). Thus, \(A^o\) is a graded Hopf algebra isomorphic to \(A\).

Induct on the length of \(\mu\). If \(\lambda\) has length \(\ell\), so that \(\lambda_\ell^t = \ell\), then
\[e_\mu^\perp h_\lambda = e_\mu^\perp (h_{\lambda_1} \cdots h_{\lambda_\ell}) \overset{(*)}{=} e_\mu^\perp (h_{\lambda_1} \cdots h_{\lambda_\ell}) = e_{(\mu_2, \mu_3, \ldots)}^\perp h_{(\lambda_1, \ldots, \lambda_\ell)}\]
where the second-to-last equality used Proposition 2.73(iii) along with (the iterates of) the coproduct formula for \(\Delta(e_n)\) in Theorem 3.18(d), and the last equality used
\[e_\mu^\perp (h_n) = \begin{cases} h_{n-1} & \text{if } k = 1, \\ 0 & \text{if } k \geq 2. \end{cases}\]
Now apply induction, since \((\lambda_1 - 1, \ldots, \lambda_\ell - 1)^t = (\lambda_2^t, \lambda_3^t, \ldots)\).
For the last assertion, note \((e_\mu^t, h_\lambda) = (e_\mu^\perp (h_\lambda), 1) = e_\mu^\perp (h_\lambda)\) for \(\lambda, \mu\) in \(\text{Par}_n\). \(\square\)

The following proposition will be the crux of the proof of Theorem 3.20(f) and (g), and turns out to be closely related to Kerov’s asymptotic theory of characters of the symmetric groups [80].

**Proposition 3.22.** Given a PSH \(A\) with PSH-basis \(\Sigma\) containing only one primitive \(\rho\), the two maps \(A \to \mathbb{Z}\) defined on \(A = \bigoplus_{n \geq 0} A_n\) via
\[\delta_h = \bigoplus_n h_n^\perp, \quad \delta_e = \bigoplus_n e_n^\perp\]
are characterized as the only two \(\mathbb{Z}\)-linear maps \(A \to \mathbb{Z}\) with the three properties of being
- **positive:** \(\delta(\mathbb{N} \Sigma) \subseteq \mathbb{N}\),
- **multiplicative:** \(\delta(a_1 a_2) = \delta(a_1) \delta(a_2)\), and
Proof. Notice that $A$ is of finite type (as shown in the proof of Theorem 3.18). Thus, $A^e$ is a graded Hopf algebra isomorphic to $A$.

It should be clear from their definitions that $\delta_h, \delta_e$ are $\mathbb{Z}$-linear, positive and normalized. To see that $\delta_h$ is multiplicative, by $\mathbb{Z}$-linearity, it suffices to check that for $a_1, a_2$ in $A_{n_1}, A_{n_2}$ with $n_1 + n_2 = n$, one has

$$\delta_h(a_1 a_2) = h_n(a_1 a_2) = \sum_{i_1 + i_2 = n} h_{i_1}(a_1) h_{i_2}(a_2) = h_{i_1}(a_1) h_{i_2}(a_2) = \delta_h(a_1) \delta_h(a_2)$$

in which the second equality used Proposition 2.73(iii) and Theorem 3.18(d). The argument for $\delta_e$ is symmetric.

Conversely, given $A \xrightarrow{\delta} \mathbb{Z}$ which is $\mathbb{Z}$-linear, positive, multiplicative, and normalized, note that

$$\delta(h_2) + \delta(e_2) = \delta(h_2 + e_2) = \delta(\rho^2) = \delta(\rho)^2 = 1^2 = 1$$

and hence positivity implies that either $\delta(h_2) = 0$ or $\delta(e_2) = 0$. Assume the latter holds, and we will show that $\delta = \delta_h$.

Given any $\sigma$ in $\Sigma \cap A_n \setminus \{h_n\}$, note that $e_2^2 \sigma \neq 0$ by Theorem 3.18(b), and hence $0 \neq (e_2^2 \sigma, \rho^n - 2) = (\sigma, e_2 \rho^{n-2})$. Thus $\sigma$ is a constituent of $e_2 \rho^{n-2}$, so positivity implies

$$0 \leq \delta(\sigma) \leq \delta(e_2 \rho^{n-2}) = \delta(e_2) \delta(\rho^{n-2}) = 0.$$ 

Thus $\delta(\sigma) = 0$ for $\sigma$ in $\Sigma \cap A_n \setminus \{h_n\}$. Since $\delta(\rho^n) = \delta(\rho)^n = 1^n = 1$, this forces $\delta(h_n) = 1$, for each $n \geq 0$ (including $n = 0$, as $1 = \delta(\rho) = \delta(\rho \cdot 1) = \delta(\rho) \delta(1) = 1 \cdot \delta(1) = \delta(1)$). Thus $\delta = \delta_h$. The argument when $\delta(h_2) = 0$ showing $\delta = \delta_e$ is symmetric.

Proof of Theorem 3.20. Many of the assertions of parts (e) and (f) will come from constructing the unique nontrivial PSH-automorphism $\omega$ of $A$ from the antipode $S$: for homogeneous $a$ in $A_n$, define $\omega(a) := (-1)^n S(a)$. We now study some of the properties of $S$ and $\omega$.

Notice that $A$ is of finite type (as shown in the proof of Theorem 3.18). Thus, $A^e$ is a graded Hopf algebra isomorphic to $A$.

Since $A$ is a PSH, it is commutative by Theorem 3.8 (applied to $A \otimes_{\mathbb{Z}} \mathbb{Q}$). This implies both that $S, \omega$ are actually algebra endomorphisms by Proposition 1.32, and that $S^2 = \text{id}_A = \omega^2$ by Corollary 1.34.

Since $A$ is self-dual and the defining diagram (1.12) satisfied by the antipode $S$ is sent to itself when one replaces $A$ by $A^e$ and all maps by their adjoints, one concludes that $S = S^*$ (where $S^*$ means the restricted adjoint $S^* : A^e \to A^e$), i.e., $S$ is self-adjoint. Since $S$ is an algebra endomorphism, and $S = S^*$, in fact $S$ is also a coalgebra endomorphism, a bialgebra endomorphism, and a Hopf endomorphism (by Proposition 1.46(c)). The same properties are shared by $\omega$.

Since $\text{id}_A = S^2 = SS^*$, one concludes that $S$ is an isometry, and hence so is $\omega$.

Since $\rho$ is primitive, one has $S(\rho) = -\rho$ and $\omega(\rho) = \rho$. Therefore $\omega(\rho^n) = \rho^n$ for $n = 1, 2, \ldots$. Use this as follows to check that $\omega$ is a PSH-automorphism, which amounts to checking that every $\sigma$ in $\Sigma$ has $\omega(\sigma)$ in $\Sigma$:

$$(\omega(\sigma), \omega(\sigma)) = (\sigma, \sigma) = 1$$

so that $\pm \omega(\sigma)$ lies in $\Sigma$, but also if $\sigma$ lies in $A_n$, then

$$(\omega(\sigma), \rho^n) = (\sigma, \omega(\rho^n)) = (\sigma, \rho^n) > 0.$$ 

In summary, $\omega$ is a PSH-automorphism of $A$, an isometry, and an involution.

Let us try to determine the action of $\omega$ on the $\{h_n\}$. By similar reasoning as in (3.7), one has

$$\Delta(h_2) = h_2 \otimes 1 + \rho \otimes \rho + 1 \otimes h_2.$$ 

Thus $0 = S(h_2) + S(\rho) \rho + h_2$, and combining this with $S(\rho) = -\rho$, one has $S(h_2) = e_2$. Thus also $\omega(h_2) = (-1)^2 S(h_2) = e_2$. 

\[\text{bullet} \quad \text{normalized: } \delta(\rho) = 1.\]
We claim that this forces $\omega(h_n) = e_n$, because $h_n^\perp \omega(h_n) = 0$ via the following calculation: for any $a$ in $A$ one has
\[
(h_n^\perp \omega(h_n), a) = (\omega(h_n), h_n a) = (h_n, \omega(h_n a)) = (h_n, e_2 \omega(a)) = (e_n, \omega(a)) = (0, \omega(a)) = 0.
\]

Consequently the involution $\omega$ swaps $h_n$ and $e_n$, while the antipode $S$ has $S(h_n) = (-1)^n e_n$ and $S(e_n) = (-1)^n h_n$. Thus the coproduct formulas in (d) and definition of the antipode $S$ imply the relation (2.17) between $\{h_n\}$ and $\{e_n\}$.

This relation (2.17) also lets one recursively express the $h_n$ as polynomials with integer coefficients in the $\{e_n\}$, and vice-versa, so that $\{h_n\}$ and $\{e_n\}$ each generate the same $\mathbb{Z}$-subalgebra $A'$ of $A$. We wish to show that $A'$ exhausts $A$.

We argue that Lemma 3.21 implies the Gram matrix $[(h_\mu, h_\lambda)]_{\mu, \lambda \in \text{Par}_n}$ has determinant $\pm 1$ as follows. Since $\{h_n\}$ and $\{e_n\}$ both generate $A'$, there exists a $\mathbb{Z}$-matrix $(a_{\mu, \lambda})$ expressing $e_\mu = \sum_\lambda a_{\mu, \lambda} h_\lambda$, and one has
\[
[(e_\mu, h_\lambda)] = [a_{\mu, \lambda}] \cdot [(h_\mu, h_\lambda)].
\]

Taking determinants of these three $\mathbb{Z}$-matrices, and using the fact that the determinant on the left is $\pm 1$, both determinants on the right must also be $\pm 1$.

Now we will show that every $\sigma \in \Sigma \cap A_n$ lies in $A'$. Uniquely express $\sigma = \sigma' + \sigma''$ in which $\sigma'$ lies in the $\mathbb{R}$-span $\mathbb{R}A_n'$ and $\sigma''$ lies in the real perpendicular space $(\mathbb{R}A_n')^\perp$ inside $\mathbb{R} \otimes \mathbb{Z} A_n$. One can compute $\mathbb{R}$-coefficients $(c_\mu)_{\mu \in \text{Par}_n}$ that express $\sigma' = \sum_\mu c_\mu h_\mu$ by solving the system
\[
\left( \sum_\mu c_\mu h_\mu, h_\lambda \right) = (\sigma, h_\lambda) \quad \text{for } \lambda \in \text{Par}_n.
\]
This linear system is governed by the Gram matrix $[(h_\mu, h_\lambda)]_{\mu, \lambda \in \text{Par}_n}$ with determinant $\pm 1$, and its right side has $\mathbb{Z}$-entries since $\sigma, h_\lambda$ lie in $A$. Hence the solution $(c_\mu)_{\mu \in \text{Par}_n}$ will have $\mathbb{Z}$-entries, so $\sigma'$ lies in $A'$. Furthermore, $\sigma'' = \sigma - \sigma'$ will lie in $A$, and hence by the orthogonality of $\sigma', \sigma''$, $1 = (\sigma, \sigma) = (\sigma', \sigma') + (\sigma'', \sigma'')$.

One concludes that either $\sigma'' = 0$, or $\sigma' = 0$. The latter cannot occur since it would mean that $\sigma = \sigma''$ is perpendicular to all of $A'$. But $\rho^n = h_n^n$ lies in $A'$, and $(\sigma, \rho^n) \neq 0$. Thus $\sigma'' = 0$, meaning $\sigma = \sigma'$ lies in $A'$.

This completes the proof of assertion (e). Note that in the process, having shown $\det(h_\mu, h_\lambda)_{\mu, \lambda \in \text{Par}_n} = \pm 1$, one also knows that $\{h_\lambda\}_{\lambda \in \text{Par}_n}$ are $\mathbb{Z}$-linearly independent, so that $\{h_1, h_2, \ldots\}$ are algebraically independent, and $A = \mathbb{Z}[h_1, h_2, \ldots]$ is the polynomial algebra generated by $\{h_1, h_2, \ldots\}$.

For assertion (f), we have seen that $\omega$ given such a PSH-automorphism $A \to A$, swapping $h_n \leftrightarrow e_n$. Conversely, given a PSH-automorphism $A \xrightarrow{\omega} A$, consider the positive, multiplicative, normalized $\mathbb{Z}$-linear map $\delta := \delta_\omega \circ \varphi : A \to \mathbb{Z}$. Proposition 3.22 shows that either
\[
\delta = \delta_n, \text{ which then forces } \varphi(h_n) = h_n \text{ for all } n, \text{ so } \varphi = \text{id}_A, \text{ or }
\]
\[
\delta = \delta_e, \text{ which then forces } \varphi(e_n) = h_n \text{ for all } n, \text{ so } \varphi = \omega.
\]

For assertion (g), given a PSH $A$ with PSH-basis $\Sigma$ having exactly one primitive $\rho$, since we have seen $A = \mathbb{Z}[h_1, h_2, \ldots]$, where $h_n$ in $A$ is as defined in Theorem 3.18, one can uniquely define an algebra morphism $A \xrightarrow{\varphi} A$ that sends the element $h_n$ to the complete homogeneous symmetric function $h_n(x)$. Assertions (d) and (e) show that $\varphi$ is a bialgebra isomorphism, and hence it is a Hopf isomorphism. To show that it is a PSH-isomorphism, we first note that it is an isometry because one can iterate Proposition 2.73(iii) together with assertions (c) and (d) to compute all inner products
\[
(h_\mu, h_\lambda)_A = (1, h_\mu^\perp h_\lambda)_A = (1, h_\mu^\perp h_\mu^\perp \cdots \cdots (h_\lambda, h_\lambda, \cdots))_A
\]
for $\mu, \lambda \in \text{Par}_n$. Hence
\[
(h_\mu, h_\lambda)_A = (h_\mu(x), h_\lambda(x))_\Lambda = (\varphi(h_\mu), \varphi(h_\lambda))_\Lambda
\]
Once one knows $\varphi$ is an isometry, then elements $\omega$ in $\Sigma \cap A_n$ are characterized in terms of the form $\langle \cdot, \cdot \rangle$ by $\langle \omega, \omega \rangle = 1$ and $\langle \omega, \rho^n \rangle > 0$. Hence $\varphi$ sends each $\sigma$ in $\Sigma$ to a Schur function $s_\lambda$, and is a PSH-isomorphism. \(\square\)
4. Complex representations for $\mathfrak{S}_n$, wreath products, $GL_n(\mathbb{F}_q)$

After reviewing the basics that we will need from representation and character theory of finite groups, we give Zelevinsky’s three main examples of PSH’s arising as spaces of virtual characters for three towers of finite groups:

- symmetric groups,
- their wreath products with any finite group, and
- the finite general linear groups.

Much in this chapter traces its roots to Zelevinsky’s book [173]. The results concerning the symmetric groups, however, are significantly older and spread across the literature: see, e.g., [157, §7.18], [54, §7.3], [109, §I.7], [141, §4.7], [89], for proofs using different tools.

4.1. Review of complex character theory. We shall now briefly discuss some basics of representation (and character) theory that will be used below. A good source for this material, including the crucial Mackey formula, is Serre [149, Chaps. 1-7].

4.1.1. Basic definitions, Maschke, Schur. For a group $G$, a representation is a homomorphism $G \xrightarrow{\varphi} GL(V)$ for some vector space $V$ over a field. We will take the field to be $\mathbb{C}$ from now on, and we will also assume that $V$ is finite-dimensional over $\mathbb{C}$. Thus a representation of $G$ is the same as a finite-dimensional (left-) $\mathbb{C}G$-module $V$. (We use the notations $\mathbb{C}G$ and $\mathbb{C}[G]$ synonymously for the group algebra of $G$ over $\mathbb{C}$. More generally, if $S$ is a set, then $\mathbb{C}S = \mathbb{C}[S]$ denotes the free $\mathbb{C}$-module with basis $S$.)

We also assume that $G$ is finite, so that Maschke’s Theorem\footnote{\ldots which has a beautiful generalization to finite-dimensional Hopf algebras due to Larson and Sweedler; see Montgomery [120, §2.2].} says that $\mathbb{C}G$ is semisimple, meaning that every $\mathbb{C}G$-module $U \subset V$ has a $\mathbb{C}G$-module complement $U'$ with $V = U \oplus U'$. Equivalently, indecomposable $\mathbb{C}G$-modules are the same thing as simple (=irreducible) $\mathbb{C}G$-modules.

Schur’s Lemma implies that for two simple $\mathbb{C}G$-modules $V_1, V_2$, one has
\[
\text{Hom}_{\mathbb{C}G}(V_1, V_2) \cong \begin{cases} \mathbb{C} & \text{if } V_1 \cong V_2, \\ 0 & \text{if } V_1 \not\cong V_2. \end{cases}
\]

4.1.2. Characters and Hom spaces. A $\mathbb{C}G$-module $V$ is completely determined up to isomorphism by its character
\[
\chi_V : G \to \mathbb{C}, \quad g \mapsto \chi_V(g) := \text{trace}(g : V \to V).
\]
This character $\chi_V$ is a class function, meaning it is constant on $G$-conjugacy classes. The space $R_{\mathbb{C}}(G)$ of class functions $G \to \mathbb{C}$ has a Hermitian, positive definite form
\[
(f_1, f_2)_G := \frac{1}{|G|} \sum_{g \in G} f_1(g) \overline{f_2(g)}.
\]
For any two $\mathbb{C}G$-modules $V_1, V_2$,
\[
(\chi_{V_1}, \chi_{V_2})_G = \dim_{\mathbb{C}} \text{Hom}_{\mathbb{C}G}(V_1, V_2).
\]
The set of all irreducible characters
\[
\text{Irr}(G) = \{ \chi_V : V \text{ is a simple } \mathbb{C}G\text{-module} \}
\]forms an orthonormal basis of $R_{\mathbb{C}}(G)$ with respect to this form, and spans a $\mathbb{Z}$-sublattice
\[
R(G) := \mathbb{Z}\text{Irr}(G) \subset R_{\mathbb{C}}(G)
\]
sometimes called the virtual characters of $G$. For every $\mathbb{C}G$-module $V$, the character $\chi_V$ belongs to $R(G)$.

Instead of working with the Hermitian form $\langle \cdot, \cdot \rangle_G$ on $G$, we could also (and some authors do) define a $\mathbb{C}$-bilinear form $\langle \cdot, \cdot \rangle_G$ on $R_{\mathbb{C}}(G)$ by
\[
\langle f_1, f_2 \rangle_G := \frac{1}{|G|} \sum_{g \in G} f_1(g) f_2(g^{-1}).
\]
This form is not identical with \( (\cdot,\cdot)_G \) (indeed, \((\cdot,\cdot)_G\) is bilinear while \((\cdot,\cdot)_G\) is Hermitian), but it still satisfies (4.1), and thus is identical with \((\cdot,\cdot)_G\) on \( R(G) \times R(G) \). Hence, for all we are going to do until Section 4.9, we could just as well use the form \((\cdot,\cdot)_G\) instead of \((\cdot,\cdot)_G\).

4.1.3. Tensor products. Given two groups \( G_1, G_2 \) and \( CG_i \)-modules \( V_i \) for \( i = 1, 2 \), their tensor product \( V_1 \otimes_{CG_1} V_2 \) becomes a \( CG(G_1 \times G_2) \)-module via \( (g_1, g_2)(v_1 \otimes v_2) = g_1(v_1) \otimes g_2(v_2) \). This module is called the \textit{outer tensor product} of \( V_1 \) and \( V_2 \). When \( V_1, V_2 \) are both simple, then so is \( V_1 \otimes V_2 \), and every simple \( CG(G_1 \times G_2) \)-module arises this way (with \( V_1 \) and \( V_2 \) determined uniquely up to isomorphism).\(^{145}\) Thus one has identifications and isomorphisms

\[
\text{Irr}(G_1 \times G_2) = \text{Irr}(G_1) \times \text{Irr}(G_2),
\]

\[
R(G_1 \times G_2) \cong R(G_1) \otimes_{CG_1} R(G_2);
\]

here, \( \chi_{V_1} \otimes \chi_{V_2} \in R(G_1) \otimes_{CG_1} R(G_2) \) is being identified with \( \chi_{V_1 \otimes V_2} \in R(G_1 \times G_2) \) for all \( CG_1 \)-modules \( V_1 \) and \( CG_2 \)-modules \( V_2 \). The latter isomorphism is actually a restriction of the isomorphism \( R_C(G_1 \times G_2) \cong R_C(G_1) \otimes_{CG_1} R_C(G_2) \) under which every pure tensor \( \phi_1 \otimes \phi_2 \in R_C(G_1) \otimes_{CG_1} R_C(G_2) \) corresponds to the class function \( G_1 \times G_2 \to C \), \( (g_1, g_2) \mapsto \phi_1(g_1) \otimes \phi_2(g_2) \).

Given two \( CG_1 \)-modules \( V_1 \) and \( W_1 \) and two \( CG_2 \)-modules \( V_2 \) and \( W_2 \), we have

\[
(\chi_{V_1 \otimes V_2}, \chi_{W_1 \otimes W_2})_{G_1 \times G_2} = (\chi_{V_1}, \chi_{W_1})_{G_1} \cdot (\chi_{V_2}, \chi_{W_2})_{G_2}.
\]

4.1.4. Induction and restriction. Given a subgroup \( H \leq G \) and \( CH \)-module \( U \), one can use the fact that \( CG \) is a \((CG, CH)\)-bimodule to form the \textit{induced} \( CG \)-module

\[
\text{Ind}_H^G U := CG \otimes_{CH} U.
\]

The fact that \( CG \) is free as a \((right-)CH\)-module\(^{146}\) on basis elements \( \{ t_g \}_{g \in H \subseteq G} \) makes this tensor product easy to analyze. For example one can compute its character

\[
\chi_{\text{Ind}_H^G U}(g) = \frac{1}{|H|} \sum_{k \in H: k g k^{-1} \in H} \chi_U(k g k^{-1}).
\]

\(^{147}\) One can also recognize when a \( CG \)-module \( V \) is isomorphic to \( \text{Ind}_H^G U \) for some \( CH \)-module \( U \); this happens if and only if there is an \( H \)-stable subspace \( V \subset W \) having the property that \( V = \bigoplus_{g \in H \subseteq G} g U \).

The above construction of a \( CG \)-module \( \text{Ind}_H^G U \) corresponding to any \( CH \)-module \( U \) is part of a functor \( \text{Ind}_H^G \) from the category of \( CH \)-modules to the category of \( CG \)-modules\(^{148}\); this functor is called \textit{induction}.

Besides induction on \( CH \)-modules, one can define induction on class functions of \( H \):

**Exercise 4.1.** Let \( G \) be a finite group, and \( H \) a subgroup of \( G \). Let \( f \in R_C(H) \) be a class function. We define the \textit{induction} \( \text{Ind}_H^G f \) of \( f \) to be the function \( G \to C \) given by

\[
\left( \text{Ind}_H^G f \right)(g) = \frac{1}{|H|} \sum_{k \in H: k g k^{-1} \in H} f(k g k^{-1}) \quad \text{ for all } g \in G.
\]

(a) Prove that this induction \( \text{Ind}_H^G f \) is a class function on \( G \), hence belongs to \( R_C(G) \).

(b) Let \( J \) be a system of right coset representatives for \( H \backslash G \), so that \( G = \bigsqcup_{j \in J} H j \). Prove that

\[
\left( \text{Ind}_H^G f \right)(g) = \sum_{j \in J: j g j^{-1} \in H} f(j g j^{-1}) \quad \text{ for all } g \in G.
\]
The induction $\text{Ind}_H^G$ defined in Exercise 4.1 is a $\mathbb{C}$-linear map $R_\mathbb{C}(H) \to R_\mathbb{C}(G)$. Since every $CH$-module $U$ satisfies
\begin{equation}
\chi_{\text{Ind}_H^G U} = \text{Ind}_H^G (\chi_U)
\end{equation}
149, this $\mathbb{C}$-linear map $\text{Ind}_H^G$ restricts to a $\mathbb{Z}$-linear map $R(H) \to R(G)$ (also denoted $\text{Ind}_H^G$) which sends the character $\chi_U$ of any $CH$-module $U$ to the character $\chi_{\text{Ind}_H^G U}$ of the induced $CG$-module $\text{Ind}_H^G U$.

**Exercise 4.2.** Let $G$, $H$ and $I$ be three finite groups such that $I < H < G$. Let $U$ be a $CI$-module. Prove that $\text{Ind}_H^I \text{Ind}_I^G U \cong \text{Ind}_H^G U$. (This fact is often referred to as the transitivity of induction.)

**Exercise 4.3.** Let $G_1$ and $G_2$ be two groups. Let $H_1 < G_1$ and $H_2 < G_2$ be two subgroups. Let $U_1$ be a $CH_1$-module, and $U_2$ be a $CH_2$-module. Show that
\begin{equation}
\text{Ind}_{H_1 \times H_2}^{G_1 \times G_2} (U_1 \otimes U_2) \cong \left( \text{Ind}_{H_1}^{G_1} U_1 \right) \otimes \left( \text{Ind}_{H_2}^{G_2} U_2 \right)
\end{equation}
as $\mathbb{C}[G_1 \times G_2]$-modules.

The restriction operation $V \mapsto \text{Res}_H^G V$ restricts a $CG$-module $V$ to a $CH$-module. Frobenius reciprocity asserts the adjointness between $\text{Ind}_H^G$ and $\text{Res}_H^G$
\begin{equation}
\text{Hom}_{\mathbb{C}G}(\text{Ind}_H^G U, V) \cong \text{Hom}_{\mathbb{C}H}(U, \text{Res}_H^G V),
\end{equation}
as a special case ($S = A = \mathbb{C}G$, $R = \mathbb{C}H$, $B = U$, $C = V$) of the general adjoint associativity
\begin{equation}
\text{Hom}_S(A \otimes_R B, C) \cong \text{Hom}_B(B, \text{Hom}_S(A, C))
\end{equation}
for $S$, $R$ two rings, $A$ an $(S, R)$-bimodule, $B$ a left $R$-module, $C$ a left $S$-module.

We can define not just the restriction of a $CG$-module, but also the restriction of a class function $f \in R_\mathbb{C}(G)$. When $H$ is a subgroup of $G$, the restriction $\text{Res}_H^G f$ of an $f \in R_\mathbb{C}(G)$ is defined as the result of restricting the map $f : G \to \mathbb{C}$ to $H$. This $\text{Res}_H^G f$ is easily seen to belong to $R_\mathbb{C}(H)$, and so $\text{Res}_H^G$ is a $\mathbb{C}$-linear map $R_\mathbb{C}(G) \to R_\mathbb{C}(H)$. This map restricts to a $\mathbb{Z}$-linear map $R(G) \to R(H)$, since we have $\text{Res}_H^G \chi_V = \chi_{\text{Res}_H^G V}$ for any $\mathbb{C}G$-module $V$. Taking characters in (4.7) (and recalling $\text{Res}_H^G \chi_V = \chi_{\text{Res}_H^G V}$ and (4.5)), we obtain
\begin{equation}
(\text{Ind}_H^G \chi_U, \chi_V)_H = (\chi_U, \text{Res}_H^G \chi_V)_H.
\end{equation}
By bilinearity, this yields the equality
\begin{equation}
\left( \text{Ind}_H^G \alpha, \beta \right)_H = (\alpha, \text{Res}_H^G \beta)_H
\end{equation}
for any class functions $\alpha \in R_\mathbb{C}(H)$ and $\beta \in R_\mathbb{C}(G)$ (since $R(G)$ spans $R_\mathbb{C}(G)$ as a $\mathbb{C}$-vector space).

**Exercise 4.4.** Let $G$ be a finite group, and let $H < G$. Let $U$ be a $CH$-module. If $A$ and $B$ are two algebras, $P$ is a $(B, A)$-bimodule and $Q$ is a left $B$-module, then $\text{Hom}_B(P, Q)$ is a left $A$-module (since $CG$ is a $(CH, CG)$-bimodule). As a consequence, $\text{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ is a $CG$-module. Prove that this $CG$-module is isomorphic to $\text{Ind}_H^G U$.

**Remark 4.5.** Some texts define the induction $\text{Ind}_H^G U$ of a $CH$-module $U$ to be $\text{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ (rather than to be $\mathbb{C}G \otimes_{\mathbb{C}H} U$, as we did).150 As Exercise 4.4 shows, this definition is equivalent to ours as long as $G$ is finite (but not otherwise).

Exercise 4.4 yields the following “wrong-way” version of Frobenius reciprocity:

**Exercise 4.6.** Let $G$ be a finite group; let $H < G$. Let $U$ be a $CG$-module, and let $V$ be a $CH$-module. Prove that $\text{Hom}_{\mathbb{C}G}(U, \text{Ind}_H^G V) \cong \text{Hom}_{\mathbb{C}H}(\text{Res}_H^G U, V)$.

149 This follows by comparing the value of $\chi_{\text{Ind}_H^G U}(g)$ obtained from (4.3) with the value of $(\text{Ind}_H^G (\chi_U))(g)$ found using (4.4).

150 Or they define it as a set of morphisms of $H$-sets from $G$ to $U$ (this is how [50, Def. 5.8.1 (Def. 4.28 in the arXiv version)] defines it); this is easily seen to be equivalent to $\text{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$.
4.1.5. Mackey’s formula. Mackey gave an alternate description of a module which has been induced and then restricted. To state it, for a subgroup $H < G$ and $g$ in $G$, let $^gH := g^{-1}Hg$ and $^gH := gHg^{-1}$. Given a $\mathbb{C}H$-module $U$, say defined by a homomorphism $H \not\to GL(U)$, let $U^g$ denote the $\mathbb{C}^gH$-module on the same $\mathbb{C}$-vector space $U$ defined by the composite homomorphism

\[ ^gH \rightarrow H \rightarrow ^gH. \]

**Theorem 4.7.** (Mackey’s formula) Consider subgroups $H, K < G$, and any $\mathbb{C}H$-module $U$. If $\{g_1, \ldots, g_t\}$ are double coset representatives for $K \backslash G / H$, then

\[ \text{Res}^G_H \text{Ind}^G_K U \cong \bigoplus_{i=1}^t \text{Ind}^K_{K \cap \psi_i} \left( \left( \text{Res}^H_{K \cap \psi_i} U \right)^g \right) \]

**Proof.** In this proof, all tensor product symbols $\otimes$ should be interpreted as $\otimes_{\mathbb{C}H}$. Recall $\mathbb{C}G$ has $\mathbb{C}$-basis $\{t_g\}_{g \in G}$. For subsets $S \subset G$, let $\mathbb{C}[S]$ denote the $\mathbb{C}$-span of $\{t_g\}_{g \in S}$ in $\mathbb{C}G$.

Note that each double coset $KgH$ gives rise to a sub-$\text{Ind}^K_{K \cap \psi_i}$-module $\mathbb{C}[KgH]$ within $\mathbb{C}G$, and one has a $\mathbb{C}K$-module direct sum decomposition

\[ \text{Ind}^G_H U = \mathbb{C}G \otimes U = \bigoplus_{i=1}^t \mathbb{C}[KgH] \otimes U. \]

Hence it suffices to check for any element $g$ in $G$ that

\[ \mathbb{C}[KgH] \otimes U \cong \text{Ind}^K_{K \cap \psi_i} \left( \left( \text{Res}^H_{K \cap \psi_i} U \right)^g \right). \]

Note that $^gH \cap K$ is the subgroup of $K$ consisting of the elements $k$ in $K$ for which $kgH = gH$. Hence by picking $\{k_1, \ldots, k_s\}$ to be coset representatives for $K / (^gH \cap K)$, one disjointly decomposes the double coset

\[ KgH = \bigsqcup_{j=1}^s k_j (^gH \cap K)gH, \]

giving a $\mathbb{C}$-vector space direct sum decomposition

\[ \mathbb{C}[KgH] \otimes U = \bigoplus_{j=1}^s \mathbb{C}[k_j (^gH \cap K)gH] \otimes U \]
\[ \cong \text{Ind}^K_{K \cap \psi_i} \left( \mathbb{C}[(^gH \cap K)gH] \otimes U \right). \]

So it remains to check that one has a $\mathbb{C}[^gH \cap K]$-module isomorphism

\[ \mathbb{C}[(^gH \cap K)gH] \otimes U \cong \left( \text{Res}^H_{K \cap \psi_i} U \right)^g. \]

Bearing in mind that, for each $k$ in $^gH \cap K$ and $h$ in $H$, one has $g^{-1}kg$ in $H$ and hence

\[ t_{kgh} \otimes u = t_g \cdot t_{g^{-1}kg} \cdot u = t_g \otimes g^{-1}kgh \cdot u \]

one sees that this isomorphism can be defined by mapping

\[ t_{kgh} \otimes u \mapsto g^{-1}kgh \cdot u. \]

\[ \square \]

4.1.6. Inflation and fixed points. There are two (adjoint) constructions on representations that apply when one has a normal subgroup $K \triangleleft G$. Given a $\mathbb{C}[G/K]$-module $U$, say defined by the homomorphism $G/K \not\to GL(U)$, the inflation of $U$ to a $\mathbb{C}G$-module $\text{Infl}^G_U$ has the same underlying space $U$, and is defined by the composite homomorphism $G \to G/K \not\to GL(U)$. We will later use the easily-checked fact that when $H < G$ is any other subgroup, one has

\[ \text{Res}^G_H \text{Infl}^G_U = \text{Infl}^H_{H/H \cap K} \text{Res}^G_H \text{Infl}^G_U. \]

(We regard $H/H \cap K$ as a subgroup of $G/K$, since the canonical homomorphism $H/H \cap K \to G/K$ is injective.)
Inflation turns out to be adjoint to the $K$-fixed space construction sending a $\mathbb{C}G$-module $V$ to the $\mathbb{C}[G/K]$-module

$$V^K := \{ v \in V : k(v) = v \text{ for } k \in K \}$$

Note that $V^K$ is indeed a $G$-stable subspace: for any $v \in V^K$ and $g \in G$, one has that $g(v)$ lies in $V^K$ since an element $k$ in $K$ satisfies $kg(v) = g \cdot g^{-1}kg(v) = g(v)$ as $g^{-1}kg$ lies in $K$. One has this adjointness

$$\text{Hom}_{\mathbb{C}G}(\text{Infl}^G_{G/K} U, V) = \text{Hom}_{\mathbb{C}[G/K]}(U, V^K)$$

because any $\mathbb{C}G$-module homomorphism $\varphi$ on the left must have the property that $k\varphi(u) = \varphi(k(u)) = \varphi(u)$ for all $k \in K$, so that $\varphi$ actually lies on the right.

We will also need the following formula for the character $\chi_{V^K}$ in terms of the character $\chi_V$:

$$\chi_{V^K}(gK) = \frac{1}{|K|} \sum_{k \in K} \chi_V(gk).$$

To see this, note that when one has a $\mathbb{C}$-linear endomorphism $\varphi$ on a space $V$ that preserves some $\mathbb{C}$-subspace $W \subset V$, if $V \xrightarrow{\pi} W$ is any idempotent projection onto $W$, then the trace of the restriction $\varphi|_W$ equals the trace of $\varphi \circ \pi$ on $V$. Applying this to $W = V^K$ and $\varphi = g$, with $\pi = \frac{1}{|K|} \sum_{k \in K} k$, gives (4.12).\footnote{For another proof of (4.12), see Exercise 4.14(1).}

Another way to restate (4.12) is:

$$\chi_{V^K}(gK) = \frac{1}{|K|} \sum_{h \in gK} \chi_V(h).$$

Inflation and $K$-fixed space construction can also be defined on class functions. For inflation, this is particularly easy: Inflation $\text{Infl}^G_{G/K} f$ of an $f \in R_C(G/K)$ is defined as the composition $G \xrightarrow{f} \mathbb{C} \xrightarrow{\chi_V} R_C(G/K)$.

This is a class function of $G$ and thus lies in $R_C(G)$. Thus, inflation $\text{Infl}^G_{G/K}$ is a $\mathbb{C}$-linear map $R_C(G/K) \to R_C(G)$. It restricts to a $\mathbb{Z}$-linear map $R(G/K) \to R(G)$, since it is clear that every $\mathbb{C}(G/K)$-module $U$ satisfies $\text{Infl}^G_{G/K} \chi_U = \chi_{\text{Infl}^G_{G/K} U}$.

We can also use (4.12) (or (4.13)) as inspiration for defining a “$K$-fixed space construction” on class functions. Explicitly, for every class function $f \in R_C(G)$, we define a class function $f^K \in R_C(G/K)$ by

$$f^K(gK) = \frac{1}{|K|} \sum_{k \in K} f(gk) = \frac{1}{|K|} \sum_{h \in gK} f(h).$$

The map $(\cdot)^K : R_C(G) \to R_C(G/K)$, $f \mapsto f^K$ is $\mathbb{C}$-linear, and restricts to a $\mathbb{Z}$-linear map $R(G) \to R(G/K)$. Again, we have a compatibility with the $K$-fixed point construction on modules: We have $\chi_{V^K} = (\chi_V)^K$ for every $\mathbb{C}G$-module $V$.

Taking characters in (4.11), we obtain

$$(\text{Infl}^G_{G/K} \chi_U, \chi_V)_G = (\chi_U, \chi_V^K)_G$$

for any $\mathbb{C}[G/K]$-module $U$ and any $\mathbb{C}G$-module $V$ (since $\chi_{\text{Infl}^G_{G/K} U} = \text{Infl}^G_{G/K} \chi_U$ and $\chi_{V^K} = (\chi_V)^K$). By $\mathbb{Z}$-linearity, this implies that

$$\left( \text{Infl}^G_{G/K} \alpha, \beta \right)_G = (\alpha, \beta^K)_G$$

for any class functions $\alpha \in R_C(G/K)$ and $\beta \in R_C(G)$.

There is also an analogue of (4.6):

**Lemma 4.8.** Let $G_1$ and $G_2$ be two groups, and $K_1 < G_1$ and $K_2 < G_2$ be two respective subgroups. Let $U_i$ be a $\mathbb{C}G_i$-module for each $i \in \{1, 2\}$. Then,

$$\left( U_1 \otimes U_2 \right)^{K_1 \times K_2} = U_1^{K_1} \otimes U_2^{K_2}$$

(as subspaces of $U_1 \otimes U_2$).
Proof. The subgroup $K_1 = K_1 \times 1$ of $G_1 \times G_2$ acts on $U_1 \otimes U_2$, and its fixed points are $\left(U_1 \otimes U_2\right)^{K_1} = U_1^{K_1} \otimes U_2$ (because for a $\mathbb{C}K_1$-module, tensoring with $U_2$ is the same as taking a direct power, which clearly commutes with taking fixed points). Similarly, $\left(U_1 \otimes U_2\right)^{K_2} = U_1 \otimes U_2^{K_2}$. Now,

\[(U_1 \otimes U_2)^{K_1 \times K_2} = (U_1 \otimes U_2)^{K_1} \cap (U_1 \otimes U_2)^{K_2} = \left(U_1^{K_1} \otimes U_2\right) \cap \left(U_1 \otimes U_2^{K_2}\right) = U_1^{K_1} \otimes U_2^{K_2}\]

according to the known linear-algebraic fact stating that if $P$ and $Q$ are subspaces of two vector spaces $U$ and $V$, respectively, then $(P \otimes V) \cap (U \otimes Q) = P \otimes Q$. □

Exercise 4.9. (a) Let $G_1$ and $G_2$ be two groups. Let $V_i$ and $W_i$ be finite-dimensional $\mathbb{C}G_i$-modules for every $i \in \{1, 2\}$. Prove that the $\mathbb{C}$-linear map

$$\text{Hom}_{\mathbb{C}G_1} (V_1, W_1) \otimes \text{Hom}_{\mathbb{C}G_2} (V_2, W_2) \rightarrow \text{Hom}_{\mathbb{C}[G_1 \times G_2]} (V_1 \otimes V_2, W_1 \otimes W_2)$$

sending each tensor $f \otimes g$ to the tensor product $f \otimes g$ of homomorphisms is a vector space isomorphism.

(b) Use part (a) to give a new proof of (4.2).

As an aside, (4.10) has a “dual” analogue:

Exercise 4.10. Let $G$ be a finite group, and let $K \triangleleft H < G$. Let $U$ be a $\mathbb{C}H$-module. As usual, regard $H/(H \cap K)$ as a subgroup of $G/K$. Show that

$$\left(\text{Ind}_{H}^{G} U\right)^{K} \cong \text{Ind}_{H/(H \cap K)}^{G/K} \left(U^{H \cap K}\right)$$

as $\mathbb{C}[G/K]$-modules.

Inflation also “commutes” with induction:

Exercise 4.11. Let $G$ be a finite group, and let $K < H < G$ be such that $K \triangleleft G$. Thus, automatically, $K \triangleleft H$, and we regard the quotient $H/K$ as a subgroup of $G/K$. Let $V$ be a $\mathbb{C}[H/K]$-module. Show that

$$\text{Infl}_{G/K}^{\mathbb{C}} \text{Ind}_{H/K}^{\mathbb{C}} V \cong \text{Ind}_{H/(H \cap K)}^{G/K} \text{Infl}_{H/K}^{\mathbb{C}} V$$

as $\mathbb{C}$-modules.

Exercise 4.12. Let $G$ be a finite group, and let $K \triangleleft G$. Let $V$ be a $\mathbb{C}G$-module. Let $I_{V,K}$ denote the $\mathbb{C}$-vector subspace of $V$ spanned by all elements of the form $v - kv$ for $k \in K$ and $v \in V$.

(a) Show that $I_{V,K}$ is a $\mathbb{C}G$-submodule of $V$.

(b) Let $V_K$ denote the quotient $\mathbb{C}G$-module $V/I_{V,K}$. (This module is occasionally called the $K$-coinvariant module of $V$, a name it sadly shares with at least two other non-equivalent constructions in algebra.) Show that $V_K \cong \text{Infl}_{G/K}^{\mathbb{C}} (V^K)$ as $\mathbb{C}G$-modules. (Use char $\mathbb{C} = 0$.)

In the remainder of this subsection, we shall briefly survey generalized notions of induction and restriction, defined in terms of a group homomorphism $\rho$ rather than in terms of a group $G$ and a subgroup $H$. These generalized notions (defined by van Leeuwen in [97, §2.2]) will not be used in the rest of these notes, but they shed some new light on the facts about induction, restriction, inflation and fixed point construction discussed above. (In particular, they reveal that some of said facts have common generalizations.)

The reader might have noticed that the definitions of inflation and of restriction (both for characters and for modules) are similar. In fact, they both are particular cases of the following construction:

Remark 4.13. Let $G$ and $H$ be two finite groups, and let $\rho : H \rightarrow G$ be a group homomorphism.

- If $f \in R_{C} (G)$, then the $\rho$-restriction $\text{Res}_{\rho} f$ of $f$ is defined as the map $f \circ \rho : H \rightarrow \mathbb{C}$. This map is easily seen to belong to $R_{C} (H)$.

- If $V$ is a $\mathbb{C}G$-module, then the $\rho$-restriction $\text{Res}_{\rho} V$ of $V$ is the $\mathbb{C}H$-module with ground space $V$ and action given by

$$h \cdot v = \rho (h) \cdot v \quad \text{for every } h \in H \text{ and } v \in V.$$"
• If \( f \in R_C(H) \), then the \( \rho \)-induction \( \text{Ind}_\rho f \) of \( f \) is a map \( G \to C \) which is defined as follows:

\[
(\text{Ind}_\rho f)(g) = \frac{1}{|H|} \sum_{(h,k) \in H \times G; \ k^{-1}h = g} f(h)
\]

for every \( g \in G \).

• If \( U \) is a \( CH \)-module, then the \( \rho \)-induction \( \text{Ind}_\rho U \) of \( U \) is defined as the \( CG \)-module \( CG \otimes_{CH} U \), where \( CG \) is regarded as a \( (CG, CH) \)-bimodule according to the following rule: The left \( CG \)-module structure on \( CG \) is plain multiplication inside \( CG \); the right \( CH \)-module structure on \( CG \) is induced by the \( C \)-algebra homomorphism \( C[\rho] : CH \to CG \) (thus, it is explicitly given by \( \gamma \eta = \gamma \cdot (C[\rho]) \eta \) for all \( \gamma \in CG \) and \( \eta \in CH \)).

Prove the following properties of this construction:

(a) For every \( f \in R_C(H) \), we have \( \text{Ind}_\rho f \in R_C(G) \).

(b) For any finite-dimensional \( CH \)-module \( U \), we have \( \chi_{\text{Ind}_\rho U} = \text{Ind}_\rho \chi_U \).

(c) If \( H \) is a subgroup of \( G \), and if \( \rho : H \to G \) is the inclusion map, then \( \text{Ind}_\rho f = \text{Ind}_H^G f \) for every \( f \in R_C(H) \).

(d) If \( H \) is a subgroup of \( G \), and if \( \rho : H \to G \) is the inclusion map, then \( \text{Ind}_\rho U = \text{Ind}_H^G U \) for every \( CH \)-module \( U \).

(e) If \( G = H/K \) for some normal subgroup \( K \) of \( H \), and if \( \rho : H \to G \) is the projection map, then \( \text{Ind}_\rho f = f^K \) for every \( f \in R_C(H) \).

(f) If \( G = H/K \) for some normal subgroup \( K \) of \( H \), and if \( \rho : H \to G \) is the projection map, then \( \text{Ind}_\rho U = U^K \) for every \( CH \)-module \( U \).

(g) Any class functions \( \alpha \in R_C(H) \) and \( \beta \in R_C(G) \) satisfy

\[
(\text{Ind}_\rho \alpha, \beta)_G = (\alpha, \text{Res}_\rho \beta)_H
\]

and

\[
(\text{Ind}_\rho \alpha, \beta)_G = (\alpha, \text{Res}_\rho \beta)_H.
\]

(See Remark 4.13 for the definition of \( \text{Res}_\rho \beta \).)

(h) We have \( \text{Hom}_{CG}(\text{Ind}_\rho U, V) \cong \text{Hom}_{CH}(U, \text{Res}_\rho V) \) for every \( CH \)-module \( U \) and every \( CG \)-module \( V \). (See Remark 4.13 for the definition of \( \text{Res}_\rho V \).)

(i) Similarly to how we made \( CG \) into a \( (CG, CH) \)-bimodule, let us make \( CG \) into a \( (CH, CG) \)-bimodule (so the left \( CG \)-module structure is plain multiplication inside \( CG \), whereas the right \( CH \)-module structure is induced by the \( C \)-algebra homomorphism \( C[\rho] : CH \to CG \)). If \( U \) is any \( CH \)-module, then the \( CG \)-module \( \text{Hom}_{CH}(CG, U) \) (defined as in Exercise 4.4 using the \( (CH, CG) \)-bimodule structure on \( CG \)) is isomorphic to \( \text{Ind}_\rho U \).

(j) We have \( \text{Hom}_{CG}(U, \text{Ind}_\rho V) \cong \text{Hom}_{CH}(\text{Res}_\rho U, V) \) for every \( CG \)-module \( U \) and every \( CH \)-module \( V \). (See Remark 4.13 for the definition of \( \text{Res}_\rho V \).)

Furthermore:

(k) Use the above to prove the formula (4.3).

(l) Use the above to prove the formula (4.12).

[Hint: Part (b) of this exercise is hard. To solve it, it is useful to have a way of computing the trace of a linear operator without knowing a basis of the vector space it is acting on. There is a way to do this using a "finite dual generating system", which is a somewhat less restricted notion than that of a basis\(^{152}\). Try to create a finite dual generating system for \( \text{Ind}_\rho U \) from one for \( U \) (and from the group \( G \)), and then use it to compute \( \chi_{\text{Ind}_\rho U} \).]

\(^{152}\)More precisely: Let \( K \) be a field, and \( V \) be a \( K \)-vector space. A \textit{finite dual generating system} for \( V \) means a triple \((I, (a_i)_{i \in I}, (f_i)_{i \in I})\), where

- \( I \) is a finite set;
- \((a_i)_{i \in I}\) is a family of elements of \( V \);
- \((f_i)_{i \in I}\) is a family of elements of \( V^* \) (where \( V^* \) means \( \text{Hom}_K(V, K) \))

such that every \( v \in V \) satisfies \( v = \sum_{i \in I} f_i(v)a_i \). For example, if \((e_j)_{j \in J}\) is a finite basis of the vector space \( V \), and if \( (e_j^*)_{j \in J} \) is the basis of \( V^* \) dual to this basis \((e_j)_{j \in J}\), then \((J, (e_j)_{j \in J}, (e_j^*)_{j \in J})\) is a finite dual generating system for \( V \); however, most finite dual generating systems are not obtained this way.
The solution of part (i) is a modification of the solution of Exercise 4.4, but complicated by the fact that $H$ is no longer (necessarily) a subgroup of $G$. Part (f) can be solved by similar arguments, or using part (i), or using Exercise 4.12(b).

The result of Exercise 4.14(h) generalizes (4.7) (because of Exercise 4.14(d)), but also generalizes (4.11) (due to Exercise 4.14(f)). Similarly, Exercise 4.14(g) generalizes both (4.9) and (4.14). Similarly, Exercise 4.14(i) generalizes Exercise 4.4, and Exercise 4.14(j) generalizes Exercise 4.6.

Similarly, Exercise 4.3 is generalized by the following exercise:

**Exercise 4.15.** Let $G_1, G_2, H_1$ and $H_2$ be four finite groups. Let $\rho_1 : H_1 \to G_1$ and $\rho_2 : H_2 \to G_2$ be two group homomorphisms. These two homomorphisms clearly induce a group homomorphism $\rho_1 \times \rho_2 : H_1 \times H_2 \to G_1 \times G_2$. Let $U_1$ be a $\mathbb{C}H_1$-module, and $U_2$ be a $\mathbb{C}H_2$-module. Show that

$$\operatorname{Ind}_{\rho_1 \times \rho_2} (U_1 \otimes U_2) \cong (\operatorname{Ind}_{\rho_1} U_1) \otimes (\operatorname{Ind}_{\rho_2} U_2)$$

as $\mathbb{C}[G_1 \times G_2]$-modules.

The $\operatorname{Ind}_\rho$ and $\operatorname{Res}_\rho$ operators behave “functorially” with respect to composition. Here is what this means:

**Exercise 4.16.** Let $G$, $H$ and $I$ be three finite groups. Let $\rho : H \to G$ and $\tau : I \to H$ be two group homomorphisms.

(a) We have $\operatorname{Ind}_\tau \operatorname{Ind}_\rho U \cong \operatorname{Ind}_{\rho \circ \tau} U$ for every $\mathbb{C}I$-module $U$.

(b) We have $\operatorname{Ind}_\tau \operatorname{Ind}_\rho f = \operatorname{Ind}_{\rho \circ \tau} f$ for every $f \in \mathbb{C}_G(I)$.

(c) We have $\operatorname{Res}_\rho \operatorname{Res}_\tau V = \operatorname{Res}_{\tau \circ \rho} V$ for every $\mathbb{C}G$-module $V$.

(d) We have $\operatorname{Res}_\tau \operatorname{Res}_\rho f = \operatorname{Res}_{\rho \circ \tau} f$ for every $f \in \mathbb{C}_G(G)$.

Exercise 4.16(a), of course, generalizes Exercise 4.2.

4.1.7. **Semidirect products.** Recall that a *semidirect product* is a group $G \ltimes K$ having two subgroups $G, K$ with

- $K \trianglelefteq (G \ltimes K)$ is a normal subgroup,
- $G \ltimes K = GK = KG$, and
- $G \cap K = \{e\}$.

In this setting one has two interesting adjoint constructions, applied in Section 4.5.

**Proposition 4.17.** Fix a $\mathbb{C}[G \ltimes K]$-module $V$.

(i) For any $\mathbb{C}G$-module $U$, one has $\mathbb{C}[G \ltimes K]$-module structure

$$\Phi(U) := U \otimes V,$$

determined via

$$k(u \otimes v) = u \otimes k(v),$$
$$g(u \otimes v) = g(u) \otimes g(v).$$

(ii) For any $\mathbb{C}[G \ltimes K]$-module $W$, one has $\mathbb{C}G$-module structure

$$\Psi(W) := \operatorname{Hom}_{\mathbb{C}K}(\operatorname{Res}_{\rho \circ \tau}^G V, \operatorname{Res}_{\rho \circ \tau}^G W),$$

determined via $g(\varphi) = g \circ \varphi \circ g^{-1}$.

(iii) The maps

$$\mathbb{C}G - \text{mods} \xrightarrow{\Phi} \mathbb{C}[G \ltimes K] - \text{mods}$$

The crucial observation is now that if $(I, (a_i)_{i \in I}, (f_i)_{i \in I})$ is a finite dual generating system for a vector space $V$, and if $T$ is an endomorphism of $V$, then

$$\text{trace } T = \sum_{i \in I} f_i(Ta_i).$$

Prove this!
are adjoint in the sense that one has an isomorphism
\[
\text{Hom}_{\mathbb{C}G}(U, \Psi(W)) \cong \text{Hom}_{\mathbb{C}[G \ltimes K]}(\Phi(U), W)
\]
\[
\text{Hom}_{\mathbb{C}G}(U, \text{Hom}_{\mathbb{C}K}(\text{Res}_K^{G \ltimes K} V, \text{Res}_K^{G \ltimes K} W)) \cong \text{Hom}_{\mathbb{C}[G \ltimes K]}(U \otimes V, W)
\]
\[
\varphi \mapsto \varphi(u \otimes v) := \varphi(u)(v)
\]

(iv) One has a \(\mathbb{C}G\)-module isomorphism
\[
(\Psi \circ \Phi)(U) \cong U \otimes \text{End}_{\mathbb{C}K}(\text{Res}_K^{G \ltimes K} V).
\]
In particular, if \(\text{Res}_K^{G \ltimes K} V\) is a simple \(\mathbb{C}K\)-module, then \((\Psi \circ \Phi)(U) \cong U\).

Proof. These are mostly straightforward exercises in the definitions. To check assertion (iv), for example, note that \(K\) acts only in the right tensor factor in \(\text{Res}_K^{G \ltimes K} (U \otimes V)\), and hence as \(\mathbb{C}G\)-modules one has
\[
(\Psi \circ \Phi)(U) = \text{Hom}_{\mathbb{C}K}(\text{Res}_K^{G \ltimes K} V, \text{Res}_K^{G \ltimes K} (U \otimes V))
\]
\[
= \text{Hom}_{\mathbb{C}K}(\text{Res}_K^{G \ltimes K} V, U \otimes \text{Res}_K^{G \ltimes K} V)
\]
\[
= U \otimes \text{Hom}_{\mathbb{C}K}(\text{Res}_K^{G \ltimes K} V, \text{Res}_K^{G \ltimes K} V)
\]
\[
= U \otimes \text{End}_{\mathbb{C}K}(\text{Res}_K^{G \ltimes K} V)
\]

\[
\square
\]

4.2. Three towers of groups. Here we consider three towers of groups
\[
G_\ast = (G_0 < G_1 < G_2 < G_3 < \cdots)
\]
where either
- \(G_n = \mathfrak{S}_n\), the symmetric group\(^{154}\), or
- \(G_n = \mathfrak{S}_n[G]\), the wreath product of the symmetric group with some arbitrary finite group \(\Gamma\), or
- \(G_n = \text{GL}_n(\mathbb{F}_q)\), the finite general linear group\(^{154}\).

Here the wreath product \(\mathfrak{S}_n[G]\) can be thought of informally as the group of monomial \(n \times n\) matrices whose nonzero entries lie in \(\Gamma\), that is, \(n \times n\) matrices having exactly one nonzero entry in each row and column, and that entry is an element of \(\Gamma\). E.g.
\[
\begin{bmatrix}
g_2 & 0 & 0 
g_1 & 0 & 0 
g_3 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
0 & 0 & g_6 
0 & g_5 & 0 
g_4 & 0 & 0
\end{bmatrix}
= \begin{bmatrix}
g_{2g_5} & 0 & 0 
0 & g_1 & g_6 
g_3g_4 & 0 & 0
\end{bmatrix}.
\]

More formally, \(\mathfrak{S}_n[G]\) is the semidirect product \(\mathfrak{S}_n \ltimes \Gamma^n\) in which \(\mathfrak{S}_n\) acts on \(\Gamma^n\) via \(\sigma(\gamma_1, \ldots, \gamma_n) = (\gamma_{\sigma^{-1}(1)}, \ldots, \gamma_{\sigma^{-1}(n)})\).

For each of the three towers \(G_\ast\), there are embeddings \(G_i \times G_j \hookrightarrow G_{i+j}\) and we introduce maps \(\text{ind}_{i,j}^{i+j}\) taking \(\mathbb{C}[G_i \times G_j]\)-modules to \(\mathbb{C}G_{i+j}\)-modules, as well as maps \(\text{res}_{i,j}^{i+j}\) carrying modules in the reverse direction which are adjoint:
\[
(4.18) \quad \text{Hom}_{\mathbb{C}G_{i+j}}(\text{ind}_{i,j}^{i+j} U, V) = \text{Hom}_{\mathbb{C}[G_i \times G_j]}(U, \text{res}_{i,j}^{i+j} V).
\]

Definition 4.18. For \(G_n = \mathfrak{S}_n\), one embeds \(\mathfrak{S}_i \times \mathfrak{S}_j\) into \(\mathfrak{S}_{i+j}\) as the permutations that permute \(\{1, 2, \ldots, i\}\) and \(\{i+1, i+2, \ldots, i+j\}\) separately. Here one defines
\[
\text{ind}_{i,j}^{i+j} := \text{Ind}_{\mathfrak{S}_i \times \mathfrak{S}_j}^{\mathfrak{S}_{i+j}},
\]
\[
\text{res}_{i,j}^{i+j} := \text{Res}_{\mathfrak{S}_i \times \mathfrak{S}_j}^{\mathfrak{S}_{i+j}}.
\]

\(^{153}\)The symmetric group \(\mathfrak{S}_0\) is the group of all permutations of the empty set \(\{1, 2, \ldots, 0\} = \emptyset\). It is a trivial group. (Note that \(\mathfrak{S}_1\) is also a trivial group.)

\(^{154}\)The group \(\text{GL}_0(\mathbb{F}_q)\) is a trivial group, consisting of the empty \(0 \times 0\) matrix.
For $G_n = S_n[\Gamma]$, similarly embed $S_i[\Gamma] \times S_j[\Gamma]$ into $S_i+j[\Gamma]$ as block monomial matrices whose two diagonal blocks have sizes $i, j$ respectively, and define

\[
\begin{align*}
\text{ind}_{i,j}^{i+j} & := \text{Ind}_{S_i[\Gamma] \times S_j[\Gamma]}^{S_{i+j}[\Gamma]}, \\
\text{res}_{i,j}^{i+j} & := \text{Res}_{S_i[\Gamma] \times S_j[\Gamma]}^{S_{i+j}[\Gamma]}.
\end{align*}
\]

For $G_n = GL_n(\mathbb{F}_q)$, which we will denote just $GL_n$, similarly embed $GL_i \times GL_j$ into $GL_{i+j}$ as block diagonal matrices whose two diagonal blocks have sizes $i, j$ respectively. However, one also introduces as an intermediate the parabolic subgroup $P_{i,j}$ consisting of the block upper-triangular matrices of the form

\[
\begin{pmatrix}
g_i & \ell \\
0 & g_j
\end{pmatrix}
\]

where $g_i, g_j$ lie in $GL_i, GL_j$, respectively, and $\ell$ in $\mathbb{F}_q^{i \times j}$ is arbitrary. One has a quotient map $P_{i,j} \twoheadrightarrow GL_i \times GL_j$ whose kernel $K_{i,j}$ is the set of matrices of the form

\[
\begin{pmatrix}
I_i & \ell \\
0 & I_j
\end{pmatrix}
\]

with $\ell$ again arbitrary. Here one defines

\[
\begin{align*}
\text{ind}_{i,j}^{i+j} & := \text{Ind}_{P_{i,j}}^{GL_i \times GL_j}, \\
\text{res}_{i,j}^{i+j} & := \left(\text{Res}_{P_{i,j}}^{GL_i \times GL_j}(-)\right)_{K_{i,j}}.
\end{align*}
\]

In the case $G_n = GL_n$, the operation $\text{ind}_{i,j}^{i+j}$ is sometimes called parabolic induction or Harish-Chandra induction. The operation $\text{res}_{i,j}^{i+j}$ is essentially just the $K_{i,j}$-fixed point construction $V \mapsto V^{K_{i,j}}$. However writing it as the above two-step composite makes it more obvious, (via (4.7) and (4.11)) that $\text{res}_{i,j}^{i+j}$ is again adjoint to $\text{ind}_{i,j}^{i+j}$.

**Definition 4.19.** For each of the three towers $G_\bullet$, define a graded $\mathbb{Z}$-module

\[
A := A(G_\bullet) = \bigoplus_{n \geq 0} R(G_n)
\]

with a bilinear form $(\cdot, \cdot)_A$ whose restriction to $A_n := R(G_n)$ is the usual form $(\cdot, \cdot)_{R(G_n)}$, and such that $\Sigma := \bigsqcup_{n \geq 0} \text{Irr}(G_n)$ gives an orthonormal $\mathbb{Z}$-basis. Notice that $A_0 = \mathbb{Z}$ has its basis element 1 equal to the unique irreducible character of the trivial group $G_0$.

Bearing in mind that $A_n = R(G_n)$ and

\[
A_i \otimes A_j = R(G_i) \otimes R(G_j) \cong R(G_i \times G_j)
\]

one then has candidates for product and coproduct defined by

\[
\begin{align*}
m := \text{ind}_{i,j}^{i+j} : & \quad A_i \otimes A_j \longrightarrow A_{i+j}, \\
\Delta := \bigoplus_{i+j=n} \text{res}_{i,j}^{i+j} : & \quad A_n \longrightarrow \bigoplus_{i+j=n} A_i \otimes A_j.
\end{align*}
\]

The coassociativity of $\Delta$ is an easy consequence of transitivity of the constructions of restriction and fixed points\footnote{More precisely, using this transitivity, it is easily reduced to proving that $K_{i+j,k} : (K_{i,j} \times \{I_k\}) = K_{i+j+k} : (\{I_i\} \times K_{j,k})$ (an equality between subgroups of $GL_{i+j+k}$). This equality can be proven by realizing that both of its sides equal the set of all block matrices of the form

\[
\begin{pmatrix}
I_i & \ell & m \\
0 & I_j & n \\
0 & 0 & I_k
\end{pmatrix}
\]

with $\ell, m$ and $n$ being matrices of sizes $i \times j$, $i \times k$ and $j \times k$, respectively.} we will instead prove it differently.

We first show that the maps $m$ and $\Delta$ are adjoint with respect to the forms $(\cdot, \cdot)_A$ and $(\cdot, \cdot)_{A \otimes A}$. In fact, if $U$, $V$, $W$ are modules over $CG_i$, $CG_j$, $CG_{i+j}$, respectively, then we can write the $\mathbb{C}[G_i \times G_j]$-module $\text{res}_{i,j}^{i+j} W$ as a direct sum $\bigoplus_k X_k \otimes Y_k$ with $X_k$ being $CG_i$-modules and $Y_k$ being $CG_j$-modules; we then have

\[
(4.19) \quad \text{res}_{i,j}^{i+j} \chi W = \sum_k \chi X_k \otimes \chi Y_k
\]

\footnote{See Exercise 4.31(c) for such a derivation.}
and

\[
(m (\chi_U \otimes \chi_V), \chi_W)_A = \left( \text{Ind}_{G_{i+j}}^{G_{i+j+k}} (\chi_U \otimes \chi_V), \chi_W \right)_{G_{i+j}} = \left( \text{Ind}_{G_{i+j}}^{G_{i+j+k}} (\chi_U \otimes \chi_V), \chi_W \right)_{G_{i+j+k}} = \left( \chi_U \otimes \chi_V, \sum_k \chi x_k \otimes \chi y_k \right)_{G_i \times G_j} = \sum_k \left( \chi_U, \chi x_k \right)_{G_i} \left( \chi V, \chi y_k \right)_{G_j}.
\]

(the third equality sign follows by taking dimensions in (4.18) and recalling (4.2); the fourth equality sign follows from (4.19); the sixth one follows from (4.2)) and

\[
(\chi_U \otimes \chi_V, \Delta (\chi_W))_{A \otimes A} = \left( \chi_U \otimes \chi_V, \text{res}_{G_i}^{G_{i+j+k}} \chi W \right)_{A \otimes A} = \left( \chi_U \otimes \chi_V, \sum_k \chi x_k \otimes \chi y_k \right)_{A \otimes A} = \sum_k \left( \chi_U, \chi x_k \right)_{G_i} \left( \chi V, \chi y_k \right)_{G_j}.
\]

(\text{the first equality sign follows by removing all terms in } \Delta (\chi_W) \text{ whose scalar product with } \chi_U \otimes \chi_V \text{ vanishes for reasons of gradedness; the second equality sign follows from (4.19)), which in comparison yield } (m (\chi_U \otimes \chi_V), \chi_W)_A = \left( \chi_U \otimes \chi_V, \Delta (\chi_W) \right)_{A \otimes A}, \text{ thus showing that } m \text{ and } \Delta \text{ are adjoint maps. Therefore, } m \text{ is associative (since } \Delta \text{ is coassociative).}

\]

Endowing \( A = \bigoplus_{q \geq 0} R G_n \) with the obvious unit and counit maps, it thus becomes a graded, finite-type \( Z \)-algebra and \( Z \)-coalgebra.

The next section addresses the issue of why they form a bialgebra. However, assuming this for the moment, it should be clear that each of these algebras \( A \) is a PSH having \( \Sigma = \bigcup_{n \geq 0} \text{Irr}(G_n) \) as its PSH-basis. \( \Sigma \) is self-dual because \( m, \Delta \) are defined by adjoint maps, and it is positive because \( m, \Delta \) take irreducible representations to genuine representations not just virtual ones, and hence have characters which are nonnegative sums of irreducible characters.

**Exercise 4.20.** Let \( i, j \) and \( k \) be three nonnegative integers. Let \( U \) be a \( \mathbb{C} \mathbb{S}_i \)-module, let \( V \) be a \( \mathbb{C} \mathbb{S}_j \)-module, and let \( W \) be a \( \mathbb{C} \mathbb{S}_k \)-module. Show that there are canonical \( \mathbb{C} [\mathbb{S}_i \times \mathbb{S}_j \times \mathbb{S}_k] \)-module isomorphisms

\[
\text{Ind}_{\mathbb{S}_{i+j} \times \mathbb{S}_k}^{\mathbb{S}_{i+j+k}} \left( \text{Ind}_{\mathbb{S}_i \times \mathbb{S}_j \times \mathbb{S}_k}^{\mathbb{S}_{i+j+k}} (U \otimes V) \otimes W \right) \cong \text{Ind}_{\mathbb{S}_i \times \mathbb{S}_j \times \mathbb{S}_k}^{\mathbb{S}_{i+j+k}} (U \otimes V \otimes W)
\]

\[
\cong \text{Ind}_{\mathbb{S}_i \times \mathbb{S}_j \times \mathbb{S}_k}^{\mathbb{S}_{i+j+k}} \left( U \otimes \text{Ind}_{\mathbb{S}_j \times \mathbb{S}_k}^{\mathbb{S}_{i+j+k}} (V \otimes W) \right).
\]

(Similar statements hold for the other two towers of groups and their respective ind functors, although the one for the \( GL_n \) tower is harder to prove. See Exercise 4.31(a) for a more general result.)

4.3. **Bialgebra and double cosets.** To show that the algebra and coalgebras \( A = A(G_n) \) are bialgebras, the central issue is checking the pentagonal diagram in (1.8), that is, as maps \( A \otimes A \to A \otimes A \), one has

\[
\Delta \circ m = (m \otimes m) \circ (\text{id} \otimes T \otimes \text{id}) \circ (\Delta \otimes \Delta).
\]

In checking this, it is convenient to have a lighter notation for various subgroups of the groups \( G_n \) corresponding to compositions \( \alpha \).

**Definition 4.21.**

(a) An **almost-composition** is a (finite) tuple \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \) of nonnegative integers. Its **length** is defined to be \( \ell \) and denoted by \( \ell(\alpha) \); its **size** is defined to be \( \alpha_1 + \alpha_2 + \cdots + \alpha_\ell \) and denoted by \( |\alpha| \); its **parts** are its entries \( \alpha_1, \alpha_2, \ldots, \alpha_\ell \). The almost-compositions of size \( n \) are called the **almost-compositions of \( n \)**.

(b) A **composition** is a finite tuple of positive integers. Of course, any composition is an almost-composition, and so all notions defined for almost-compositions (like size and length) make sense for compositions.

Note that any partition of \( n \) (written without trailing zeroes) is a composition of \( n \). We write \( \emptyset \) (and sometimes, sloppily, \( (0) \), when there is no danger of mistaking it for the almost-composition \( (0) \)) for the empty composition \( () \).
Definition 4.22. Given an almost-composition \( \alpha = (\alpha_1, \ldots, \alpha_\ell) \) of \( n \), define a subgroup
\[
G_\alpha \cong G_{\alpha_1} \times \cdots \times G_{\alpha_\ell} \leq G_n
\]
via the block-diagonal embedding with diagonal blocks of sizes \( (\alpha_1, \ldots, \alpha_\ell) \). This \( G_\alpha \) is called a Young subgroup when \( G_n = \mathfrak{S}_n \), and a Levi subgroup when \( G_n = GL_n \). In the case when \( G_n = \mathfrak{S}_n[\Gamma] \), we also denote \( G_\alpha \) by \( \mathfrak{S}_n[\alpha] \). In the case where \( G_n = GL_n \), also define the parabolic subgroup \( P_\alpha \) to be the subgroup of \( G_n \) consisting of block-upper triangular matrices whose diagonal blocks have sizes \( (\alpha_1, \ldots, \alpha_\ell) \), and let \( K_\alpha \) be the kernel of the obvious surjection \( P_\alpha \to G_\alpha \) which sends a block upper-triangular matrix to the tuple of its diagonal blocks whose sizes are \( \alpha_1, \alpha_2, \ldots, \alpha_\ell \). Notice that \( P_{(i,j)} \to P_{i,j} \) for any \( i \) and \( j \) with \( i + j \leq n \); similarly, \( K_{(i,j)} = K_{i,j} \) for any \( i \) and \( j \) with \( i + j = n \). We will also abbreviate \( G_{(i,j)} = G_i \times G_j \) by \( G_{i,j} \).

When \( (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \) is an almost-composition, we abbreviate \( G_{(\alpha_1, \alpha_2, \ldots, \alpha_\ell)} \) (and similarly for the \( P_\alpha \)).

Definition 4.23. Let \( K \) and \( H \) be two groups, \( \tau : K \to H \) a group homomorphism, and \( U \) a \( CH \)-module. Then, \( U^\tau \) is defined as the \( CK \)-module with ground space \( U \) and action given by \( k \cdot u = \tau(k) \cdot u \) for all \( k \in K \) and \( u \in U \). This very simple construction generalizes the definition of \( U^g \) for an element \( g \in G \), where \( G \) is a group containing \( H \) as a subgroup; in fact, in this situation we have \( U^g = U^{\tau(g)} \), where \( K = gH \) and \( \tau : K \to H \) is the map \( k \mapsto g^{-1}kg \).

Using homogeneity, checking the bialgebra condition (4.20) in the homogeneous component \( (A \otimes A)_n \) amounts to the following: for each pair of representations \( U_1, U_2 \) or \( G_{r_1}, G_{r_2} \) with \( r_1 + r_2 = n \), and for each \( (c_1, c_2) \) with \( c_1 + c_2 = n \), one must verify that
\[
\operatorname{res}_A^{c_1,c_2} \left( \operatorname{ind}_A^{r_1,r_2} (U_1 \otimes U_2) \right)
\]
\[
\cong \bigoplus_{A} \left( \operatorname{ind}_{a_1,a_2}^{c_1} \otimes \operatorname{ind}_{a_1,a_2}^{c_2} \right) \left( \left( \operatorname{res}_{a_1,a_2}^{r_1} U_1 \otimes \operatorname{res}_{a_1,a_2}^{r_2} U_2 \right)^{\tau_A^{-1}} \right)
\]
where the direct sum is over all matrices \( A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \) in \( \mathbb{H}^{2 \times 2} \) with row sums \( (r_1, r_2) \) and column sums \( (c_1, c_2) \), and where \( \tau_A \) is the obvious isomorphism between the subgroups
\[
G_{a_{11},a_{12},a_{21},a_{22}} \quad (< G_{r_1,r_2})
\]
\[
G_{a_{11},a_{12},a_{21},a_{22}} \quad (< G_{c_1,c_2})
\]
(we are using the inverse \( \tau_A^{-1} \) of this isomorphism \( \tau_A \) to identify modules for the first subgroup with modules for the second subgroup, according to Definition 4.23).

As one might guess, (4.21) comes from the Mackey formula (Theorem 4.7), once one identifies the appropriate double coset representatives. This is just as easy to do in a slightly more general setting.

Definition 4.24. Given almost-compositions \( \alpha, \beta \) of \( n \) having lengths \( \ell, m \) and a matrix \( A \) in \( \mathbb{N}^{\ell \times m} \) with row sums \( \alpha \) and column sums \( \beta \), define a permutation \( w_A \) in \( \mathfrak{S}_n \) as follows. Disjointly decompose \( [n] = \{1, 2, \ldots, n\} \) into consecutive intervals of numbers
\[
[n] = I_1 \sqcup \cdots \sqcup I_\ell
\]
\[
[n] = J_1 \sqcup \cdots \sqcup J_m
\]
such that \( |I_i| = \alpha_i, |J_j| = \beta_j \). For every \( j \in [m] \), disjointly decompose \( J_j \) into consecutive intervals of numbers \( J_j = J_{j,1} \sqcup J_{j,2} \sqcup \cdots \sqcup J_{j,\ell} \) such that every \( i \in [\ell] \) satisfies \( |J_{j,i}| = \alpha_{ij} \). For every \( i \in [\ell] \), disjointly decompose \( I_i \) into consecutive intervals of numbers \( I_i = I_{i,1} \sqcup I_{i,2} \sqcup \cdots \sqcup I_{i,m} \) such that every \( j \in [m] \) satisfies \( |I_{i,j}| = \alpha_{ij} \). Now, for every \( i \in [\ell] \) and \( j \in [m] \), let \( \pi_{i,j} \) be the increasing bijection from \( J_{j,i} \) to \( I_{i,j} \) (this is well-defined since these two sets both have cardinality \( \alpha_{ij} \)). The disjoint union of these bijections \( \pi_{i,j} \) over all \( i \) and \( j \) is a bijection \( [n] \to [n] \) (since the disjoint union of the sets \( J_{j,i} \) over all \( i \) and \( j \) is \( [n] \)), and so is the disjoint union of the sets \( I_{i,j} \), that is, a permutation of \( [n] \); this permutation is what we call \( w_A \).

Example 4.25. Taking \( n = 9 \) and \( \alpha = (4, 5), \beta = (3, 4, 2) \), one has
\[
I_1 = \{1, 2, 3, 4\}, \quad I_2 = \{5, 6, 7, 8, 9\}
\]
\[
J_1 = \{1, 2, 3\}, \quad J_2 = \{4, 5, 6, 7\}, \quad J_3 = \{8, 9\}.
\]

\[^{157}\text{We have already met this CK-module } U^\tau \text{ in Remark 4.13, where it was called Res}_U U.\]
Then one possible matrix $A$ having row and column sums $\alpha, \beta$ is $A = \begin{bmatrix} 2 & 2 & 0 \\ 1 & 2 & 2 \end{bmatrix}$, and its associated permutation $w_A$ written in two-line notation is

$$\begin{array}{c|c|c|c|c|c|c|c|c|}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
\hline
1 & 2 & 5 & 3 & 4 & 6 & 7 & 8 & 9 \\
\end{array}$$

with vertical lines dividing the sets $J_j$ on top, and with elements of $I_i$ underlined $i$ times on the bottom.

**Remark 4.26.** Given almost-compositions $\alpha$ and $\beta$ of $n$ having lengths $\ell$ and $m$, and a permutation $w \in S_n$. It is easy to see that there exists a matrix $A \in \mathbb{N}^{\ell \times m}$ satisfying $w_A = w$ if and only if the restriction of $w$ to each $J_j$ and the restriction of $w^{-1}$ to each $I_i$ are increasing. In this case, the matrix $A$ is determined by $a_{ij} = |w(J_j) \cap I_i|$. In particular, almost-compositions are easily seen to be constant on double cosets.

Among our three towers $G_*$ of groups, the symmetric group tower $(G_n = S_n)$ is the simplest one. We will now see that it also embeds into the two others, in the sense that $S_n$ embeds into $S_n[\Gamma]$ for every $\Gamma$ and into $GL_n(\mathbb{F}_q)$ for every $q$.

First, for every $n \in \mathbb{N}$ and any group $\Gamma$, we embed the group $S_n$ into $S_n[\Gamma]$ by means of the canonical embedding $S_n \to S_n \times \Gamma^n = S_n[\Gamma]$. If we regard elements of $S_n[\Gamma]$ as $n \times n$ monomial matrices with nonzero entries in $\Gamma$, then this boils down to identifying every $\pi \in S_n$ with the permutation matrix of $\pi$ (in which the 1’s are read as the neutral element of $\Gamma$). If $\alpha$ is an almost-composition of $n$, then this embedding $S_n \to S_n[\Gamma]$ makes the subgroup $S_\alpha$ of $S_n$ become a subgroup of $S_n[\Gamma]$, more precisely a subgroup of $S_\alpha[\Gamma] < S_n[\Gamma]$.

For every $n \in \mathbb{N}$ and every $q$, we embed the group $S_n$ into $GL_n(\mathbb{F}_q)$ by identifying every permutation $\pi \in S_n$ with its coset in $GL_n(\mathbb{F}_q)$. If $\alpha$ is an almost-composition of $n$, then this embedding makes the subgroup $S_\alpha$ of $S_n$ become a subgroup of $GL_n(\mathbb{F}_q)$. If we let $G_n = GL_n(\mathbb{F}_q)$, then $S_\alpha < G_\alpha < P_\alpha$.

The embeddings we just defined commute with the group embeddings $G_n < G_{n+1}$ on both sides.

**Proposition 4.27.** The permutations $\{w_A\}$ as $A$ runs over all matrices in $\mathbb{N}^{\ell \times m}$ having row, column sums $\alpha, \beta$ give a system of double coset representatives for

$$S_\alpha \setminus S_n / S_\beta$$

$$S_\alpha[\Gamma] \setminus S_n[\Gamma] / S_\beta[\Gamma]$$

$$P_\alpha \setminus GL_n / P_\beta$$

**Proof.** First note that double coset representatives for $S_\alpha \setminus S_n / S_\beta$ should also provide double coset representatives for $S_\alpha[\Gamma] \setminus S_n[\Gamma] / S_\beta[\Gamma]$, since

$$S_\alpha[\Gamma] = S_\alpha \Gamma^n = \Gamma^n S_\alpha.$$ 

We give an algorithm to show that every double coset $S_n w S_\beta$ contains some $w_A$. Start by altering $w$ within its coset $w S_\beta$, that is, by permuting the *positions* within each set $J_j$, to obtain a representative $w'$ for $w S_\beta$ in which each set $w'(J_j)$ appears in increasing order in the second line of the two-line notation for $w'$. Then alter $w'$ within its coset $S_n w'$, that is, by permuting the *values* within each set $I_i$, to obtain a representative $w_A$ having the elements of each set $I_i$ appearing in increasing order in the second line; because the values within each set $I_i$ are consecutive, this alteration will not ruin the property that one had each set $w'(J_j)$ appearing in increasing order. For example, one might have

$$w = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 8 & 2 & 5 & 3 & 9 & 1 & 7 & 6 \end{bmatrix} \in S_\beta$$

$$w' = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 4 & 8 & 1 & 3 & 5 & 9 & 6 & 7 \end{bmatrix} \in w S_\beta$$

$$w_A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 2 & 5 & 3 & 4 & 6 & 7 & 8 & 9 \end{bmatrix} \in S_\alpha w' \subset S_\alpha w' S_\beta = S_\alpha w S_\beta$$

Next note that $S_\alpha w_A S_\beta = S_\alpha w S_\beta$ implies $A = B$, since the quantities

$$a_{ij}(w) := |w(J_j) \cap I_i|$$

are easily seen to be constant on double cosets $S_\alpha w S_\beta$. 
A similar argument shows that \( P_A w_A P\beta = P_A w_B P\beta \) implies \( A = B \): for \( g \) in \( GL_n \), the rank \( r_{ij}(g) \) of the matrix obtained by restricting \( g \) to rows \( I_i \sqcup I_{i+1} \sqcup \cdots \sqcup I_k \) and columns \( J_1 \sqcup J_2 \sqcup \cdots \sqcup J_j \) is constant on double cosets \( P_A g P\beta \), and for a permutation matrix \( w \) one can recover \( a_{i,j}(w) \) from the formula

\[
a_{i,j}(w) = r_{i,j}(w) - r_{i,j-1}(w) - r_{i+1,j}(w) + r_{i+1,j-1}(w).
\]

Thus it only remains to show that every double coset \( P_A g P\beta \) contains some \( w_A \). Since \( \mathcal{S}_n < P_n \), and we have already that every double coset \( \mathcal{S}_n w \mathcal{S}_n \) contains some \( w_A \), it suffices to show that every double coset \( P_A g P\beta \) contains some permutation \( w \). However, we claim that this is already true for the smaller double cosets \( B g B \) where \( B = P_1 \) is the Borel subgroup of upper triangular invertible matrices, that is, one has the usual Bruhat decomposition

\[
GL_n = \bigsqcup_{w \in \mathcal{S}_n} B w B.
\]

To prove this decomposition, we show how to find a permutation \( w \) in each double coset \( B g B \). The freedom to alter \( g \) within its coset \( g B \) allows one to scale columns and add scalar multiples of earlier columns to later columns. We claim that using such column operations, one can always find a representative \( g' \) for coset \( g B \) in which

- the bottommost nonzero entry of each column is 1 (call this entry a pivot),
- the entries to right of each pivot within its row are all 0, and
- there is one pivot in each row and each column, so that their positions are the positions of the 1’s in some permutation matrix \( w \).

In fact, we will see below that \( B g B = B w B \) in this case. The algorithm which produces \( g' \) from \( g \) is simple: starting with the leftmost column, find its bottommost nonzero entry, and scale the column to make this entry a 1, creating the pivot in this column. Now use this pivot to clear out all entries in its row to its right, using column operations that subtract multiples of this column from later columns. Having done this, move on to the next column to the right, and repeat, scaling to create a pivot, and using it to eliminate entries to its right.\(^{158}\)

For example, the typical matrix \( g \) lying in the double coset \( B g B \) where

\[
w = \begin{pmatrix} 1 & 2 & 3 & | & 4 & 5 & 6 & 7 & | & 8 & 9 \\ 4 & 8 & 2 & | & 5 & 3 & 9 & 1 & | & 7 & 6 \end{pmatrix}
\]

from before is one that can be altered within its coset \( g B \) to look like this:

\[
g' = \begin{bmatrix}
* & * & * & * & * & 1 & 0 & 0 \\
* & * & 1 & 0 & 0 & 0 & 0 & 0 \\
* & * & 0 & * & 1 & 0 & 0 & 0 \\
0 & * & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & * & 0 & 0 & 0 & * & 0 & 1 \\
0 & * & 0 & 0 & 0 & * & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0
\end{bmatrix} \in g B.
\]

Having found this \( g' \) in \( g B \), a similar algorithm using left multiplication by \( B \) shows that \( w \) lies in \( B g' B = B g B \). This time no scalings are required to create the pivot entries: starting with the bottom row, one uses its pivot to eliminate all the entries above it in the same column (shown by stars * above) by adding

---

\(^{158}\)To see that this works, we need to check three facts:

(a) We will find a nonzero entry in every column during our algorithm.

(b) Our column operations preserve the zeroes lying to the right of already existing pivots.

(c) Every row contains exactly one pivot at the end of the algorithm.

But fact (a) simply says that our matrix can never have an all-zero column during the algorithm; this is clear (since the rank of the matrix remains constant during the algorithm and was \( n \) at its beginning). Fact (b) holds because all our operations either scale columns (which clearly preserves zero entries) or subtract a multiple of the column \( c \) containing the current pivot from a later column \( d \) (which will preserve every zero lying to the right of an already existing pivot, because any already existing pivot must lie in a column \( b < c \) and therefore both columns \( c \) and \( d \) have zeroes in its row). Fact (c) follows from noticing that there are \( n \) pivots altogether at the end of the algorithm, but no row can contain two of them (since the entries to the right of a pivot in its row are 0).
multiples of the bottom row to higher rows. Then do the same using the pivot in the next-to-bottom row, etc. The result is the permutation matrix for $w$.

Remark 4.28. The Bruhat decomposition $GL_n = \bigcup_{w \in S_n} BwB$ is related to the so-called LPU factorization—one of a myriad of matrix factorizations appearing in linear algebra. It is actually a fairly general phenomenon, and requires neither the finiteness of $F$, nor the invertibility, nor even the squareness of the matrices (see Exercise 4.29(b) for an analogue holding in a more general setup).

Exercise 4.29. Let $F$ be any field.

(a) For any $n \in \mathbb{N}$ and any $A \in GL_n(F)$, prove that there exist a lower-triangular matrix $L \in GL_n(F)$, an upper-triangular matrix $U \in GL_n(F)$ and a permutation matrix $P \in S_n \subset GL_n(F)$ (here, we identify permutations with the corresponding permutation matrices) such that $A = LPU$.

(b) Let $n, m \in \mathbb{N}$. Let $F_{n,m}$ denote the set of all $n \times m$-matrices $B \in \{0, 1\}^{n \times m}$ such that each row of $B$ contains at most one 1 and each column of $B$ contains at most one 1. We regard $F_{n,m}$ as a subset of $F^{n \times m}$ by means of regarding $(0, 1)$ as a subset of $F$.

For every $k \in \mathbb{N}$, we let $B_k$ denote the subgroup of $GL_k(F)$ consisting of all upper-triangular matrices.

Corollary 4.30. For each of the three towers of groups $G_*$, the product and coproduct structures on $A = A(G_*)$ endow it with a bialgebra structure, and hence they form PSH's.

Proof. The first two towers $G_n = S_n$ and $G_n = S_n[1]$ have product, coproduct defined by induction, restriction along embeddings $G_i \times G_j < G_{i+j}$. Hence the desired bialgebra equality (4.21) follows from Mackey's Theorem 4.7, taking $G = G_n, H = G(r_1, r_2), K = G(c_1, c_2), U = U_1 \otimes U_2$ with double coset representatives \[ \{g_1, \ldots, g_t\} = \{w_{A'} : A \in \mathbb{N}^{2 \times 2}, A \text{ has row sums } (r_1, r_2) \text{ and column sums } (c_1, c_2)\} \]

and checking for a given double coset \[ KgH = (G(c_1, c_2)w_{A'})(G(r_1, r_2)) \]

indexed by a matrix $A$ in $\mathbb{N}^{2 \times 2}$ with row sums $(r_1, r_2)$ and column sums $(c_1, c_2)$, that the two subgroups appearing on the left in (4.22) are exactly \[ H \cap K^w_{A'} = G_{r_1, r_2} \cap (G_{c_1, c_2})^{w_{A'}}, \]

\[ w_{A'}H \cap K = w_{A'}(G_{r_1, r_2}) \cap G_{c_1, c_2}, \]

respectively. One should also apply (4.6) and check that the isomorphism $\tau_A$ between the two subgroups in (4.22) is the conjugation isomorphism by $w_{A'}$ (that is, $\tau_A(g) = w_{A'}gw_{A'}^{-1}$ for every $g \in H \cap K^w_{A'}$). We leave all of these bookkeeping details to the reader to check. \[ ^{161} \]

\[ ^{159} \]Specifically, an LPU factorization of a matrix $A \in GL_n(F)$ (for an arbitrary field $F$) means a way to write $A$ as a product $A = LPU$ with $L \in GL_n(F)$ being lower-triangular, $U \in GL_n(F)$ being upper-triangular, and $P \in S_n \subset GL_n(F)$ being a permutation matrix. Such a factorization always exists (although it is generally not unique). This can be derived from the Bruhat decomposition (see Exercise 4.29(a) for a proof). See also [162] for related discussion.

\[ ^{160} \]Proposition 4.27 gives as a system of double coset representatives for $G_{c_1, c_2}\setminus G_n/(G_{r_1, r_2})$ the elements \[ \{w_A : A \in \mathbb{N}^{2 \times 2}, A \text{ has row sums } (c_1, c_2) \text{ and column sums } (r_1, r_2)\} \]

\[ = \{w_{A'} : A \in \mathbb{N}^{2 \times 2}, A \text{ has row sums } (r_1, r_2) \text{ and column sums } (c_1, c_2)\} \]

where $A'$ denotes the transpose matrix of $A$.

\[ ^{161} \]It helps to recognize $w_{A'}$ as the permutation written in two-line notation as

\[
\begin{pmatrix}
1 & 2 & \ldots & a_{11} & a_{11} + 1 & a_{11} + 2 & \ldots & r_1 & r_1 + 1 & r_1 + 2 & \ldots & a'_{22} & a'_{22} + 1 & a'_{22} + 2 & \ldots & n \\
1 & 2 & \ldots & a_{11} & c_1 + 1 & c_1 + 2 & \ldots & a'_{22} & a_{11} + 1 & a_{11} + 2 & \ldots & c_1 & c_2 & c_3 & \ldots & n \\
\end{pmatrix},
\]

where $a'_{22} = r_1 + a_{21} = c_1 + a_{12} = n - a_{22}$. In matrix form, $w_{A'}$ is the block matrix

\[
\begin{pmatrix}
I_{a_{11}} & 0 & 0 & 0 \\
0 & 0 & I_{a_{21}} & 0 \\
0 & I_{a_{12}} & 0 & 0 \\
0 & 0 & 0 & I_{a_{22}} \\
\end{pmatrix}.
\]
For the tower with $G_n = GL_n$, there is slightly more work to be done to check the equality (4.21). Via Mackey’s Theorem 4.7 and Proposition 4.27, the left side is

\[ \text{res}^n_{c_1,c_2} \left( \text{ind}^n_{r_1,r_2} (U_1 \otimes U_2) \right) = \left( \text{Res}^{G_n}_{P_{r_1,r_2}} \text{Ind}^{G_n}_{P_{r_1,r_2}} \text{Infl}^{P_{r_1,r_2}}_{G_{r_1,r_2}} (U_1 \otimes U_2) \right)^{K_{c_1,c_2}} \]

\[ = \bigoplus_A \left( \text{Ind}^P_{A} \cap P_{r_1,r_2} \cap P_{r_1,c_2} \left( \left( \text{Res}^{P_{r_1,r_2}}_{P_{r_1,r_2} \cap P_{r_1,c_2}} \text{Infl}^{P_{r_1,r_2}}_{G_{r_1,r_2}} (U_1 \otimes U_2) \right)^{\tau_A^{-1}} \right)^{K_{c_1,c_2}} \right) \]

where $A$ runs over the usual $2 \times 2$ matrices. The right side is a direct sum over this same set of matrices $A$:

\[ \bigoplus_A \left( \text{Ind}^g_{a_1,a_2} \otimes \text{Ind}^g_{a_2,a_2} \left( \text{res}^g_{r_1,r_2} (U_1 \otimes \text{res}^g_{a_2,a_2} U_2) \right)^{\tau_A^{-1}} \right) \]

\[ = \bigoplus_A \left( \text{Ind}^G_{P_{a_1,a_2}} \otimes \text{Ind}^G_{P_{a_2,a_2}} \circ \left( \text{Infl}^{P_{a_1,a_2}}_{P_{a_1,a_2}} \otimes \text{Infl}^{P_{a_2,a_2}}_{P_{a_2,a_2}} \right) \left( \left( \text{Res}^{G_{a_1,a_2}}_{P_{a_1,a_2}} (U_1) \otimes \text{Res}^{G_{a_2,a_2}}_{P_{a_2,a_2}} (U_2) \right)^{K_{a_1,a_2} \otimes K_{a_2,a_2}} \right)^{\tau_A^{-1}} \right) \]

\[ = \bigoplus_A \text{Ind}^G_{P_{a_1,a_2} \times P_{a_2,a_2}} \left( \left( \text{Res}^{G_{a_1,a_2}}_{P_{a_1,a_2} \times P_{a_2,a_2}} (U_1 \otimes U_2) \right)^{K_{a_1,a_2} \otimes K_{a_2,a_2}} \right)^{\tau_A^{-1}} \]

(by (4.6), (4.15) and their obvious analogues for restriction and inflation). Thus it suffices to check for each $2 \times 2$ matrix $A$ that any $CG_{c_1,c_2}$-module of the form $V_1 \otimes V_2$ has the same inner product with the $A$-summands of (4.23) and (4.24). Abbreviate $w := w_A$ and $\tau := \tau_A^{-1}$.

Notice that $wP_{r_1,r_2}$ is the group of all matrices having the block form

\[ \begin{bmatrix} g_{11} & h & i & j \\ 0 & g_{21} & 0 & k \\ d & e & g_{12} & \ell \\ 0 & f & 0 & g_{22} \end{bmatrix} \]

in which the diagonal blocks $g_{ij}$ for $i, j = 1, 2$ are invertible of size $a_{ij} \times a_{ij}$, while the blocks $h, i, j, k, \ell, d, e, f$ are all arbitrary matrices\(^\text{162}\) of the appropriate (rectangular) block sizes. Hence, $wP_{r_1,r_2} \cap P_{c_1,c_2}$ is the group of all matrices having the block form

\[ \begin{bmatrix} g_{11} & h & i & j \\ 0 & g_{21} & 0 & k \\ 0 & 0 & g_{12} & \ell \\ 0 & 0 & 0 & g_{22} \end{bmatrix} \]

in which the diagonal blocks $g_{ij}$ for $i, j = 1, 2$ are invertible of size $a_{ij} \times a_{ij}$, while the blocks $h, i, j, k, \ell$ are all arbitrary matrices of the appropriate (rectangular) block sizes; then $wP_{r_1,r_2} \cap G_{c_1,c_2}$ is the subgroup where the blocks $i, j, k$ all vanish. The canonical projection $wP_{r_1,r_2} \cap P_{c_1,c_2} \rightarrow wP_{r_1,r_2} \cap G_{c_1,c_2}$ (obtained by restricting the projection $P_{c_1,c_2} \rightarrow G_{c_1,c_2}$) has kernel $wP_{r_1,r_2} \cap P_{c_1,c_2} \cap K_{c_1,c_2}$. Consequently,

\[ \left( wP_{r_1,r_2} \cap P_{c_1,c_2} \right) / \left( wP_{r_1,r_2} \cap P_{c_1,c_2} \cap K_{c_1,c_2} \right) = wP_{r_1,r_2} \cap G_{c_1,c_2}. \]

Similarly,

\[ \left( P_{r_1,r_2} \cap P_{c_1,c_2} \right) / \left( P_{r_1,r_2} \cap P_{c_1,c_2} \cap K_{c_1,c_2} \right) = G_{r_1,r_2} \cap P_{c_1,c_2}. \]

\(^{162}\)The blocks $i$ and $j$ have nothing to do with the indices $i, j$ in $g_{ij}$. 


Computing first the inner product of $V_1 \otimes V_2$ with the $A$-summand of (4.23), and using adjointness properties, one gets

$$
\left( \left( \Res^{P_{r_1},r_2}_{P_{c_1},c_2} \Inf^{P_{r_1},c_2}_{G_{r_1},c_2} (U_1 \otimes U_2) \right)^\tau, \\
\Res^{P_{r_1},r_2}_{P_{c_1},c_2} \Inf^{P_{r_1},c_2}_{G_{r_1},c_2} (V_1 \otimes V_2) \right)_{w_{P_{r_1},r_2} \cap P_{c_1},c_2}
$$

(4.10)

(by (4.28) and (4.27)). One can compute this inner product by first recalling that $w_{P_{r_1},r_2} \cap P_{c_1},c_2$ is the group of matrices having the block form (4.26) in which the diagonal blocks $g_{ij}$ for $i, j = 1, 2$ are invertible of size $a_{ij} \times a_{ij}$, while the blocks $h, i, j, k, \ell$ are all arbitrary matrices of the appropriate (rectangular) block sizes; then $w_{P_{r_1},r_2} \cap G_{c_1},c_2$ is the subgroup where the blocks $i, j, k$ all vanish. The inner product above then becomes

$$
\frac{1}{w_{P_{r_1},r_2} \cap P_{c_1},c_2} \sum_{(g_{ij}) \in (h,i,j,k, \ell)} \chi_U \left( \begin{array}{cc} g_{11} & i \\ 0 & g_{12} \end{array} \right) \chi_U \left( \begin{array}{cc} g_{21} & k \\ 0 & g_{22} \end{array} \right) \\
\chi_V \left( \begin{array}{cc} g_{11} & h \\ 0 & g_{21} \end{array} \right) \chi_V \left( \begin{array}{cc} g_{12} & \ell \\ 0 & g_{22} \end{array} \right).
$$

(4.29)

If one instead computes the inner product of $V_1 \otimes V_2$ with the $A$-summand of (4.24), using adjointness properties and (4.13) one gets

$$
\left( \left( \Res^{G_{r_1},c_2}_{P_{a_{11},a_{12}} \times P_{a_{21},a_{22}}} (U_1 \otimes U_2) \right)^{K_{a_{11},a_{12}} \times K_{a_{21},a_{22}}}, \\
\Res^{G_{r_1},c_2}_{P_{a_{11},a_{12}} \times P_{a_{21},a_{22}}} (V_1 \otimes V_2) \right)^{K_{a_{11},a_{12}} \times K_{a_{12},a_{22}}}
$$

$$
= \frac{1}{[G_{a_{11},a_{21},a_{12},a_{22}]}} \sum_{(g_{ij})} \frac{1}{[K_{a_{11},a_{12}} \times K_{a_{21},a_{22}}]} \sum_{(i,k)} \chi_U \left( \begin{array}{cc} g_{11} & i \\ 0 & g_{12} \end{array} \right) \chi_U \left( \begin{array}{cc} g_{21} & k \\ 0 & g_{22} \end{array} \right) \\
\frac{1}{[K_{a_{11},a_{21}} \times K_{a_{12},a_{22}}]} \sum_{(h, \ell)} \chi_V \left( \begin{array}{cc} g_{11} & h \\ 0 & g_{21} \end{array} \right) \chi_V \left( \begin{array}{cc} g_{12} & \ell \\ 0 & g_{22} \end{array} \right).
$$

But this right hand side can be seen to equal (4.29), after one notes that

$$
|w_{P_{r_1},r_2} \cap P_{c_1},c_2| = |G_{a_{11},a_{21},a_{12},a_{22}}| \cdot |K_{a_{11},a_{12}} \times K_{a_{21},a_{22}}| \cdot |K_{a_{11},a_{21}} \times K_{a_{12},a_{22}}| \cdot \# \{ j \in \mathbb{F}_q^{a_{11} \times a_{22}} \}
$$

and that the summands in (4.29) are independent of the matrix $j$ in the summation. \qed

We can also define a $\mathbb{C}$-vector space $A_{C}$ as the direct sum $\bigoplus_{n \geq 0} R_{C}(G_n)$. In the same way as we have made $A = \bigoplus_{n \geq 0} R(G_n)$ into a $\mathbb{Z}$-bialgebra, we can turn $A_{C} = \bigoplus_{n \geq 0} R_{C}(G_n)$ into a $\mathbb{C}$-bialgebra.\footnote{The definitions of $m$ and $\Delta$ for this $\mathbb{C}$-bialgebra look the same as for $A$: For instance, $m$ is still defined to be $\text{ind}_{i,j}^{i+j}$ on $(A_C)_i \otimes (A_C)_j$, where $\text{ind}_{i,j}^{i+j}$ is defined by the same formulas as in Definition 4.18. However, the operators of induction, restriction, inflation and $K$-fixed space construction appearing in these formulas now act on class functions as opposed to modules. The fact that these maps $m$ and $\Delta$ satisfy the axioms of a $\mathbb{C}$-bialgebra is easy to check: they are merely the $\mathbb{C}$-linear extensions of the maps $m$ and $\Delta$ of the $\mathbb{Z}$-bialgebra $A$ (this is because, for instance, induction of class functions and induction of modules are related by the identity (4.5)), and thus satisfy the same axioms as the latter.} There is a $\mathbb{C}$-bilinear form $(\cdot, \cdot)_C$ on $A_{C}$ which can be defined either as the $\mathbb{C}$-bilinear extension of the $\mathbb{Z}$-bilinear form $(\cdot, \cdot)_A : A \times A \to \mathbb{Z}$ to $A_{C}$, or (equivalently) as the $\mathbb{C}$-bilinear form on $A_{C}$ which restricts to $(\cdot, \cdot)_{S_n}$ on every homogeneous component $R_{C}(G_n)$ and makes different homogeneous components mutually orthogonal. The obvious embedding of $A$ into the $\mathbb{C}$-bialgebra $A_{C}$ (obtained from the embeddings $R(G_n) \to R_{C}(G_n)$ for...}
all \( n \) respects the bialgebra operations\(^{164}\), and the \( \mathbb{C} \)-bialgebra \( A_\mathbb{C} \) can be identified with \( A \otimes_{\mathbb{Z}} \mathbb{C} \) (the result of extending scalars to \( \mathbb{C} \) in \( A \)), because every finite group \( G \) satisfies \( R_G(\mathbb{C}) \cong R(\mathbb{C}) \otimes_{\mathbb{Z}} \mathbb{C} \). The embedding of \( A \) into \( A_\mathbb{C} \) also respects the bilinear forms.

**Exercise 4.31.** Let \( G_* \) be one of the three towers.

For every almost-composition \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \) of \( n \in \mathbb{N} \), let us define a map \( \text{ind}_n^a \) which takes \( \mathbb{C} G_\alpha \)-modules to \( \mathbb{C} G_n \)-modules as follows: If \( G_* = \mathcal{S}_n \) or \( G_* = \mathcal{S}_n [\Gamma] \), we set

\[
\text{ind}_n^a := \text{Ind}^{G_n}_{G_\alpha}.
\]

If \( G_* = GL_n \), then we set

\[
\text{ind}_n^a := \text{Ind}^*_{P_\alpha} \text{Infl}^{P_\alpha}_{G_\alpha}.
\]

(Note that \( \text{ind}_n^a = \text{ind}^*_{ij} \) if \( \alpha \) has the form \((i, j)\).)

Similarly, for every almost-composition \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \) of \( n \in \mathbb{N} \), let us define a map \( \text{res}_n^a \) which takes \( \mathbb{C} G_n \)-modules to \( \mathbb{C} G_\alpha \)-modules as follows: If \( G_* = \mathcal{S}_n \) or \( G_* = \mathcal{S}_n [\Gamma] \), we set

\[
\text{res}_n^a := \text{Res}^{G_\alpha}_{G_n}.
\]

If \( G_* = GL_n \), then we set

\[
\text{res}_n^a := (\text{Res}^{G_n}_{P_\alpha} (-))^{K_\alpha}.
\]

(Note that \( \text{res}_n^a = \text{res}^*_{ij} \) if \( \alpha \) has the form \((i, j)\).)

(a) If \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \) is an almost-composition of an integer \( n \in \mathbb{N} \) satisfying \( \ell \geq 1 \), and if \( V_i \) is a \( \mathbb{C} G_{\alpha_i} \)-module for every \( i \in \{1, 2, \ldots, \ell\} \), then show that

\[
\text{ind}_n^a(\alpha_1 + \alpha_2 + \cdots + \alpha_{\ell-1}, \alpha_\ell) = \text{ind}^n_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})} \left( V_1 \otimes V_2 \otimes \cdots \otimes V_{\ell-1} \otimes V_\ell \right)
\]

\[
\cong \text{ind}_n^a(V_1 \otimes V_2 \otimes \cdots \otimes V_\ell)
\]

\[
\cong \text{ind}_n^{\alpha_1, \alpha_2 + \cdots + \alpha_\ell}(V_1 \otimes \text{ind}_n^{\alpha_2 + \cdots + \alpha_\ell}(V_2 \otimes V_3 \otimes \cdots \otimes V_\ell))
\]

(b) Solve Exercise 4.20 again using Exercise 4.31(a).

(c) We proved above that the map \( m : A \otimes A \to A \) (where \( A = A(G_*) \)) is associative, by using the adjointness of \( m \) and \( \Delta \). Give a new proof of this fact, which makes no use of \( \Delta \).

(d) If \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \) is an almost-composition of an \( n \in \mathbb{N} \), and if \( \chi_i \in R(G_{\alpha_i}) \) for every \( i \in \{1, 2, \ldots, \ell\} \), then show that

\[
\chi_1 \chi_2 \cdots \chi_\ell = \text{ind}_n^a(\chi_1 \otimes \chi_2 \otimes \cdots \otimes \chi_\ell)
\]

in \( A = A(G_*) \).

(e) If \( n \in \mathbb{N}, \ell \in \mathbb{N} \) and \( \chi \in R(G_n) \), then show that

\[
\Delta^{(\ell-1)}(\chi) = \sum \text{res}_n^a(\chi)
\]

in \( A^{\otimes \ell} \), where \( A = A(G_*) \). Here, the sum on the right hand side runs over all almost-compositions \( \alpha \) of \( n \) having length \( \ell \).

### 4.4 Symmetric groups

Finally, some payoff. Consider the tower of symmetric groups \( G_n = \mathcal{S}_n \), and \( A = A(G_*) =: A(\mathcal{S}) \). Denote by \( 1_{\mathcal{S}_n}, \text{sgn}\mathcal{S}_n \) the *trivial* and *sign* characters on \( \mathcal{S}_n \). For a partition \( \lambda \) of \( n \), denote by \( 1_{\mathcal{S}_\lambda}, \text{sgn}\mathcal{S}_\lambda \) the trivial and sign characters restricted to the Young subgroup \( \mathcal{S}_\lambda = \mathcal{S}_{\lambda_1} \times \mathcal{S}_{\lambda_2} \times \cdots \), and denote by \( 1_\lambda \) the class function which is the characteristic function for the \( \mathcal{S}_n \)-conjugacy class of permutations of cycle type \( \lambda \).

**Theorem 4.32.** (a) *Irreducible complex characters* \( \{\chi^\lambda\} \) of \( \mathcal{S}_n \) are indexed by partitions \( \lambda \) in \( \text{Par}_n \), and one has a PSH-isomorphism, the *Frobenius characteristic map*\(^{165}\),

\[
A = A(\mathcal{S}) \xrightarrow{\text{ch}} \Lambda
\]

---

\(^{164}\) This is because, for example, induction of class functions harmonizes with induction of modules (i.e., the equality (4.5) holds).

\(^{165}\) It is unrelated to the Frobenius endomorphisms from Exercise 2.85.
that for $n \geq 0$ and $\lambda \in \text{Par}_n$ sends

\[
\begin{align*}
1_{\mathfrak{S}_n} & \mapsto h_n \\
\text{sgn}_{\mathfrak{S}_n} & \mapsto e_n \\
\chi & \mapsto s_\lambda \\
\text{Ind}_{\mathfrak{S}_\lambda \times \mathfrak{S}_\lambda}^{\mathfrak{S}_n} & \mapsto h_\lambda \\
\text{Ind}_{\mathfrak{S}_\lambda \times \mathfrak{S}_\lambda}^{\mathfrak{S}_n} \text{sgn}_{\mathfrak{S}_\lambda} & \mapsto e_\lambda \\
1_\lambda & \mapsto \frac{p_\lambda}{z_\lambda}
\end{align*}
\]

(where $\text{ch}$ is extended to a $\mathbb{C}$-linear map $AC \to \Lambda_C$) and for $n \geq 1$ sends

\[
1_{(n)} \mapsto \frac{p_n}{n}
\]

Here, $z_\lambda$ is defined as in Proposition 2.43.

(b) For each $n \geq 0$, the involution on class functions $f : \mathfrak{S}_n \to \mathbb{C}$ sending $f \mapsto \text{sgn}_{\mathfrak{S}_n} * f$ where

\[
(\text{sgn}_{\mathfrak{S}_n} * f)(g) := \text{sgn}(g)f(g)
\]

preserves the $\mathbb{Z}$-sublattice $R(\mathfrak{S}_n)$ of genuine characters. The direct sum of these involutions induces an involution on $A = A(\mathfrak{S}) = \bigoplus_{n \geq 0} R(\mathfrak{S}_n)$ that corresponds under $\text{ch}$ to the involution $\omega$ on $\Lambda$.

Proof. (a) Corollary 4.30 implies that the set $\Sigma = \bigsqcup_{n \geq 0} \text{Irr}(\mathfrak{S}_n)$ gives a PSH-basis for $A$. Since a character $\chi$ of $\mathfrak{S}_n$ has

\[
(4.30) \quad \Delta(\chi) = \bigoplus_{i+j=n} \text{Res}_{\mathfrak{S}_i \times \mathfrak{S}_j} \chi,
\]

such an element $\chi \in \Sigma \cap A_n$ is never primitive for $n \geq 2$. Hence the unique irreducible character $\rho = 1_{\mathfrak{S}_1}$ of $\mathfrak{S}_1$ is the only element of $\mathcal{C} = \Sigma \cap p$.

Thus Theorem 3.30(g) tells us that there are two PSH-isomorphisms $A \to \Lambda$, each of which sends $\Sigma$ to the PSH-basis of Schur functions $\{s_\lambda\}$ for $\Lambda$. It also tells us that we can pin down one of the two isomorphisms to call $\text{ch}$, by insisting that it map the two characters $1_{\mathfrak{S}_2}, \text{sgn}_{\mathfrak{S}_2}$ in $\text{Irr}(\mathfrak{S}_2)$ to $h_2, e_2$ (and not $e_2, h_2$).

Bearing in mind the coproduct formula (4.30), and the fact that $1_{\mathfrak{S}_n}, \text{sgn}_{\mathfrak{S}_n}$ restrict, respectively, to trivial and sign characters of $\mathfrak{S}_i \times \mathfrak{S}_j$ for $i + j = n$, one finds that for $n \geq 2$ one has $\text{sgn}_{\mathfrak{S}_n}$ annihilating $1_{\mathfrak{S}_n}$, and $1_{\mathfrak{S}_2}$ annihilating $\text{sgn}_{\mathfrak{S}_n}$. Therefore Theorem 3.18(b) (applied to $\Lambda$) implies $1_{\mathfrak{S}_n}, \text{sgn}_{\mathfrak{S}_n}$ are sent under $\text{ch}$ to $h_n, e_n$. Then the fact that $\text{Ind}_{\mathfrak{S}_\lambda \times \mathfrak{S}_\lambda}^{\mathfrak{S}_n} 1_{\mathfrak{S}_n}$, $\text{Ind}_{\mathfrak{S}_\lambda \times \mathfrak{S}_\lambda}^{\mathfrak{S}_n} \text{sgn}_{\mathfrak{S}_\lambda}$ are sent to $h_\lambda, e_\lambda$ follows via induction products.

Recall that the $\mathbb{C}$-vector space $AC = \bigoplus_{n \geq 0} R_C(\mathfrak{S}_n)$ is a $\mathbb{C}$-bialgebra, and can be identified with $A \otimes_{\mathbb{Z}} \mathbb{C}$. The multiplication and the comultiplication of $AC$ are $\mathbb{C}$-linear extensions of those of $A$, and are still given by the same formulas $m = \text{ind}_{i+j}^{i+j}$ and $\Delta = \bigoplus_{i+j=n} \text{res}_{i+j}^{i+j}$ as those of $A$ (but now, induction and restriction are defined for class functions, not just for representations). The $\mathbb{C}$-bilinear form $(\cdot, \cdot)_A$ on $AC$ extends both the $\mathbb{Z}$-bilinear form $(\cdot, \cdot)_A$ on $A$ and the $\mathbb{C}$-bilinear forms $(\cdot, \cdot)_{\mathfrak{S}_n}$ on all $R_C(\mathfrak{S}_n)$.

For the assertion about $1_{(n)}$, note that it is primitive in $AC$ for $n \geq 1$, because as a class function, the indicator function of $n$-cycles vanishes upon restriction to $\mathfrak{S}_i \times \mathfrak{S}_j$ for $i + j = n$ if both $i, j \geq 1$; these subgroups contain no $n$-cycles. Hence Corollary 3.9 implies that $\text{ch}(1_{(n)})$ is a scalar multiple of $p_n$. To pin down the scalar, note $p_n = m(n)$ so $(h_n, p_n)_\lambda = (h_n, m_n)_\lambda = 1$, while $\text{ch}^{-1}(h_n) = 1_{\mathfrak{S}_n}$ has

\[
(1_{\mathfrak{S}_n}, 1_{(n)}) = \frac{1}{n!} \cdot (n-1)! = \frac{1}{n}.
\]

\[166\] Thus $\text{ch}(1_{(n)}) = \frac{p_n}{n}$. The fact that $\text{ch}(1_{\lambda}) = \frac{p_\lambda}{z_\lambda}$ then follows via induction product calculations\[167\]. Part (b) follows from Exercise 4.35 below.

\[\square\]

Remark 4.33. The paper of Liulevicius [101] gives a very elegant alternate approach to the Frobenius map as a Hopf isomorphism $A(\mathfrak{S}) \xrightarrow{\text{ch}} \Lambda$, inspired by equivariant $K$-theory and vector bundles over spaces which are finite sets of points!

\[166\] The first equality sign in this computation uses the fact that the number of all $n$-cycles in $\mathfrak{S}_n$ is $(n-1)!$. This is because any $n$-cycle in $\mathfrak{S}_n$ can be uniquely written in the form $(i_1, i_2, \ldots, i_{n-1}, n)$ (in cycle notation) with $(i_1, i_2, \ldots, i_{n-1})$ being a permutation in $\mathfrak{S}_{n-1}$ (written in one-line notation).

\[167\] For instance, one can use (4.3) to show that $z_\lambda 1_\lambda = 1_{\lambda_1} \lambda_2 \cdots \lambda_\ell \cdot 1_{(\lambda_1)} 1_{(\lambda_2)} \cdots 1_{(\lambda_\ell)}$ if $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ with $\ell = \ell(\lambda)$. See Exercise 4.34(d) for the details.
Exercise 4.34. If \( P \) is a subset of a group \( G \), we denote by \( 1_P \), the map \( G \to \mathbb{C} \) which sends every element of \( P \) to \( 1 \) and all remaining elements of \( G \) to \( 0 \).\(^{168}\) For any finite group \( G \) and any \( h \in G \), we introduce the following notations:

- Let \( Z_G(h) \) denote the centralizer of \( h \) in \( G \).
- Let \( \text{Conj}_G(h) \) denote the conjugacy class of \( h \) in \( G \).
- Define a map \( \alpha_{G,h} : G \to \mathbb{C} \) by \( \alpha_{G,h} = |Z_G(h)|^{-1} \chi_{Conj_G(h)} \). This map \( \alpha_{G,h} \) is a class function\(^{169}\).

(a) Prove that \( \alpha_{G,h}(g) = \sum_{k \in G} [khk^{-1} = g] \) for every finite group \( G \) and any \( h \in G \) and \( g \in G \). Here, we are using the Iverson bracket notation (that is, for any statement \( A \), we define \([A]\) to be the integer 1 if \( A \) is true, and 0 otherwise).

(b) Prove that if \( H \) is a subgroup of a finite group \( G \), and if \( h \in H \), then \( \text{Ind}_h^G \alpha_{H,h} = \alpha_{G,h} \).

(c) Prove that if \( G_1 \) and \( G_2 \) are finite groups, and if \( h_1 \in G_1 \) and \( h_2 \in G_2 \), then the canonical isomorphism \( R_\mathbb{C}(G_1) \otimes R_\mathbb{C}(G_2) \to R_\mathbb{C}(G_1 \times G_2) \) sends \( \alpha_{G_1,h_1} \otimes \alpha_{G_2,h_2} \) to \( \alpha_{G_1 \times G_2,(h_1,h_2)} \).

(d) Fill in the details of the proof of \( \text{ch}(1_G) = \frac{1}{|G|} \) in the proof of Theorem 4.32.

(e) Obtain an alternative proof of Remark 2.44.

(f) If \( G \) and \( H \) are two finite groups, and if \( \rho : H \to G \) is a group homomorphism, then prove that

\[
\text{Ind}_\rho \alpha_{H,h} = \alpha_{G,\rho(h)} \quad \text{for every} \ h \in H, \text{where} \ \text{Ind}_\rho \text{is defined as in Exercise 4.14}.
\]

Exercise 4.35. If \( G \) is a group and \( U_1 \) and \( U_2 \) are two \( \mathbb{C} G \)-modules, then the tensor product \( U_1 \otimes U_2 \) is a \( \mathbb{C}[G \times G] \)-module, which can be made into a \( G \)-module by letting \( g \in G \) act as \( (g, g) \in G \times G \). This \( \mathbb{C} G \)-module \( U_1 \otimes U_2 \) is called the inner tensor product\(^{170}\) of \( U_1 \) and \( U_2 \), and is a restriction of the outer tensor product \( U_1 \otimes U_2 \) using the inclusion map \( G \to G \times G, g \mapsto (g, g) \).

Let \( n \geq 0 \), and let \( sgn_{\mathbb{S}_n} \) be the 1-dimensional \( \mathbb{C} \mathbb{S}_n \)-module \( \mathbb{C} \) on which every \( g \in \mathbb{S}_n \) acts as multiplication by \( sgn(g) \). If \( V \) is a \( \mathbb{C} \mathbb{S}_n \)-module, show that the involution on \( A(\mathbb{S}) = \mathbb{S}_{n \geq 0} R(\mathbb{S}_n) \) defined in Theorem 4.32(b) sends \( \chi_V \mapsto \chi_{sgn_{\mathbb{S}_n} \otimes V} \) where \( \chi_{sgn_{\mathbb{S}_n} \otimes V} \) is the inner tensor product of \( sgn_{\mathbb{S}_n} \) and \( V \). Use this to show that this involution is a nontrivial PSH-automorphism of \( A(\mathbb{S}) \), and deduce Theorem 4.32(b).

Exercise 4.36. (a) Show that for every \( n \geq 0 \), every \( g \in \mathbb{S}_n \) and every finite-dimensional \( \mathbb{C} \mathbb{S}_n \)-module \( V \), we have \( \chi_V(1) \in \mathbb{Z} \).

(b) Show that for every \( n \geq 0 \) and every finite-dimensional \( \mathbb{C} \mathbb{S}_n \)-module \( V \), there exists a \( \mathbb{C} \mathbb{S}_n \)-module \( W \) such that \( V \cong \mathbb{C} \otimes_\mathbb{Q} W \). (In the representation theorists’ parlance, this says that all representations of \( \mathbb{S}_n \) are defined over \( \mathbb{Q} \). This part of the exercise requires some familiarity with representation theory.

Remark 4.37. Parts (a) and (b) of Exercise 4.36 both follow from an even stronger result: For every \( n \geq 0 \) and every finite-dimensional \( \mathbb{C} \mathbb{S}_n \)-module \( V \), there exists a \( \mathbb{Z} \mathbb{S}_n \)-module \( W \) which is finitely generated and free as a \( \mathbb{Z} \)-module and satisfies \( V \cong \mathbb{C} \otimes_\mathbb{Z} W \) as \( \mathbb{Z} \mathbb{S}_n \)-modules. This follows from the combinatorial approach to the representation theory of \( \mathbb{S}_n \), in which the irreducible representations of \( \mathbb{S}_n \) (the Specht modules) are constructed using Young tableaux and tabloids. See the literature on the symmetric group, e.g., [141], [54], [7] or [170] for this approach.

The connection between \( \Lambda \) and \( A(\mathbb{S}) \) as established in Theorem 4.32 benefits both the study of \( \Lambda \) and that of \( A(\mathbb{S}) \). The following two exercises show some applications to \( \Lambda \):

Exercise 4.38. If \( G \) is a group and \( U_1 \) and \( U_2 \) are two \( \mathbb{C} G \)-modules, then let \( U_1 \boxtimes U_2 \) denote the inner tensor product of \( U_1 \) and \( U_2 \) (as defined in Exercise 4.35). Consider also the binary operation \( * \) on \( \Lambda_\mathbb{Q} \) defined in Exercise 2.80(h).

(a) Show that \( \chi_{U_1 \boxtimes U_2} = \chi_{U_1} * \chi_{U_2} \) for any \( n \in \mathbb{N} \) and any two \( \mathbb{C} \mathbb{S}_n \)-modules \( U_1 \) and \( U_2 \).

(b) Use this to obtain a new solution for Exercise 2.80(h).

(c) Show that \( s_\mu \ast s_\nu = \sum_{\lambda \in \text{Par}} n_{\lambda\mu} s_\lambda \) for any two partitions \( \mu \) and \( \nu \).

[Hint: For any group \( G \), introduce a binary operation \( * \) on \( R_\mathbb{C}(G) \) which satisfies \( \chi_{U_1 \boxtimes U_2} = \chi_{U_1} \chi_{U_2} \) for any two \( \mathbb{C} G \)-modules \( U_1 \) and \( U_2 \).]

\(^{168}\)This is not in conflict with the notation \( 1_G \) for the trivial character of \( G \), since \( 1_P = 1_G \) for \( P = G \). Note that \( 1_P \) is a class function when \( P \) is a union of conjugacy classes of \( G \).

\(^{169}\)In fact, \( 1_{\text{Conj}_G(h)} \) is a class function (since \( \text{Conj}_G(h) \) is a conjugacy class), and so \( \alpha_{G,h} \) (being the scalar multiple \( |Z_G(h)| 1_{\text{Conj}_G(h)} \)) must also be a class function.

\(^{170}\)Do not confuse this with the inner product of characters.
Exercise 4.39. Define a \(\mathbb{Q}\)-bilinear map \(\square : \Lambda_Q \times \Lambda_Q \to \Lambda_Q\), which will be written in infix notation (that is, we will write \(a \square b\) instead of \(\square(a, b)\), by setting
\[
p_\chi \square p_\mu = \prod_{i=1}^{\ell(\lambda)} \prod_{j=1}^{\ell(\mu)} p_{\gcd(\lambda_i, \mu_j)}
\]
for any partitions \(\lambda\) and \(\mu\).

(a) Show that \(\Lambda_Q\), equipped with the binary operation \(\square\), becomes a commutative \(\mathbb{Q}\)-algebra with unity \(p_1\).

(b) For every \(r \in \mathbb{Z}\), define the \(\mathbb{Q}\)-algebra homomorphism \(\epsilon_r : \Lambda_Q \to \mathbb{Q}\) as in Exercise 2.80(c). Show that \(1 \square f = \epsilon_1(f) 1\) for every \(f \in \Lambda_Q\) (where 1 denotes the unity of \(\Lambda\)).

(c) Show that \(s_\mu \square s_\nu = \sum_{\lambda \in \text{Par}_n} N_{\lambda\mu\lambda}\) for any two partitions \(\mu\) and \(\nu\).

(d) Show that \(f \square g \in \Lambda\) for any \(f \in \Lambda\) and \(g \in \Lambda\).

[Hint: For every set \(X\), let \(\mathfrak{S}_X\) denote the group of all permutations of \(X\). For two sets \(X\) and \(Y\), there is a canonical group homomorphism \(\mathfrak{S}_X \times \mathfrak{S}_Y \to \mathfrak{S}_{X \times Y}\), which is injective if \(X\) and \(Y\) are nonempty. For positive integers \(n\) and \(m\), this yields an embedding \(\mathfrak{S}_n \times \mathfrak{S}_m \to \mathfrak{S}_{1,2\ldots,n \times 1,2\ldots,m}\), which, once \(n \times m\) is nonempty. For positive integers \(n\) and \(m\), this yields an embedding \(\mathfrak{S}_n \times \mathfrak{S}_m \to \mathfrak{S}_{nm}\), and thus allows defining a \(\mathbb{C}\mathfrak{S}_{nm}\)-module \(\text{Ind}_{\mathfrak{S}_n \times \mathfrak{S}_m}^{\mathfrak{S}_{nm}}(U \otimes V)\) for any \(\mathbb{C}\mathfrak{S}_n\)-module \(U\) and any \(\mathbb{C}\mathfrak{S}_m\)-module \(V\). This gives a binary operation on \(A(\mathfrak{S})\). Show that this operation corresponds to \(\square\) under the PSH-isomorphism \(\text{ch} : A(\mathfrak{S}) \to \Lambda\).]

Remark 4.40. The statements (and the idea of the solution) of Exercise 4.39 are due to Manuel Maia and Miguel Méndez (see [111] and, more explicitly, [118]), who call the operation \(\square\) the arithmetic product.

4.5. Wreath products. Next consider the tower of groups \(G_n = \mathfrak{S}_n[\Gamma]\) for a finite group \(\Gamma\), and the Hopf algebra \(A = A(G_\infty) = A(\mathfrak{S}[\Gamma])\). Recall (from Theorem 4.32) that irreducible complex representations \(\chi^\lambda\) of \(\mathfrak{S}_n\) are indexed by partitions \(\lambda\) in \(\text{Par}_n\). Index the irreducible complex representations of \(\Gamma\) as \(\text{Irr}(\Gamma) = \{\rho_1, \ldots, \rho_d\}\).

Definition 4.41. Define for a partition \(\lambda\) in \(\text{Par}_n\) and \(\rho\) in \(\text{Irr}(\Gamma)\) a representation \(\chi^{\lambda, \rho}\) of \(\mathfrak{S}_n[\Gamma]\) in which \(\sigma\) in \(\mathfrak{S}_n\) and \(\gamma = (\gamma_1, \ldots, \gamma_n)\) in \(\Gamma^n\) act on the space \(\chi^{\lambda, \rho}\) as follows
\[
\sigma(u \otimes (v_1 \otimes \cdots \otimes v_n)) = \sigma(u) \otimes (v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(n)})
\]
\[
\gamma(u \otimes (v_1 \otimes \cdots \otimes v_n)) = u \otimes (\gamma_1 v_1 \otimes \cdots \otimes \gamma_n v_n)
\]
\[(4.31)\]

Theorem 4.42. The irreducible \(\mathbb{C}\mathfrak{S}_n[\Gamma]\)-modules are the induced characters
\[
\chi^{\lambda} := \text{Ind}_{\mathfrak{S}_n[\Gamma]}^{\mathfrak{S}_n} \left( \chi^{\lambda(1), \rho_1 \otimes \cdots \otimes \chi^{(d), \rho_d}} \right)
\]
as \(\lambda\) runs through all functions
\[
\text{Irr}(\Gamma) \xrightarrow{\lambda \mapsto \text{Par}} \Gamma
\]
with the property that \(\sum_{i=1}^d |\lambda(i)| = n\). Here, \(\text{deg}(\lambda)\) denotes the \(d\)-tuple \((|\lambda(1)|, |\lambda(2)|, \ldots, |\lambda(d)|) \in \mathbb{N}^d\), and \(\mathfrak{S}_{\text{deg}(\lambda)}\) is defined as the subgroup \(\mathfrak{S}_{|\lambda(1)|} \times \mathfrak{S}_{|\lambda(2)|} \times \cdots \times \mathfrak{S}_{|\lambda(d)|}\) of \(\mathfrak{S}_n\).

Furthermore, one has a PSH-isomorphism
\[
A(\mathfrak{S}[\Gamma]) \xrightarrow{\chi^{\lambda} \mapsto \text{Ind}_{\mathfrak{S}_n[\Gamma]}^{\mathfrak{S}_n} \chi^{\lambda(1) \otimes \cdots \otimes \chi^{\lambda(d)}}} \text{Ind}_{\mathfrak{S}_n[\Gamma]}^{\mathfrak{S}_n} \chi^{\lambda(1) \otimes \cdots \otimes \chi^{\lambda(d)}}.
\]

Proof. We know from Corollary 4.30 that \(A(\mathfrak{S}[\Gamma])\) is a PSH, with PSH-basis \(\Sigma\) given by the union of all irreducible characters of all groups \(\mathfrak{S}_n[\Gamma]\). Therefore Theorem 3.12 tells us that \(A(\mathfrak{S}[\Gamma]) \cong \bigotimes_{\rho \in \text{Irr}(\Gamma)} A(\mathfrak{S}[\Gamma])(\rho)\) where \(C\) is the set of irreducible characters which are also primitive. Just as in the case of \(\mathfrak{S}_n\), it is clear from the definition of the coproduct that an irreducible character \(\rho\) of \(\mathfrak{S}_n[\Gamma]\) is primitive if and only if \(n = 1\), that in this case \(\mathfrak{S}_n[\Gamma] = \Gamma\), and \(\rho\) lies in \(\text{Irr}(\Gamma) = \{\rho_1, \ldots, \rho_d\}\).

\[171\]This is well-defined, since \((p_{\lambda})_{\lambda \in \text{Par}}\) is a \(\mathbb{Q}\)-module basis of \(\Lambda_Q\).
The remaining assertions of the theorem will then follow from the definition of the induction product algebra structure on $A(\mathcal{S}[\Gamma])$, once we have shown that, for every $\rho \in \text{Irr}(\Gamma)$, there is a PSH-isomorphism sending

$$
A(\mathcal{S}) \xrightarrow{\chi^\Lambda} A(\mathcal{S}[\Gamma])(\rho)
$$

(4.32)

Such an isomorphism comes from applying Proposition 4.17 to the semidirect product $\mathcal{S}_n[\Gamma] = \mathcal{S}_n \rtimes \Gamma^n$, so that $K = \Gamma^n, G = \mathcal{S}_n$, and fixing $V = \rho^{\otimes n}$ as $\mathbb{C}\mathcal{S}_n[\Gamma]$-module with structure as defined in (4.31) (but with $\lambda$ set to $(n)$, so that $\chi^\lambda$ is the trivial 1-dimensional $\mathbb{C}\mathcal{S}_n$-module). One obtains for each $n$, maps

$$
R(\mathcal{S}_n) \xrightarrow{\Phi} R(\mathcal{S}_n[\Gamma])
$$

where

$$
\chi \xrightarrow{\Phi} \chi \otimes (\rho^{\otimes n})
$$

$$
\alpha \xrightarrow{\Psi} \text{Hom}_{\mathbb{C}\Gamma^n}(\rho^{\otimes n}, \alpha).
$$

Taking the direct sum of these maps for all $n$ gives maps $A(\mathcal{S}) \xrightarrow{\Phi} A(\mathcal{S}[\Gamma])$.

These maps are coalgebra morphisms because of their interaction with restriction to $\mathcal{S}_i \times \mathcal{S}_j$. Since Proposition 4.17(iii) gives the adjointness property that

$$
(\chi, \Psi(\alpha))_{A(\mathcal{S})} = (\Phi(\chi), \alpha)_{A(\mathcal{S}[\Gamma])},
$$

one concludes from the self-duality of $A(\mathcal{S}), A(\mathcal{S}[\Gamma])$ that $\Phi, \Psi$ are also algebra morphisms. Since they take genuine characters to genuine characters, they are PSH-morphisms. Since $\rho$ being a simple $\mathbb{C}\Gamma$-module implies that $V = \rho^{\otimes n}$ is a simple $\mathbb{C}\Gamma^n$-module, Proposition 4.17(iv) shows that

$$
(\psi \circ \Phi)(\chi) = \chi
$$

(4.33)

for all $\mathcal{S}_n$-characters $\chi$. Hence $\Phi$ is an injective PSH-morphism. Using adjointness, (4.33) also shows that $\Phi$ sends $\mathbb{C}\mathcal{S}_n$-simples $\chi$ to $\mathbb{C}[\mathcal{S}_n[\Gamma]]$-simples $\Phi(\chi)$:

$$
(\Phi(\chi), \Phi(\chi))_{A(\mathcal{S}[\Gamma])} = (\psi \circ \Phi)(\chi, \chi)_{A(\mathcal{S})} = (\chi, \chi)_{A(\mathcal{S})} = 1.
$$

Since $\Phi(\chi) = \chi \otimes (\rho^{\otimes n})$ has $V = \rho^{\otimes n}$ as a constituent upon restriction to $\Gamma^n$, Frobenius Reciprocity shows that the irreducible character $\Phi(\chi)$ is a constituent of $\text{Ind}_{\mathbb{C}\Gamma^n}(\rho^{\otimes n} = \rho^n$. Hence the entire image of $\Phi$ lies in $A(\mathcal{S}[\Gamma])(\rho)$ (due to how we defined $A(\rho)$ in the proof of Theorem 3.12), and so $\Phi$ must restrict to an isomorphism as desired in (4.32).

One of Zelevinsky’s sample applications of the theorem is this branching rule.

**Corollary 4.43.** Given $\Lambda = (\lambda^{(1)}, \ldots, \lambda^{(d)})$ with $\sum_{i=1}^d |\lambda^{(i)}| = n$, one has

$$
\text{Res}^{\mathcal{S}_n[\Gamma]}_{\mathcal{S}_{n-1}[\Gamma] \times \Gamma} (\chi^\Lambda) = \sum_{i=1}^d \sum_{\lambda^{(i)} \subseteq \lambda, |\lambda^{(i)}| = 1} \chi^{(\lambda^{(1)}, \ldots, \lambda^{(i)}, \ldots, \lambda^{(d)}) \otimes \rho_i}.
$$

(We are identifying functions $\Lambda : \text{Irr}(\Gamma) \to \text{Par}$ with the corresponding $d$-tuples $(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(d)})$ here.)

**Example 4.44.** For $\Gamma$ a two-element group, so $\text{Irr}(\Gamma) = \{\rho_1, \rho_2\}$ and $d = 2$, then

$$
\text{Res}^{\mathcal{S}_n[\Gamma]}_{\mathcal{S}_n - 1[\Gamma] \times \Gamma} (\chi^{((3,1),(1,1))}) = \chi^{((3,1),(1,1)) \otimes \rho_1 + \chi^{((2,1),(1,1)) \otimes \rho_1 + \chi^{((3,1),(1)) \otimes \rho_2}.\n$$

**Proof of Corollary 4.43.** By Theorem 4.42, this is equivalent to computing in the Hopf algebra $A := \Lambda^{\otimes d}$ the component of the coproduct of $s_{\lambda^{(1)}} \otimes \cdots \otimes s_{\lambda^{(d)}}$ that lies in $A_{n-1} \otimes A_1$. Working within each tensor factor $\Lambda$, the Pieri formula implies that the $\Lambda^{(\lambda_i)}_{\lambda_i-1} \otimes A_1$-component of $\Delta(s_{\lambda})$ is

$$
\sum_{\lambda \subseteq \lambda, |\lambda| = 1} s_{\lambda_i} \otimes \rho.
$$

One must apply this in each of the $d$ tensor factors of $A = \Lambda^{\otimes d}$, then sum on $i$. \qed
4.6. **General linear groups.** We now consider the tower of finite general linear groups $G_n = GL_n = GL_n(\mathbb{F}_q)$ and $A = A(G_1) := A(GL)$. Corollary 4.30 tells us that $A(GL)$ is a PSH, with PSH-basis $\Sigma$ given by the union of all irreducible characters of all groups $GL_n$. Therefore Theorem 3.12 tells us that

\begin{equation}
A(GL) \cong \bigotimes_{\rho \in C} A(GL)(\rho)
\end{equation}

where $C = \Sigma \cap p$ is the set of primitive irreducible characters.

**Definition 4.45.** Call an irreducible representation $\rho$ of $GL_n$ cuspidal for $n \geq 1$ if it lies in $C$, that is, its restriction to proper parabolic subgroups $P_{i,j}$ with $i+j = n$ and $i, j > 0$ contain no nonzero vectors which are $K_{i,j}$-invariant. Given an irreducible character $\sigma$ of $GL_n$, say that $d(\sigma) = n$, and let $C_n := \{ \rho \in C : d(\rho) = n \}$ for $n \geq 1$ denote the subset of cuspidal characters of $GL_n$.

Just as was the case for $\mathfrak{S}_1$ and $\mathfrak{S}_1[\Gamma] = \Gamma$, every irreducible character $\rho$ of $GL_1(\mathbb{F}_q) = \mathbb{F}_q^\times$ is cuspidal. However, this does not exhaust the cuspidal characters. In fact, one can predict the number of cuspidal characters in $C_n$, using knowledge of the number of conjugacy classes in $GL_n$. Let $F$ denote the set of all nonconstant monic irreducible polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$. Let $F_n := \{ f \in F : \deg(f) = n \}$ for $n \geq 1$.

**Proposition 4.46.** The number $|C_n|$ of cuspidal characters of $GL_n(\mathbb{F}_q)$ is the number of $|F_n|$ of irreducible monic degree $n$ polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$ with nonzero constant term.

**Proof.** We show $|C_n| = |F_n|$ for $n \geq 1$ by strong induction on $n$. For the base case $n = 1$, just as with the families $G_n = \mathfrak{S}_n$ and $G_n = \mathfrak{S}_n[\Gamma]$, when $n = 1$ any irreducible character $\chi$ of $G_1 = GL_1(\mathbb{F}_q)$ gives a primitive element of $A = A(GL)$, and hence is cuspidal. Since $GL_1(\mathbb{F}_q) = \mathbb{F}_q^\times$ is abelian, there are $|\mathbb{F}_q^\times| = q - 1$ such cuspidal characters in $\mathfrak{C}_1$, which agrees with the fact that there are $q - 1$ monic (irreducible) linear polynomials $f(x) \neq x$ in $\mathbb{F}_q[x]$, namely $F_1 := \{ f(x) = x - c : c \in \mathbb{F}_q^\times \}$.

In the inductive step, use the fact that the number $|\Sigma_n|$ of irreducible complex characters $\chi$ of $GL_n(\mathbb{F}_q)$ equals its number of conjugacy classes. These conjugacy classes are uniquely represented by rational canonical forms, which are parametrized by functions $\Lambda : F \rightarrow \text{Par}$ with the property that $\sum_{f \in F} \deg(f)|\Lambda(f)| = n$. On the other hand, (4.34) tells us that $|\Sigma_n|$ is similarly parametrized by the functions $\Lambda : C \rightarrow \text{Par}$ having the property that $\sum_{\rho \in C} d(\rho)|\Lambda(\rho)| = n$. Thus we have parallel disjoint decompositions

\begin{align*}
\mathcal{C} &= \bigcup_{n \geq 1} C_n \\
\mathcal{F} &= \bigcup_{n \geq 1} F_n
\end{align*}

and hence an equality for all $n \geq 1$

\begin{align*}
\left| \left\{ C \xrightarrow{\Lambda} \text{Par} : \sum_{\rho \in C} d(\rho)|\Lambda(\rho)| = n \right\} \right| &= |\Sigma_n| = \left| \left\{ F \xrightarrow{\Lambda} \text{Par} : \sum_{f \in F} \deg(f)|\Lambda(f)| = n \right\} \right|.
\end{align*}

Since there is only one partition $\lambda$ having $|\lambda| = 1$ (namely, $\lambda = (1)$), this leads to parallel recursions

\begin{align*}
|C_n| &= |\Sigma_n| - \left| \bigcup_{i=1}^{n-1} C_i \xrightarrow{\Lambda} \text{Par} : \sum_{\rho \in C} d(\rho)|\Lambda(\rho)| = n \right| \\
|F_n| &= |\Sigma_n| - \left| \bigcup_{i=1}^{n-1} F_i \xrightarrow{\Lambda} \text{Par} : \sum_{f \in F} \deg(f)|\Lambda(f)| = n \right|
\end{align*}

and induction implies that $|C_n| = |F_n|$.

We shall use the notation $1_H$ for the trivial character of a group $H$ whenever $H$ is a finite group. This generalizes the notations $1_{\mathfrak{S}_n}$ and $1_{\mathfrak{S}_\lambda}$ introduced above.

---

172 Actually, we don’t need any base case for our strong induction. We nevertheless handle the case $n = 1$ as a warmup.
Example 4.47. Taking $q = 2$, let us list the sets $\mathcal{F}_n$ of monic irreducible polynomials $f(x) \neq x$ in $\mathbb{F}_2[x]$ of degree $n$ for $n \leq 3$, so that we know how many cuspidal characters of $GL_n(\mathbb{F}_q)$ in $\mathcal{C}_n$ to expect:

$$
\mathcal{F}_1 = \{x + 1\}
$$

$$
\mathcal{F}_2 = \{x^2 + x + 1\}
$$

$$
\mathcal{F}_3 = \{x^3 + x + 1, x^3 + x^2 + 1\}
$$

Thus we expect

- one cuspidal character of $GL_1(\mathbb{F}_2)$, namely $\rho_1 := 1_{GL_1(\mathbb{F}_2)}$,
- one cuspidal character $\rho_2$ of $GL_2(\mathbb{F}_2)$, and
- two cuspidal characters $\rho_3, \rho_3'$ of $GL_3(\mathbb{F}_2)$.

We will say more about $\rho_2, \rho_3, \rho_3'$ in the next section.

Exercise 4.48. (a) Show that for $n \geq 2$,

$$
|\mathcal{C}_n| = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d
$$

where $\mu(m)$ is the number-theoretic Möbius function of $m$, that is $\mu(m) = (-1)^d$ if $m = p_1 \cdots p_d$ for $d$ distinct primes, and $\mu(m) = 0$ if $m$ is not squarefree. (Here, the summation sign $\sum_{d|n}$ means a sum over all positive divisors $d$ of $n$.)

(b) Show that (4.35) also counts the necklaces with $n$ beads of $q$ colors (=equivalence classes under the $\mathbb{Z}/n\mathbb{Z}$-action of cyclic rotation on sequences $(a_1, \ldots, a_n)$ in $\mathbb{F}_q^n$) which are primitive in the sense that no nontrivial rotation fixes any of the sequences within the equivalence class. For example, when $q = 2$, here are representatives of these primitive necklaces for $n = 2, 3, 4$:

$$
n = 2 : \{(0, 1)\}
$$

$$
n = 3 : \{(0, 0, 1), (0, 1, 1)\}
$$

$$
n = 4 : \{(0, 0, 0, 1), (0, 0, 1, 1), (0, 1, 1, 1)\}
$$

4.7. Steinberg’s unipotent characters. Not surprisingly, the (cuspidal) character $\iota := 1_{GL_1}$ of $GL_1(\mathbb{F}_q)$ plays a distinguished role. The parabolic subgroup $P_{(1^n)}$ of $GL_n(\mathbb{F}_q)$ is the Borel subgroup $B$ of upper triangular matrices, and we have $\iota^n = \text{Ind}^{GL_n}_{B} 1_B = \mathbb{C}[GL_n/B]$ (identifying representations with their characters as usual). The subalgebra $A(GL)(\iota)$ of $A(GL)$ is the $\mathbb{Z}$-span of the irreducible characters $\sigma$ that appear as constituents of $\iota^n = \text{Ind}^{GL_n}_{B} 1_B = \mathbb{C}[GL_n/B]$ for some $n$.

Definition 4.49. An irreducible character $\sigma$ of $GL_n$ appearing as a constituent of $\text{Ind}^{GL_n}_{B} 1_B = \mathbb{C}[GL_n/B]$ is called a unipotent character. Equivalently, by Frobenius reciprocity, $\sigma$ is unipotent if it contains a nonzero $B$-invariant vector.

In particular, $1_{GL_n}$ is a unipotent character of $GL_n$ for each $n$.

Proposition 4.50. One can choose $\Lambda \cong A(GL)(\iota)$ in Theorem 3.20(g) so that $h_n \mapsto 1_{GL_n}$.

Proof. Theorem 3.18(a) tells us $\iota^2 = \text{Ind}^{GL_2}_{B} 1_B$ must have exactly two irreducible constituents, one of which is $1_{GL_2}$; call the other one $\text{St}_2$. Choose the isomorphism so as to send $h_2 \mapsto 1_{GL_2}$. Then $h_n \mapsto 1_{GL_n}$ follows from the claim that $\text{St}_2^n (1_{GL_n}) = 0$ for $n \geq 2$; one has

$$
\Delta(1_{GL_n}) = \sum_{i+j=n} \left(\text{Res}_{P_i,j}^{G_n} 1_{GL_n}\right)^{K_{i,j}} = \sum_{i+j=n} 1_{GL_i} \otimes 1_{GL_j}
$$

173Proof. Exercise 4.31(d) (applied to $G = GL_\ast$, $\ell = n$, $\alpha = (1^n) = \left(\begin{array}{c} 1, 1, \ldots, 1 \\ n \text{ times} \end{array}\right)$ and $\chi_\ell = \iota$) gives

$$
\iota^n = \text{Ind}_{B}^{GL_n} 1_B = \text{Ind}_{B}^{GL_n} \text{Ind}_{B}^{P_{(1^n)}} 1_B = \text{Ind}_{B}^{GL_n} 1_B = \mathbb{C}[GL_n/B],
$$

where the last equality follows from the general fact that if $G$ is a finite group and $H$ is a subgroup of $G$, then $\text{Ind}_{H}^{G} 1_H \cong \mathbb{C}[G/H]$ as $CG$-modules.
so that \( \text{St}_2^+ (1_{GL_n}) = (\text{St}_2, 1_{GL_2}) 1_{GL_{n-2}} = 0 \) since \( \text{St}_2 \neq 1_{GL_2} \).

This subalgebra \( A(GL)(\iota) \), and the unipotent characters \( \chi^\lambda_q \) corresponding under this isomorphism to the Schur functions \( s_\lambda \), were introduced by Steinberg \[159\]. He wrote down \( \chi^\lambda_q \) as a virtual sum of induced characters \( \text{Ind}^{GL_n}_{P_\alpha} 1_{P_\alpha} (= \text{Ind}^{GL_n}_{V_\alpha} \cdots \text{Ind}^{GL_n}_{V_{\alpha_1}}) \), modelled on the Jacobi-Trudi determinantal expression for \( s_\lambda = \det(h_{\lambda, -i+j}) \). Note that \( \text{Ind}^{GL_n}_{P_\alpha} 1_{P_\alpha} \) is the transitive permutation representation \( \mathbb{C}[G/P_\alpha] \) for \( GL_n \) permuting the finite partial flag variety \( G/P_\alpha \), that is, the set of \( \alpha \)-flags of subspaces

\[
\{0\} \subset V_{\alpha_1} \subset V_{\alpha_1 + \alpha_2} \subset \cdots \subset V_{\alpha_1 + \alpha_2 + \cdots + \alpha_{t-1}} \subset \mathbb{F}^n
\]

where \( \dim_{\mathbb{F}_q} V_\alpha = d \) in each case. This character has dimension equal to \( |G/P_\alpha| \), with formula given by the \( q \)-multinomial coefficient (see e.g. Stanley \[157, \S 1.7\]):

\[
\begin{bmatrix} n \\ \alpha \end{bmatrix}_q = \frac{[n]_q!}{[\alpha_1]_q! \cdots [\alpha_t]_q!}
\]

where \( [n]_q := [n]_q [n-1]_q \cdots [2]_q [1]_q \) and \( [n]_q := 1 + q + \cdots + q^{n-1} = \frac{q^n - 1}{q - 1} \).

Our terminology \( \text{St}_2 \) is motivated by the \( n = 2 \) special case of the Steinberg character \( \text{St}_n \), which is the unipotent character corresponding under the isomorphism in Proposition 4.50 to \( e_n = s_{(1^n)} \). It can be defined by the virtual sum

\[
\text{St}_n := \chi^{(1^n)}_q = \sum_{\alpha} (-1)^{n-\ell(\alpha)} \text{Ind}_{P_\alpha}^{GL_n} 1_{P_\alpha}
\]

in which the sum runs through all compositions \( \alpha \) of \( n \). This turns out to be the genuine character for \( GL_n(\mathbb{F}_q) \) acting on the top homology group of its Tits building: the simplicial complex whose vertices are nonzero proper subspaces \( V \) of \( \mathbb{F}_q^n \), and whose simplices correspond to flags of nested subspaces. One needs to know that this Tits building has only top homology, so that one can deduce the above character formula from the Hopf trace formula; see Björner \[21\].

4.8. Examples: \( GL_2(\mathbb{F}_2) \) and \( GL_3(\mathbb{F}_2) \). Let’s get our hands dirty.

**Example 4.51.** For \( n = 2 \), there are two unipotent characters, \( \chi^{(2)}_q = 1_{GL_2} \) and

\[
(4.36) \quad \text{St}_2 := \chi^{(1,1)}_q = \frac{1}{2} 1_{GL_1} - 1_{GL_2} = \text{Ind}_{B}^{GL_2} 1_B - 1_{GL_2}
\]

since the Jacobi-Trudi formula (2.22) gives \( s_{(1,1)} = \det \begin{bmatrix} h_1 & h_2 \\ 1 & h_1 \end{bmatrix} = h_1^2 - h_2 \). The description (4.36) for this Steinberg character \( \text{St}_2 \) shows that it has dimension

\[
|GL_2(B)| - 1 = (q + 1) - 1 = q
\]

and that one can think of it as follows: consider the permutation action of \( GL_2 \) on the \( q+1 \) lines \( \{\ell_0, \ell_1, \ldots, \ell_q\} \) in the projective space \( \mathbb{P}^1_{\mathbb{F}_q} = GL_2(\mathbb{F}_q)/B \), and take the invariant subspace perpendicular to the sum of basis elements \( e_{\ell_0} + \cdots + e_{\ell_q} \).

**Example 4.52.** Continuing the previous example, but taking \( q = 2 \), we find that we have constructed two unipotent characters: \( 1_{GL_2} = \chi^{(2)}_{q=2} \) of dimension 1, and \( \text{St}_2 = \chi^{(1,1)}_{q=2} \) of dimension \( q = 2 \). This lets us identify the unique cuspidal character \( \rho_2 \) of \( GL_2(\mathbb{F}_2) \), using knowledge of the character table of \( GL_2(\mathbb{F}_2) \cong S_3 \):

<table>
<thead>
<tr>
<th>( 1_{GL_2} = \chi^{(2)}_{q=2} )</th>
<th>\begin{bmatrix} 1 &amp; 0 \ 0 &amp; 1 \end{bmatrix}</th>
<th>\begin{bmatrix} 1 &amp; 1 \ 0 &amp; 1 \end{bmatrix}</th>
<th>\begin{bmatrix} 1 &amp; 0 \ 1 &amp; 1 \end{bmatrix}</th>
<th>\begin{bmatrix} 0 &amp; 1 \ 1 &amp; 0 \end{bmatrix}</th>
<th>\begin{bmatrix} 1 &amp; 1 \ 1 &amp; 0 \end{bmatrix}</th>
<th>\begin{bmatrix} 0 &amp; 1 \ 1 &amp; 1 \end{bmatrix}</th>
</tr>
</thead>
<tbody>
<tr>
<td>unipotent</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

In other words, the cuspidal character \( \rho_2 \) of \( GL_2(\mathbb{F}_2) \) corresponds under the isomorphism \( GL_2(\mathbb{F}_2) \cong S_3 \) to the sign character \( \text{sgn}_{S_3} \).
Example 4.53. Continuing the previous example to \( q = 2 \) and \( n = 3 \) lets us analyze the irreducible characters of \( \text{GL}_3(\mathbb{F}_2) \). Recalling our labelling \( \rho_1, \rho_2, \rho_3, \rho_3' \) from Example 4.47 of the cuspidal characters of \( \text{GL}_n(\mathbb{F}_2) \) for \( n = 1, 2, 3 \), Zelevinsky’s Theorem 3.12 tells us that the \( \text{GL}_3(\mathbb{F}_2) \)-irreducible characters should be labelled by functions \( \{ \rho_1, \rho_2, \rho_3, \rho_3' \} \xrightarrow{\Delta} \text{Par} \) for which

\[
1 \cdot |\Delta(\rho_1)| + 2 \cdot |\Delta(\rho_2)| + 3 \cdot |\Delta(\rho_3)| + 3 \cdot |\Delta(\rho_3')| = 3
\]

We will label such an irreducible character \( \chi^\Delta = \chi(\Delta(\rho_1), \Delta(\rho_2), \Delta(\rho_3), \Delta(\rho_3')) \).

Three of these irreducibles will be the unipotent characters, mapping under the isomorphism from Proposition 4.50 as follows:

- \( s_{(3)} = h_3 \xrightarrow{\chi((3),\emptyset,\emptyset,\emptyset)} \text{GL}_3 \) of dimension 1.
- \( s_{(2,1)} = \det \begin{bmatrix} h_2 & h_3 \\ 1 & h_1 \end{bmatrix} = h_2 h_3 - h_1 \xrightarrow{\chi((2,1),\emptyset,\emptyset,\emptyset)} \text{Ind}_{\text{GL}_2}^{\text{GL}_3} 1_{\text{GL}_2} - 1_{\text{GL}_3}, \) of dimension \( \frac{3}{q} - 1 = q^3 + q \xrightarrow{q=2} 6. \)
- Lastly,

\[
s_{(1,1,1)} = \det \begin{bmatrix} h_1 & h_2 & h_3 \\ 1 & h_1 & h_2 \\ 0 & 1 & h_1 \end{bmatrix} = h_1^3 - h_2 h_1 - h_1 h_2 + h_3 \xrightarrow{\text{St}_3} \chi((1,1,1),\emptyset,\emptyset,\emptyset) = \text{Ind}_{\text{GL}_2}^{\text{GL}_3} 1_B - \text{Ind}_{\text{GL}_2}^{\text{GL}_3} 1_{\text{GL}_2} - \text{Ind}_{\text{GL}_2}^{\text{GL}_3} 1_{\text{GL}_2} + 1_{\text{GL}_3}, \) of dimension

\[
\left[ \frac{3}{1,1,1} \right] - \left[ \frac{3}{2,1} \right] - \left[ \frac{3}{1,2} \right] + \left[ \frac{3}{3} \right] = [3]_q - 1 = q^3 + q \xrightarrow{q=2} 8.
\]

There should also be one non-unipotent, non-cuspidal character, namely

\[
\chi^{(1)}((1),\emptyset,\emptyset,\emptyset) = \rho_1 \rho_2 = \text{Ind}_{\text{GL}_2}^{\text{GL}_3} \text{Ind}_{\text{GL}_1 \times \text{GL}_2}^{\text{GL}_3} (1_{\text{GL}_1} \otimes \rho_2), \) having dimension \( \left[ \frac{3}{1,2} \right] \cdot 1 \cdot 1 = [3]_q \xrightarrow{q=2} 7. \)

Finally, we expect cuspidal characters \( \rho_3 = \chi((\emptyset,\emptyset,\emptyset,1),\emptyset), \rho_3' = \chi((\emptyset,\emptyset,\emptyset,1)), \) whose dimensions \( d_3, d_3' \) can be deduced from the equation

\[
1^2 + 6^2 + 8^2 + 7^2 + d_3^2 + (d_3')^2 = |\text{GL}_3(\mathbb{F}_2)| = [(q^3 - q^0)(q^3 - q^1)(q^3 - q^2)]_{q=2} = 168.
\]

This forces \( d_3^2 + (d_3')^2 = 18, \) whose only solution in positive integers is \( d_3 = d_3' = 3. \)

We can check our predictions of the dimensions for the various \( \text{GL}_3(\mathbb{F}_2) \)-irreducible characters since \( \text{GL}_3(\mathbb{F}_2) \) is the finite simple group of order 168 (also isomorphic to \( \text{PSL}_2(\mathbb{F}_7) \)), with known character table (see James and Liebeck [79, p. 318]):

<table>
<thead>
<tr>
<th>Character</th>
<th>Centralizer Order</th>
<th>Unipotent/Cuspidal</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 1_{\text{GL}_3} )</td>
<td>( \chi((3),\emptyset,\emptyset,\emptyset) )</td>
<td>Unipotent</td>
</tr>
<tr>
<td>( \chi((2,1),\emptyset,\emptyset,\emptyset) )</td>
<td>Unipotent</td>
<td></td>
</tr>
<tr>
<td>( \text{St}_3 )</td>
<td>( \chi((1,1,1),\emptyset,\emptyset,\emptyset) )</td>
<td>Unipotent</td>
</tr>
<tr>
<td>( \chi((1),(1),\emptyset,\emptyset,\emptyset) )</td>
<td>Unipotent</td>
<td></td>
</tr>
<tr>
<td>( \rho_3 )</td>
<td>( \chi((\emptyset,\emptyset,\emptyset,1),\emptyset) )</td>
<td>Cuspidal</td>
</tr>
<tr>
<td>( \rho_3' )</td>
<td>( \chi((\emptyset,\emptyset,\emptyset,1),\emptyset) )</td>
<td>Cuspidal</td>
</tr>
</tbody>
</table>

Here \( a := -1/2 + i\sqrt{7}/2. \)

Remark 4.54. It is known (see e.g. Bump [25, Cor. 7.4]) that, for \( n \geq 2 \), the dimension of any cuspidal irreducible character \( \rho \) of \( \text{GL}_n(\mathbb{F}_q) \) is

\[
(q^n - 1)(q^{n-1} - 1) \cdots (q^2 - 1)(q - 1).
\]
Note that when \( q = 2 \),
- for \( n = 2 \) this gives \( 2^1 - 1 = 1 \) for the dimension of \( \rho_2 \), and
- for \( n = 3 \) it gives \((2^2 - 1)(2 - 1) = 3\) for the dimensions of \( \rho_3, \rho'_3 \),
agreeing with our calculations above. Much more is known about the character table of \( GL_n(\mathbb{F}_q) \); see Remark 4.68 below, Zelevinsky [173, Chap. 11], and Macdonald [109, Chap. IV].

4.9. The Hall algebra. There is another interesting Hopf subalgebra (and quotient Hopf algebra) of \( A(GL) \), related to unipotent conjugacy classes in \( GL_n(\mathbb{F}_q) \).

**Definition 4.55.** Say that an element \( g \) in \( GL_n(\mathbb{F}_q) \) is unipotent if its eigenvalues are all equal to 1. Equivalently, \( g \in GL_n(\mathbb{F}_q) \) is unipotent if and only if \( g - id_{\mathbb{F}_q} \) is nilpotent. A conjugacy class in \( GL_n(\mathbb{F}_q) \) is unipotent if its elements are unipotent.

Denote by \( H_n \) the \( \mathbb{C} \)-subspace of \( R_C(GL_n) \) consisting of those class functions which are supported only on unipotent conjugacy classes, and let \( H = \bigoplus_{n \geq 0} H_n \) as a \( \mathbb{C} \)-subspace of \( A_C(GL) = \bigoplus_{n \geq 0} R_C(GL_n) \).

**Proposition 4.56.** The subspace \( H \) is a Hopf subalgebra of \( A_C(GL) \), which is graded, connected, and of finite type, and self-dual with respect to the inner product on class functions inherited from \( A_C(GL) \). It is also a quotient Hopf algebra of \( A_C(GL) \), as the \( \mathbb{C} \)-linear surjection \( A_C(GL) \to H \) restricting class functions to unipotent classes is a Hopf algebra homomorphism. This surjection has kernel \( H^\perp \), which is both an ideal and a two-sided coideal.

**Proof.** It is immediately clear that \( H^\perp \) is a graded \( \mathbb{C} \)-vector subspace of \( A_C(GL) \), whose \( n \)-th homogeneous component consists of those class functions on \( GL_n \) whose values on all unipotent classes are 0. (This holds no matter whether the perpendicular space is taken with respect to the Hermitian form \( \langle \cdot, \cdot \rangle_G \) or with respect to the bilinear form \( \langle \cdot, \cdot \rangle_G \). In other words, \( H^\perp \) is the kernel of the surjection \( A_C(GL) \to H \) defined in the proposition.

Given two class functions \( \chi_i, \chi_j \) on \( GL_i, GL_j \) and \( g \) in \( GL_{i+j} \), one has
\[
(4.37) \quad (\chi_i \cdot \chi_j)(g) = \frac{1}{|P_{i,j}|} \sum_{h \in GL_{i+j}} \chi_i(g_i) \chi_j(g_j).
\]
Since \( g \) is unipotent if and only if \( h^{-1}g h \) is unipotent if and only if both \( g_i, g_j \) are unipotent, the formula \( (4.37) \) shows both that \( H \) is a subalgebra\(^{174} \) and that \( H^\perp \) is a two-sided ideal\(^{175} \). It also shows that the surjection \( A_C(GL) \to H \) restricting every class function to unipotent classes is an algebra homomorphism\(^{176} \).

Similarly, for class functions \( \chi \) on \( GL_n \) and \( (g_i, g_j) \) in \( GL_{i,j} = GL_i \times GL_j \), one has
\[
\Delta(\chi)(g_i, g_j) = \frac{1}{q^{ij}} \sum_{k \in \mathbb{F}_q^{i+j}} \chi \begin{bmatrix} g_i & k \\ 0 & g_j \end{bmatrix}
\]
using \( (4.13) \). This shows both that \( H \) is a sub-coalgebra of \( A = A_C(GL) \) (that is, it satisfies \( \Delta H \subset H \otimes H \)) and that \( H^\perp \) is a two-sided coideal (that is, we have \( \Delta(H^\perp) \subset H^\perp \otimes A + A \otimes H^\perp \)), since it shows that if \( \chi \) is supported only on unipotent classes, then \( \Delta(\chi) \) vanishes on \( (g_1, g_2) \) that have either \( g_1 \) or \( g_2 \) non-unipotent. It also shows that the surjection \( A_C(GL) \to H \) restricting every class function to unipotent classes is a coalgebra homomorphism. The rest follows.

The subspace \( H \) is called the Hall algebra. It has an obvious orthogonal \( \mathbb{C} \)-basis, with interesting structure constants.

**Definition 4.57.** Given a partition \( \lambda \) of \( n \), let \( J_\lambda \) denote the \( GL_n \)-conjugacy class of unipotent matrices whose Jordan type (that is, the list of the sizes of the Jordan blocks, in decreasing order) is given by \( \lambda \). Furthermore, let \( z_\lambda(g) \) denote the size of the centralizer of any element of this conjugacy class \( J_\lambda \).

\(^{174} \) Indeed, if \( \chi_i \) and \( \chi_j \) are both supported only on unipotent classes, then the same holds for \( \chi_i \cdot \chi_j \).

\(^{175} \) In fact, if one of \( \chi_i \) and \( \chi_j \) annihilates all unipotent classes, then so does \( \chi_i \cdot \chi_j \).

\(^{176} \) Because if \( g \) is unipotent, then the only values of \( \chi_i \) and \( \chi_j \) appearing on the right hand side of \( (4.37) \) are those on unipotent elements.
The indicator class functions \( \{ 1_{J_i} \}_{i \in \text{Par}} \) form a \( \mathbb{C} \)-basis for \( \mathcal{H} \) whose multiplicative structure constants are called the Hall coefficients \( g_{\mu, \nu}^\lambda(q) \):

\[
1_{J_i} 1_{J_j} = \sum_{\lambda} g_{\mu, \nu}^\lambda(q) 1_{J_\lambda}.
\]

Because the dual basis to \( \{ 1_{J_i} \} \) is \( \{ z_\lambda(q) 1_{J_i} \} \), self-duality of \( \mathcal{H} \) shows that the Hall coefficients are (essentially) also structure constants for the comultiplication:

\[
\Delta 1_{J_\lambda} = \sum_{\mu, \nu} g_{\mu, \nu}^\lambda(q) \frac{z_\mu(q) z_\nu(q)}{z_\lambda(q)} 1_{J_\mu} \otimes 1_{J_\nu}.
\]

The Hall coefficient \( g_{\mu, \nu}^\lambda(q) \) has the following interpretation.

**Proposition 4.58.** Fix any \( g \in GL_n(\mathbb{F}_q) \) acting unipotently on \( \mathbb{F}_q^n \) with Jordan type \( \lambda \). Then \( g_{\mu, \nu}^\lambda(q) \) counts the \( g \)-stable \( \mathbb{F}_q \)-subspaces \( V \subset \mathbb{F}_q^n \) for which the restriction \( g|_V \) acts with Jordan type \( \mu \), and the induced map \( \bar{g} \) on the quotient space \( \mathbb{F}_q^n/V \) has Jordan type \( \nu \).

**Proof.** Given \( \mu, \nu \) partitions of \( i, j \) with \( i + j = n \), taking \( \chi_i, \chi_j \) equal to \( 1_{J_\mu}, 1_{J_\nu} \) in (4.37) shows that for any \( g \) in \( GL_n \), the value of \( \left( 1_{J_\mu} \cdot 1_{J_\nu} \right)(g) \) is given by

\[
\left( \frac{1}{|P_{i,j}|} \right) \left\{ h \in GL_n : h^{-1} gh = \begin{bmatrix} g_i & \ast \\ 0 & g_j \end{bmatrix} \text{ with } g_i \in J_\mu, g_j \in J_\nu \right\}.
\]

Let \( S \) denote the set appearing in (4.38), and let \( \mathbb{F}_q^i \) denote the \( i \)-dimensional subspace of \( \mathbb{F}_q^n \) spanned by the first \( i \) standard basis vectors. Note that the condition on an element \( h \) in \( S \) saying that \( h^{-1}gh \) is in block upper-triangular form can be re-expressed by saying that the subspace \( V := h(\mathbb{F}_q^i) \) is \( g \)-stable. One then sees that the map \( h \mapsto V = h(\mathbb{F}_q^i) \) surjects \( S \) onto the set of \( i \)-dimensional \( g \)-stable subspaces \( V \) of \( \mathbb{F}_q^n \) for which \( g|_V \) and \( \bar{g} \) are unipotent of types \( \mu, \nu \), respectively. Furthermore, for any particular such \( V \), its fiber \( \varphi^{-1}(V) \) in \( S \) is a coset of the stabilizer within \( GL_n \) of \( V \), which is conjugate to \( P_{i,j} \), and hence has cardinality \( |\varphi^{-1}(V)| = |P_{i,j}| \). This proves the assertion of the proposition. \( \square \)

The Hall algebra \( \mathcal{H} \) will turn out to be isomorphic to the ring \( \Lambda_C \) of symmetric functions with \( C \)-coefficients, via a composite \( \varphi \) of three maps

\[ \Lambda_C \longrightarrow A(GL)(i)_C \longrightarrow A(GL)_C \longrightarrow \mathcal{H} \]

in which the first map is the isomorphism from Proposition 4.50, the second is inclusion, and the third is the quotient map from Proposition 4.56.

**Theorem 4.59.** The above composite \( \varphi \) is a Hopf algebra isomorphism, sending

\[
\begin{align*}
h_n &\mapsto \sum_{\lambda \in \text{Par}_n} 1_{J_\lambda}, \\
e_n &\mapsto q^{\binom{n}{2}} 1_{J_{(1^n)}}, \\
p_n &\mapsto \sum_{\lambda \in \text{Par}_n} (q; q)_\lambda e(\lambda)^{-1} 1_{J_\lambda} \quad \text{(for } n > 0),
\end{align*}
\]

where

\[
(x; q)_m := (1-x)(1-qx)(1-q^2x)\cdots(1-q^{m-1}x) \quad \text{for all } m \in \mathbb{N}.
\]

**Proof.** That \( \varphi \) is a graded Hopf morphism follows because it is a composite of three such morphisms. We claim that once one shows the formula for the (nonzero) image \( \varphi(p_n) \) given above is correct, then this will already show \( \varphi \) is an isomorphism, by the following argument. Note first that \( \Lambda_C \) and \( \mathcal{H} \) both have dimension \( |\text{Par}_n| \) for their \( n \)th homogeneous components, so it suffices to show that the graded map \( \varphi \) is injective. On the other hand, both \( \Lambda_C \) and \( \mathcal{H} \) are (graded, connected, finite type) self-dual Hopf algebras (although with respect to a sesquilinear form), so Theorem 3.8 says that each is the symmetric algebra on its space of primitive elements. Thus it suffices to check that \( \varphi \) is injective when restricted to their subspaces

\[\text{Note that } 1_P = 1_G \text{ for } P = G.\]
of primitives.\footnote{An alternative way to see that it suffices to check this is by recalling Exercise 1.54(c).} For $\Lambda$, by Corollary 3.9 the primitives are spanned by $\{p_1, p_2, \ldots\}$, with only one basis element in each degree $n \geq 1$. Hence $\varphi$ is injective on the subspace of primitives if and only if it does not annihilate any $p_n$.

Thus it only remains to show the above formulas for the images of $h_n, e_n, p_n$ under $\varphi$. This is clear for $h_n$, since Proposition 4.50 shows that it maps under the first two composites to the indicator function $1_{GL_n}$ which then restricts to the sum of indicators $\sum_{\lambda \in \Par_n} 1_{J_{\lambda}}$ in $\mathcal{H}$. For $e_n, p_n$, we resort to generating functions. Let $\hat{h}_n, \hat{e}_n, \hat{p}_n$ denote the three putative images in $\mathcal{H}$ of $h_n, e_n, p_n$, appearing on the right side in the theorem, and define generating functions

$$
\hat{H}(t) := \sum_{n \geq 0} \hat{h}_n t^n, \quad \hat{E}(t) := \sum_{n \geq 0} \hat{e}_n t^n, \quad \hat{P}(t) := \sum_{n \geq 0} \hat{p}_n t^n \quad \text{in } \mathcal{H}[t].
$$

We wish to show that the map $\varphi[[t]] : \Lambda[[t]] \to \mathcal{H}[[t]]$ (induced by $\varphi$) maps $H(t), E(t), P(t)$ in $\Lambda[[t]]$ to these three generating functions.\footnote{See (2.14), (2.15), (2.30) for the definitions of $H(t), E(t), P(t)$.} Since we have already shown this is correct for $H(t)$, by (2.16), (2.30), it suffices to check that in $\mathcal{H}[[t]]$ one has

$$
\hat{H}(t) \hat{E}(-t) = 1, \quad \text{or equivalently,} \quad \sum_{k=0}^{n} (-1)^k \hat{e}_k \hat{h}_{n-k} = \delta_{0,n};
$$

$$
\hat{H}'(t) \hat{E}(-t) = \hat{P}(t), \quad \text{or equivalently,} \quad \sum_{k=0}^{n} (-1)^k (n-k) \hat{e}_k \hat{h}_{n-k} = \hat{p}_n.
$$

Thus it would be helpful to evaluate the class function $\hat{e}_k \hat{h}_{n-k}$. Note that a unipotent $g$ in $GL_n$ having $\ell$ Jordan blocks has an $\ell$-dimensional 1-eigenspace, so that the number of $k$-dimensional $q$-stable $\mathbb{F}_q$-subspaces of $\mathbb{F}_q^n$ on which $g$ has Jordan type $(1^k)$ (that is, on which $g$ acts as the identity) is the $q$-binomial coefficient

$$
\binom{\ell}{k}_q = \frac{(q; q)_\ell}{(q; q)_k (q; q)_{\ell-k}}.
$$

counting $k$-dimensional $\mathbb{F}_q$-subspaces $V$ of an $\ell$-dimensional $\mathbb{F}_q$-vector space; see, e.g., [157, §1.7]. Hence, for a unipotent $g$ in $GL_n$ having $\ell$ Jordan blocks, we have

$$
(\hat{e}_k \hat{h}_{n-k})(g) = q^{(\ell)} \cdot \left( \frac{1}{\ell_{(1^k)}} \cdot \hat{h}_{n-k} \right)(g) = q^{(\ell)} \cdot \sum_{\nu \in \Par_{n-k}} \left( \frac{1}{\ell_{(\nu)}} \cdot \hat{\nu} \right)(g) = q^{(\ell)} \binom{\ell}{k}_q
$$

(by Proposition 4.58). Thus one needs for $\ell \geq 1$ that

$$
\sum_{k=0}^{\ell} (-1)^k q^{(\ell)} \binom{\ell}{k}_q = 0, \quad \text{(4.39)}
$$

$$
\sum_{k=0}^{\ell} (-1)^k (n-k) q^{(\ell)} \binom{\ell}{k}_q = (q; q)_{\ell-1}. \quad \text{(4.40)}
$$

Identity (4.39) comes from setting $x = 1$ in the $q$-binomial theorem [157, Exer. 3.119]:

$$
\sum_{k=0}^{\ell} (-1)^k q^{(\ell)} \binom{\ell}{k}_q x^{\ell-k} = (x-1)(x-q)(x-q^2) \cdots (x-q^{\ell-1}).
$$

Identity (4.40) comes from taking $\frac{d}{dx}$ in the $q$-binomial theorem, then setting $x = 1$, and finally adding $(n-\ell)$ times (4.39). \hfill \Box

**Exercise 4.60.** Fix a prime power $q$. For any $k \in \mathbb{N}$, and any $k$ partitions $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}$, we define a family $\left( g^{\lambda} \right)_{\lambda \in \Par}$ of elements of $\mathbb{C}$ by the equation

$$
\prod_{\lambda^{(1)}} 1_{J_{\lambda^{(1)}}} \cdots \prod_{\lambda^{(k)}} 1_{J_{\lambda^{(k)}}} = \sum_{\lambda \in \Par} g^{\lambda} \prod_{\lambda^{(1)}} 1_{J_{\lambda^{(1)}}} \cdots \prod_{\lambda^{(k)}} 1_{J_{\lambda^{(k)}}}
$$

in $\mathcal{H}$. This notation generalizes the notation $g^{\lambda}_{\nu}(q)$ we introduced in Definition 4.57. Note that $g^{\lambda}_{\mu}(q) = \delta_{\lambda, \mu}$ for any two partitions $\lambda$ and $\mu$, and that $g^{\lambda}(q) = \delta_{\lambda, \varnothing}$ for any partition $\lambda$ (where $g^{\lambda}(q)$ is to be understood as $g^{\lambda}_{\varnothing}(q)$ for $k = 0$).
(a) Let $\lambda \in \operatorname{Par}$, and let $n = |\lambda|$. Let $V$ be an $n$-dimensional $\mathbb{F}_q$-vector space, and let $g$ be a unipotent endomorphism of $V$ having Jordan type $\lambda$. Let $k \in \mathbb{N}$, and let $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}$ be $k$ partitions. A $(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)})$-compatible $g$-flag will mean a sequence $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_k = V$ of $g$-invariant $\mathbb{F}_q$-vector subspaces $V_i$ of $V$ such that for every $i \in \{1, 2, \ldots, k\}$, the endomorphism of $V_i/V_{i-1}$ induced by $g$ has Jordan type $\lambda^{(i)}$.

Show that $g_{\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}}(q)$ is the number of $(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)})$-compatible $g$-flags.\footnote{This is well-defined. In fact, both $V_i$ and $V_{i-1}$ are $g$-invariant, so that $g$ restricts to an endomorphism of $V_i$, which further restricts to an endomorphism of $V_{i-1}$, and thus gives rise to an endomorphism of $V_i/V_{i-1}$.}  

(b) Let $\lambda \in \operatorname{Par}$. Let $k \in \mathbb{N}$, and let $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}$ be $k$ partitions. Show that $g_{\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}}(q) = 0$ unless $|\lambda^{(1)}| + |\lambda^{(2)}| + \cdots + |\lambda^{(k)}| = |\lambda|$ and $\lambda^{(1)} + \lambda^{(2)} + \cdots + \lambda^{(k)} \geq \lambda$. (Here and in the following, we are using the notations of Exercise 2.93.)

(c) Let $\lambda \in \operatorname{Par}$, and let us write the transpose partition $\lambda'$ as $\lambda' = ((\lambda')_1, (\lambda')_2, \ldots, (\lambda')_k)$. Show that $g_{\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}}(q) \neq 0$.

(d) Let $n \in \mathbb{N}$ and $\lambda \in \operatorname{Par}_n$. Show that

$$\varphi(e_{\lambda}) = \sum_{\mu \in \operatorname{Par}_n; \lambda \vdash \mu} \alpha_{\lambda, \mu} J_{\mu}$$

for some coefficients $\alpha_{\lambda, \mu} \in \mathbb{C}$ satisfying $\alpha_{\lambda, \lambda'} \neq 0$.

(e) Give another proof of the fact that the map $\varphi$ is injective.\footnote{This can be seen as a generalization of Proposition 4.58. In fact, if $\mu$ and $\nu$ are two partitions, then a $(\mu, \nu)$-compatible $g$-flag is a sequence $0 = V_0 \subset V_1 \subset V_2 = V$ of $g$-invariant $\mathbb{F}_q$-vector subspaces $V_i$ of $V$ such that the endomorphism of $V_i/V_{i-1}$ induced by $g$ has Jordan type $\mu$, and the endomorphism of $V_0/V_1 \cong V/V_1$ induced by $g$ has Jordan type $\nu$. Choosing such a sequence amounts to choosing $V_1$ (since there is only one choice for each of $V_0$ and $V_2$), and the conditions on this $V_1$ are precisely the conditions on $V$ in Proposition 4.58.}

[Hint: For (b), use Exercise 2.98(b).]

We next indicate, without proof, how $\mathcal{H}$ relates to the classical Hall algebra.

**Definition 4.61.** Let $p$ be a prime. The usual Hall algebra, or what Schiffmann [143, §2.3] calls Steinitz’s classical Hall algebra (see also Macdonald [109, Chap. II]), has $\mathbb{Z}$-basis elements $\{u_\lambda\}_{\lambda \in \operatorname{Par}}$, with the multiplicative structure constants $g_{\mu, \nu}^\lambda(p)$ in

$$u_\mu u_\nu = \sum_{\lambda} g_{\mu, \nu}^\lambda(p) u_\lambda$$

defined as follows: fix a finite abelian $p$-group $L$ of type $\lambda$, meaning that

$$L \cong \bigoplus_{i=1}^{(\ell(\lambda))} \mathbb{Z}/p^{i} \mathbb{Z},$$

and let $g_{\mu, \nu}^\lambda(p)$ be the number of subgroups $M$ of $L$ of type $\mu$, for which the quotient $N := L/M$ is of type $\nu$.

In other words, $g_{\mu, \nu}^\lambda(p)$ counts, for a fixed abelian $p$-group $L$ of type $\lambda$, the number of short exact sequences $0 \to M \to L \to N \to 0$ in which $M, N$ have types $\mu, \nu$, respectively (modulo isomorphism of short exact sequences restricting to the identity on $L$).

We claim that when one takes the finite field $\mathbb{F}_q$ of order $q = p$ a prime, the $\mathbb{Z}$-linear map

$$u_\lambda \longrightarrow \prod_{i=1}^{\ell(\lambda)} L_i$$

gives an isomorphism from this classical Hall algebra to the $\mathbb{Z}$-algebra $\mathcal{H}_\mathbb{Z} \subset \mathcal{H}$. The key point is Hall’s Theorem, a non-obvious statement for which Macdonald includes two proofs in [109, Chap. II], one of them due to Zelevinsky.\footnote{See also [143, Thm. 2.6, Prop. 2.7] for quick proofs of part of it, similar to Zelevinsky’s. Another proof, based on a recent category-theoretical paradigm, can be found in [46, Theorem 3.53].}

We next recall some notions about discrete valuation rings.

**Definition 4.62.** A discrete valuation ring (DVR) $\mathfrak{o}$ is a principal ideal domain having only one maximal ideal $m \neq 0$, with quotient $k = \mathfrak{o}/m$ called its residue field.
Theorem 4.65. Assume for all partitions with respect to its basis (composition series of composition length \(N\), cardinality \(o\)) residue field \(o\) of type \(\lambda\) and that its residue field \(k = o/m = \mathbb{F}[t]/(t^h)\) of \(L\) has an \(\mathbb{F}\)-basis \(\{1, t, t^2, \ldots, t^{h-1}\}\) on which the map \(T\) that multiplies by \(t\) acts as a nilpotent Jordan block of size \(\lambda_i\). Note also that, in this setting, \(o\)-submodules are the same as \(T\)-stable (or \(g\)-stable) \(\mathbb{F}\)-subspaces.

Example 4.64. The ring of \(p\)-adic integers \(o = \mathbb{Z}_p\) is a DVR with maximal ideal \(m = (p)\) and residue field \(k = o/m = \mathbb{Z}_p/p\mathbb{Z}_p \cong \mathbb{F}_p\). An \(o\)-module \(L\) of type \(\lambda\) is an \(\mathbb{F}\)-vector space together with an \(\mathbb{F}\)-linear transformation \(T \in \text{End} \ L\) that acts on \(L\) nilpotently (so that \(g := T + 1\) acts unipotently, where \(1 = \text{id}_L\) with Jordan blocks of sizes given by \(\lambda\): each summand \(o/m^{\lambda_i}\) of \(L\) has an \(\mathbb{F}\)-basis \(\{1, t, t^2, \ldots, t^{\lambda_i-1}\}\) on which the map \(T\) that multiplies by \(t\) acts as a nilpotent Jordan block of size \(\lambda_i\). Note also that, in this setting, \(o\)-submodules are the same as subgroups.

One last notation: \(n(\lambda) := \sum_{i \geq 1} (i - 1)\lambda_i\) for \(\lambda\) in Par. Hall’s Theorem is as follows.

Theorem 4.65. Assume \(o\) is a DVR with maximal ideal \(m\), and that its residue field \(k = o/m\) is finite of cardinality \(q\). Fix an \(o\)-module \(L\) of type \(\lambda\). Then the number of \(o\)-submodules \(M\) of type \(\mu\) for which the quotient \(N = L/M\) is of type \(\nu\) can be written as the specialization

\[
[g_{\mu,\nu}^\lambda(t)]_{t=q}
\]
of a polynomial \(g_{\mu,\nu}^\lambda(t)\) in \(\mathbb{Z}[t]\), called the Hall polynomial.

Furthermore, the Hall polynomial \(g_{\mu,\nu}^\lambda(t)\) has degree at most \(n(\lambda) - (n(\mu) + n(\nu))\), and its coefficient of \(t^{n(\lambda)-n(\mu)-n(\nu)}\) is the Littlewood-Richardson coefficient \(c_{\mu,\nu}^\lambda\).

Comparing what Hall’s Theorem says in Examples 4.63 and 4.64, shows that the map (4.41) gives the desired isomorphism from the classical Hall algebra to \(\mathcal{H}\).

We close this section with some remarks on the vast literature on Hall algebras that we will not discuss here.

Remark 4.66. Macdonald’s version of Hall’s Theorem [109, (4.3)] is stronger than Theorem 4.65, and useful for certain applications: he shows that \(g_{\mu,\nu}^\lambda(t)\) is the zero polynomial whenever the Littlewood-Richardson coefficient \(c_{\mu,\nu}^\lambda\) is zero.

Remark 4.67. In general, not all coefficients of the Hall polynomials \(g_{\mu,\nu}^\lambda(t)\) are nonnegative (see Butler/Hales [27] for a study of when they are): it often happens that \(g_{\mu,\nu}^\lambda(1) = 0\) despite \(g_{\mu,\nu}^\lambda(t)\) not being the zero polynomial\(^\text{183}\). However, in [81, Thm. 4.2], Klein showed that the polynomial values \(g_{\mu,\nu}^\lambda(p)\) for \(p\) prime are always positive when \(c_{\mu,\nu}^\lambda \neq 0\). (This easily yields the same result for \(p\) a prime power.)

Remark 4.68. Zelevinsky in [173, Chaps 10, 11] uses the isomorphism \(\Lambda_C \rightarrow \mathcal{H}\) to derive J. Green’s formula for the value of any irreducible character \(\chi\) of \(GL_n\) on any unipotent class \(J_\lambda\). The answer involves values of irreducible characters of \(S_n\) along with Green’s polynomials \(Q_{\mu}^\lambda(q)\) (see Macdonald [109, §III.7]; they are denoted \(Q(\lambda, \mu)\) by Zelevinsky), which express the images under the isomorphism of Theorem 4.59 of the symmetric function basis \(\{p_\lambda\}\) in terms of the basis \(\{1_{J_\lambda}\}\).

Remark 4.69. The Hall polynomials \(g_{\mu,\nu}^\lambda(t)\) also essentially give the multiplicative structure constants for \(\Lambda(x)[t]\) with respect to its basis of Hall-Littlewood symmetric functions \(P_\lambda = P_\lambda(x; t)\):

\[
P_\mu P_\nu = \sum_{\lambda} t^{n(\lambda) - (n(\mu) + n(\nu))} g_{\mu,\nu}^\lambda(t^{-1}) P_\lambda.
\]

See Macdonald [109, §III.3].

\(^{183}\)Actually, Butler/Hales show in [27, proof of Prop. 2.4] that the values \(g_{\mu,\nu}^\lambda(1)\) are the structure constants of the ring \(\Lambda\) with respect to its basis \(\{m_\lambda\}_{\lambda \in Par}\), we have

\[
m_\mu m_\nu = \sum_{\lambda \in Par} g_{\mu,\nu}^\lambda(1) m_\lambda
\]

for all partitions \(\mu\) and \(\nu\).
Remark 4.70. Schiffmann [143] discusses self-dual Hopf algebras which vastly generalize the classical Hall algebra called Ringel-Hall algebras, associated to abelian categories which are hereditary. Examples come from categories of nilpotent representations of quivers; the quiver having exactly one node and one arc recovers the classical Hall algebra $H_Z$ discussed above.

Remark 4.71. The general linear groups $GL_n(\mathbb{F}_q)$ are one of four families of so-called classical groups. Progress has been made on extending Zelevinsky’s PSH theory to the other families:

(a) Work of Thiem and Vinroot [167] shows that the tower $\{G_*\}$ of finite unitary groups $U_n(\mathbb{F}_{q^2})$ give rise to another positive self-dual Hopf algebra $A = \bigoplus_{n \geq 0} R(U_n(\mathbb{F}_{q^2}))$, in which the role of Harish-Chandra induction is played by Deligne-Lusztig induction. In this theory, character and degree formulas for $U_n(\mathbb{F}_{q^2})$ are related to those of $GL_n(\mathbb{F}_q)$ by substituting $q \mapsto -q$, along with appropriate scalings by $\pm 1$, a phenomenon sometimes called Ennola duality. See also [158, §4].

(b) van Leeuwen [97] has studied $\bigoplus_{n \geq 0} R(Sp_{2n}(\mathbb{F}_q))$, $\bigoplus_{n \geq 0} R(O_{2n}(\mathbb{F}_q))$, and $\bigoplus_{n \geq 0} R(U_n(\mathbb{F}_{q^2}))$ not as Hopf algebras, but rather as so-called twisted PSH-modules over the PSH $A(GL)$ (a “deformed” version of the older notion of Hopf modules). He classified these PSH-modules axiomatically similarly to Zelevinsky’s above classification of PSH’s.

(c) In a recent honors thesis [152], Shelley-Abrahamson defined yet another variation of the concept of Hopf modules, named 2-compatible Hopf modules, and identified $\bigoplus_{n \geq 0} R(Sp_{2n}(\mathbb{F}_q))$ and $\bigoplus_{n \geq 0} R(O_{2n+1}(\mathbb{F}_q))$ as such modules over $A(GL)$. 
5. Quasisymmetric functions and \( P \)-partitions

We discuss here our next important example of a Hopf algebra arising in combinatorics: the quasisymmetric functions of Gessel [60], with roots in work of Stanley [154] on \( P \)-partitions.

5.1. Definitions, and Hopf structure. The definitions of quasisymmetric functions require a totally ordered variable set. Usually we will use a variable set denoted \( x = (x_1, x_2, \ldots) \) with the usual ordering \( x_1 < x_2 < \ldots \). However, it is good to have some flexibility in changing the ordering, which is why we make the following definition.

**Definition 5.1.** Given any totally ordered set \( I \), create a totally ordered variable set \( \{x_i\}_{i \in I} \), and then let \( R(\{x_i\}_{i \in I}) \) denote the power series of bounded degree in \( \{x_i\}_{i \in I} \) having coefficients in \( k \).

The quasisymmetric functions \( \text{QSym}(\{x_i\}_{i \in I}) \) over the alphabet \( \{x_i\}_{i \in I} \) will be the \( k \)-submodule consisting of the elements \( f \in R(\{x_i\}_{i \in I}) \) that have the same coefficient on the monomials \( x_i^{a_1} \cdots x_i^{a_\ell} \) and \( x_j^{a_1} \cdots x_j^{a_\ell} \) whenever both \( i_1 < \cdots < i_\ell \) and \( j_1 < \cdots < j_\ell \) in the total order on \( I \). We write \( \text{QSym}_k(\{x_i\}_{i \in I}) \) instead of \( \text{QSym}(\{x_i\}_{i \in I}) \) to stress the choice of base ring \( k \).

It immediately follows from this definition that \( \text{QSym}(\{x_i\}_{i \in I}) \) is a free \( k \)-submodule of \( R(\{x_i\}_{i \in I}) \), having as \( k \)-basis elements the monomial quasisymmetric functions

\[
M_\alpha(\{x_i\}_{i \in I}) := \sum_{i_1 < \cdots < i_\ell \text{ in } I} x_i^{a_1} \cdots x_i^{a_\ell}
\]

for all compositions \( \alpha \) satisfying \( \ell(\alpha) \leq |I| \). When \( I \) is infinite, this means that the \( M_\alpha \) for all compositions \( \alpha \) form a basis of \( \text{QSym}(\{x_i\}_{i \in I}) \).

Note that \( \text{QSym}(\{x_i\}_{i \in I}) = \bigoplus_{n \geq 0} \text{QSym}_n(\{x_i\}_{i \in I}) \) is a graded \( k \)-module of finite type, where \( \text{QSym}_n(\{x_i\}_{i \in I}) \) is the \( k \)-submodule of quasisymmetric functions which are homogeneous of degree \( n \). Letting \( \text{Comp} \) denote the set of all compositions \( \alpha \), and \( \text{Comp}_n \) the compositions \( \alpha \) of \( n \) (that is, compositions whose parts sum to \( n \)), the subset \( \{M_\alpha\}_{\alpha \in \text{Comp}_n : \ell(\alpha) \leq |I|} \) gives a \( k \)-basis for \( \text{QSym}_n(\{x_i\}_{i \in I}) \).

**Example 5.2.** Taking the variable set \( x = (x_1 < x_2 < \cdots) \) to define \( \text{QSym}(x) \), for \( n = 0, 1, 2, 3 \), one has these basis elements in \( \text{QSym}_n(x) \):

\[
M(1) = M_0 = 1 \quad M(1,1,1) = x_1 x_2 x_3 + x_1 x_2 x_4 + x_1 x_3 x_4 + \cdots = m(1,1,1) = e_3
\]

It is not obvious that \( \text{QSym}(x) \) is a subalgebra of \( R(x) \), but we will show this momentarily. For example,

\[
M_\alpha(\{x_i\}_{i \in I}) M_{\beta}(\{x_i\}_{i \in I}) = (x_1^a + x_2^a + x_3^a + \cdots)(x_1^b + x_2^b + x_3^b + \cdots)
\]

It is not obvious that \( \text{QSym}(x) \) is a subalgebra of \( R(\{x_i\}_{i \in I}) \), with multiplication in the \( \{M_\alpha\} \)-basis as follows: Fix three pairwise disjoint chain posets \( (i_1 < \cdots < i_\ell) \) and \( (j_1 < \cdots < j_m) \), and \( (k_1 < k_2 < \cdots) \). Now, if \( \alpha = (\alpha_1, \ldots, \alpha_\ell), \beta = (\beta_1, \ldots, \beta_m) \) then

\[
M_\alpha M_\beta = \sum_f M_{\alpha f} M_{\beta f}
\]

where the sum is over all \( p \in \mathbb{N} \) and all maps \( f \) from the disjoint union of two chains to a chain

\[
(i_1 < \cdots < i_\ell) \cup (j_1 < \cdots < j_m) \xrightarrow{f} (k_1 < \cdots < k_p)
\]
which are both surjective and strictly order-preserving \((x < y)\) implies \(f(x) < f(y)\), and where the composition \(wt(f) := (wt_1(f), \ldots, wt_p(f))\) is defined by \(wt_s(f) := \sum_{i_u \in f^{-1}(k_u)} \alpha_u + \sum_{j_v \in f^{-1}(k_v)} \beta_v\).

**Example 5.4.** For this example, set \(\alpha = (2, 1)\) and \(\beta = (3, 4, 2)\). Let us compute \(M_\alpha M_\beta\) using (5.1). Indeed, the length of \(\alpha\) is \(\ell = 2\), and the length of \(\beta\) is \(m = 3\), so the sum on the right hand side of (5.1) is over all \(p \in \mathbb{N}\) and all surjective strictly order-preserving maps \(f\) from the disjoint union \((i_1 < j_2) \sqcup (j_1 < j_2 < j_3)\) of two chains to the chain \((k_1 < k_2 < \cdots < k_p)\). Such maps can exist only when \(p \leq 5\) (due to having to be surjective) and only for \(p \geq 3\) (since, being strictly order-preserving, they have to be injective when restricted to \((j_1 < j_2 < j_3)\)). Hence, enumerating them is a finite problem. The reader can check that the value obtained for \(M_\alpha M_\beta\) is

\[
M_{(2,1,3,4,2)} + M_{(2,3,1,4,2)} + M_{(2,3,4,1,2)} + M_{(2,3,4,2,1)} + M_{(3,2,1,4,2)} \\
+ M_{(3,2,4,1,2)} + M_{(3,4,2,1,2)} + M_{(3,4,2,2,1)} + M_{(4,3,2,2,1)} \\
+ M_{(2,3,4,3)} + M_{(2,3,5,2)} + M_{(2,4,4,2)} + M_{(3,2,4,3)} + M_{(3,2,5,2)} + M_{(3,4,2,3)} \\
+ M_{(3,4,4,1)} + M_{(3,6,1,2)} + M_{(3,6,2,1)} + M_{(5,1,4,2)} + M_{(5,4,1,2)} + M_{(5,4,2,1)} \\
+ M_{(5,4,3)} + M_{(5,5,2)} + M_{(3,6,3)}.
\]

Here, we have listed the addends corresponding to \(p = 5\) on the first two rows, the addends corresponding to \(p = 4\) on the next two rows, and those corresponding to \(p = 3\) on the third row. The reader might notice that the first two rows (i.e., the addends with \(p = 5\)) are basically a list of shuffles of \(\alpha\) and \(\beta\). In general, the maps (5.2) for \(p = \ell + m\) are in bijection with the elements of \(\text{Sh}_{\ell+m}\)

\footnote{The bijection takes a map \(f\) to the inverse of the permutation \(\sigma \in S_p\) which sends every \(x \in \{1, 2, \ldots, \ell\}\) to the index \(y\) satisfying \(f(i_x) = k_y\), and sends every \(x \in \{\ell + 1, \ell + 2, \ldots, \ell + m\}\) to the index \(y\) satisfying \(f(j_{x-\ell}) = k_y\).}

and the corresponding compositions \(wt\) are the shuffles of \(\alpha\) and \(\beta\). Therefore the name “overlapping shuffle product”.

**Proof of Proposition 5.3.** Formula (5.1) comes from considering a typical product of two monomials in the expansion of \(M_\alpha M_\beta\):

\[
(x_{i_1} x_{i_2} \cdots x_{i_\ell}) (x_{j_1} x_{j_2} \cdots x_{j_m}) = x_{k_1} x_{k_2} \cdots x_{k_p}
\]

for subscripts sequences \(i_1 < \cdots < i_\ell\) and \(j_1 < \cdots < j_m\) and \(k_1 < \cdots < k_p\) with

\[
\{k_1, \ldots, k_p\} = \{i_1, \ldots, i_\ell\} \sqcup \{j_1, \ldots, j_m\}.
\]

Thinking of \(\{i_s\}\) and \(\{j_t\}\) as disjoint sets, multiplication gives a surjective map \(f\) as in (5.2), with \(\gamma_s = wt_s(f)\). □

The multiplication rule (5.1) shows that the \(k\)-algebra \(QSym_{\ell}((x_i)_{i \in I})\) does not depend much on \(I\), as long as \(I\) is infinite. More precisely, all such \(k\)-algebras are mutually isomorphic. We can use this to define a \(k\)-algebra of quasisymmetric functions without any reference to \(I\):

**Definition 5.5.** Let \(QSym\) be the \(k\)-algebra defined as having \(k\)-basis \(\{Q_{\alpha}\}_{\alpha \in \text{Comp}}\) and with multiplication defined \(k\)-linearly by (5.1). This is called the \(k\)-algebra of quasisymmetric functions. We write \(QSym_k\) instead of \(QSym\) to stress the choice of base ring \(k\).

The \(k\)-algebra \(QSym\) is graded, and its \(n\)-th graded component \(QSym_n\) has \(k\)-basis \(\{Q_{\alpha}\}_{\alpha \in \text{Comp}_n}\).

For every infinite totally ordered set \(I\), the \(k\)-algebra \(QSym_{\ell}\) is isomorphic to the \(k\)-algebra \(QSym((x_i)_{i \in I})\). The isomorphism sends \(M_n \mapsto M_n((x_i)_{i \in I})\).

In particular, we obtain the isomorphism \(QSym \cong QSym_{\infty}\) for \(x\) being the infinite chain \((x_1 < x_2 < x_3 < \cdots)\). We will identify \(QSym\) with \(QSym_{\infty}\) along this isomorphism. This allows us to regard quasisymmetric functions either as power series in a specific set of variables (“alphabet”), or as formal linear combinations of \(M_n\)’s, whatever is more convenient.

For any infinite alphabet \((x_i)_{i \in I}\) and any \(f \in QSym\), we denote by \(f((x_i)_{i \in I})\) the image of \(f\) under the algebra isomorphism \(QSym \to QSym((x_i)_{i \in I})\) defined in Definition 5.5.

The comultiplication of \(QSym\) will extend the one that we defined for \(\Lambda\), but we need to take care about the order of the variables this time. We consider the linear order from (2.12) on two sets of variables \((x, y) = (x_1 < x_2 < \cdots < y_1 < y_2 < \cdots)\), and we embed the \(k\)-algebra \(QSym(x) \otimes QSym(y)\) into the
This k-algebra \( R(x, y) \) by identifying every \( f \otimes g \in \text{QSym}(x) \otimes \text{QSym}(y) \) with \( fg \in R(x, y) \) (this embedding is indeed injective\(^{185}\)). It can then be seen that

\[
\text{QSym}(x, y) \subset \text{QSym}(x) \otimes \text{QSym}(y)
\]

(where the right hand side is viewed as k-subalgebra of \( R(x, y) \) via said embedding)\(^{186}\), so that one can define \( \text{QSym} \xrightarrow{\Delta} \text{QSym} \otimes \text{QSym} \) as the composite of the maps in the bottom row here:

\[
\begin{align*}
\text{QSym} & \cong R(x, y) & = & R(x, y) \\
\Delta & \cong \text{QSym}(x, y) & \\n\beta \cdot f & \to f(x, y) = f(x_1, x_2, \ldots, y_1, y_2, \ldots)
\end{align*}
\]

(Recall that \( f(x, y) \) is formally defined as the image of \( f \) under the algebra isomorphism \( \text{QSym} \to \text{QSym}(x, y) \) defined in Definition 5.5.)

**Example 5.6.** For example,

\[
\Delta M_{(a,b,c)} = M_{(a,b,c)}(x_1, x_2, \ldots, y_1, y_2, \ldots) \\
= x_1^a y_1^b z_1^c + x_1^a y_2^b z_1^c + \cdots \\
+ x_1^a y_1^b y_2^c + x_1^a y_1^b y_3^c + \cdots \\
+ y_1^a y_2^b y_3^c + y_1^a y_2^b y_4^c + \cdots \\
= M_{(a,b,c)}(x_1) + M_{(a,b)}(x_1) M_{(c)}(y_1) + M_{(a)}(x_1) M_{(b,c)}(y_1) + M_{(a,b,c)}(y_1)
\]

Defining the concatenation \( \beta \cdot \gamma \) of two compositions \( \beta = (\beta_1, \ldots, \beta_r), \gamma = (\gamma_1, \ldots, \gamma_s) \) to be the composition \( (\beta_1, \ldots, \beta_r, \gamma_1, \ldots, \gamma_s), \) one has the following description of the coproduct in the \( \{M_{a}\} \) basis.

**Proposition 5.7.** For a composition \( \alpha = (\alpha_1, \ldots, \alpha_{\ell}) \), one has

\[
\Delta M_{\alpha} = \sum_{k=0}^{\ell} M_{(\alpha_1, \ldots, \alpha_k)} \otimes M_{(\alpha_{k+1}, \ldots, \alpha_{\ell})} = \sum_{(\beta, \gamma) \colon \beta \cdot \gamma = \alpha} M_{\beta} \otimes M_{\gamma}
\]

**Proof.** This comes from expressing a monomial in \( \Delta M_{\alpha} = M_{\alpha}(x, y) \) uniquely in the form \( x_{i_1}^{\alpha_{i_1}} \cdots x_{i_k}^{\alpha_{i_k}} y_{j_1}^{\alpha_{j_1}} \cdots y_{j_{\ell-k}}^{\alpha_{j_{\ell-k}}} \) for some \( k \in \{0, 1, \ldots, n\} \) and some subscripts \( i_1 < \cdots < i_k \) and \( j_1 < \cdots < j_{\ell-k} \). \qed

**Proposition 5.8.** The quasisymmetric functions QSym form a graded connected Hopf algebra of finite type, which is commutative, and contains the symmetric functions \( \Lambda \) as a Hopf subalgebra.

**Proof.** To prove coassociativity of \( \Delta \), we need to be slightly careful. It seems reasonable to argue by \( (\Delta \otimes \text{id}) \Delta f = (\text{id} \otimes \Delta) \Delta f \) as in the case of \( \Lambda \), but this would now require further justification, as terms like \( f(x, y) \) and \( f(x, y, z) \) are no longer directly defined as evaluations of \( f \) on some sequences (but rather are defined as images of \( f \) under certain homomorphisms). However, it is very easy to see that \( \Delta \) is

\(^{185}\)This is because it sends the basis elements \( M_{\beta}(x) \otimes M_{\gamma}(y) \) of the former k-algebra to the linearly independent power series \( M_{\beta}(x) M_{\gamma}(y) \).

\(^{186}\)This is not completely obvious, but can be easily checked by verifying that \( M_{\alpha}(x, y) = \sum_{(\beta, \gamma) \colon \beta \cdot \gamma = \alpha} M_{\beta}(x) \otimes M_{\gamma}(y) \) for every composition \( \alpha \) (see the proof of Proposition 5.7 for why this holds).
coassociative by checking \((\Delta \otimes \text{id}) \circ \Delta = (\text{id} \otimes \Delta) \circ \Delta\) on the \(\{M_{\alpha}\}\) basis: Proposition 5.7 yields

\[
((\Delta \otimes \text{id}) \circ \Delta) M_{\alpha} = \sum_{k=0}^{\ell} \Delta(M_{(\alpha_1, \ldots, \alpha_k)}) \otimes M_{(\alpha_{k+1}, \ldots, \alpha_{\ell})}
\]

\[
= \sum_{k=0}^{\ell} \left( \sum_{i=0}^{k} M_{(\alpha_1, \ldots, \alpha_i)} \otimes M_{(\alpha_{i+1}, \ldots, \alpha_k)} \right) \otimes M_{(\alpha_{k+1}, \ldots, \alpha_{\ell})}
\]

\[
= \sum_{k=0}^{\ell} \sum_{i=0}^{k} M_{(\alpha_1, \ldots, \alpha_i)} \otimes M_{(\alpha_{i+1}, \ldots, \alpha_k)} \otimes M_{(\alpha_{k+1}, \ldots, \alpha_{\ell})}
\]

and the same expression for \(((\text{id} \otimes \Delta) \circ \Delta) M_{\alpha}\).

The coproduct \(\Delta\) of \(\text{QSym}\) is an algebra morphism because it is defined as a composite of algebra morphisms in the bottom row of (5.3). To prove that the restriction of \(\Delta\) to the subring \(\Lambda\) of \(\text{QSym}\) is the comultiplication of \(\Lambda\), it thus is enough to check that it sends the elementary symmetric function \(e_n\) to \(\sum_{i=0}^{n} e_i \otimes e_{n-i}\) for every \(n \in \mathbb{N}\). This again follows from Proposition 5.7, since \(e_n = M_{(1,1,\ldots,1)}\) (with \(n\) times 1).

The counit is as usual for a graded connected coalgebra, and just as in the case of \(\Lambda\), sends a quasisymmetric function \(f(x)\) to its constant term \(f(0,0,\ldots)\). This is an evaluation, and hence an algebra map. Hence \(\text{QSym}\) forms a bialgebra, and as it is graded and connected, also a Hopf algebra by Proposition 1.36. It is clearly of finite type and contains \(\Lambda\) as a Hopf subalgebra.

We will identify the antipode in \(\text{QSym}\) shortly, but we first deal with another slightly subtle issue. In addition to the counit evaluation \(\epsilon(f) = f(0,0,\ldots)\), starting in Section 7.1, we will want to specialize elements in \(\text{QSym}(x)\) by making other variable substitutions, in which all but a finite list of variables are set to zero. We justify this here.

**Proposition 5.9.** Fix a totally ordered set \(I\), a commutative \(k\)-algebra \(A\), a finite list of variables \(x_{i_1}, \ldots, x_{i_m}\), say with \(i_1 < \ldots < i_m\) in \(I\), and an ordered list of elements \((a_1, \ldots, a_m) \in A^m\).

Then there is a well-defined evaluation homomorphism

\[
\text{QSym}(\{x_{i}\}_{i \in I}) \rightarrow A
\]

\[
f \rightarrow [f]_{x_{i_1} = a_1, \ldots, x_{i_m} = a_m, x_j = 0 \text{ for } j \notin \{i_1, \ldots, i_m\}}
\]

Furthermore, this homomorphism depends only upon the list \((a_1, \ldots, a_m)\), as it coincides with the following:

\[
\text{QSym}(\{x_{i}\}_{i \in I}) \cong \text{QSym}(x_1, x_2, \ldots) \rightarrow A
\]

\[
f(x_1, x_2, \ldots) \rightarrow f(a_1, \ldots, a_m, 0, 0, \ldots)
\]

(This latter statement is stated for the case when \(I\) is infinite; otherwise, read “\(x_{1,2,\ldots,x_{|I|}}\)” for “\(x_1, x_2, \ldots, x_{|I|}\),” and interpret \((a_1, \ldots, a_m, 0, 0, \ldots)\) as an \(|I|\)-tuple.)

**Proof.** One already can make sense of evaluating \(x_{i_1} = a_1, \ldots, x_{i_m} = a_m\) and \(x_j = 0\) for \(j \notin \{i_1, \ldots, i_m\}\) in the ambient ring \(R(\{x_{i}\}_{i \in I})\) containing \(\text{QSym}(\{x_{i}\}_{i \in I})\), since a power series \(f\) of bounded degree will have finitely many monomials that only involve the variables \(x_{i_1}, \ldots, x_{i_m}\). The last assertion follows from quasisymmetry of \(f\), and is perhaps checked most easily when \(f = M_{\alpha}(\{x_{i}\}_{i \in I})\) for some \(\alpha\).

The antipode in \(\text{QSym}\) has a reasonably simple expression in the \(\{M_{\alpha}\}\) basis, but requiring a definition.

**Definition 5.10.** For \(\alpha, \beta\) in \(\text{Comp}_n\), say that \(\alpha\) refines \(\beta\) or \(\beta\) coarsens \(\alpha\) if, informally, one can obtain \(\beta\) from \(\alpha\) by combining some of its adjacent parts. Alternatively, one has a bijection \(\text{Comp}_n \rightarrow 2^{[n-1]}\) where \([n-1] := \{1,2,\ldots,n-1\}\) which sends \(\alpha = (\alpha_1, \ldots, \alpha_\ell)\) having length \(\ell(\alpha) = \ell\) to its subset of partial sums

\[
D(\alpha) := \{\alpha_1, \alpha_1 + \alpha_2, \ldots, \alpha_1 + \cdots + \alpha_{\ell-1}\}
\]

and this sends the refinement ordering to the inclusion ordering on the Boolean algebra \(2^{[n-1]}\) (to be more precise: a composition \(\alpha\) refines a composition \(\beta\) if and only if \(D(\alpha) \supset D(\beta)\)). There is also a bijection sending \(\alpha\) to its ribbon diagram: the skew diagram \(\lambda/\mu\) having rows of sizes \((\alpha_1, \ldots, \alpha_\ell)\) read from bottom to
top with exactly one column of overlap between adjacent rows. These bijections and the refinement partial order are illustrated here for $n = 4$:

Given $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, its reverse composition is $\text{rev}(\alpha) = (\alpha_\ell, \alpha_{\ell-1}, \ldots, \alpha_2, \alpha_1)$. Note that $\alpha \mapsto \text{rev}(\alpha)$ is a poset automorphism for the refinement ordering.

**Theorem 5.11.** For any composition $\alpha$ in $\text{Comp}$,

$$S(M_\alpha) = (-1)^{\ell(\alpha)} \sum_{\gamma \in \text{Comp}: \gamma \text{ coarsens } \text{rev}(\alpha)} M_\gamma.$$

For example,

$$S(M_{(a,b,c)}) = - (M_{(c,b,a)} + M_{(b+c,a)} + M_{(c,a+b)} + M_{(a+b+c)}).$$

**Proof.** We give Ehrenborg’s proof\(^{187}\) [48, Prop. 3.4] via induction on $\ell = \ell(\alpha)$. One has easy base cases when $\ell(\alpha) = 0$, where $S(M_\emptyset) = S(1) = 1 = (-1)^0 M_{\text{rev}(\emptyset)}$, and when $\ell(\alpha) = 1$, where $M_\alpha$ is primitive by Proposition 5.7, so Proposition 1.37 shows $S(M_\alpha) = -M_\alpha = (-1)^1 M_{\text{rev}(\alpha)}$.

For the inductive step, apply the inductive definition of $S$ from the proof of Proposition 1.36:

$$S(M_{(\alpha_1, \ldots, \alpha_\ell)}) = - \sum_{i=0}^{\ell-1} S(M_{(\alpha_1, \ldots, \alpha_i)}) M_{(\alpha_{i+1}, \ldots, \alpha_\ell)}$$

$$= \sum_{i=0}^{\ell-1} \sum_{\beta \text{ coarsening } (\alpha_1, \alpha_{i+1}, \ldots, \alpha_\ell)} (-1)^{i+1} M_\beta M_{(\alpha_{i+1}, \ldots, \alpha_\ell)}.$$

The idea will be to cancel terms of opposite sign that appear in the expansions of the products $M_\beta M_{(\alpha_{i+1}, \ldots, \alpha_\ell)}$. Note that each composition $\beta$ appearing above has first part $\beta_1$ of the form $\alpha_i + \alpha_{i-1} + \cdots + \alpha_h$ for some $h \leq i$ (unless $\beta = \emptyset$), and hence each term $M_\beta$ in the expansion of the product $M_\beta M_{(\alpha_{i+1}, \ldots, \alpha_\ell)}$ has $\gamma_1$ (that is, the first entry of $\gamma$) a sum that can take one of these three forms:

- $\alpha_i + \alpha_{i-1} + \cdots + \alpha_h$,
- $\alpha_{i+1} + (\alpha_i + \alpha_{i-1} + \cdots + \alpha_h)$,
- $\alpha_{i+1}$.

Say that the type of $\gamma$ is $i$ in the first case, and $i + 1$ in the second two cases\(^{188}\); in other words, the type is the largest subscript $k$ on a part $\alpha_k$ which was combined in the sum $\gamma_1$. It is not hard to see that a given $\gamma$ for which the type $k$ is strictly smaller than $\ell$ arises from exactly two pairs $(\beta, \gamma), (\beta', \gamma)$, having opposite signs $(-1)^k$ and $(-1)^{k+1}$ in the above sum\(^{189}\). For example, if $\alpha = (\alpha_1, \ldots, \alpha_8)$, then the composition

\(^{187}\)A different proof was given by Malvenuto and Reutenauer [113, Cor. 2.3], and is sketched in Remark 5.36 below.

\(^{188}\)We imagine that we label the terms obtained by expanding $M_\beta M_{(\alpha_{i+1}, \ldots, \alpha_\ell)}$ by distinct labels, so that each term knows how exactly it was created (i.e., which $i$, which $\beta$ and which map $f$ as in (5.2) gave rise to it). Strictly speaking, it is these triples $(i, \beta, f)$ that we should be assigning types to, not terms.

\(^{189}\)Strictly speaking, this means that we have an involution on the set of our $(i, \beta, f)$ triples having type smaller than $\ell$, and this involution switches the sign of $(-1)^k M_{w_{i,f}}$. 


Thus one can cancel almost all the terms, excepting those with \(\gamma\) 
\(\beta = (\alpha_6 + \alpha_5 + \alpha_4, \alpha_3, \alpha_7, \alpha_8 + \alpha_2 + \alpha_1)\) of type 6 can arise from either of 
\(\beta' = (\alpha_4 + \alpha_3, \alpha_2 + \alpha_1)\) with \(i = 6 \) and sign \((-1)^7\)

Similarly, \(\gamma = (\alpha_6, \alpha_5 + \alpha_4, \alpha_3, \alpha_7, \alpha_8 + \alpha_2 + \alpha_1)\) can arise from either of 
\(\beta = (\alpha_6, \alpha_5 + \alpha_4, \alpha_3, \alpha_2 + \alpha_1)\) with \(i = 6 \) and sign \((-1)^7\)
\(\beta' = (\alpha_5 + \alpha_4, \alpha_3, \alpha_2 + \alpha_1)\) with \(i = 5 \) and sign \((-1)^6\).

Thus one can cancel almost all the terms, excepting those with \(\gamma\) of type \(\ell\) among the terms \(M_\gamma\) in the expansion of the last \((i = \ell - 1)\) summand \(M_\beta M_{\alpha_1}\). A bit of thought shows that these are the \(\gamma\) coarsening rev(\(\alpha\)), and all have sign \((-1)^\ell\). \(\square\)

5.2. The fundamental basis and \(P\)-partitions. There is a second important basis for QSym which arose originally in Stanley’s \(P\)-partition theory [154].

Definition 5.12. A labelled poset will here mean a partially ordered set \(P\) whose underlying set is some finite subset of the integers. A \(P\)-partition is a function \(P \mapsto \{1, 2, \ldots\}\) with the property that

- \(i < P j\) and \(i < Z j\) implies \(f(i) \leq f(j)\), and
- \(i < P j\) and \(i > Z j\) implies \(f(i) < f(j)\).

Denote by \(A(P)\) the set of all \(P\)-partitions \(f\), and let \(F_P(x) := \sum_{f \in A(P)} x_f\) where \(x_f := \Pi_{i \in P} x_{f(i)}\). This \(F_P(x)\) is an element of \(k[[x]] := k[[x_1, x_2, \ldots]]\).

Example 5.13. Depicted is a labelled poset \(P\), along with the relations among the four values \(f = (f(1), f(2), f(3), f(4))\) that define its \(P\)-partitions \(f\):

\[
\begin{array}{ccc}
4 & 2 & f(4) \\
3 & 1 & f(3) \\
& f(2) & f(1)
\end{array}
\]

Remark 5.14. Stanley’s treatment of \(P\)-partitions in [157, §3.15 and §7.19] uses a language different from ours. First, Stanley works not with labelled posets \(P\), but with pairs \((P, \omega)\) of a poset \(P\) and a bijective labelling \(\omega : P \to [n]\). Thus, the relation \(<_P\) is not given on \(P\) a priori, but has to be pulled back from \([n]\) using \(\omega\) (and it depends on \(\omega\), whence Stanley speaks of “\((P, \omega)\)-partitions”). Furthermore, Stanley uses the notations \(F_P\) and \(F_{P, \omega}\) for something different from what we denote by \(F_P\), whereas what we call \(F_P\) is dubbed \(K_{P, \omega}\) in [157, §7.19].

The so-called fundamental quasisymmetric functions are an important special case of the \(F_P(x)\). We shall first define them directly and then see how they are obtained as \(P\)-partition enumerators \(F_P(x)\) for some special labelled posets \(P\).

Definition 5.15. Let \(n \in \mathbb{N}\) and \(\alpha \in \text{Comp}_n\). We define the fundamental quasisymmetric function \(L_\alpha = L_\alpha(x) \in \text{QSym}\) by

\[
L_\alpha := \sum_{\beta \in \text{Comp}_n: \beta \text{ refines } \alpha} M_\beta.
\]

Example 5.16. The extreme cases for \(\alpha\) in \(\text{Comp}_n\) give quasisymmetric functions \(L_\alpha\) which are symmetric:

\[
\begin{align*}
L_{(1^n)} &= M_{(1^n)} = e_n, \\
L_{(n)} &= \sum_{\alpha \in \text{Comp}_n} M_\alpha = h_n
\end{align*}
\]

---

190 See [61] for a history of \(P\)-partitions; our notations, however, strongly differ from those in [61].
One sees that the \( \{L_{\alpha}\}_{\alpha \in \text{Comp}} \) are a \( k \)-basis for \( \text{QSym} \), as inclusion-exclusion applied to (5.4) gives

\[
M_\alpha = \sum_{\beta \in \text{Comp}_n : \beta \text{ refines } \alpha} (-1)^{l(\beta) - l(\alpha)} L_\beta.
\]

**Proposition 5.17.** Let \( n \in \mathbb{N} \). Let \( \alpha \) be a composition of \( n \). Let \( I \) be an infinite totally ordered set. Then,

\[
L_\alpha (\{x_i\}_{i \in I}) = \sum_{i_1 \leq i_2 \leq \cdots \leq i_n \in I; \quad i_j < i_{j+1} \text{ if } j \in D(\alpha)} x_{i_1} x_{i_2} \cdots x_{i_n},
\]

where \( L_\alpha (\{x_i\}_{i \in I}) \) is defined as the image of \( L_\alpha \) under the isomorphism \( \text{QSym} \rightarrow \text{QSym} (\{x_i\}_{i \in I}) \) obtained in Definition 5.5. In particular, for the standard (totally ordered) variable set \( x = (x_1 < x_2 < \ldots) \), we obtain

\[
L_\alpha = L_\alpha (x) = \sum_{(1 \leq i_1 \leq i_2 \leq \cdots \leq i_n; \quad i_j < i_{j+1} \text{ if } j \in D(\alpha)} x_{i_1} x_{i_2} \cdots x_{i_n}.
\]

**Proof.** Every composition \( \beta = (\beta_1, \ldots, \beta_\ell) \) of \( n \) satisfies

\[
M_\beta (\{x_i\}_{i \in I}) = \sum_{k_1 < \cdots < k_\ell \in I} x_{\beta_1}^{k_1} \cdots x_{\beta_\ell}^{k_\ell} = \sum_{i_1 \leq i_2 \leq \cdots \leq i_n \in I; \quad i_j < i_{j+1} \text{ if and only if } j \in D(\beta)} x_{i_1} x_{i_2} \cdots x_{i_n}.
\]

Applying the ring homomorphism \( \text{QSym} \rightarrow \text{QSym} (\{x_i\}_{i \in I}) \) to (5.4), we obtain

\[
L_\alpha (\{x_i\}_{i \in I}) = \sum_{\beta \in \text{Comp}_n : \beta \text{ refines } \alpha} M_\beta (\{x_i\}_{i \in I}) = \sum_{\beta \in \text{Comp}_n : \beta \text{ refines } \alpha} \sum_{i_1 \leq i_2 \leq \cdots \leq i_n \in I; \quad i_j < i_{j+1} \text{ if and only if } j \in D(\beta)} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{\beta \in \text{Comp}_n : \beta \text{ refines } \alpha} \sum_{i_1 \leq i_2 \leq \cdots \leq i_n \in I; \quad i_j < i_{j+1} \text{ if and only if } j \in D(\beta)} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{\beta \in \text{Comp}_n : \beta \text{ refines } \alpha} \sum_{i_1 \leq i_2 \leq \cdots \leq i_n \in I; \quad i_j < i_{j+1} \text{ if and only if } j \in D(\beta)} x_{i_1} x_{i_2} \cdots x_{i_n}.
\]

\[\square\]

**Proposition 5.18.** When the labelled poset \( P \) is a total or linear order \( w = (w_1 < \ldots < w_n) \), the generating function \( F_w(x) \) depends only upon the descent set

\[
\text{Des}(w) := \{i : w_i > w_{i+1}\} \subset \{1, 2, \ldots, n - 1\}
\]

and its associated composition \( \alpha \in \text{Comp}_n \) having partial sums \( D(\alpha) = \text{Des}(w) \): one has that \( F_w(x) \) equals the fundamental quasisymmetric function \( L_\alpha \).

E.g., total order \( w = 35142 \) has \( \text{Des}(w) = \{2, 4\} \) and composition \( \alpha = (2, 2, 1) \), so

\[
F_{35142} (x) = \sum_{f(3) \leq f(5) < f(1) \leq f(4) < f(2)} x_f(3) x_f(5) x_f(1) x_f(4) x_f(2) = \sum_{i_1 \leq i_2 < i_3 \leq i_4 < i_5} x_{i_1} x_{i_2} x_{i_3} x_{i_4} x_{i_5} = L_{(2,2,1)} = M_{(2,2,1)} + M_{(1,1,2,1)} + M_{(1,1,1,1,1,1)}.
\]

**Proof of Proposition 5.18.** Write \( F_w(x) \) as a sum of monomials \( x_{f(w_1)} \cdots x_{f(w_n)} \) over all \( P \)-partitions \( f \). These \( P \)-partitions are exactly the sequences \( f(w_1) \leq \cdots \leq f(w_n) \) having strict inequalities \( f(w_i) < f(w_{i+1}) \) whenever \( i \) is in \( \text{Des}(w) \). Hence, comparison with (5.5) reveals that this sum equals \( L_\alpha \). \[\square\]

The next proposition ([157, Cor. 7.19.5], [107, Cor. 3.3.24]) is an algebraic shadow of Stanley’s main lemma [157, Thm. 7.19.4] in \( P \)-partition theory. It expands any \( F_P(x) \) in the \( \{L_\alpha\} \) basis, as a sum over the set \( \mathcal{L}(P) \) of all linear extensions \( w \) of \( P \).

\[191\]Let us explain what we mean by linear extensions and how we represent them.
Theorem 5.19. For any labelled poset $P$,

$$F_P(x) = \sum_{w \in \mathcal{L}(P)} F_w(x).$$

Proof. We give Gessel’s proof [60, Thm. 1], via induction on the number of pairs $i, j$ which are incomparable in $P$. When this quantity is 0, then $P$ is itself a linear order $w$, so that $\mathcal{L}(P) = \{w\}$ and there is nothing to prove.

In the inductive step, let $i, j$ be incomparable elements. Consider the two posets $P_{<j}$ and $P_{<i}$ which are obtained from $P$ by adding in an order relation between $i$ and $j$, and then taking the transitive closure; it is not hard to see that these transitive closures cannot contain a cycle, so that these really do define two posets. The result then follows by induction applied to $P_{<j}, P_{<i}$, once one notices that $\mathcal{L}(P) = \mathcal{L}(P_{<j}) \sqcup \mathcal{L}(P_{<i})$ since every linear extension $w$ of $P$ either has $i$ before $j$ or vice-versa, and $A(P) = A(P_{<j}) \sqcup A(P_{<i})$ since, assuming that $i <_\mathbb{Z} j$ without loss of generality, every $f$ in $A(P)$ either satisfies $f(i) \leq f(j)$ or $f(i) > f(j)$. \[ \square \]

Example 5.20. To illustrate the induction in the above proof, consider the poset $P$ from Example 5.13, having $\mathcal{L}(P) = \{3124, 3142, 3412\}$. Then choosing as incomparable pair $(i, j) = (1, 4)$, one has

\[
\begin{align*}
P_{<j} &= (4, 3, 2, 1) \quad \text{with } \mathcal{L}(P_{<j}) = \{3124, 3142\} \\
P_{<i} &= (2, 4, 1, 3) \quad \text{with } \mathcal{L}(P_{<i}) = \{3412\}
\end{align*}
\]


[Hint: For every $f : P \to \{1, 2, 3, \ldots\}$, we can define a binary relation $\prec_f$ on the set $P$ by letting $i \prec_f j$ hold if and only if

$$(f(i) < f(j) \text{ or } f(i) = f(j) \text{ and } i <_\mathbb{Z} j).$$

Show that this binary relation $\prec_f$ is (the smaller relation of) a total order. When $f$ is a $P$-partition, then endowing the set $P$ with this total order yields a linear extension of $P$. Use this to show that the set $A(P)$ is the union of its pairwise disjoint subsets $A(w)$ with $w \in \mathcal{L}(P)$.]

We next wish to describe the structure maps for the Hopf algebra QSym in the basis $\{L_\alpha\}$ of fundamental quasisymmetric functions. For this purpose, two more definitions are useful.

If $P$ is a finite poset, then a linear extension of $P$ denotes a total order $w$ on the set $P$ having the property that $i <_P j$ (for two elements $i$ and $j$ of $P$) implies $i <_w j$. (In other words, it is a linear order on the ground set $P$ which extends $P$ as a poset; therefore the name.) We identify such a total order $w$ with the list $(p_1, p_2, \ldots, p_n)$ containing all elements of $P$ in $w$-increasing order (that is, $p_1 <_w p_2 <_w \cdots <_w p_n$).

(Stanley, in [157, §3.5], defines linear extensions in a slightly different way: For him, a linear extension of a finite poset $P$ is an order-preserving bijection from $P$ to the subposet $\{1, 2, \ldots, |P|\}$ of $\mathbb{Z}$. But this is equivalent to our definition, since a bijection like this can be used to transport the order relation of $\{1, 2, \ldots, |P|\}$ back to $P$, thus resulting in a total order on $P$ which is a linear extension of $P$ in our sense.)
Definition 5.22. Given two nonempty compositions $\alpha = (\alpha_1, \ldots, \alpha_\ell), \beta = (\beta_1, \ldots, \beta_m)$, their near-concatenation is

$$\alpha \odot \beta := (\alpha_1, \ldots, \alpha_{\ell - 1}, \alpha_\ell + \beta_1, \beta_2, \ldots, \beta_m)$$

For example, the figure below depicts for $\alpha = (1, 3, 3)$ (black squares) and $\beta = (4, 2)$ (white squares) the concatenation and near-concatenation as ribbons:

![Diagram](image_url)

Lastly, given $\alpha$ in $\text{Comp}_n$, let $\omega(\alpha)$ be the unique composition in $\text{Comp}_n$ whose partial sums $D(\omega(\alpha))$ form the complementary set within $[n - 1]$ to the partial sums $D(\text{rev}(\alpha))$; alternatively, one can check this means that the ribbon for $\omega(\alpha)$ is obtained from that of $\alpha$ by conjugation or transposing, that is, if $\alpha = \lambda/\mu$ then $\omega(\alpha) = \lambda^t/\mu^t$. E.g. if $\alpha = (4, 2, 2)$ so that $n = 8$, then $\text{rev}(\alpha) = (2, 2, 4)$ has $D(\text{rev}(\alpha)) = \{2, 4\} \subset \{7\}$, complementary to the set $\{1, 3, 5, 6, 7\}$ which are the partial sums for $\omega(\alpha) = (1, 2, 2, 1, 1, 1)$, and the ribbon diagrams of $\alpha, \omega(\alpha)$ are

$$\alpha = \begin{array}{ccccccc} \square & & & & & & \\
\square & & & & & & \\
\square & & & & & & \\
\square & & & & & & \\
\square & & & & & & \\
\square & & & & & & \\
\end{array} \quad \text{and} \quad \omega(\alpha) = \begin{array}{ccccccc} \square & & & & & & \\
\square & & & & & & \\
\square & & & & & & \\
\square & & & & & & \\
\square & & & & & & \\
\square & & & & & & \\
\end{array}$$

Proposition 5.23. The structure maps for the Hopf algebra $\text{QSym}$ in the basis $\{L_\alpha\}$ of fundamental quasisymmetric functions are as follows:

$$\Delta L_\alpha = \sum_{\beta \gamma = \alpha \text{ or } \beta \odot \gamma = \alpha} L_\beta \otimes L_\gamma$$  \hspace{1cm} (5.7)

$$L_\alpha L_\beta = \sum_{w \in w_\alpha \downarrow w_\beta} L_{\gamma(w)}$$  \hspace{1cm} (5.8)

$$S(L_\alpha) = (-1)^{|\alpha|} L_{\omega(\alpha)}.$$  \hspace{1cm} (5.9)

Here we are making use of the following notations in (5.8) (recall also Definition 1.68):

- A labelled linear order will mean a labelled poset $P$ whose order $<_P$ is a total order. We will identify any labelled linear order $P$ with the word (over the alphabet $\mathbb{Z}$) obtained by writing down the elements of $P$ in increasing order (with respect to the total order $<_P$). This way, every word (over the alphabet $\mathbb{Z}$) which has no two equal letters becomes identified with a labelled linear order.
- $w_\alpha$ is any labelled linear order with underlying set $\{1, 2, \ldots, |\alpha|\}$ such that $\text{Des}(w_\alpha) = D(\alpha)$.
- $w_\beta$ is any labelled linear order with underlying set $\{|\alpha| + 1, |\alpha| + 2, \ldots, |\alpha| + |\beta|\}$ such that $\text{Des}(w_\beta) = D(\beta)$.
- $\gamma(w)$ is the unique composition of $|\alpha| + |\beta|$ with $D(\gamma(w)) = \text{Des}(w)$.

At first glance the formula (5.7) for $\Delta L_\alpha$ might seem more complicated than the formula of Proposition 5.7 for $\Delta M_\alpha$. However, it is equally simple when viewed in terms of ribbon diagrams: it cuts the ribbon diagram $\alpha$ into two smaller ribbons $\beta$ and $\gamma$, in all $|\alpha| + 1$ possible ways, via horizontal cuts ($\beta \cdot \gamma = \alpha$) or vertical cuts ($\beta \odot \gamma = \alpha$). For example,
Example 5.24. To multiply $L_{(1,1)}L_{(2)}$, one could pick $w_\alpha = 21$ and $w_\beta = 34$, and then

$$L_{(1,1)}L_{(2)} = \sum_{w \in L_{(1,1)} \cup L_{(2)}} L_\gamma(w) = L_\gamma(2134) + L_\gamma(2314) + L_\gamma(3214) + L_\gamma(2341) + L_\gamma(3241) + L_\gamma(3421).$$

Proof of Proposition 5.23. To prove formula (5.7) for $\alpha$ in Comp$_n$, note that

$$(5.10) \Delta L_\alpha = L_\alpha(x, y) = \sum_{k=0}^{n} \sum_{1 \leq i_1 \leq \cdots \leq i_k \text{ with } i_r < i_{r+1} \text{ for } r \in D(\alpha) \setminus \{k\}} x_{i_1} \cdots x_{i_k} \cdot y_{i_{k+1}} \cdots y_n$$

by Proposition 5.17 (where we identify QSym $\otimes$ QSym with a $k$-subalgebra of $R(x, y)$ by means of the embedding QSym $\otimes$ QSym $\rightarrow$ QSym as in the definition of the comultiplication on QSym). One then realizes that the inner sums corresponding to values of $k$ that lie (resp. do not lie) in $D(\alpha) \cup \{0, n\}$ correspond to the terms $L_\beta(x)L_\gamma(y)$ for pairs $(\beta, \gamma)$ in which $\beta \cdot \gamma = \alpha$ (resp. $\beta \circ \gamma = \alpha$).

For formula (5.8), let $P$ be the labelled poset which is the disjoint union of linear orders $w_\alpha, w_\beta$. Then

$$L_\alpha L_\beta = F_{w_\alpha}(x)F_{w_\beta}(x) = F_P(x) = \sum_{w \in \mathcal{L}(P)} F_w(x) = \sum_{w \in w_\alpha \sqcup w_\beta} L_\gamma(w)$$

where the first equality used Proposition 5.18, the second equality comes from the definition of a $P$-partition, the third equality from Theorem 5.19, and the fourth from the equality $\mathcal{L}(P) = w_\alpha \sqcup w_\beta$.

To prove formula (5.9), compute using Theorem 5.11 that

$$S(L_\alpha) = \sum_{\beta \text{ refining } \alpha} S(M_\beta) = \sum_{(\beta, \gamma) \text{ with } \beta \text{ refining } \alpha, \gamma \text{ coarsens } \text{rev}(\beta)} (-1)^{\ell(\beta)} M_\gamma = \sum_{\gamma} M_\gamma \sum_{\beta \text{ refining } \omega(\alpha)} (-1)^{\ell(\beta)}$$

in which the last inner sum is over $\beta$ for which

$$D(\beta) \supset D(\alpha) \cup D(\text{rev}(\gamma)).$$

The alternating signs make such inner sums vanish unless they have only the single term where $D(\beta) = [n-1]$ (that is, $\beta = (1^n)$). This happens exactly when $D(\text{rev}(\gamma)) \cup D(\alpha) = [n-1]$ or equivalently, when $D(\text{rev}(\gamma))$ contains the complement of $D(\alpha)$, that is, when $D(\gamma)$ contains the complement of $D(\text{rev}(\alpha))$, that is, when $\gamma$ refines $\omega(\alpha)$. Thus

$$S(L_\alpha) = \sum_{\gamma \in \text{Comp}_n, \gamma \text{ refining } \omega(\alpha)} M_\gamma \cdot (-1)^n = (-1)^{|\alpha|} L_{\omega(\alpha)}.$$

The antipode formula (5.9) for $L_\alpha$ leads to a general interpretation for the antipode of QSym acting on $P$-partition enumerators $F_P(x)$.

Definition 5.25. Given a labelled poset $P$ on $\{1, 2, \ldots, n\}$, let the opposite or dual labelled poset $P^{opp}$ have $i < P^{opp} j$ if and only if $j < P i$.

For example,

$$P = \begin{array}{ccc} 1 & 2 & 3 \\ 4 & 3 & 2 \end{array} \quad P^{opp} = \begin{array}{ccc} 4 & 3 & 2 \\ 1 & 2 & 3 \end{array}$$

The following observation is straightforward.
Proposition 5.26. When \( P \) is a linear order corresponding to some permutation \( w = (w_1, \ldots, w_n) \) in \( \mathfrak{S}_n \), then \( w^{\text{opp}} = w w_0 \) where \( w_0 \in \mathfrak{S}_n \) is the permutation that swaps \( i \leftrightarrow n + 1 - i \) (this is the so-called longest permutation, thus named due to it having the highest “Coxeter length” among all permutations in \( \mathfrak{S}_n \)). Furthermore, in this situation one has \( F_w(x) = L_\alpha \), that is, \( \text{Des}(w) = D(\alpha) \) if and only if \( \text{Des}(w^{\text{opp}}) = D(\omega(\alpha)) \), that is \( F_{w^{\text{opp}}}(x) = L_{\omega(\alpha)} \). Thus,

\[
S(F_w(x)) = (-1)^n F_{w^{\text{opp}}}(x).
\]

For example, given the compositions considered earlier

\[
\alpha = (4, 2, 2) = \begin{array}{ccc}
\circ & \circ & \\
\circ & \circ & \\
\circ & \circ & \\
\end{array}
\qquad \omega(\alpha) = (1, 2, 2, 1, 1) = \begin{array}{ccccc}
\circ & \circ & \circ & \circ & \\
\circ & \circ & \circ & \circ & \\
\circ & \circ & \circ & \circ & \\
\end{array}
\]

if one picks \( w = 1235 \cdot 47 \cdot 68 \) (with descent positions marked by dots) having \( \text{Des}(w) = \{4, 6\} = D(\alpha) \), then \( w^{\text{opp}} = w w_0 = 8 \cdot 67 \cdot 45 \cdot 3 \cdot 2 \cdot 1 \) has \( \text{Des}(w^{\text{opp}}) = \{1, 3, 5, 6, 7\} = D(\omega(\alpha)) \).

Corollary 5.27. For any labelled poset \( P \) on \( \{1, 2, \ldots, n\} \), one has

\[
S(F_P(x)) = (-1)^n F_{P^{\text{opp}}}(x).
\]

Proof. Since \( S \) is linear, one can apply Theorem 5.19 and Proposition 5.26

\[
S(F_P(x)) = \sum_{w \in \mathcal{L}(P)} S(F_w(x)) = \sum_{w \in \mathcal{L}(P)} (-1)^n F_{w^{\text{opp}}}(x) = (-1)^n F_{P^{\text{opp}}}(x).
\]

as \( \mathcal{L}(P^{\text{opp}}) = \{w^{\text{opp}} : w \in \mathcal{L}(P)\} \).

Remark 5.28. Malvenuto and Reutenauer, in [114, Theorem 3.1], prove an even more general antipode formula, which encompasses our Corollary 5.27, Proposition 5.26, Theorem 5.11 and (5.9).

We remark on a special case of Corollary 5.27 to which we alluded earlier, related to skew Schur functions.

Corollary 5.29. In \( \Lambda \), the action of \( \omega \) and the antipode \( S \) on skew Schur functions \( s_{\lambda/\mu} \) are as follows:

\[
\omega(s_{\lambda/\mu}) = s_{\lambda'/\mu'}, \quad S(s_{\lambda/\mu}) = (-1)^{|\lambda/\mu|} s_{\lambda'/\mu'}.
\]

Proof. Given a skew shape \( \lambda/\mu \), one can always create a labelled poset \( P \) which is its skew Ferrers poset, together with one of many column-strict labellings, in such a way that \( F_P(x) = s_{\lambda/\mu}(x) \). An example is shown here for \( \lambda/\mu = (4, 4, 2)/(1, 1, 0) \):

\[
\lambda/\mu = \begin{array}{cccc}
\circ & \circ & \\
\circ & \circ & \\
\end{array} \quad P = \begin{array}{cccccccc}
5 & 4 & 2 & \overset{f(5)}{\lessgtr} & \overset{f(4)}{\lessgtr} & \overset{f(3)}{\lessgtr} & \overset{f(2)}{\lessgtr} & \overset{f(1)}{\lessgtr} \\
8 & 3 & 1 & \overset{f(8)}{\lessgtr} & \overset{f(7)}{\lessgtr} & \overset{f(6)}{\lessgtr} \\
7 & 6 & \\
\end{array}
\]

The general definition is as follows: Let \( P \) be the set of all boxes of the skew diagram \( \lambda/\mu \). Label these boxes by the numbers \( 1, 2, \ldots, n \) (where \( n = |\lambda/\mu| \) row by row from bottom to top (reading every row from left to right), and then define an order relation \( \prec_P \) on \( P \) by requiring that every box be smaller (in \( P \)) than its right neighbor and smaller (in \( P \)) than its lower neighbor. It is not hard to see that in this situation,
$F_{\text{opp}}(\mathbf{x}) = \sum_T x^\text{cont}(T)$ as $T$ ranges over all reverse semistandard tableaux or column-strict plane partitions of $\lambda/\mu$: 

\[
\begin{array}{cccc}
\lambda/\mu = & \begin{array}{c}
\square \ \square \\
\square \ \square \\
\square \ \square
\end{array} & P&\text{opp} = \\
& \begin{array}{cccc}
7 & 6 & 3 & 1 \\
8 & 4 & 2 & \text{ } \\
5 & \text{ } & \text{ } & \text{ }
\end{array}
\end{array}
\]

But this means that $F_{\text{opp}}(\mathbf{x}) = s_{\lambda/\mu}(\mathbf{x})$, since the fact that skew Schur functions lie in $\Lambda$ implies that they can be defined either as generating functions for column-strict tableaux or reverse semistandard tableaux; see Remark 2.12 above, or [157, Prop. 7.10.4].

Thus we have

\[
F_P(\mathbf{x}) = s_{\lambda/\mu}(\mathbf{x}) \\
F_{\text{opp}}(\mathbf{x}) = s_{\lambda'/\mu'}(\mathbf{x}).
\]

Proposition 1.46(c) tell us that the antipode for QSym must specialize to the antipode for $\Lambda$ (see also Remark 5.43 below), so (5.12) is a special case of Corollary 5.27. Then (5.11) follows from the relation (2.20) that $S(f) = (-1)^n\omega(f)$ for $f$ in $\Lambda_n$. \hfill $\Box$

Remark 5.30. Before leaving $P$-partitions temporarily, we mention two open questions about them.

The first is a conjecture of Stanley from his thesis [154]. As mentioned in the proof of Corollary 5.29, each skew Schur function $s_{\lambda/\mu}(\mathbf{x})$ is a special instance of $P$-partition enumerator $F_P(\mathbf{x})$.

Conjecture 5.31. A labelled poset $P$ has $F_P(\mathbf{x})$ symmetric, and not just quasisymmetric, if and only if $P$ is a column-strict labelling of some skew Ferrers poset $\lambda/\mu$.

A somewhat weaker result in this direction was proven by Malvenuto in her thesis [112, Thm. 6.4], showing that if a labelled poset $P$ has the stronger property that its set of linear extensions $\mathcal{L}(P)$ is a union of plactic or Knuth equivalence classes, then $P$ must be a column-strict labelling of a skew Ferrers poset.

The next question is due to P. McNamara, and is suggested by the obvious factorizations of $P$-partition enumerators $F_{P_1 \cup P_2}(\mathbf{x}) = F_{P_1}(\mathbf{x})F_{P_2}(\mathbf{x})$.

Question 5.32. If $k$ is a field, does a connected labelled poset $P$ always have $F_P(\mathbf{x})$ irreducible within the ring QSym?

The phrasing of this question requires further comment. It is assumed here that $\mathbf{x} = (x_1, x_2, \ldots)$ is infinite; for example when $P$ is a 2-element chain labelled “against the grain” (i.e., the bigger element of the chain has the smaller label), then $F_P(\mathbf{x}) = e_2(\mathbf{x})$ is irreducible, but its specialization to two variables $\mathbf{x} = (x_1, x_2)$ is $e_2(x_1, x_2) = x_1x_2$, which is reducible. If one wishes to work in finitely many variables $\mathbf{x} = (x_1, \ldots, x_m)$ one can perhaps assume that $m$ is at least $|P| + 1$.

When working in $\text{QSym} = \text{QSym}(\mathbf{x})$ in infinitely many variables, it is perhaps not so clear where factorizations occur. For example, if $f$ lies in $\text{QSym}$ and factors $f = g \cdot h$ with $g, h$ in $\text{R}(\mathbf{x})$, does this imply that $g, h$ also lie in $\text{QSym}$? The answer is “Yes” (for $k = \mathbb{Z}$), but this is not obvious, and was proven by P. Pylyavskyy in [133, Chap. 11].

One also might wonder whether $\text{QSym}_2$ is a unique factorization domain, but this follows from the result of M. Hazewinkel ([65] and [69, Thm. 6.7.5], and Theorem 6.82 further below) who proved a conjecture of Ditters that $\text{QSym}_2$ is a polynomial algebra; earlier Malvenuto and Reutenauer [113, Cor. 2.2] had shown that $\text{QSym}_2$ is a polynomial algebra. In fact, one can find polynomial generators $\{P_\alpha\}$ for $\text{QSym}_2$ as a subset of the dual basis to the $\mathbb{Q}$-basis $\{\xi_\alpha\}$ for $\text{NSym}_2$ which comes from taking products $\xi_\alpha := \xi_{\alpha_1} \cdots \xi_{\alpha_k}$ of the elements $\{\xi_\alpha\}$ defined in Remark 5.36 below. Specifically, one takes those $P_\alpha$ for which the composition $\alpha$ is a Lyndon composition; see the First proof of Proposition 6.83 for a mild variation on this construction.
Hazewinkel’s proof [69, Thm. 6.7.5] of the polynomiality of $QSym$ also shows that $QSym$ is a polynomial ring over $\Lambda$ (see Corollary 6.127): in particular, this yields that $QSym$ is a free $\Lambda$-module.\footnote{The latter statement has an analogue in finitely many indeterminates, proven by Lauve and Mason in [94, Corollary 13]: The quasisymmetric functions $QSym \{x_i\}_{i \in I}$ are free as a $\Lambda \{x_i\}_{i \in I}$-module for any finite totally ordered set $I$, infinite or not. In the case of finite $I$, this cannot be derived from Hazewinkel’s arguments, as the ring $QSym \{x_i\}_{i \in I}$ is not in general a polynomial ring (e.g., when $k = \mathbb{Q}$ and $I = \{1, 2\}$, this ring is not even a UFD, as witnessed by $(x_1^2 x_2) \cdot (x_1 x_2^2) = (x_1 x_2)^3$).

An affirmative answer to Question 5.32 is known at least in the special case where $P$ is a connected column-strict labelling of a skew Ferrers diagram, that is, when $F_P(x) = s_{\lambda/\mu}(x)$ for some connected skew diagram $\lambda/\mu$; see [12].

5.3. The Hopf algebra $NSym$ dual to $QSym$. We introduce here the (graded) dual Hopf algebra to $QSym$. This is well-defined, as $QSym$ is connected graded of finite type.

**Definition 5.33.** Let $NSym := QSym^\circ$, with dual pairing $NSym \otimes QSym \overset{\cdot}{\longrightarrow} k$. Let $\{H_\alpha\}$ be the $k$-basis of $NSym$ dual to the $k$-basis $\{M_\alpha\}$ of $QSym$, so that

\[
(H_\alpha, M_\beta) = \delta_{\alpha, \beta}.
\]

When the base ring $k$ is not clear from the context, we write $NSym_k$ in lieu of $NSym$.

The Hopf algebra $NSym$ is known as the *Hopf algebra of noncommutative symmetric functions*. Its study goes back to [58].

**Theorem 5.34.** Letting $H_n := H_{(n)}$ for $n = 0, 1, 2, \ldots$, with $H_0 = 1$, one has that

\[
(5.13) \quad NSym \cong k\langle H_1, H_2, \ldots \rangle,
\]

the free associative (but not commutative) algebra on generators $\{H_1, H_2, \ldots\}$ with coproduct determined by\footnote{The abbreviated summation indexing $\sum_{i+j=n} t_{i,j} \cdot M_{i,j}$ used here is intended to mean $\sum_{i,j \geq 0, i+j=n} t_{i,j} \cdot M_{i,j}$.}

\[
(5.14) \quad \Delta H_n = \sum_{i+j=n} H_i \otimes H_j.
\]

**Proof.** Since Proposition 5.7 asserts that $\Delta M_\alpha = \sum_{\beta, \gamma} \beta; \gamma = \alpha M_\beta \otimes M_\gamma$, and since $\{H_\alpha\}$ are dual to $\{M_\alpha\}$, one concludes that for any compositions $\beta, \gamma$, one has

\[
H_\beta H_\gamma = H_{\beta; \gamma}
\]

Iterating this gives

\[
(5.15) \quad H_\alpha = H_{(\alpha_1, \ldots, \alpha_\ell)} = H_{\alpha_1} \cdots H_{\alpha_\ell}.
\]

Since the $H_\alpha$ are a $k$-basis for $NSym$, this shows $NSym \cong k\langle H_1, H_2, \ldots \rangle$.

Note that $H_n = H_{(n)}$ is dual to $M_{(n)}$, so to understand $\Delta H_n$, one should understand how $M_{(n)}$ can appear as a term in the product $M_\alpha M_\beta$. By (5.1) this occurs only if $\alpha = (i), \beta = (j)$ where $i + j = n$, where

\[
M_{(i)} M_{(j)} = M_{(i+j)} + M_{(i,j)} + M_{(j,i)}
\]

(where the $M_{(i,j)}$ and $M_{(j,i)}$ addends have to be disregarded if one of $i$ and $j$ is 0). By duality, this implies the formula (5.14).

**Corollary 5.35.** The algebra homomorphism defined by

\[
NSym \xrightarrow{\pi} \Lambda
\]


is a Hopf algebra surjection, and adjoint to the inclusion $\Lambda \hookrightarrow QSym$ (with respect to the dual pairing $NSym \otimes QSym \overset{\cdot}{\longrightarrow} k$).

---

192 The latter statement has an analogue in finitely many indeterminates, proven by Lauve and Mason in [94, Corollary 13]: The quasisymmetric functions $QSym \{x_i\}_{i \in I}$ are free as a $\Lambda \{x_i\}_{i \in I}$-module for any totally ordered set $I$, infinite or not. In the case of finite $I$, this cannot be derived from Hazewinkel's arguments, as the ring $QSym \{x_i\}_{i \in I}$ is not in general a polynomial ring (e.g., when $k = \mathbb{Q}$ and $I = \{1, 2\}$, this ring is not even a UFD, as witnessed by $(x_1^2 x_2) \cdot (x_1 x_2^2) = (x_1 x_2)^3$).

193 The abbreviated summation indexing $\sum_{i+j=n} t_{i,j} \cdot M_{i,j}$ used here is intended to mean $\sum_{i,j \geq 0, i+j=n} t_{i,j} \cdot M_{i,j}$. 

---

---
As (5.14) and Proposition 2.24(iii) assert that the composition \( \alpha \) of the map \( \pi \) on a graded free \( k \)-module \( V \) with basis \( \{ H_1, H_2, \ldots \} \) to the symmetric algebra on \( V \), since
\[
\Lambda \cong k[h_1, h_2, \ldots ]
\]
As (5.14) and Proposition 2.24(iii) assert that
\[
\Delta H_n = \sum_{i+j=n} H_i \otimes H_j
\]
\[
\Delta h_n = \sum_{i+j=n} h_i \otimes h_j
\]
this map \( \pi \) is also a bialgebra morphism, and hence a Hopf morphism by Proposition 1.46(c).

To check \( \pi \) is adjoint to \( i \), let \( \lambda(\alpha) \) denote the partition which is the weakly decreasing rearrangement of the composition \( \alpha \), and note that the bases \( \{ H_\alpha \} \) of NSym and \( \{ m_\lambda \} \) of \( \Lambda \) satisfy
\[
(\pi(H_\alpha), m_\lambda) = (h_{\lambda(\alpha)}, m_\lambda) = \begin{cases} 1 & \text{if } \lambda(\alpha) = \lambda \\ 0 & \text{otherwise} \end{cases} = \left( H_\alpha, \sum_{\beta : \lambda(\beta) = \lambda} M_{\beta} \right) = (H_\alpha, i(m_\lambda)).
\]

\[\square\]

Remark 5.36. For those who prefer generating functions to sign-reversing involutions, we sketch here Malvenuto and Reutenauer’s elegant proof [113, Cor. 2.3] of the antipode formula (Theorem 5.11). One needs to know that when \( Q \) is a subring of \( k \), and \( A \) is a \( k \)-algebra (possibly noncommutative), in the ring of power series \( A[[t]] \) where \( t \) commutes with all of \( A \), one still has familiar facts, such as
\[
a(t) = \log b(t) \quad \text{if and only if} \quad b(t) = \exp a(t)
\]
and whenever \( a(t), b(t) \) commute in \( A[[t]] \), one has
\[
\exp (a(t) + b(t)) = \exp a(t) \exp b(t)
\]
\[
\log (a(t)b(t)) = \log a(t) + \log b(t)
\]
Start by assuming WLOG that \( k = Z \) (as \( \text{NSym}_k = \text{NSym}_Z \otimes_k k \) in the general case). Now, define in \( \text{NSym}_Q = \text{NSym} \otimes_k Q \) the elements \( \{ \xi_1, \xi_2, \ldots \} \) via generating functions in \( \text{NSym}_Q[[t]] \):
\[
\tilde{H}(t) := \sum_{n \geq 0} H_n t^n,
\]
\[
\xi(t) := \sum_{n \geq 1} \xi_n t^n = \log \tilde{H}(t)
\]
One first checks that this makes each \( \xi_n \) primitive, via a computation in the ring \( (\text{NSym}_Q \otimes \text{NSym}_Q)[[t]] \) (into which we “embed” the ring \( (\text{NSym}_Q[[t]]) \otimes_Q[[t]] (\text{NSym}_Q[[t]]) \) via the canonical ring homomorphism from the latter into the former \[\text{NSym}_Q[[t]] \to (\text{NSym}_Q \otimes \text{NSym}_Q)[[t]] \):}
\[
\Delta \xi(t) = \Delta \left( \sum_{n \geq 0} H_n t^n \right) = \sum_{n \geq 0} \Delta (H_n) t^n = \sum_{n \geq 0} \left( \sum_{i+j=n} H_i \otimes H_j \right) t^n
\]
\[
= \log \left( \sum_{i \geq 0} H_i t^i \otimes \sum_{j \geq 0} H_j t^j \right) = \log \left( \sum_{i \geq 0} H_i t^i \otimes 1 \right) \left( 1 \otimes \sum_{j \geq 0} H_j t^j \right)
\]
\[
\overset{(5.17)}{=} \log \tilde{H}(t) \otimes 1 + 1 \otimes \log \tilde{H}(t) = \xi(t) \otimes 1 + 1 \otimes \xi(t).
\]

\[\text{This ring homomorphism might fail to be injective, whence the “embed” stands in quotation marks. This does not need to worry us, since we will not draw any conclusions in \( (\text{NSym}_Q[[t]]) \otimes_Q[[t]] (\text{NSym}_Q[[t]]) \) from our computation.}

\[\text{We are also somewhat cavalier with the notation } \Delta : \text{ we use it both for the comultiplication } \Delta : \text{NSym}_Q \to \text{NSym}_Q \otimes \text{NSym}_Q \text{ of the Hopf algebra } \text{NSym}_Q \text{ and for the continuous } k\text{-algebra homomorphism } \text{NSym}_Q[[t]] \to (\text{NSym}_Q \otimes \text{NSym}_Q)[[t]] \text{ it induces.}\]
Comparing coefficients in this equality yields $\Delta(\xi_n) = \xi_n \otimes 1 + 1 \otimes \xi_n$. Thus $S(\xi_n) = -\xi_n$, by Proposition 1.37. This allows one to determine $S(H_0)$ and $S(H_0)$, after one first inverts the relation (5.18) to get that $\tilde{H}(t) = \exp \xi(t)$, and hence

$$S(\tilde{H}(t)) = S(\exp \xi(t)) = \exp S(\xi(t)) = \exp (-\xi(t)) \overset{(5.16)}{=} (\exp \xi(t))^{-1}$$

$$= \tilde{H}(t)^{-1} = (1 + H_1 t + H_2 t^2 + \cdots)^{-1}.$$ 

Upon expanding the right side, and comparing coefficients of $t^n$, this gives

$$S(H_n) = \sum_{\beta \in \Comp_n} (-1)^{\ell(\beta)} H_\beta$$

and hence

$$S(H_n) = S(H_{\alpha_1}) \cdots S(H_{\alpha_2}) S(H_{\alpha_1}) = \sum_{\gamma \text{ coarsens } \text{rev}(\alpha)} (-1)^{\ell(\gamma)} H_\gamma.$$ 

As $S_{\NSym}, S_{\QSym}$ are adjoint, and $\{H_\alpha\}, \{M_\alpha\}$ are dual bases, this is equivalent to Theorem 5.11:

$$S(M_\alpha) = (-1)^{\ell(\alpha)} \sum_{\gamma \text{ coarsens } \text{rev}(\alpha)} M_\gamma$$

(because if $\mu$ and $\nu$ are two compositions, then $\mu$ coarsens $\nu$ if and only if $\text{rev}(\mu)$ coarsens $\text{rev}(\nu)$). Thus, Theorem 5.11 is proven once again.

Let us say a bit more about the elements $\xi_n$ defined in (5.18) above. The elements $n! \xi_n$ are noncommutative analogues of the power sum symmetric functions $p_n$ (and, indeed, are lifts of the latter to $\NSym$, as Exercise 5.37 below shows). They are called the noncommutative power sums of the second kind in [58]195, and their products form a basis of $\NSym$. They are furthermore useful in studying the so-called Eulerian idempotent of a cocommutative Hopf algebra, as shown in Exercise 5.38 below.

**Exercise 5.37.** Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Define a sequence of elements $\xi_1, \xi_2, \xi_3, \ldots$ of $\NSym = \NSym_{\mathbf{k}}$ by (5.18).

(a) For every $n \geq 1$, show that $\xi_n$ is a primitive homogeneous element of $\NSym$ of degree $n$.

(b) For every $n \geq 1$, show that $\pi(n! \xi_n)$ is the $n$-th power sum symmetric function $p_n \in \Lambda$.

(c) For every $n \geq 1$, show that

$$\xi_n = \sum_{\alpha \in \Comp_n} (-1)^{\ell(\alpha)-1} \frac{1}{\ell(\alpha)} H_\alpha.$$ 

(d) For every composition $\alpha$, define an element $\xi_\alpha$ of $\NSym$ by $\xi_\alpha = \xi_{\alpha_1} \xi_{\alpha_2} \cdots \xi_{\alpha_\ell}$, where $\alpha$ is written in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ with $\ell = \ell(\alpha)$. Show that

$$H_n = \sum_{\alpha \in \Comp_n} \frac{1}{\ell(\alpha)!} \xi_\alpha$$

for every $n \in \mathbb{N}$.

Use this to prove that $(\xi_\alpha)_{\alpha \in \Comp_n}$ is a $\mathbf{k}$-basis of $\NSym_n$ for every $n \in \mathbb{N}$.

**Exercise 5.38.** Assume that $\mathbb{Q}$ is a subring of $\mathbf{k}$. Let $A$ be a cocommutative connected graded $\mathbf{k}$-bialgebra. Let $A = \bigoplus_{n \geq 0} A_n$ be the decomposition of $A$ into homogeneous components. If $f$ is any $\mathbf{k}$-linear map $A \to A$ annihilating $A_0$, then $f$ is locally $\star$-nilpotent196, and so the sum $\sum^\star (f + \epsilon) := \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^n$ is a well-defined endomorphism of $A$. Let $\epsilon$ denote the endomorphism $\log^\star (\id_A)$ of $A$ (obtained by setting $f = \id_A - \epsilon : A \to A$). Show that $\epsilon$ is a projection from $A$ to the $\mathbf{k}$-submodule $p$ of all primitive elements of $A$ (and thus, in particular, is idempotent).

**Hint:** For every $n \geq 0$, let $\pi_n : A \to A$ be the projection onto the $n$-th homogeneous component $A_n$. Since $\NSym$ is the free $\mathbf{k}$-algebra with generators $H_1, H_2, H_3, \ldots$, we can define a $\mathbf{k}$-algebra homomorphism $29 : \NSym \to (\End A, \star)$ by sending $H_n$ to $\pi_n$. Show that:

195See Exercise 5.44 for the ones of the first kind.
196See the proof of Proposition 1.44 for what this means.
(a) The map $\epsilon : A \to A$ is graded. For every $n \geq 0$, we will denote the map $\pi_n \circ \epsilon = \epsilon \circ \pi_n : A \to A$ by $\epsilon_n$.

(b) We have $\mathfrak{M}(\xi_n) = \epsilon_n$ for all $n \geq 1$, where $\xi_n$ is defined as in Exercise 5.37.

(c) If $w$ is an element of NSym, and if we write $\Delta(w) = \sum_{(w)} w_1 \otimes w_2$ using Sweedler’s notation, then

$$\Delta \circ (\mathfrak{M}(w)) = \left( \sum_{(w)} \mathfrak{M}(w_1) \otimes \mathfrak{M}(w_2) \right) \circ \Delta.$$

(d) We have $\epsilon_n(A) \subset \mathfrak{p}$ for every $n \geq 0$.

(e) We have $\epsilon(A) \subset \mathfrak{p}$.

(f) The map $\epsilon$ fixes any element of $\mathfrak{p}$.

**Remark 5.39.** The endomorphism $\epsilon$ of Exercise 5.38 is known as the *Eulerian idempotent* of $A$, and can be contrasted with the Dynkin idempotent of Remark 1.66. It has been studied in [126], [129], [26] and [45], and relates to the Hochschild cohomology of commutative algebras [102, §4.5.2].

**Exercise 5.40.** Assume that $\mathfrak{Q}$ is a subring of $k$. Let $A$, $A_n$ and $\epsilon$ be as in Exercise 5.38.

(a) Show that $\epsilon^m \circ \epsilon = m! \delta_{n,m} \epsilon^n$ for all $n \in \mathbb{N}$ and $m \in \mathbb{N}$.

(b) Show that $\epsilon^m \circ \text{id}_A = \text{id}_A \circ \epsilon^n = m^n \epsilon^n$ for all $n \in \mathbb{N}$ and $m \in \mathbb{N}$.

We next explore the basis for NSym dual to the $\{L_\alpha\}$ in QSym.

**Definition 5.41.** Define the *noncommutative ribbon functions* $\{R_\alpha\}$ to be the $k$-basis of NSym dual to the fundamental basis $\{L_\alpha\}$ of QSym, so that $(R_\alpha, L_\beta) = \delta_{\alpha,\beta}$.

**Theorem 5.42.** (a) One has that

$$H_\alpha = \sum_{\beta \text{ coarsens } \alpha} R_\beta \quad \text{and} \quad R_\alpha = \sum_{\beta \text{ coarsens } \alpha} (-1)^{\ell(\beta) - \ell(\alpha)} H_\beta$$

(b) The surjection $\text{NSym} \twoheadrightarrow \Lambda$ sends $R_\alpha \mapsto s_\alpha$, the skew Schur function associated to the ribbon $\alpha$.

(c) Furthermore,

$$R_\alpha R_\beta = R_{\alpha \beta} + R_{\alpha \circ \beta} \quad \text{if } \alpha \text{ and } \beta \text{ are nonempty}$$

$$S(R_\alpha) = (-1)^{|\alpha|} R_{\omega(\alpha)}$$

Finally, $R_{\beta}$ is the multiplicative identity of NSym.

**Proof.** (a) For the first assertion, note that

$$H_\alpha = \sum_{\beta} (H_\alpha, L_\beta) R_\beta = \sum_{\beta} \left( H_\alpha, \sum_{\gamma \text{ refines } \beta} M_\gamma \right) R_\beta = \sum_{\beta} R_\beta.$$

The second assertion follows from the first by inclusion-exclusion.

(b) Write $\alpha$ as $(\alpha_1, \ldots, \alpha_\ell)$. To show that $\pi(R_\alpha) = s_\alpha$, we instead examine $\pi(H_\alpha)$:

$$\pi(H_\alpha) = \pi(H_{\alpha_1} \cdots H_{\alpha_\ell}) = h_{\alpha_1} \cdots h_{\alpha_\ell} = s_{\alpha_1} \cdots s_{\alpha_\ell} = s_{\alpha_1 \oplus \cdots \oplus \alpha_\ell}$$

where $(\alpha_1 \oplus \cdots \oplus \alpha_\ell)$ is some skew shape which is a horizontal strip having rows of lengths $\alpha_1, \ldots, \alpha_\ell$ from bottom to top. We claim

$$s_{\alpha_1 \oplus \cdots \oplus \alpha_\ell} = \sum_{\beta \text{ coarsens } \alpha} \sum_{\beta} s_\beta$$

because column-strict tableaux $T$ of shape $(\alpha_1 \oplus \cdots \oplus \alpha_\ell)$ biject to column-strict tableaux $T'$ of some ribbon $\beta$ coarsening $\alpha$, as follows: let $a_i, b_i$ denote the leftmost, rightmost entries of the $i^{th}$ row from the bottom in $T$, of length $\alpha_i$, and

- if $b_i \leq a_{i+1}$, merge parts $\alpha_i, \alpha_{i+1}$ in $\beta$, and concatenate the rows of length $\alpha_i, \alpha_{i+1}$ in $T'$, or
- if $b_i > a_{i+1}$, do not merge parts $\alpha_i, \alpha_{i+1}$ in $\beta$, and let these two rows overlap in one column in $T'$.
E.g., if \( \alpha = (3, 3, 2, 3) \), this \( T \) of shape \((\alpha_1) \oplus \cdots \oplus (\alpha_\ell)\) maps to this \( T' \) of shape \( \beta = (3, 8, 2): \)

\[
T = \begin{array}{cccc}
3 & 4 & 4 & 5 \\
2 & 4 & 4 & 3 \\
1 & 2 & 3 & 1 \\
\end{array}
\quad \mapsto \quad \begin{array}{cccc}
3 & 4 & 4 & 5 \\
2 & 3 & 4 & 4 \\
1 & 1 & 3 & 1 \\
\end{array}
\]

The reverse bijection breaks the rows of \( T' \) into the rows of \( T \) of lengths dictated by the parts of \( \alpha \). Having shown \( \pi(H_\alpha) = \sum \beta: \beta \text{ coarsens } \alpha \ s_\beta \), the relation (5.21) and inclusion-exclusion show \( \pi(R_\alpha) = s_\alpha \).

(c) Finally, (5.23) and (5.24) follow from (5.7) and (5.9) by duality. \( \square \)

**Remark 5.43.** Since the maps \( \pi: \text{NSym} \rightarrow \text{QSym} \), \( \Lambda \rightarrow i \), \( \text{NSym} \rightarrow \text{NSym} \), but it is interesting to compare them explicitly using the fundamental basis for \( \text{QSym} \) and the ribbon basis for \( \text{NSym} \).

On one hand (5.9) shows that \( S_{\text{QSym}}(L_\alpha) = (-1)^{|\alpha|} L_{\omega(\alpha)} \) extends the map \( S_\Lambda \) since \( L(1^n) = e_n \) and \( L(n) = h_n \), as observed in Example 5.16, and \( \omega((n)) = (1^n) \).

On the other hand, (5.24) shows that \( S_{\text{NSym}}(R_\alpha) = (-1)^{|\alpha|} R_{\omega(\alpha)} \) lifts the map \( S_\Lambda \). Theorem 5.42(b) showed that \( R_\alpha \) lifts the skew Schur function \( s_\alpha \), while (2.21) asserted that \( S(s_\lambda/\mu) = (-1)^{|\lambda/\mu|} s_\lambda t^{\lambda/\mu} \), and a ribbon \( \alpha = \lambda/\mu \) has \( \omega(\alpha) = \lambda/\mu \).

**Exercise 5.44.**

(a) Show that any integers \( n \) and \( i \) with \( 0 \leq i < n \) satisfy

\[
R_{(1^i, n-i)} = \sum_{j=0}^{i} (-1)^{i-j} R_{(1^j)} H_{n-j}.
\]

(Here, as usual, \( 1^i \) stands for the number 1 repeated \( i \) times.)

(b) Show that any integers \( n \) and \( i \) with \( 0 \leq i < n \) satisfy

\[
(-1)^{i} R_{(1^i, n-i)} = \sum_{j=0}^{i} S(H_j) H_{n-j}.
\]

(c) For every positive integer \( n \), define an element \( \Psi_n \) of \( \text{NSym} \) by

\[
\Psi_n = \sum_{i=0}^{n-1} (-1)^{i} R_{(1^i, n-i)}.
\]

Show that \( \Psi_n = (S \ast E)(H_n) \), where the map \( E: \text{NSym} \rightarrow \text{NSym} \) is defined as in Exercise 1.65 (for \( A = \text{NSym} \)). Conclude that \( \Psi_n \) is primitive.

(d) Prove that

\[
\sum_{k=0}^{n-1} H_k \Psi_{n-k} = n H_n
\]

for every \( n \in \mathbb{N} \).

(e) Define two power series \( \psi(t) \) and \( \tilde{H}(t) \) in \( \text{NSym}[[t]] \) by

\[
\psi(t) = \sum_{n \geq 1} \Psi_n t^{n-1};
\]

\[
\tilde{H}(t) = \sum_{n \geq 0} H_n t^n.
\]
Show that\[^{197}\frac{d}{dt} \tilde{H}(t) = \tilde{H}(t) \cdot \psi(t)\].

(The functions $\Psi_n$ are called \textit{noncommutative power sums of the first kind}; they are studied in [58]. The power sums of the second kind are the $n_k$ in Remark 5.36.)

(f) Show that $\pi(\Psi_n)$ equals the power sum symmetric function $p_n$ for every positive integer $n$.

(g) Show that every positive integer $n$ satisfies

$$p_n = \sum_{i=0}^{n-1} (-1)^i s_{(n-i, 1^i)} \quad \text{in } \Lambda.$$ 

(h) For every nonempty composition $\alpha$, define a positive integer $\ell_\alpha (\alpha)$ by $\ell_\alpha (\alpha) = \ell(\alpha)$, where $\alpha$ is written in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ with $\ell = \ell(\alpha)$. (Thus, $\ell_\alpha (\alpha)$ is the last part of $\alpha$.) Show that every positive integer $n$ satisfies

$$\Psi_n = \sum_{\alpha \in \text{Comp}_n} (-1)^{\ell(\alpha)-1} \ell_\alpha (\alpha) H_\alpha.$$ 

(i) Assume that $\mathbb{Q}$ is a subring of $\mathbb{k}$. For every composition $\alpha$, define an element $\Psi_\alpha$ of $\text{NSym}$ by $\Psi_\alpha = \Psi_{\alpha_1} \Psi_{\alpha_2} \cdots \Psi_{\alpha_\ell}$, where $\alpha$ is written in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ with $\ell = \ell(\alpha)$. For every composition $\alpha$, define $\pi_\alpha (\alpha)$ to be the positive integer $\alpha_1 (\alpha_1 + \alpha_2) \cdots (\alpha_1 + \alpha_2 + \cdots + \alpha_\ell)$, where $\alpha$ is written in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ with $\ell = \ell(\alpha)$. Show that

$$H_n = \sum_{\alpha \in \text{Comp}_n} \frac{1}{\pi_\alpha (\alpha)} \Psi_\alpha$$

for every $n \in \mathbb{N}$.

Use this to prove that $\{\Psi_\alpha \}_{\alpha \in \text{Comp}_n}$ is a $\mathbb{k}$-basis of $\text{NSym}_n$ for every $n \in \mathbb{N}$.

(j) Assume that $\mathbb{Q}$ is a subring of $\mathbb{k}$. Let $V$ be the free $\mathbb{k}$-module with basis $(b_n)_{n \in \{1, 2, 3, \ldots\}}$. Define a $\mathbb{k}$-module homomorphism $f : V \to \text{NSym}$ by requiring that $f(b_n) = \Psi_n$ for every $n \in \{1, 2, 3, \ldots\}$. Let $F$ be the $\mathbb{k}$-algebra homomorphism $T(V) \to \text{NSym}$ induced by this $f$ (using the universal property of the tensor algebra $T(V)$). Show that $F$ is a Hopf algebra isomorphism (where the Hopf algebra structure on $T(V)$ is as in Example 1.38).

(k) Assume that $\mathbb{Q}$ is a subring of $\mathbb{k}$. Let $V$ be as in Exercise 5.44(j). Show that $\text{QSym}$ is isomorphic to the shuffle algebra $\text{Sh}(V)$ (defined as in Proposition 1.73) as Hopf algebras.

(l) Solve parts (a) and (b) of Exercise 2.90 again using the ribbon basis functions $R_n$.

One might wonder whether the Frobenius endomorphisms of $\Lambda$ (defined in Exercise 2.85) and the Verschiebung endomorphisms of $\Lambda$ (defined in Exercise 2.86) generalize to analogous operators on either $\text{QSym}$ or $\text{NSym}$. The next two exercises (whose claims mostly come from [66, §13]) answer this question: The Frobenius endomorphisms extend to $\text{QSym}$, and the Verschiebung ones lift to $\text{NSym}$.

**Exercise 5.45.** For every $n \in \{1, 2, 3, \ldots\}$, define a map $F_n : \text{QSym} \to \text{QSym}$ by setting

$$F_n (a) = a(x^n_1, x^n_2, x^n_3, \ldots)$$

for every $a \in \text{QSym}$.

(So what $F_n$ does to a quasi-symmetric function is replacing all variables $x_1, x_2, x_3, \ldots$ by their $n$-th powers.)

(a) Show that $F_n : \text{QSym} \to \text{QSym}$ is a $\mathbb{k}$-algebra homomorphism for every $n \in \{1, 2, 3, \ldots\}$.

(b) Show that $F_n \circ F_m = F_{nm}$ for any two positive integers $n$ and $m$.

(c) Show that $F_1 = \text{id}$.

(d) Prove that $F_n (M(\beta_1, \beta_2, \ldots, \beta_s)) = M(n\beta_1, n\beta_2, \ldots, n\beta_s)$ for any $n \in \{1, 2, 3, \ldots\}$ and $(\beta_1, \beta_2, \ldots, \beta_s) \in \text{Comp}$.

(e) Prove that $F_n : \text{QSym} \to \text{QSym}$ is a Hopf algebra homomorphism for every $n \in \{1, 2, 3, \ldots\}$.

(f) Consider the maps $f_n : \Lambda \to \Lambda$ defined in Exercise 2.85. Show that $F_n |_{\Lambda} = f_n$ for every $n \in \{1, 2, 3, \ldots\}$.

(g) Assume that $\mathbb{k} = \mathbb{Z}$. Prove that $f_p (a) \equiv a^p \mod p \text{QSym}$ for every $a \in \text{QSym}$ and every prime number $p$.

(h) Give a new solution to Exercise 2.85(d).

\[^{197}\]The derivative $\frac{d}{dt} Q(t)$ of a power series $Q(t) \in R[[t]]$ over a noncommutative ring $R$ is defined just as in the case of $R$ commutative: by setting $\frac{d}{dt} Q(t) = \sum_{i \geq 1} \eta_i t^{i-1}$, where $Q(t)$ is written in the form $Q(t) = \sum_{i \geq 0} \eta_i t^i$.\]
Exercise 5.46. For every $n \in \{1, 2, 3, \ldots\}$, define a $k$-algebra homomorphism $V_n : \text{NSym} \rightarrow \text{NSym}$ by

$$V_n(H_m) = \begin{cases} H_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases}$$

for every positive integer $m$.

(a) Show that any positive integers $n$ and $m$ satisfy

$$V_n(\Psi_m) = \begin{cases} n\Psi_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases},$$

where the elements $\Psi_m$ and $\Psi_{m/n}$ of $\text{NSym}$ are as defined in Exercise 5.44(c).

(b) Show that if $\mathbb{Q}$ is a subring of $k$, then any positive integers $n$ and $m$ satisfy

$$V_n(\xi_m) = \begin{cases} \xi_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases},$$

where the elements $\xi_m$ and $\xi_{m/n}$ of $\text{NSym}$ are as defined in Exercise 5.37.

(c) Prove that $V_n \circ V_m = V_{nm}$ for any two positive integers $n$ and $m$.

(d) Prove that $V_1 = \text{id}$.

(e) Prove that $V_n : \text{NSym} \rightarrow \text{NSym}$ is a Hopf algebra homomorphism for every $n \in \{1, 2, 3, \ldots\}$.

Now, consider also the maps $F_n : \text{QSym} \rightarrow \text{QSym}$ defined in Exercise 2.85. Fix a positive integer $n$.

(f) Prove that the maps $F_n : \text{QSym} \rightarrow \text{QSym}$ and $V_n : \text{NSym} \rightarrow \text{NSym}$ are adjoint with respect to the dual pairing $\text{NSym} \otimes \text{QSym} \rightarrow k$.

(g) Consider the maps $v_n : \Lambda \rightarrow \Lambda$ defined in Exercise 2.86. Show that the surjection $\pi : \text{NSym} \rightarrow \Lambda$ satisfies $v_n \circ \pi = \pi \circ V_n$ for every $n \in \{1, 2, 3, \ldots\}$.

(h) Give a new solution to Exercise 2.86(f).

---

\footnote{This is well-defined, since $\text{NSym}$ is (isomorphic to) the free associative algebra with generators $H_1, H_2, H_3, \ldots$ (according to (5.13)).}
6. Polynomial generators for QSym and Lyndon words

In this chapter, we shall construct an algebraically independent generating set for QSym as a \( k \)-algebra, thus showing that QSym is a polynomial ring over \( k \). This has been done by Malvenuto [112, Cor. 4.19] when \( k \) is a field of characteristic 0, and by Hazewinkel [65] in the general case. We will begin by introducing the notion of Lyndon words (Section 6.1), on which both of these constructions rely; we will then (Section 6.2) elucidate the connection of Lyndon words with shuffles, and afterwards (Section 6.3) apply it to prove Radford’s theorem stating that the shuffle algebra of a free \( k \)-module over a commutative \( \mathbb{Q} \)-algebra is a polynomial algebra (Theorem 6.71). The shuffle algebra is not yet QSym, but Radford’s theorem on the shuffle algebra serves as a natural stepping stone for the study of the more complicated algebra QSym. We will prove – in two ways – that QSym is a polynomial algebra when \( \mathbb{Q} \) is a subring of \( k \) in Section 6.4, and then we will finally prove the general case in Section 6.5.

Strictly speaking, this whole Chapter 6 is a digression, as it involves almost no coalgebraic or Hopf-algebraic structures, and its results will not be used in further chapters (which means it can be skipped if so desired). However, it sheds additional light on QSym and serves as an excuse to study Lyndon words, which are a combinatorial object of independent interest (and are involved in the study of free algebras and Hopf algebras, apart from QSym – see [134] and [137]199).

We will take a scenic route to the proof of Hazewinkel’s theorem. A reader only interested in the proof proper can restrict themselves to reading only the following:

- from Section 6.1, everything up to Corollary 6.6, then from Definition 6.13 up to Proposition 6.18, then from Definition 6.25 up to Lemma 6.28, and finally Theorem 6.30. (Proposition 6.19 and Theorem 6.20 are also relevant if one wants to use a different definition of Lyndon words, as they prove the equivalence of most such definitions.)
- from Section 6.2, everything except for Exercise 6.67.
- from Section 6.3, Definition 6.68, Lemma 6.73, and Lemma 6.76.
- from Section 6.4, Definition 6.80, Theorem 6.82, then from Proposition 6.84 up to Definition 6.88, and Lemma 6.90.
- all of Section 6.5.

6.1. Lyndon words. Lyndon words have been independently defined by Shirshov [153], Lyndon [108], Radford [134, \S2] and de Bruijn/Klarner [24] (though using different and sometimes incompatible notations). They have since been surfacing in various places in noncommutative algebra (particularly the study of free Lie algebras); expositions of their theory can be found in [106, \S5], [137, \S5.1] and [93, \S1] (in German). We will follow our own approach to the properties of Lyndon words that we need.

**Definition 6.1.** We fix a totally ordered set \( \mathfrak{A} \), which we call the alphabet. Throughout Section 6.1 and Section 6.2, we will understand “word” to mean a word over \( \mathfrak{A} \).

We recall that a *word* is just a (finite) tuple of elements of \( \mathfrak{A} \). In other words, a word is an element of the set \( \bigcup_{n \geq 0} \mathfrak{A}^n \). We denote this set by \( \mathfrak{A}^* \).

The *empty word* is the unique tuple with 0 elements. It is denoted by \( \emptyset \). If \( w \in \mathfrak{A}^n \) is a word and \( i \in \{1, 2, \ldots, n\} \), then the *i*-th letter of \( w \) means the *i*-th entry of the *n*-tuple \( w \). This *i*-th letter will be denoted by \( w_i \).

The *length* \( \ell (w) \) of a word \( w \in \bigcup_{n \geq 0} \mathfrak{A}^n \) is defined to be the \( n \in \mathbb{N} \) satisfying \( w \in \mathfrak{A}^n \). Thus, \( w = (w_1, w_2, \ldots, w_{\ell (w)}) \) for every word \( w \).

Given two words \( u \) and \( v \), we say that \( u \) is *longer* than \( v \) (or, equivalently, \( v \) is *shorter* than \( u \)) if and only if \( \ell (u) > \ell (v) \).

The *concatenation* of two words \( u \) and \( v \) is defined to be the word \( (u_1, u_2, \ldots, u_{\ell (u)}, v_1, v_2, \ldots, v_{\ell (v)}) \). This concatenation is denoted by \( uv \) or \( u \cdot v \). The set \( \mathfrak{A}^* \) of all words is a monoid with respect to concatenation, with neutral element \( \emptyset \). It is precisely the free monoid on generators \( \mathfrak{A} \). If \( u \) is a word and \( i \in \mathbb{N} \), we will understand \( u^i \) to mean the *i*-th power of \( u \) in this monoid (that is, the word \( uuu \cdots u \) \( i \) times).

The elements of \( \mathfrak{A} \) are called *letters*, and will be identified with elements of \( \mathfrak{A}^1 = \bigcup_{n \geq 0} \mathfrak{A} = \mathfrak{A}^* \). This identification equates every letter \( u \in \mathfrak{A} \) with the one-letter word \( (u) \in \mathfrak{A}^1 \). Thus, every word

199 They also are involved in indexing basis elements of combinatorial Hopf algebras other than QSym. See Bergeron/Zabrocki [17].
$(u_1, u_2, \ldots, u_n) \in \mathfrak{A}^*$ equals the concatenation $u_1 u_2 \cdots u_n$ of letters, hence allowing us to use $u_1 u_2 \cdots u_n$ as a brief notation for the word $(u_1, u_2, \ldots, u_n)$.

If $w$ is a word, then:

- a prefix of $w$ means a word of the form $(w_1, w_2, \ldots, w_i)$ for some $i \in \{0, 1, \ldots, \ell (w)\}$;
- a suffix of $w$ means a word of the form $(w_{i+1}, w_{i+2}, \ldots, w_{\ell (w)})$ for some $i \in \{0, 1, \ldots, \ell (w)\}$;
- a proper suffix of $w$ means a word of the form $(w_{i+1}, w_{i+2}, \ldots, w_{\ell (w)})$ for some $i \in \{1, 2, \ldots, \ell (w)\}$.

In other words,

- a prefix of $w \in \mathfrak{A}^*$ is a word $u \in \mathfrak{A}^*$ such that there exists a $v \in \mathfrak{A}^*$ satisfying $w = uv$;
- a suffix of $w \in \mathfrak{A}^*$ is a word $v \in \mathfrak{A}^*$ such that there exists a $u \in \mathfrak{A}^*$ satisfying $w = uv$;
- a proper suffix of $w \in \mathfrak{A}^*$ is a word $v \in \mathfrak{A}^*$ such that there exists a nonempty $u \in \mathfrak{A}^*$ satisfying $w = uv$.

Clearly, any proper suffix of $w \in \mathfrak{A}^*$ is a suffix of $w$. Moreover, if $w \in \mathfrak{A}^*$ is any word, then a proper suffix of $w$ is the same thing as a suffix of $w$ distinct from $w$.

We define a relation $\leq$ on the set $\mathfrak{A}^*$ as follows: For two words $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$, we set $u \leq v$ to hold if and only if

either there exists an $i \in \{1, 2, \ldots, \min \{\ell (u), \ell (v)\}\}$ such that $(u_i < v_i$, and every $j \in \{1, 2, \ldots, i-1\}$ satisfies $u_j = v_j)$,

or the word $u$ is a prefix of $v$.

This order relation (taken as the smaller-or-equal relation) makes $\mathfrak{A}^*$ into a poset (by Proposition 6.2(a) below), and we will always be regarding $\mathfrak{A}^*$ as endowed with this poset structure (thus, notations such as $<, \leq, >$ and $\geq$ will be referring to this poset structure). This poset is actually totally ordered (see Proposition 6.2(a)).

Here are some examples of words compared by the relation $\leq$:

\begin{align*}
113 & \leq 114, & 113 & \leq 132, & 19 & \leq 195, & 41 & \leq 412, \\
41 & \leq 421, & 539 & \leq 54, & \emptyset & \leq 21, & \emptyset & \leq \emptyset
\end{align*}

(where $\mathfrak{A}$ is the alphabet $\{1 < 2 < 3 < \cdots\}$).

Notice that if $u$ and $v$ are two words of the same length (i.e., we have $u, v \in \mathfrak{A}^n$ for one and the same $n$), then $u \leq v$ holds if and only if $u$ is lexicographically smaller-or-equal to $v$. In other words, the relation $\leq$ is an extension of the lexicographic order on every $\mathfrak{A}^n$ to $\mathfrak{A}^*$. This is the reason why this relation $\leq$ is usually called the lexicographic order on $\mathfrak{A}^*$. In particular, we will be using this name. However, unlike the lexicographic order on $\mathfrak{A}^n$, it does not always respect concatenation from the right: It can happen that $u, v, w \in \mathfrak{A}^*$ satisfy $u \leq v$ but not $uw \leq vw$. (For example, $u = 1$, $v = 13$ and $w = 4$, again with $\mathfrak{A} = \{1 < 2 < 3 < \cdots\}$.) We will see in Proposition 6.2 that this is rather an exception than the rule and the relation $\leq$ still behaves mostly predictably with respect to concatenation.

Some basic properties of the order relation $\leq$ just defined are collected in the following proposition:

**Proposition 6.2.** (a) The order relation $\leq$ is (the smaller-or-equal relation of) a total order on the set $\mathfrak{A}^*$.

(b) If $a, c, d \in \mathfrak{A}^*$ satisfy $c \leq d$, then $ac \leq ad$.

(c) If $a, c, d \in \mathfrak{A}^*$ satisfy $ac \leq ad$, then $c \leq d$.

(d) If $a, b, c \in \mathfrak{A}^*$ satisfy $a \leq c$, then either we have $ab \leq cd$ or the word $a$ is a prefix of $c$.

(e) If $a, b, c \in \mathfrak{A}^*$ satisfy $ab \leq cd$, then either we have $a \leq c$ or the word $c$ is a prefix of $a$.

(f) If $a, b, c \in \mathfrak{A}^*$ satisfy $ab \leq cd$ and $\ell (a) \leq \ell (c)$, then $a \leq c$.

(g) If $a, b, c \in \mathfrak{A}^*$ satisfy $a \leq b \leq ac$, then $a$ is a prefix of $b$.

(h) If $a \in \mathfrak{A}^*$ is a prefix of $b \in \mathfrak{A}^*$, then $a \leq b$.

(i) If $a$ and $b$ are two prefixes of $c \in \mathfrak{A}^*$, then either $a$ is a prefix of $b$, or $b$ is a prefix of $a$.

(j) If $a, b, c \in \mathfrak{A}^*$ are such that $a \leq b$ and $\ell (a) \geq \ell (b)$, then $ac \leq bc$.

(k) If $a \in \mathfrak{A}^*$ and $b \in \mathfrak{A}^*$ are such that $b$ is nonempty, then $a < ab$.  

The relation $\leq$ is also known as the dictionary order, due to the fact that it is the order in which words appear in a dictionary.
**Exercise 6.3.** Prove Proposition 6.2.

*Hint:* No part of Proposition 6.2 requires more than straightforward case analysis. However, the proof of (a) can be simplified by identifying the order relation ≤ on $A^*$ as a restriction of the lexicographic order on the set $A^\infty$, where $\mathfrak{A}$ is a suitable extension of the alphabet $\mathfrak{A}$. What is this extension, and how to embed $\mathfrak{A}^*$ into $\mathfrak{A}^\infty$?

Proposition 6.2 provides a set of tools for working with the lexicographic order without having to refer to its definition; we shall use it extensively. Proposition 6.2(h) (and its equivalent form stating that $a \leq ac$ for every $a \in A^*$ and $c \in A^*$) and Proposition 6.2(k) will often be used without explicit mention.

Before we define Lyndon words, let us show two more facts about words which will be used later. First, when do words commute?

**Proposition 6.4.** Let $u, v \in A^*$ satisfy $uv = vu$. Then, there exist a $t \in A^*$ and two nonnegative integers $n$ and $m$ such that $u = t^n$ and $v = t^m$.

*Proof.* We prove this by strong induction on $\ell(u) + \ell(v)$. We assume WLOG that $\ell(u)$ and $\ell(v)$ are positive (because otherwise, one of $u$ and $v$ is the empty word, and everything is trivial). It is easy to see that either $u$ is a prefix of $v$, or $v$ is a prefix of $u$.

If the word $u$ is a prefix of $v$, then the statement of Proposition 6.5 can be easily deduced from Proposition 6.4 to $w$ instead of $v$ (by the induction assumption, since $\ell(u) + \ell(w) < \ell(u) + \ell(v)$), and obtain that there exist a $t \in A^*$ and two nonnegative integers $n$ and $m$ such that $u = t^n$ and $w = t^m$. Consider this $t$ and these $n$ and $m$. Of course, $u = t^n$ and $v = u = t^m$. Cancelling $u$ from this equality, we obtain $u = wu$. Now, we can apply Proposition 6.4 to $w$ instead of $v$ (by the induction assumption, since $\ell(u) + \ell(w) < \ell(u) + \ell(v)$), and obtain that there exist a $t \in A^*$ and two nonnegative integers $n$ and $m$ such that $u = t^n$ and $w = t^m$. Consider this $t$ and these $n$ and $m$. Of course, $u = t^n$ and $v = u = t^m$. Hence, Proposition 6.4 is proven.

**Proposition 6.5.** Let $u, v, w \in A^*$ be nonempty words satisfying $uv \geq vu$, $vw \geq uv$ and $wu \geq uw$. Then, there exist a $t \in A^*$ and three nonnegative integers $n$, $m$, and $p$ such that $u = t^n$, $v = t^m$ and $w = t^p$.

*Proof.* We prove this by strong induction on $\ell(u) + \ell(v) + \ell(w)$. Clearly, $\ell(u), \ell(v)$ and $\ell(w)$ are positive (since $u$, $v$ and $w$ are nonempty). We assume WLOG that $\ell(u) = \min\{\ell(u), \ell(v), \ell(w)\}$ (because there is a cyclic symmetry in our situation). Thus, $\ell(u) \leq \ell(v)$ and $\ell(u) \leq \ell(w)$. But $uv \leq wu$. Hence, Proposition 6.2(e) (applied to $a = v$, $b = u$, $c = u$ and $d = v$) yields that either we have $v \leq u$ or the word $u$ is a prefix of $v$. But Proposition 6.2(f) (applied to $a = u$, $b = w$, $c = w$ and $d = u$) yields $u \leq w$ (since $wu \leq uw$ and $\ell(u) \leq \ell(w)$). Furthermore, $wu \leq wv$. Hence, Proposition 6.2(e) (applied to $a = w$, $b = v$, $c = v$ and $d = w$) yields that either we have $w \leq u$ or the word $v$ is a prefix of $w$.

From what we have found so far, it is easy to see that $u$ is a prefix of $v$. In other words, there exists a $v' \in A^*$ such that $v = uv'$. Consider this $v'$.

If the word $v'$ is empty, then the statement of Proposition 6.5 can be easily deduced from Proposition 6.4. Thus, we assume WLOG that this is not the case. Hence, $v'$ is nonempty.

Using $v = uv'$, we can rewrite $uv \geq vu$ as $uwv' \geq vu$. That is, $uv' \leq uwv'$, so that $v'u \leq uv'$ (by Proposition 6.2(c), applied to $a = u$, $c = v'u$ and $d = wv'$). That is, $uv' \geq v'u$. But $\ell(uv') = \ell(u) + \ell(v) = \ell(v')$. Hence, Proposition 6.2(i) (applied to $a = uv'$, $b = w$ and $c = v'$) yields the desired result.

---

201Proof. The word $u$ is a prefix of $uv$. But the word $v$ is also a prefix of $w$ (since $uv = vu$). Hence, Proposition 6.2(i) (applied to $a = u$, $b = v$ and $c = uv'$) yields that either $u$ is a prefix of $v$, or $v$ is a prefix of $u$, qed.

202Proof. Assume the contrary. Then, $u$ is not a prefix of $v$. Hence, we must have $v \leq u$ (since either we have $v \leq u$ or the word $u$ is a prefix of $v$), and in fact $v < u$ (because $u = u$ would contradict to $u$ not being a prefix of $v$). Thus, $v < u \leq w$. But recall that either we have $w \leq v$ or the word $v$ is a prefix of $w$. Thus, $v$ must be a prefix of $w$ (because $v < w$ rules out $w \leq v$). In other words, there exists a $q \in A^*$ such that $w = vq$. Consider this $q$. We have $v < w \leq wq$. Thus, Proposition 6.2(g) (applied to $a = v$, $b = u$ and $c = q$) yields that $v$ is a prefix of $u$. This contradicts $v < u$. This contradiction completes this proof.

203Proof. Assume that the word $v'$ is empty. Then, $v = uv'$ becomes $v = u$. Therefore, $vu \geq wu$ becomes $uv \geq vu$. Combined with $wu \geq vu$, this yields $uv = vu$. Hence, Proposition 6.4 (applied to $w$ instead of $v$) yields that there exist a $t \in A^*$ and two nonnegative integers $n$ and $m$ such that $u = t^n$ and $w = t^m$. Clearly, $v = u = t^n$ as well, and so the statement of Proposition 6.5 is true.
Corollary 6.6. Let \( u,v,w \in \mathfrak{A}^* \) be words satisfying \( uv \geq vu \) and \( vw \geq vw \). Assume that \( v \) is nonempty. Then, \( uv \geq uw \).

Proof. Assume the contrary. Thus, \( uw < wu \), so that \( wu \geq uv \).

If \( u \) or \( w \) is empty, then everything is obvious. We thus WLOG assume that \( u \) and \( w \) are nonempty. Thus, Proposition 6.5 shows that there exist a \( t \in \mathfrak{A}^* \) and three nonnegative integers \( n,m \), and \( p \) such that \( u = t^n \), \( v' = t^m \) and \( w = t^p \). Clearly, this \( t \) and these \( n,m,p \) satisfy \( v = u^{n \cdot m} v' = t^n t^m = t^{n+m} \), and so the statement of Proposition 6.5 is satisfied. The induction step is thus complete.

Exercise 6.7. Find an alternative proof of Corollary 6.6 which does not use Proposition 6.5.

The above results have a curious consequence, which we are not going to use:

Corollary 6.8. We can define a preorder on the set \( \mathfrak{A}^* \setminus \{\emptyset\} \) of all nonempty words by defining a nonempty word \( u \) to be greater-or-equal to a nonempty word \( v \) (with respect to this preorder) if and only if \( uv \geq vu \).

Two nonempty words \( u,v \) are equivalent with respect to the equivalence relation induced by this preorder if and only if there exist a \( t \in \mathfrak{A}^* \) and two nonnegative integers \( n,m \) such that \( u = t^n \) and \( v = t^m \).

Proof. The alleged preorder is transitive (by Corollary 6.6) and reflexive (obviously), and hence is really a preorder. The claim in the second sentence follows from Proposition 6.4.

As another consequence of Proposition 6.5, we obtain a classical property of words [106, Proposition 1.3.1]:

Exercise 6.9. Let \( u \) and \( v \) be words and \( n \) and \( m \) be positive integers such that \( u^n = v^m \). Prove that there exists a word \( t \) and positive integers \( i,j \) such that \( u = t^i \) and \( v = t^j \).

Here is another application of Corollary 6.6:

Exercise 6.10. Let \( n \) and \( m \) be positive integers. Let \( u \in \mathfrak{A}^* \) and \( v \in \mathfrak{A}^* \) be two words. Prove that \( uv \geq vu \) holds if and only if \( u^n v^m \geq v^m u^n \) holds.

Exercise 6.11. Let \( n \) and \( m \) be positive integers. Let \( u \in \mathfrak{A}^* \) and \( v \in \mathfrak{A}^* \) be two words satisfying \( n \ell (u) = m \ell (v) \). Prove that \( uv \geq vu \) holds if and only if \( u^n \geq v^m \) holds.

We can also generalize Propositions 6.4 and 6.5:

Exercise 6.12. Let \( u_1,u_2, \ldots, u_k \) be nonempty words such that every \( i \in \{1,2,\ldots,k\} \) satisfies \( u_i u_{i+1} \geq u_{i+1} u_i \), where \( u_{k+1} \) means \( u_1 \). Show that there exist a word \( t \) and nonnegative integers \( n_1,n_2, \ldots, n_k \) such that \( u_1 = t^{n_1}, u_2 = t^{n_2}, \ldots, u_k = t^{n_k} \).

Now, we define the notion of a Lyndon word. There are several definitions in literature, some of which will be proven equivalent in Theorem 6.20.

Definition 6.13. A word \( w \in \mathfrak{A}^* \) is said to be Lyndon if it is nonempty and satisfies the following property: Every nonempty proper suffix of \( w \) satisfies \( v > w \).

For example, the word 113 is Lyndon (because its nonempty proper suffixes are 13 and 3, and these are both \( > 113 \)), and the word 242427 is Lyndon (its nonempty proper suffixes are 422427, 2427, 427, 27 and 7, and again these are each \( > 242427 \)). The words 2424 and 35346 are not Lyndon (the word 2424 has a
nonempty proper suffix $24 \leq 2424$, and the word $35346$ has a nonempty proper suffix $346 \leq 35346$). Every word of length $1$ is Lyndon (since it has no nonempty proper suffixes). A word $w = (w_1, w_2)$ with two letters is Lyndon if and only if $w_1 < w_2$. A word $w = (w_1, w_2, w_3)$ of length $3$ is Lyndon if and only if $w_1 < w_3$ and $w_1 \leq w_2$. A four-letter word $w = (w_1, w_2, w_3, w_4)$ is Lyndon if and only if $w_1 < w_4$, $w_1 \leq w_3$, $w_1 \leq w_2$ and (if $w_1 = w_4$ then $w_2 < w_4$). (These rules only get more complicated as the words grow longer.)

We will show several properties of Lyndon words now. We begin with trivialities which will make some arguments a bit shorter:

**Proposition 6.14.** Let $w$ be a Lyndon word. Let $u$ and $v$ be words such that $w = uv$.

(a) If $v$ is nonempty, then $v \geq u$.
(b) If $v$ is nonempty, then $v > u$.
(c) If $u$ and $v$ are nonempty, then $vu > uv$.
(d) We have $uv \geq u$.

*Proof.* (a) Assume that $v$ is nonempty. Clearly, $v$ is a suffix of $w$ (since $w = uv$). If $v$ is a proper suffix of $w$, then the definition of a Lyndon word yields that $v > w$ (since $w$ is a Lyndon word); otherwise, $v$ must be $w$ itself. In either case, we have $v \geq w$. Hence, Proposition 6.14(a) is proven.

(b) Assume that $v$ is nonempty. From Proposition 6.14(a), we obtain $v \geq w = uv > u$ (since $v$ is nonempty). This proves Proposition 6.14(b).

(c) Assume that $u$ and $v$ are nonempty. Since $u$ is nonempty, we have $vu > v \geq w$ (by Proposition 6.14(a)). Since $w = uv$, this becomes $vu > uv$. This proves Proposition 6.14(c).

(d) We need to prove that $vu \geq u$. If either $u$ or $v$ is empty, $vu$ and $uv$ are obviously equal, and thus $vu \geq u$ is true in this case. Hence, we can WLOG assume that $u$ and $v$ are nonempty. Assume this. Then, $vu \geq uv$ follows from Proposition 6.14(c). This proves Proposition 6.14(d). $\square$

**Corollary 6.15.** Let $w$ be a Lyndon word. Let $v$ be a nonempty suffix of $w$. Then, $v \geq w$.

*Proof.* Since $v$ is a nonempty suffix of $w$, there exists $u \in \mathfrak{A}^*$ such that $w = uv$. Thus, $v \geq w$ follows from Proposition 6.14(a). $\square$

Our next proposition is [69, Lemma 6.5.4]; its part (a) is also [137, (5.1.2)]:

**Proposition 6.16.** Let $u$ and $v$ be two Lyndon words such that $u < v$. Then:

(a) The word $uv$ is Lyndon.
(b) We have $w < v$.

*Proof.* (b) The word $u$ is Lyndon and thus nonempty. Hence, $uv \neq v$.

204. If $uv \leq v\varnothing$, then Proposition 6.16(b) easily follows. Hence, for the rest of this proof, we can WLOG assume that $w \neq v \varnothing$. Assume this.

We have $u < v$. Hence, Proposition 6.2(d) (applied to $a = u$, $b = v$, $c = v$ and $d = \varnothing$) yields that either we have $uv \leq v\varnothing$ or the word $u$ is a prefix of $v$. Since we don’t have $uv \leq v\varnothing$, we thus see that the word $u$ is a prefix of $v$. In other words, there exists a $t \in \mathfrak{A}^*$ satisfying $v = ut$. Consider this $t$. Then, $t$ is nonempty (else we would have $v = u t = u$ in contradiction to $u < v$).

Now, $v = ut$. Hence, $t$ is a proper suffix of $v$ (proper because $u$ is nonempty). Thus, $t$ is a nonempty proper suffix of $v$. Since every nonempty proper suffix of $v$ is $v$ (because $v$ is Lyndon), this shows that $t > v$. Hence, $v \leq t$. Thus, Proposition 6.2(b) (applied to $a = u$, $c = v$ and $d = t$) yields $uv \leq ut = v$. Combined with $uv \neq v$, this yields $uv < v$. Hence, Proposition 6.16(b) is proven.

(a) The word $v$ is nonempty (since it is Lyndon). Hence, $uv$ is nonempty. It thus remains to check that every nonempty proper suffix $p$ of $uv$ satisfies $p > uv$.

So let $p$ be a nonempty proper suffix of $uv$. We must show that $p > uv$. Since $p$ is a nonempty proper suffix of $uv$, we must be in one of the following two cases (depending on whether this suffix begins before the suffix $v$ of $uv$ begins or afterwards):

204.*Proof.* Assume the contrary. Then, $uv = v$. Thus, $uv = v \varnothing v$. Cancelling $v$ from this equation, we obtain $u = \varnothing$. That is, $u$ is empty. This contradicts the fact that $u$ is nonempty. This contradiction proves that our assumption was wrong, qed.

205.*Proof.* Assume that $uv \leq v\varnothing$. Thus, $uv \leq v\varnothing = v$. Since $uv \neq v$, this becomes $uv < v$, so that Proposition 6.16(b) is proven.
Case 1: The word $p$ is a nonempty suffix of $v$. (Note that $p = v$ is allowed.)

Case 2: The word $p$ has the form $qv$ where $q$ is a nonempty proper suffix of $u$.

Let us first handle Case 1. In this case, $p$ is a nonempty suffix of $v$. Since $v$ is Lyndon, this yields that $p \geq v$ (by Corollary 6.15, applied to $v$ and $p$ instead of $w$ and $v$). But Proposition 6.16(b) yields $uv < v$, thus $v > uv$. Hence, $p \geq v > uv$. We thus have proven $p > uv$ in Case 1.

Let us now consider Case 2. In this case, $p$ has the form $qv$ where $q$ is a nonempty proper suffix of $u$. Consider this $q$. Clearly, $q > u$ (since $u$ is Lyndon and since $q$ is a nonempty proper suffix of $u$), so that $u \leq q$. Thus, Proposition 6.2(d) (applied to $a = u$, $b = v$, $c = q$ and $d = v$) yields that either we have $uv \leq qv$ or the word $u$ is a prefix of $q$. Since $u$ being a prefix of $q$ is impossible (in fact, $q$ is a proper suffix of $u$, thus shorter than $u$), we thus must have $uv \leq qv$. Since $uv \neq qv$ (because otherwise we would have $uv = qv$, thus $u = q$ (because we can cancel $v$ from the equality $uv = qv$), contradicting $q > u$), this can be strengthened to $uv < qv = p$. Thus, $p > uv$ is proven in Case 2 as well.

Now that $p > uv$ is shown to hold in both cases, we conclude that $p > uv$ always holds.

Now, let us forget that we fixed $p$. We have thus shown that every nonempty proper suffix $p$ of $uv$ satisfies $p > uv$. Since $uv$ is nonempty, this yields that $uv$ is Lyndon (by the definition of a Lyndon word). Thus, the proof of Proposition 6.16(a) is complete.

Proposition 6.16(b), combined with Corollary 6.6, leads to a technical result which we will find good use for later:

\textbf{Corollary 6.17.} Let $u$ and $v$ be two Lyndon words such that $u < v$. Let $z$ be a word such that $zv \geq vz$ and $uz \geq zu$. Then, $z$ is the empty word.

\textit{Proof.} Assume the contrary. Then, $z$ is nonempty. Thus, Corollary 6.6 (applied to $z$ and $v$ instead of $v$ and $w$) yields $uv \geq vu$. But Proposition 6.16(b) yields $uv < v \leq vu$, contradicting $uv \geq vu$. This contradiction completes our proof. \hfill \Box

We notice that the preorder of Corollary 6.8 becomes particularly simple on Lyndon words:

\textbf{Proposition 6.18.} Let $u$ and $v$ be two Lyndon words. Then, $u \geq v$ if and only if $uv \geq vu$.

\textit{Proof.} We distinguish between three cases:

\textit{Case 1:} We have $u < v$.

\textit{Case 2:} We have $u = v$.

\textit{Case 3:} We have $u > v$.

Let us consider Case 1. In this case, we have $u < v$. Thus,

\[ uv < v \leq vu. \]

Hence, we have neither $u \geq v$ nor $uv \geq vu$ (because we have $u < v$ and $uv < vu$). Thus, Proposition 6.18 is proven in Case 1.

In Case 2, we have $u = v$. Therefore, in Case 2, both inequalities $u \geq v$ and $uv \geq vu$ hold (and actually are equalities). Thus, Proposition 6.18 is proven in Case 2 as well.

Let us finally consider Case 3. In this case, we have $u > v$. In other words, $v < u$. Thus,

\[ vu < u \leq uv \]

Hence, we have both $u \geq v$ and $uv \geq vu$ (because we have $v < u$ and $uv = vu$). Thus, Proposition 6.18 is proven in Case 3.

Proposition 6.18 is now proven in all three possible cases. \hfill \Box

\textbf{Proposition 6.19.} Let $w$ be a nonempty word. Let $v$ be the (lexicographically) smallest nonempty suffix of $w$. Then:

\begin{enumerate}
  \item The word $v$ is a Lyndon word.
  \item Assume that $w$ is not a Lyndon word. Then there exists a nonempty $u \in \mathbb{A}^*$ such that $w = uv$, $u \geq v$ and $uv \geq vu$.
\end{enumerate}
Proof. (a) Every nonempty proper suffix of \( v \) is \( \geq v \) (since every nonempty proper suffix of \( v \) is a nonempty suffix of \( w \), but \( v \) is the smallest such suffix) and therefore \( > v \) (since a proper suffix of \( v \) cannot be \( = v \)). Combined with the fact that \( v \) is nonempty, this yields that \( v \) is Lyndon. Proposition 6.19(a) is proven.

(b) Assume that \( w \) is not a Lyndon word. Then, \( w \neq v \) (since \( v \) is Lyndon (by Proposition 6.19(a)) while \( w \) is not). Now, \( v \) is a suffix of \( w \). Thus, there exists an \( u \in \mathfrak{A}^* \) such that \( w = uv \). Consider this \( u \). Clearly, \( u \) is nonempty (since \( uv = w \neq v \)). Assume (for the sake of contradiction) that \( u < v \). Let \( v' \) be the (lexicographically) smallest nonempty suffix of \( u \). Then, \( v' \) is a Lyndon word (by Proposition 6.19(a), applied to \( u \) and \( v' \) instead of \( w \) and \( v \)) and satisfies \( v' \leq u \) (since \( v' \) is a nonempty suffix of \( u \), whereas \( v' \) is the smallest such suffix). Thus, \( v' \) and \( v \) are Lyndon words such that \( v' \leq u < v \). Proposition 6.16(a) (applied to \( v' \) instead of \( u \)) now yields that the word \( v'v \) is Lyndon. Hence, every nonempty proper suffix of \( v'v \) is \( > v'v \). Since \( v \) is a nonempty proper suffix of \( v'v \), this yields that \( v > v'v \).

But \( v' \) is a nonempty suffix of \( u \), so that \( v'v \) is a nonempty suffix of \( uv = w \). Since \( v \) is the smallest such suffix, this yields that \( v'v \geq v \). This contradicts \( v > v'v \). Our assumption (that \( u < v \)) therefore fails. We conclude that \( u \geq v \).

It remains to prove that \( uv \geq vu \). Assume the contrary. Then, \( uv < vu \). Thus, there exists at least one suffix \( t \) of \( u \) such that \( tv < vt \) (namely, \( t = u \)). Let \( p \) be the minimum-length such suffix. Then, \( pv < vp \). Thus, \( p \) is nonempty.

Since \( p \) is a suffix of \( u \), it is clear that \( pv \) is a suffix of \( uv = w \). So we know that \( pv \) is a nonempty suffix of \( w \). Since \( v \) is the smallest such suffix, this yields that \( v \leq pv < vp \). Thus, Proposition 6.2(g) (applied to \( a = v, b = pv \) and \( c = p \)) yields that \( v \) is a prefix of \( pv \). In other words, there exists a \( q \in \mathfrak{A}^* \) such that \( pv = vq \). Consider this \( q \). This \( q \) is nonempty (because otherwise we would have \( pv = vq = v \), contradicting the fact that \( p \) is nonempty). From \( vq = pv < vp \), we obtain \( q \leq p \) (by Proposition 6.2(c), applied to \( a = v, c = q \) and \( d = p \)).

We know that \( q \) is a suffix of \( pv \) (since \( vq = pv \)), whereas \( pv \) is a suffix of \( w \). Thus, \( q \) is a suffix of \( w \). So \( q \) is a nonempty suffix of \( w \). Since \( v \) is the smallest such suffix, this yields that \( v \leq q \). We now have \( v \leq q \leq p \leq pv < vp \). Hence, \( v \) is a prefix of \( p \) (by Proposition 6.2(g), applied to \( a = v, b = p \) and \( c = p \)). In other words, there exists an \( r \in \mathfrak{A}^* \) such that \( p = vr \). Consider this \( r \). Clearly, \( r \) is a suffix of \( p \), while \( p \) is a suffix of \( u \); therefore, \( r \) is a suffix of \( u \). Also, \( pv < vp \) rewrites as \( vrv < vvr \) (because \( p = vr \)). Thus, Proposition 6.2(c) (applied to \( a = v, c = rc \) and \( d = vr \)) yields \( rv \leq vrc \). Since \( rv \neq vr \) (because otherwise, we would have \( rv = vr \), thus \( vrv = vvr \), contradicting \( vrv < vvr \)), this becomes \( rv < vr \).

Now, \( r \) is a suffix of \( u \) such that \( rv < vr \). Since \( p \) is the minimum-length such suffix, this yields \( \ell (r) \geq \ell (p) \). But this contradicts the fact that \( \ell (p) = \ell (vr) = \ell (v) + \ell (r) > \ell (r) \). This contradiction proves our assumption wrong; thus, we have shown that \( uv \geq vu \). Proposition 6.19(b) is proven.

\( \Box \)

**Theorem 6.20.** Let \( w \) be a nonempty word. The following four assertions are equivalent:

- **Assertion A:** The word \( w \) is Lyndon.
- **Assertion B:** Any nonempty words \( u \) and \( v \) satisfying \( w = uv \) satisfy \( v > w \).
- **Assertion C:** Any nonempty words \( u \) and \( v \) satisfying \( w = uv \) satisfy \( v > u \).
- **Assertion D:** Any nonempty words \( u \) and \( v \) satisfying \( w = uv \) satisfy \( vu > wv \).

**Proof.** Proof of the implication \( A \implies B \): If Assertion \( A \) holds, then Assertion \( B \) clearly holds (in fact, whenever \( u \) and \( v \) are nonempty words satisfying \( w = uv \), then \( v \) is a nonempty proper suffix of \( w \), and therefore \( > w \) by the definition of a Lyndon word).

Proof of the implication \( A \implies C \): This implication follows from Proposition 6.14(b).

Proof of the implication \( A \implies D \): This implication follows from Proposition 6.14(c).

Proof of the implication \( B \implies A \): Assume that Assertion \( B \) holds. If \( v \) is a nonempty proper suffix of \( w \), then there exists an \( u \in \mathfrak{A}^* \) satisfying \( w = uv \). This \( u \) is nonempty because \( v \) is a proper suffix, and thus Assertion \( B \) yields \( v > w \). Hence, every nonempty proper suffix \( v \) of \( w \) satisfies \( v > w \). By the definition of a Lyndon word, this yields that \( w \) is Lyndon, so that Assertion \( A \) holds.
Proof of the implication \(\mathcal{C} \implies \mathcal{A}\): Assume that Assertion \(\mathcal{C}\) holds. If \(w\) was not Lyndon, then Proposition 6.19(b) would yield nonempty words \(u\) and \(v\) such that \(w = uv\) and \(u \geq v\); this would contradict Assertion \(\mathcal{C}\). Thus, \(w\) is Lyndon, and Assertion \(\mathcal{A}\) holds.

Proof of the implication \(\mathcal{D} \implies \mathcal{A}\): Assume that Assertion \(\mathcal{D}\) holds. If \(w\) was not Lyndon, then Proposition 6.19(b) would yield nonempty words \(u\) and \(v\) such that \(w = uv\) and \(uv \geq vv\); this would contradict Assertion \(\mathcal{D}\). Thus, \(w\) is Lyndon, and Assertion \(\mathcal{A}\) holds.

Now we have proven enough implications to conclude the equivalence of all four assertions. \(\square\)

Theorem 6.20 connects our definition of Lyndon words with some of the definitions appearing in literature. For example, Lothaire [106, §5.1], Shirshov [153] and de Bruijn/Klarner [24, §4] define Lyndon words using Assertion \(\mathcal{D}\) (note, however, that Shirshov takes \(<\) instead of \(>\) and calls Lyndon words “regular words”; also, de Bruijn/Klarner call Lyndon words “normal words”). Chen-Fox-Lyndon [31, §1], Reutenauer [137] and Radford [134] use our definition (but Chen-Fox-Lyndon call the Lyndon words “standard sequences”, and Radford calls them “primes” and uses \(\leq\) instead of \(>\)).

Theorem 6.20 appears (with different notations) in Zhou-Lu [175, Proposition 1.4]. The equivalence \(\mathcal{D} \iff \mathcal{A}\) of our Theorem 6.20 is equivalent to [106, Proposition 5.12] and to [31, \(\mathfrak{X}' = \mathfrak{X}''\)].

The following exercise provides a different (laborious) approach to Theorem 6.20:

Exercise 6.21. (a) Prove that if \(u \in \mathfrak{X}^*\) and \(v \in \mathfrak{X}^*\) are two words satisfying \(uv \prec vu\), then there exists a nonempty suffix \(s\) of \(u\) satisfying \(sv \prec v\).

(b) Give a new proof of Theorem 6.20 (avoiding the use of Proposition 6.19).

[Hint: For (a), perform strong induction on \(\ell(u) + \ell(v)\), assume the contrary, and distinguish between the case when \(u \leq v\) and the case when \(v\) is a prefix of \(u\). For (b), use part (a) in proving the implication \(\mathcal{D} \implies \mathcal{B}\), and factor \(v\) as \(v = w^m v'\) with \(m\) maximal in the proof of the implication \(\mathcal{C} \implies \mathcal{B}\).]

The following two exercises are taken from [67]206.

Exercise 6.22. Let \(w\) be a nonempty word. Prove that \(w\) is Lyndon if and only if every nonempty word \(t\) and every positive integer \(n\) satisfy (if \(w \leq t^n\), then \(w \leq t\)).

Exercise 6.23. Let \(w_1, w_2, \ldots, w_n\) be \(n\) Lyndon words, where \(n\) is a positive integer. Assume that \(w_1 \leq w_2 \leq \cdots \leq w_n\) and \(w_1 < w_n\). Show that \(w_1 w_2 \cdots w_n\) is a Lyndon word.

The following exercise is a generalization (albeit not in an obvious way) of Exercise 6.23:

Exercise 6.24. Let \(w_1, w_2, \ldots, w_n\) be \(n\) Lyndon words, where \(n\) is a positive integer. Assume that \(w_i^j w_{i+1} \cdots w_n \geq w_i w_{i+1} \cdots w_n\) for every \(i \in \{1, 2, \ldots, n\}\). Show that \(w_1 w_2 \cdots w_n\) is a Lyndon word.

We are now ready to meet the one of the most important features of Lyndon words: a bijection between all words and multisets of Lyndon words207; it is clear that such a bijection is vital for constructing polynomial generating sets of commutative algebras with bases indexed by words, such as QSym or shuffle algebras. This is given by the Chen-Fox-Lyndon factorization:

Definition 6.25. Let \(w\) be a word. A Chen-Fox-Lyndon factorization (in short, CFL factorization) of \(w\) means a tuple \((a_1, a_2, \ldots, a_k)\) of Lyndon words satisfying \(w = a_1 a_2 \cdots a_k\) and \(a_1 \geq a_2 \geq \cdots \geq a_k\).

Example 6.26. The tuple \((23, 2, 14, 13323, 13, 12, 12, 1)\) is a CFL factorization of the word 23214133231312121 over the alphabet \(\{1, 2, 3, \ldots\}\) (ordered by \(1 < 2 < 3 < \cdots\)), since 23, 2, 14, 13323, 13, 12, 12 and 1 are Lyndon words satisfying 23214133231312121 = 23 · 2 · 14 · 13323 · 13 · 12 · 12 · 1 and 23 ≥ 2 ≥ 14 ≥ 13323 ≥ 13 ≥ 12 ≥ 12 ≥ 1.

The bijection is given by the following Chen-Fox-Lyndon theorem ([69, Theorem 6.5.5], [106, Thm. 5.1.5], [134, part of Thm. 2.1.4]):

Theorem 6.27. Let \(w\) be a word. Then, there exists a unique CFL factorization of \(w\).

Before we prove this, we need to state a lemma (which is [106, Proposition 5.1.6]):
Lemma 6.28. Let \((a_1, a_2, \ldots, a_k)\) be a CFL factorization of a nonempty word \(w\). Let \(p\) be a nonempty suffix of \(w\). Then, \(p \geq a_k\).

Proof of Lemma 6.28. We will prove Lemma 6.28 by induction over the (obviously) positive integer \(k\).

**Induction base:** Assume that \(k = 1\). Thus, \((a_1, a_2, \ldots, a_k) = (a_1)\) is a tuple of Lyndon words satisfying \(w = a_1a_2 \cdots a_k\). We have \(w = a_1a_2 \cdots a_k = a_1\) (since \(k = 1\)), so that \(w\) is a Lyndon word (since \(a_1\) is a Lyndon word). Thus, Corollary 6.15 (applied to \(v = p\)) yields \(p \geq a_1 = a_k\) (since \(1 = k\)). Thus, Lemma 6.28 is proven in the case \(k = 1\). The induction base is complete.

**Induction step:** Let \(K\) be a positive integer. Assume (as the induction hypothesis) that Lemma 6.28 is proven for \(k = K\). We now need to show that Lemma 6.28 holds for \(k = K + 1\).

So let \((a_1, a_2, \ldots, a_{K+1})\) be a CFL factorization of a nonempty word \(w\). Let \(p\) be a nonempty suffix of \(w\). We need to prove that \(p \geq a_{K+1}\).

By the definition of a CFL factorization, \((a_1, a_2, \ldots, a_{K+1})\) is a tuple of Lyndon words satisfying \(w = a_1a_2 \cdots a_{K+1}\) and \(a_1 \geq a_2 \geq \cdots \geq a_{K+1}\). Let \(w' = a_2a_3 \cdots a_{K+1};\) then, \(w = a_1a_2 \cdots a_{K+1} = a_1(a_2a_3 \cdots a_{K+1}) = a_1w'\). Hence, every nonempty suffix of \(w\) is either a nonempty suffix of \(w'\), or has the form \(qw'\) for a nonempty suffix \(q\) of \(a_1\). Since \(p\) is a nonempty suffix of \(w\), we thus must be in one of the following two cases:

1. **Case 1:** The word \(p\) is a nonempty suffix of \(w'\).
   - Let us first consider Case 1. In this case, \(p\) is a nonempty suffix of \(w'\). The \(K\)-tuple \((a_2, a_3, \ldots, a_{K+1})\) of Lyndon words satisfies \(w' = a_2a_3 \cdots a_{K+1}\) and \(a_2 \geq a_3 \geq \cdots \geq a_{K+1}\); therefore, \((a_2, a_3, \ldots, a_{K+1})\) is a CFL factorization of \(w'\). We can thus apply Lemma 6.28 to \(K\), \(w'\) and \((a_2, a_3, \ldots, a_{K+1})\) instead of \(k\), \(w\) and \((a_1, a_2, \ldots, a_k)\) (because we assumed that Lemma 6.28 is proven for \(k = K\)). As a result, we obtain that \(p \geq a_{K+1}\). Thus, \(p \geq a_{K+1}\) is proven in Case 1.
   - Let us now consider Case 2. In this case, \(p\) has the form \(qw'\) for a nonempty suffix \(q\) of \(a_1\). Consider this \(q\). Since \(a_1\) is a Lyndon word, we have \(q \geq a_1\) (by Corollary 6.15, applied to \(a_1\) and \(q\) instead of \(w\) and \(v\)). Thus, \(q \geq a_1 \geq a_2 \geq \cdots \geq a_{K+1}\), so that \(p = qw' \geq q \geq a_{K+1}\). Thus, \(p \geq a_{K+1}\) is proven in Case 2.
   - We have now proven \(p \geq a_{K+1}\) in all cases. This proves that Lemma 6.28 holds for \(k = K + 1\). The induction step is thus finished, and with it the proof of Lemma 6.28.

**Proof of Theorem 6.27.** Let us first prove that there exists a CFL factorization of \(w\).

Indeed, there clearly exists a tuple \((a_1, a_2, \ldots, a_k)\) of Lyndon words satisfying \(w = a_1a_2 \cdots a_k\). Fix such a tuple with minimum \(k\). We claim that \(a_1 \geq a_2 \geq \cdots \geq a_k\). Indeed, if some \(i \in \{1, 2, \ldots, k - 1\}\) would satisfy \(a_i < a_{i+1}\), then the word \(a_ia_{i+1}\) would be Lyndon (by Proposition 6.16(a), applied to \(u = a_i\) and \(v = a_{i+1}\)), whence \((a_1, a_2, \ldots, a_{i-1}, a_ia_{i+1}, a_{i+2}, a_{i+3}, \ldots, a_k)\) would also be a tuple of Lyndon words satisfying \(w = a_1a_2 \cdots a_{i-1}(a_ia_{i+1})a_{i+2}a_{i+3} \cdots a_k\) but having length \(k - 1 < k\), contradicting the fact that \(k\) is the minimum length of such a tuple. Hence, no \(i \in \{1, 2, \ldots, k - 1\}\) can satisfy \(a_i < a_{i+1}\). In other words, every \(i \in \{1, 2, \ldots, k - 1\}\) satisfies \(a_i \geq a_{i+1}\). In other words, \(a_1 \geq a_2 \geq \cdots \geq a_k\). Thus, \((a_1, a_2, \ldots, a_k)\) is a CFL factorization of \(w\), so we have shown that such a CFL factorization exists.

It remains to show that there exists at most one CFL factorization of \(w\). We shall prove this by induction over \(\ell(w)\). Thus, we fix a word \(w\) and assume that

\[(1.1) \quad \text{for every word } v \text{ with } \ell(v) < \ell(w), \text{ there exists at most one CFL factorization of } v.\]

We now have to prove that there exists at most one CFL factorization of \(w\).

Indeed, let \((a_1, a_2, \ldots, a_k)\) and \((b_1, b_2, \ldots, b_m)\) be two CFL factorizations of \(w\). We need to prove that \((a_1, a_2, \ldots, a_k) = (b_1, b_2, \ldots, b_m)\). If \(w\) is empty, then this is obvious, so we WLOG assume that it is not; thus, \(k > 0\) and \(m > 0\).

Since \((b_1, b_2, \ldots, b_m)\) is a CFL factorization of \(w\), we have \(w = b_1b_2 \cdots b_m\), and thus \(b_m\) is a nonempty suffix of \(w\). Thus, Lemma 6.28 (applied to \(p = b_m\)) yields \(b_m \geq a_k\). The same argument (but with the roles of \((a_1, a_2, \ldots, a_k)\) and \((b_1, b_2, \ldots, b_m)\) switched) shows that \(a_k \geq b_m\). Combined with \(b_m \geq a_k\), this yields \(a_k = b_m\). Now let \(v = a_1a_2 \cdots a_{k-1}\). Then, \((a_1, a_2, \ldots, a_{k-1})\) is a CFL factorization of \(v\) (since \(a_1 \geq a_2 \geq \cdots \geq a_{k-1}\)).

208 For instance, the tuple \((w_1, w_2, \ldots, w_{\ell(w)})\) of one-letter words is a valid example (recall that one-letter words are always Lyndon).
Since \((a_1, a_2, \ldots, a_k)\) is a CFL factorization of \(w\), we have \(w = a_1a_2 \cdots a_k = a_1a_2 \cdots a_{k-1} a_k = vb_m\), so that \(vb_m = w = b_1b_2 \cdots b_m = b_1b_2 \cdots b_{m-1}b_m\).

Cancelling \(b_m\) yields \(v = b_1b_2 \cdots b_{m-1}\). Thus, \((b_1, b_2, \ldots, b_{m-1})\) is a CFL factorization of \(v\) (since \(b_1 \geq b_2 \geq \cdots \geq b_{m-1}\)). Since \(\ell(v) < \ell(w)\) (because \(v = a_1a_2 \cdots a_{k-1}\) is shorter than \(w = a_1a_2 \cdots a_k\)), we can apply (6.1) to obtain that there exists at most one CFL factorization of \(v\). But we already know two such CFL factorizations: \((a_1, a_2, \ldots, a_{k-1})\) and \((b_1, b_2, \ldots, b_{m-1})\). Thus, \((a_1, a_2, \ldots, a_{k-1}) = (b_1, b_2, \ldots, b_{m-1})\), which, combined with \(a_k = b_m\), leads to \((a_1, a_2, \ldots, a_k) = (b_1, b_2, \ldots, b_m)\). This is exactly what we needed to prove.

So we have shown (by induction) that there exists at most one CFL factorization of \(w\). This completes the proof of Theorem 6.27. \(\square\)

The CFL factorization allows us to count all Lyndon words of a given length if \(\mathcal{A}\) is finite:

**Exercise 6.29.** Assume that the alphabet \(\mathcal{A}\) is finite. Let \(q = |\mathcal{A}|\). Show that the number of Lyndon words of length \(n\) equals \(\frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}\) for every positive integer \(n\) (where \(\sum_{d|n}\) means a sum over all positive divisors of \(n\), and where \(\mu\) is the number-theoretic Möbius function).

Exercise 6.29 is a well-known result and appears, e.g., in [31, Theorem 1.5].

We will now study another kind of factorization: not of an arbitrary word into Lyndon words, but of a proper suffix of \(w\). This factorization is called standard factorization ([106, 5.5.1]) or canonical factorization ([69, Lemma 6.5.33]); we only introduce it from the viewpoint we are interested in, namely its providing a way to do induction over Lyndon words\(^{210}\). Here is what we need to know:

**Theorem 6.30.** Let \(w\) be a Lyndon word of length \(>1\). Let \(v\) be the (lexicographically) smallest nonempty proper suffix of \(w\). Since \(v\) is a proper suffix of \(w\), there exists a nonempty \(u \in \mathcal{A}^*\) such that \(w = uv\). Consider this \(u\). Then:

(a) The words \(u\) and \(v\) are Lyndon.
(b) We have \(u < v < w\).

**Proof.** Every nonempty proper suffix of \(v\) is \(\geq v\) (since every nonempty proper suffix of \(v\) is a nonempty proper suffix of \(w\), but \(v\) is the smallest such suffix) and therefore > \(v\) (since a proper suffix of \(v\) cannot be \(v\)). Combined with the fact that \(v\) is nonempty, this yields that \(v\) is Lyndon.

Since \(w\) is Lyndon, we know that every nonempty proper suffix of \(w\) is > \(w\). Applied to the nonempty proper suffix \(v\) of \(w\), this yields that \(v > w\). Hence, \(w < v\). Since \(v\) is nonempty, we have \(u < uv = w < v\). This proves Theorem 6.30(b).

Let \(p\) be a nonempty proper suffix of \(u\). Then, \(pv\) is a nonempty proper suffix of \(uw = w\). Thus, \(pv > w\) (since every nonempty proper suffix of \(w\) is > \(w\)). Thus, \(pv > w = uv\), so that \(uv < pv\). Thus, Proposition 6.2(e) (applied to \(a = u, b = v, c = p\) and \(d = v\)) yields that either we have \(u \leq p\) or the word \(p\) is a prefix of \(u\).

Let us assume (for the sake of contradiction) that \(p \leq u\). Then, \(p < u\) (because \(p\) is a proper suffix of \(u\), and therefore \(p \neq u\)). Hence, we cannot have \(u \leq p\). Thus, the word \(p\) is a prefix of \(u\) (since either we have \(u \leq p\) or the word \(p\) is a prefix of \(u\)). In other words, there exists a \(q \in \mathcal{A}^*\) such that \(u = pq\). Consider this \(q\). We have \(w = u = pv = pqv = pv = qv\), and thus \(qv\) is a proper suffix of \(w\) (proper because \(p\) is nonempty).

Moreover, \(qv\) is nonempty (since \(v\) is nonempty). Hence, \(qv\) is a nonempty proper suffix of \(w\). Since \(v\) is the smallest such suffix, this entails that \(v \leq qv\). Proposition 6.2(b) (applied to \(a = p, c = v\) and \(d = qv\)) thus yields \(pv \leq pv\). Hence, \(pv \leq qv = w\), which contradicts \(pv > w\). This contradiction shows that our assumption (that \(p \leq u\)) was false. We thus have \(p > u\).

We now have shown that \(p > u\) whenever \(p\) is a nonempty proper suffix of \(u\). Combined with the fact that \(u\) is nonempty, this shows that \(u\) is a Lyndon word. This completes the proof of Theorem 6.30(a). \(\square\)

---

\(^{209}\)In particular, \(\frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}\) is an integer.

\(^{210}\)e.g., allowing to solve Exercise 6.24 in a simpler way
Another approach to the standard factorization is given in the following exercise:

**Exercise 6.31.** Let \( w \) be a Lyndon word of length \( >1 \). Let \( v \) be the longest proper suffix of \( w \) such that \( v \) is Lyndon.\(^{211}\) Since \( v \) is a proper suffix of \( w \), there exists a nonempty \( u \in \mathfrak{A}^* \) such that \( w = uv \). Consider this \( u \). Prove that:

(a) The words \( u \) and \( v \) are Lyndon.
(b) We have \( u < w < v \).
(c) The words \( u \) and \( v \) are precisely the words \( u \) and \( v \) constructed in Theorem 6.30.

Notice that a well-known recursive characterization of Lyndon words [31, \( \mathfrak{A}' = \mathfrak{A}'' \)] can be easily derived from Theorem 6.30 and Proposition 6.16(a). We will not dwell on it.

The following exercise surveys some variations on the characterizations of Lyndon words\(^{212}\):

**Exercise 6.32.** Let \( w \) be a nonempty word. Consider the following nine assertions:

- **Assertion** \( A' \): The word \( w \) is a power of a Lyndon word.
- **Assertion** \( B' \): If \( u \) and \( v \) are nonempty words satisfying \( w = uv \), then either we have \( v \geq w \) or the word \( v \) is a prefix of \( w \).
- **Assertion** \( C' \): If \( u \) and \( v \) are nonempty words satisfying \( w = uv \), then either we have \( v \geq u \) or the word \( v \) is a prefix of \( u \).
- **Assertion** \( D' \): If \( u \) and \( v \) are nonempty words satisfying \( w = uv \), then we have \( vu \geq uv \).
- **Assertion** \( E' \): If \( u \) and \( v \) are nonempty words satisfying \( w = uv \), then either we have \( v \geq u \) or the word \( v \) is a prefix of \( w \).
- **Assertion** \( F' \): The word \( w \) is a prefix of a Lyndon word in \( \mathfrak{A}^* \).
- **Assertion** \( F'' \): Let \( m \) be an object not in the alphabet \( \mathfrak{A} \). Let us equip the set \( \mathfrak{A} \cup \{m\} \) with a total order which extends the total order on the alphabet \( \mathfrak{A} \) and which satisfies \( a < m \) for every \( a \in \mathfrak{A} \). Then, the word \( wm \in (\mathfrak{A} \cup \{m\})^* \) (the concatenation of the word \( w \) with the one-letter word \( m \)) is a Lyndon word.
- **Assertion** \( G' \): There exists a Lyndon word \( t \in \mathfrak{A}^* \), a positive integer \( \ell \) and a prefix \( p \) of \( t \) (possibly empty) such that \( w = t^\ell p \).
- **Assertion** \( H' \): There exists a Lyndon word \( t \in \mathfrak{A}^* \), a nonnegative integer \( \ell \) and a prefix \( p \) of \( t \) (possibly empty) such that \( w = t^\ell p \).

(a) Prove the equivalence \( A' \iff D' \).
(b) Prove the equivalence \( B' \iff C' \iff E' \iff F'' \iff G' \iff H' \).
(c) Prove the implication \( F' \Rightarrow B' \).
(d) Prove the implication \( D' \Rightarrow B' \). (The implication \( B' \Rightarrow D' \) is false, as witnessed by the word \( 11211 \).)
(e) Prove that if there exists a letter \( \mu \in \mathfrak{A} \) such that \( (\mu > a \text{ for every letter } a \text{ of } w) \), then the equivalence \( F' \iff F'' \) holds.
(f) Prove that if there exists a letter \( \mu \in \mathfrak{A} \) such that \( (\mu > a \text{ for some letter } a \text{ of } w) \), then the equivalence \( F' \iff F'' \) holds.

The next exercise (based on work of Hazewinkel [68]) extends some of the above properties of Lyndon words (and words in general) to a more general setting, in which the alphabet \( \mathfrak{A} \) is no longer required to be totally ordered, but only needs to be a poset:

**Exercise 6.33.** In this exercise, we shall loosen the requirement that the alphabet \( \mathfrak{A} \) be a totally ordered set; Instead, we will only require \( \mathfrak{A} \) to be a poset. The resulting more general setting will be called the partial-order setting, to distinguish it from the total-order setting in which \( \mathfrak{A} \) is required to be a totally ordered set. All results in Chapter 6 so far address the total-order setting. In this exercise, we will generalize some of them to the partial-order setting.

All notions that we have defined in the total-order setting (the notion of a word, the relation \( \leq \), the notion of a Lyndon word, etc.) are defined in precisely the same way in the partial-order setting. However, the poset \( \mathfrak{A}^* \) is no longer totally ordered in the partial-order setting.

---

\(^{211}\)This is well-defined, because there exists at least one proper suffix \( v \) of \( w \) such that \( v \) is Lyndon. (Indeed, the last letter of \( w \) forms such a suffix, because it is a proper suffix of \( w \) (since \( w \) has length \( >1 \)) and is Lyndon (since it is a one-letter word, and since every one-letter word is Lyndon).)

\(^{212}\)Compare this with [88, §7.2.11, Theorem Q].
(a) Prove that Proposition 6.2 holds in the partial-order setting, as long as one replaces “a total order” by “a partial order” in part (a) of this Proposition.

(b) Prove (in the partial-order setting) that if \( a, b, c, d \in \mathfrak{A}^+ \) are four words such that the words \( ab \) and \( cd \) are comparable (with respect to the partial order \( \leq \)), then the words \( a \) and \( c \) are comparable.


(d) Find a counterexample to Exercise 6.22 in the partial-order setting.

(e) Salvage Exercise 6.22 in the partial-order setting (i.e., find a statement which is easily equivalent to this exercise in the total-order setting, yet true in the partial-order setting).

(f) In the partial-order setting, a Hazewinkel-CFL factorization of a word \( w \) will mean a tuple \( (a_1, a_2, \ldots, a_k) \) of Lyndon words such that \( w = a_1a_2\cdots a_k \) and such that no \( i \in \{1, 2, \ldots, k - 1\} \) satisfies \( a_i < a_{i+1} \).

   Prove that every word \( w \) has a unique Hazewinkel-CFL factorization (in the partial-order setting).

(g) Prove that Exercise 6.32 still holds in the partial-order setting.

The reader is invited to try extending other results to the partial-order setting (it seems that no research has been done on this except for Hazewinkel’s [68]). We shall now, however, return to the total-order setting (which has the most known applications).

Lyndon words are related to various other objects in mathematics, such as free Lie algebras (Subsection 6.1.1 below), shuffles and shuffle algebras (Sections 6.2 and 6.3 below), QSym (Sections 6.4 and 6.5), Markov chains on combinatorial Hopf algebras ([41]), de Bruijn sequences ([53], [121], [88], §7.2.11, Algorithm F), symmetric functions (specifically, the transition matrices between the bases \( (h_\lambda)_{\lambda \in \text{Par}} \), \( (e_\lambda)_{\lambda \in \text{Par}} \) and \( (m_\lambda)_{\lambda \in \text{Par}} \) see [84] for this), and the Burrows-Wheeler algorithm for data compression ([83]). They are also connected to necklaces (in the combinatorial sense) – a combinatorial object that also happens to be related to a lot of algebra ([140, Chapter 5], [39]). Let us survey the basics of this latter classical connection in an exercise:

**Exercise 6.34.** Let \( \mathfrak{A} \) be any set (not necessarily totally ordered). Let \( C \) denote the infinite cyclic group, written multiplicatively. Fix a generator \( c \) of \( C \). \(^{214}\) Fix a positive integer \( n \). The group \( C \) acts on \( \mathfrak{A}^n \) from the left according to the rule

\[
c \cdot (a_1, a_2, \ldots, a_n) = (a_2, a_3, \ldots, a_n, a_1)
\]

for all \( (a_1, a_2, \ldots, a_n) \in \mathfrak{A}^n \).

The \( n \)-necklace containing a given \( n \)-tuple \( w \in \mathfrak{A}^n \) will be denoted by \( [w] \).

(a) Prove that every \( n \)-necklace \( N \) is a finite nonempty set and satisfies \(|N| = n \). (Recall that \( N \) is an orbit, thus a set; as usual, \(|N| \) denotes the cardinality of this set.)

The period of an \( n \)-necklace \( N \) is defined as the positive integer \(|N| \). (This \(|N| \) is indeed a positive integer, since \( N \) is a finite nonempty set.)

An \( n \)-necklace is said to be aperiodic if its period is \( n \).

(b) Given any \( n \)-tuple \( w = (w_1, w_2, \ldots, w_n) \in \mathfrak{A}^n \), prove that the \( n \)-necklace \([w]\) is aperiodic if and only if every \( k \in \{1, 2, \ldots, n - 1\} \) satisfies \( (w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w \).

From now on, we assume that the set \( \mathfrak{A} \) is totally ordered. We use \( \mathfrak{A} \) as our alphabet to define the notions of words, the lexicographic order, and Lyndon words. All notations that we introduced for words will thus be used for elements of \( \mathfrak{A}^n \).

(c) Prove that every aperiodic \( n \)-necklace contains exactly one Lyndon word.

---

\( ^{213} \)This result, as well as the validity of Proposition 6.16 in the partial-order setting, are due to Hazewinkel [68].

\( ^{214} \)So \( C \) is a group isomorphic to \((\mathbb{Z}, +)\), and the isomorphism \((\mathbb{Z}, +) \to C, n \mapsto c^n \). (Recall that we write the binary operation of \( C \) as \( \cdot \) instead of +.)

\( ^{215} \)In other words, \( c \) rotates any \( n \)-tuple of elements of \( \mathfrak{A} \) cyclically to the left. Thus, \( c^n \in C \) acts trivially on \( \mathfrak{A}^n \), and so this action of \( C \) on \( \mathfrak{A}^n \) factors through \( C/\langle c^n \rangle \) (a cyclic group of order \( n \)).

\( ^{216} \)Classically, one visualizes them as necklaces of \( n \) beads of \( |\mathfrak{A}| \) colors. (The colors are the elements of \( \mathfrak{A} \).) The intuition behind this is that a necklace is an object that doesn’t really change when we rotate it in its plane. However, to make this intuition match the definition, we need to think of a necklace as being stuck in its (fixed) plane, so that we cannot lift it up and turn it around, dropping it back to its plane in a reflected state.
(d) If $N$ is an $n$-necklace which is not aperiodic, then prove that $N$ contains no Lyndon word.

(e) Show that the aperiodic $n$-necklaces are in bijection with Lyndon words of length $n$.

From now on, we assume that the set $\mathfrak{A}$ is finite. Define the number-theoretic M"obius function $\mu$ and the Euler totient function $\phi$ as in Exercise 2.82.

(f) Prove that the number of all aperiodic $n$-necklaces is
\[ \frac{1}{n} \sum_{d|n} \mu(d) |\mathfrak{A}|^{n/d}. \]

(g) Prove that the number of all $n$-necklaces is
\[ \frac{1}{n} \sum_{d|n} \phi(d) |\mathfrak{A}|^{n/d}. \]

(h) Solve Exercise 6.29 again.

(i) Forget that we fixed $\mathfrak{A}$. Show that every $q \in \mathbb{Z}$ satisfies $n | \sum_{d|n} \mu(d) q^{n/d}$ and $n | \sum_{d|n} \phi(d) q^{n/d}$. [Hint: For (c), use Theorem 6.20. For (i), either use parts (f) and (g) and a trick to extend to $q$ negative; or recall Exercise 2.84.]

6.1.1. Free Lie algebras. In this brief subsection, we shall review the connection between Lyndon words and free Lie algebras (following [93, Kap. 4], but avoiding the generality of Hall sets in favor of just using Lyndon words). None of this material shall be used in the rest of these notes. We will only prove some basic results; for more thorough and comprehensive treatments of free Lie algebras, see [137], [22, Chapter 2] and [93, Kap. 4].

We begin with some properties of Lyndon words.

**Exercise 6.35.** Let $w \in \mathfrak{A}^*$ be a nonempty word. Let $v$ be the longest Lyndon suffix of $w$. \(^{217}\) Let $t$ be a Lyndon word. Then, $t$ is the longest Lyndon suffix of $wt$ if and only if we do not have $v < t$.

(We have written “we do not have $v < t^r$” instead of “$v \geq t^r$” in Exercise 6.35 for reasons of generalizability: This way, Exercise 6.35 generalizes to the partial-order setting introduced in Exercise 6.33, whereas the version with “$v \geq t^r$” does not.)

**Exercise 6.36.** Let $w \in \mathfrak{A}^*$ be a word of length $> 1$. Let $v$ be the longest Lyndon proper suffix of $w$. \(^{218}\) Let $t$ be a Lyndon word. Then, $t$ is the longest Lyndon proper suffix of $wt$ if and only if we do not have $v < t$.

(Exercise 6.36, while being a trivial consequence of Exercise 6.35, is rather useful in the study of free Lie algebras. It generalizes both [31, Lemma (1.6)] (which is obtained by taking $w = c, v = b$ and $t = d$) and [106, Proposition 5.1.4] (which is obtained by taking $v = m$ and $t = n$).)

**Definition 6.37.** For the rest of Subsection 6.1.1, we let $\mathfrak{L}$ be the set of all Lyndon words (over the alphabet $\mathfrak{A}$).

**Definition 6.38.** Let $w$ be a Lyndon word of length $> 1$. Let $v$ be the longest proper suffix of $w$ such that $v$ is Lyndon. (This is well-defined, as we know from Exercise 6.31.) Since $v$ is a proper suffix of $w$, there exists a nonempty $u \in \mathfrak{A}^*$ such that $w = uv$. Consider this $u$. (Clearly, this $u$ is unique.) Theorem 6.30(a) shows that the words $u$ and $v$ are Lyndon. In other words, $u \in \mathfrak{L}$ and $v \in \mathfrak{L}$. Hence, $(u, v) \in \mathfrak{L} \times \mathfrak{L}$. The pair $(u, v) \in \mathfrak{L} \times \mathfrak{L}$ is called the standard factorization of $w$, and is denoted by $stf w$.

For the sake of easier reference, we gather a few basic properties of the standard factorization:

**Exercise 6.39.** Let $w$ be a Lyndon word of length $> 1$. Let $(g, h) = stf w$. Prove the following:

(a) The word $h$ is the longest Lyndon proper suffix of $w$.

(b) We have $w = gh$.

(c) We have $g < gh < h$.

(d) The word $g$ is Lyndon.

---

217 Of course, a Lyndon suffix of $w$ just means a suffix $p$ of $w$ such that $p$ is Lyndon.

218 Of course, a Lyndon proper suffix of $w$ just means a proper suffix $p$ of $w$ such that $p$ is Lyndon.
(e) We have \( g \in \mathcal{L}, h \in \mathcal{L}, \ell (g) < \ell (w) \) and \( \ell (h) < \ell (w) \).

(f) Let \( t \) be a Lyndon word. Then, \( t \) is the longest Lyndon proper suffix of \( wt \) if and only if we do not have \( h < t \).

**Exercise 6.40.** Let \( \mathfrak{g} \) be a Lie algebra. For every Lyndon word \( w \), let \( b_w \) be an element of \( \mathfrak{g} \). Assume that for every Lyndon word \( w \) of length \( > 1 \), we have

\[
(6.2) \quad b_w = [b_u, b_v], \quad \text{where } (u, v) = \text{stf } w.
\]

Let \( B \) be the \( k \)-submodule of \( \mathfrak{g} \) spanned by the family \((b_w)_{w \in \mathcal{L}}\).

(a) Prove that \( B \) is a Lie subalgebra of \( \mathfrak{g} \).

(b) Let \( \mathfrak{h} \) be a \( k \)-Lie algebra. Let \( f : B \rightarrow \mathfrak{h} \) be a \( k \)-module homomorphism. Assume that whenever \( w \) is a Lyndon word of length \( > 1 \), we have

\[
(6.3) \quad f ([b_u, b_v]) = [f (b_u), f (b_v)], \quad \text{where } (u, v) = \text{stf } w.
\]

Prove that \( f \) is a Lie algebra homomorphism.

**[Hint:]** Given two words \( w \) and \( w' \), write \( w \sim w' \) if and only if \( w' \) is a permutation of \( w \). Part (a) follows from the fact that for any \((p, q) \in \mathcal{L} \times \mathcal{L} \) satisfying \( p < q \), we have \([b_p, b_q] \in B_{pq,q} \), where \( B_{h,s} \) denotes the \( k \)-linear span of \( \{b_w \mid w \in \mathcal{L}, w \sim h \text{ and } w < s \} \) for any two words \( h \) and \( s \). Prove this fact by a double induction, first inducting over \( \ell (pq) \) and then (for fixed \( \ell (pq) \)) inducting over the rank of \( q \) in lexicographic order (i.e., assume that the fact is already proven for every \( q' < q \) instead of \( q \)). In the induction step, assume that \((p, q) \neq \text{stf } (pq) \) (since otherwise the claim is rather obvious) and conclude that \( p \) has length \( > 1 \); thus, set \((u, v) = \text{stf } p \), so that

\[
\left[ \begin{array}{c}
\frac{b_p}{\text{stf } p}, \frac{b_q}{\text{stf } q} \\
= [b_u, b_v], [b_u, b_q], [b_v, b_q], [b_v, b_u], \end{array} \right]
\]

and use Exercise 6.36 to obtain \( v < q \).

The proof of (b) proceeds by a similar induction, piggybacking on the \([b_p, b_q] \in B_{pq,q} \) claim.

**Exercise 6.41.** Let \( V \) be the free \( k \)-module with basis \((x_a)_{a \in \mathfrak{A}} \). For every word \( w \in \mathfrak{A}^* \), let \( x_w \) be the tensor \( x_{w_1} \otimes x_{w_2} \otimes \cdots \otimes x_{w_{|w|}} \). As we know from Example 1.2, the tensor algebra \( T (V) \) is a free \( k \)-module with basis \((x_w)_{w \in \mathcal{L}} \). We regard \( V \) as a \( k \)-submodule of \( T (V) \).

The tensor algebra \( T (V) \) becomes a Lie algebra via the commutator (i.e., its Lie bracket is defined by \([\alpha, \beta] = \alpha \beta - \beta \alpha \) for all \( \alpha, \beta \in T (V) \) and \( \beta \in T (V) \)).

We define a sequence \((\mathfrak{g}_1, \mathfrak{g}_2, \mathfrak{g}_3, \ldots) \) of \( k \)-submodules of \( T (V) \) as follows: Recursively, we set \( \mathfrak{g}_1 = V \), and for every \( i \in \{2, 3, 4, \ldots\} \), we set \( \mathfrak{g}_i = [V, \mathfrak{g}_{i-1}] \). Let \( \mathfrak{g} \) be the \( k \)-submodule \( \mathfrak{g}_1 + \mathfrak{g}_2 + \mathfrak{g}_3 + \cdots \) of \( T (V) \).

Prove the following:

(a) The \( k \)-submodule \( \mathfrak{g} \) is a Lie subalgebra of \( T (V) \).

(b) If \( \mathfrak{k} \) is any Lie subalgebra of \( T (V) \) satisfying \( V \subset \mathfrak{k} \), then \( \mathfrak{g} \subset \mathfrak{k} \).

Now, for every \( w \in \mathcal{L} \), we define an element \( b_w \) of \( T (V) \) as follows: We define \( b_w \) by recursion on the length of \( w \). If the length of \( w \) is 1 \( \text{[219]} \), then we have \( w = (a) \) for some letter \( a \in \mathfrak{A} \), and we set \( b_w = x_a \) for this letter \( a \). If the length of \( w \) is \( > 1 \), then we set \( b_w = [b_u, b_v] \), where \((u, v) = \text{stf } w \) \( \text{[220]} \).

Prove the following:

(c) For every \( w \in \mathcal{L} \), we have

\[
b_w \in x_w + \sum_{v \in \mathfrak{A}^{|w|}; \ell (v) > \ell (w)} kx_v.
\]

(d) The family \((b_w)_{w \in \mathcal{L}} \) is a basis of the \( k \)-module \( \mathfrak{g} \).

(e) Let \( \mathfrak{h} \) be any \( k \)-Lie algebra. Let \( \xi : \mathfrak{A} \rightarrow \mathfrak{h} \) be any map. Then, there exists a unique Lie algebra homomorphism \( \Xi : \mathfrak{g} \rightarrow \mathfrak{h} \) such that every Lie algebra \( \Xi (x_a) = \xi (a) \).
Remar 6.42. Let $V$ and $g$ be as in Exercise 6.41. In the language of universal algebra, the statement of Exercise 6.41(e) says that $g$ (or, to be more precise, the pair $(g, f)$, where $f : A \to g$ is the map sending each $a \in A$ to $x_a \in g$) satisfies the universal property of the free Lie algebra on the set $A$. Thus, this exercise allows us to call $g$ the free Lie algebra on $A$. Most authors define the free Lie algebra differently, but all reasonable definitions of a free Lie algebra lead to isomorphic Lie algebras (because the universal property determines the free Lie algebra uniquely up to canonical isomorphism).

Notice that the Lie algebra $g$ does not depend on the total order on the alphabet $A$, but the basis $(b_w)_{w \in \mathbb{C}}$ constructed in Exercise 6.41(d) does. There is no known basis of $g$ defined without ordering $A$.

It is worth noticing that our construction of $g$ proves not only that the free Lie algebra on $A$ exists, but also that this free Lie algebra can be realized as a Lie subalgebra of the (associative) algebra $T(V)$. Therefore, if we want to prove that a certain identity holds in every Lie algebra, we only need to check that this identity holds in every associative algebra (if all Lie brackets are replaced by commutators); the universal property of the free Lie algebra (i.e., Exercise 6.41(e)) will then ensure that this identity also holds in every Lie algebra $g$.

There is much more to say about free Lie algebras than what we have said here; in particular, there are connections to symmetric functions, necklaces, representations of symmetric groups and NSym. See [106, §5.3], [137], [22, Chapter 2], [93, §4] and [20] for further developments.

6.2. Shuffles and Lyndon words. We will now connect the theory of Lyndon words with the notion of shuffle products. We have already introduced the latter notion in Definition 1.68, but we will now study it more closely and introduce some more convenient notations (e.g., we will need a notation for single shuffles, not just the whole multiset).

Definition 6.43. (a) Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Then, $Sh_{n,m}$ denotes the subset

$$\{ \sigma \in S_{n+m} : \sigma^{-1}(1) < \sigma^{-1}(2) < \cdots < \sigma^{-1}(n) ; \sigma^{-1}(n+1) < \sigma^{-1}(n+2) < \cdots < \sigma^{-1}(n+m) \}$$

of the symmetric group $S_{n+m}$.

(b) Let $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_m)$ be two words. If $\sigma \in Sh_{n,m}$, then $u \shuffle \sigma v$ will denote the word $(w_\sigma(1), w_\sigma(2), \ldots, w_\sigma(n+m))$, where $(w_1, w_2, \ldots, w_{n+m})$ is the concatenation $u \cdot v = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)$. We notice that the multiset of all letters of $u \shuffle \sigma v$ is the disjoint union of the multiset of all letters of $u$ with the multiset of all letters of $v$. As a consequence, $\ell(u \shuffle \sigma v) = \ell(u) + \ell(v)$.

(c) Let $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_m)$ be two words. The multiset of shuffles of $u$ and $v$ is defined as the multiset $\{ (w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}) : \sigma \in Sh_{n,m} \}_{\text{multiset}}$, where $(w_1, w_2, \ldots, w_{n+m})$ is the concatenation $u \cdot v = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)$. In other words, the multiset of shuffles of $u$ and $v$ is the multiset

$$\left\{ u \shuffle_{\sigma} v : \sigma \in Sh_{n,m} \right\}_{\text{multiset}}.$$

It is denoted by $u \shuffle v$.

The next fact provides the main connection between Lyndon words and shuffles:

Theorem 6.44. Let $u$ and $v$ be two words.

Let $(a_1, a_2, \ldots, a_p)$ be the CFL factorization of $u$. Let $(b_1, b_2, \ldots, b_q)$ be the CFL factorization of $v$.

(a) Let $(c_1, c_2, \ldots, c_{p+q})$ be the result of sorting the list $(a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)$ in decreasing order. Then, the lexicographically highest element of the multiset $u \shuffle v$ is $c_1c_2\cdots c_{p+q}$ (and $(c_1, c_2, \ldots, c_{p+q})$ is the CFL factorization of this element).

---

221Here, we call a definition “reasonable” if the “free Lie algebra” it defines satisfies the universal property.

222The claim made in [20, page 2] that “$\{x_1, \ldots, x_n\}$ generates freely a Lie subalgebra of $A_R^*$” is essentially our Exercise 6.41(e).

223Parts (a) and (c) of the below Definition 6.43 define notions which have already been introduced in Definition 1.68. Of course, the definitions of these notions are equivalent; however, the variables are differently labelled in the two definitions (for example, the variables $u$, $v$, $w$ and $\sigma$ of Definition 6.43(c) correspond to the variables $a$, $b$, $c$ and $w$ of Definition 1.68). The labels in Definition 6.43 have been chosen to match with the rest of Section 6.2.

224With respect to the total order on $\mathbb{A}^*$ whose greater-or-equal relation is $\geq$. 
(b) Let $\mathfrak{L}$ denote the set of all Lyndon words. If $w$ is a Lyndon word and $z$ is any word, let $\text{mult}_w z$ denote the number of terms in the CFL factorization of $z$ which are equal to $w$. The multiplicity with which the lexicographically highest element of the multiset $u \cup v$ appears in the multiset $u \cup v$ is $\prod_{w \in \mathfrak{L}} \left( \text{mult}_w u + \text{mult}_w v \right)$. (This product is well-defined because almost all of its factors are 1.)

(c) If $a_i \geq b_j$ for every $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$, then the lexicographically highest element of the multiset $u \cup v$ is $w$.

(d) If $a_i > b_j$ for every $i \in \{1, 2, \ldots, p\}$ and $j \in \{1, 2, \ldots, q\}$, then the multiplicity with which the word $w$ appears in the multiset $u \cup v$ is 1.

(e) Assume that $u$ is a Lyndon word. Also, assume that $u \geq b_j$ for every $j \in \{1, 2, \ldots, q\}$. Then, the lexicographically highest element of the multiset $u \cup v$ is $w$, and the multiplicity with which this word $w$ appears in the multiset $u \cup v$ is $\text{mult}_u w + 1$.

**Example 6.45.** For this example, let $u$ and $v$ be the words $u = 23232$ and $v = 323221$ over the alphabet $\mathfrak{A} = \{1, 2, 3, \ldots\}$ with total order given by $1 < 2 < 3 < \cdots$. The CFL factorizations of $u$ and $v$ are $(23, 23, 2)$ and $(3, 23, 2, 2, 1)$, respectively. Thus, using the notations of Theorem 6.44, we have $p = 3$, $(a_1, a_2, \ldots, a_p) = (23, 23, 2)$, $q = 5$ and $(b_1, b_2, \ldots, b_q) = (3, 23, 2, 2, 1)$. Thus, Theorem 6.44(a) predicts that the lexicographically highest element of the multiset $u \cup v$ is $c_3 c_2 c_3 c_2 c_1 c_6 c_5 c_6 c_8$, where $c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8$ are the words $23, 23, 2, 3, 23, 2, 2, 1$ listed in decreasing order (in other words, $(c_1, c_2, c_3, c_4, c_5, c_6, c_7, c_8) = (23, 23, 2, 3, 23, 2, 2, 1)$). In other words, Theorem 6.44(a) predicts that the lexicographically highest element of the multiset $u \cup v$ is $323232221$. We could verify this by brute force, but this would be laborious since the multiset $u \cup v$ has $\binom{5 + 6}5 = 462$ elements (with multiplicities). Theorem 6.44(b) predicts that this lexicographically highest element $323232221$ appears in the multiset $u \cup v$ with a multiplicity of $\prod_{w \in \mathfrak{L}} \left( \text{mult}_w u + \text{mult}_w v \right)$. This product $\prod_{w \in \mathfrak{L}} \left( \text{mult}_w u + \text{mult}_w v \right)$ is infinite, but all but finitely many of its factors are 1 and therefore can be omitted; the only factors which are not 1 are those corresponding to Lyndon words $w$ which appear both in the CFL factorization of $u$ and in the CFL factorization of $v$ (since for any other factor, at least one of the numbers mult$_w u$ or mult$_w v$ equals 0, and therefore the binomial coefficient $\binom{\text{mult}_w u + \text{mult}_w v}{\text{mult}_w u}$ equals 1). Thus, in order to compute the product $\prod_{w \in \mathfrak{L}} \left( \text{mult}_w u + \text{mult}_w v \right)$, we only need to multiply these factors. In our example, these are the factors for $w = 23$ and for $w = 2$ (these are the only Lyndon words which appear both in the CFL factorization $(23, 23, 2)$ of $u$ and in the CFL factorization $(3, 23, 2, 2, 1)$ of $v$). So we have

$$\prod_{w \in \mathfrak{L}} \left( \text{mult}_w u + \text{mult}_w v \right) = \left( \text{mult}_{23} u + \text{mult}_{23} v \right) \left( \text{mult}_{2} u + \text{mult}_{2} v \right) = 3 \cdot 3 = 9.$$ 

The word $323232221$ must thus appear in the multiset $u \cup v$ with a multiplicity of 9. This, too, could be checked by brute force.

Theorem 6.44 (and Theorem 6.64 further below, which describes more precisely how the lexicographically highest element of $u \cup v$ emerges by shuffling $u$ and $v$) is fairly close to [134, Theorem 2.2.2] (and will be used for the same purposes), the main difference being that we are talking about the shuffle product of two (not necessarily Lyndon) words, while Radford (and most other authors) study the shuffle product of many Lyndon words.

In order to prove Theorem 6.44, we will need to make some stronger statements, for which we first have to introduce some more notation:

**Definition 6.46.**

(a) If $p$ and $q$ are two integers, then $\lfloor p : q \rfloor^+$ denotes the interval $\{p + 1, p + 2, \ldots, q\}$ of $\mathbb{Z}$. Note that $\lfloor p : q \rfloor^+ = q - p$ if $q \geq p$.

(b) If $I$ and $J$ are two nonempty intervals of $\mathbb{Z}$, then we say that $I < J$ if and only if every $i \in I$ and $j \in J$ satisfy $i < j$. This defines a partial order on the set of nonempty intervals of $\mathbb{Z}$. (Roughly speaking, $I < J$ if the interval $I$ ends before $J$ begins.)
(c) If $w$ is a word with $n$ letters (for some $n \in \mathbb{N}$), and $I$ is an interval of $\mathbb{Z}$ such that $I \subset [0 : n]^+$, then $w[I]$ will denote the word $(w_{p+1}, w_{p+2}, \ldots, w_q)$, where $I$ is written in the form $I = [p : q]^+$ with $q \geq p$. Obviously, \( \ell(w[I]) = |I| = q - p \). A word of the form $w[I]$ for an interval $I \subset [0 : n]^+$ (equivalently, a word which is a prefix of a suffix of $w$) is called a factor of $w$.

(d) Let $\alpha$ be a composition. Then, we define a tuple $\text{intsy} \alpha$ of intervals of $\mathbb{Z}$ as follows: Write $\alpha$ in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ (so that $\ell = \ell(\alpha)$). Then, set $\text{intsy} \alpha = (I_1, I_2, \ldots, I_\ell)$, where

$$I_i = \left[ \sum_{k=1}^{i-1} \alpha_k : \sum_{k=1}^{i} \alpha_k \right]^+ \quad \text{ for every } i \in \{1, 2, \ldots, \ell\}.$$ 

This $\ell$-tuple $\text{intsy} \alpha$ is a tuple of nonempty intervals of $\mathbb{Z}$. This tuple $\text{intsy} \alpha$ is called the interval system corresponding to $\alpha$. (This is precisely the $\ell$-tuple $(I_1, I_2, \ldots, I_\ell)$ constructed in Definition 4.24.) The length of the tuple $\text{intsy} \alpha$ is $\ell(\alpha)$.

**Example 6.47.**  
(a) We have $[2 : 4]^+ = \{3, 4\}$ and $[3 : 3]^+ = \emptyset$.  
(c) If $w$ is the word 915325, then $w[0 : 3]^+ = (w_1, w_2, w_3) = 915$ and $w[2 : 4]^+ = (w_3, w_4) = 53$.  
(d) If $\alpha$ is the composition $(4, 1, 4, 2, 3)$, then the interval system corresponding to $\alpha$ is

$$\text{intsy} \alpha = \left( \{0 : 4\}^+, \{4 : 5\}^+, \{5 : 9\}^+, \{9 : 11\}^+, \{11 : 14\}^+ \right) = \{\{1, 2, 3, 4\}, \{5\}, \{6, 7, 8, 9\}, \{10, 11\}, \{12, 13, 14\}\}.$$ 

The following properties of the notions introduced in the preceding definition are easy to check:

**Remark 6.48.**  
(a) If $I$ and $J$ are two nonempty intervals of $\mathbb{Z}$ satisfying $I < J$, then $I$ and $J$ are disjoint.  
(b) If $I$ and $J$ are two disjoint nonempty intervals of $\mathbb{Z}$, then either $I < J$ or $J < I$.  
(c) Let $\alpha$ be a composition. Write $\alpha$ in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ (so that $\ell = \ell(\alpha)$). The interval system $\text{intsy} \alpha$ can be described as the unique $\ell$-tuple $(I_1, I_2, \ldots, I_\ell)$ of nonempty intervals of $\mathbb{Z}$ satisfying the following three properties:  
  - The intervals $I_1, I_2, \ldots, I_\ell$ form a set partition of the set $[0 : n]^+$, where $n = |\alpha|$.  
  - We have $I_1 < I_2 < \cdots < I_\ell$.  
  - We have $|I_i| = \alpha_i$ for every $i \in \{1, 2, \ldots, \ell\}$.

**Exercise 6.49.** Prove Remark 6.48.

The following two lemmas are collections of more or less trivial consequences of what it means to be an element of $\text{Sh}_{n,m}$, and what it means to be a shuffle:

**Lemma 6.50.** Let $n \in \mathbb{N}$ and $m \in \mathbb{N}$. Let $\sigma \in \text{Sh}_{n,m}$.

(a) If $I$ is an interval of $\mathbb{Z}$ such that $I \subset [0 : n + m]^+$, then $\sigma(I) \cap [0 : n]^+ \text{ and } \sigma(I) \cap [n : n + m]^+$ are intervals.  
(b) Let $K$ and $L$ be nonempty intervals of $\mathbb{Z}$ such that $K \subset [0 : n]^+$ and $L \subset [0 : n]^+$ and $K < L$ and such that $K \cup L$ is an interval. Assume that $\sigma^{-1}(K)$ and $\sigma^{-1}(L)$ are intervals, but $\sigma^{-1}(K) \cup \sigma^{-1}(L)$ is not an interval. Then, there exists a nonempty interval $P \subset [n : n + m]^+$ such that $\sigma^{-1}(P)$, $\sigma^{-1}(K) \cup \sigma^{-1}(P)$ and $\sigma^{-1}(P) \cup \sigma^{-1}(L)$ are intervals and such that $\sigma^{-1}(K) < \sigma^{-1}(P) < \sigma^{-1}(L)$.  
(c) Lemma 6.50(b) remains valid if “$K \subset [0 : n]^+$ and $L \subset [0 : n]^+$” and “$P \subset [n : n + m]^+$” are replaced by “$K \subset [n : n + m]^+$ and $L \subset [n : n + m]^+$” and “$P \subset [0 : n]^+$”, respectively.

**Exercise 6.51.** Prove Lemma 6.50.

**Lemma 6.52.** Let $u$ and $v$ be two words. Let $n = \ell(u)$ and $m = \ell(v)$. Let $\sigma \in \text{Sh}_{n,m}$.

(a) If $I$ is an interval of $\mathbb{Z}$ satisfying either $I \subset [0 : n]^+$ or $I \subset [n : n + m]^+$, and if $\sigma^{-1}(I)$ is an interval, then

$$\left( u \cup v \right)_{\sigma} \left[ \sigma^{-1}(I) \right] = (uv)[I].$$
(b) Assume that $u \cup v$ is the lexicographically highest element of the multiset $u \cup v$. Let $I \subset [0: n]^+$ and $J \subset [n: n + m]^+$ be two nonempty intervals. Assume that $\sigma^{-1}(I)$ and $\sigma^{-1}(J)$ are also intervals, that $\sigma^{-1}(I) < \sigma^{-1}(J)$, and that $\sigma^{-1}(I) \cup \sigma^{-1}(J)$ is an interval as well. Then, $(uv)[I] \cdot (uv)[J] \geq (uv)[J] \cdot (uv)[I]$.

(c) Lemma 6.52(b) remains valid if “$I \subset [0: n]^+$ and $J \subset [n: n + m]^+$” is replaced by “$I \subset [n: n + m]^+$ and $J \subset [0: n]^+$”.


[Hint: For (b), show that there exists a $\tau \in Sh_{n,m}$ such that $u \cup v$ differs from $u \cup v$ only in the order of the subwords $(uv)[I]$ and $(uv)[J]$.]

We are still a few steps away from stating our results in a way that allows comfortably proving Theorem 6.44. For the latter aim, we introduce the notion of $\alpha$-clumping permutations, and characterize them in two ways:

Definition 6.54. Let $n \in \mathbb{N}$. Let $\alpha$ be a composition of $n$. Let $\ell = \ell(\alpha)$.

(a) For every set $S$ of positive integers, let $S$ denote the list of all elements of $S$ in increasing order (with each element appearing exactly once). Notice that this list $S$ is a word over the set of positive integers.

(b) For every $\tau \in \mathcal{S}_\ell$, we define a permutation $iper(\alpha, \tau) \in \mathcal{S}_n$ as follows:

The interval system corresponding to $\alpha$ is an $\ell$-tuple of intervals (since $\ell(\alpha) = \ell$); denote this $\ell$-tuple by $(I_1, I_2, \ldots, I_\ell)$. Now, define $iper(\alpha, \tau)$ to be the permutation in $\mathcal{S}_n$ which (in one-line notation) is the word $I_{\tau(1)} I_{\tau(2)} \cdots I_{\tau(\ell)}$ (a concatenation of $\ell$ words). This is well-defined\footnote{In fact, from the properties of interval systems, we know that the intervals $I_1, I_2, \ldots, I_\ell$ form a set partition of the set $[0: n]^+$. Hence, the intervals $I_{\tau(1)}, I_{\tau(2)}, \ldots, I_{\tau(\ell)}$ form a set partition of the set $[0: n]^+$. As a consequence, the word $I_{\tau(1)} I_{\tau(2)} \cdots I_{\tau(\ell)}$ is a permutation of the word $12 \ldots n$, and so there exists a permutation in $\mathcal{S}_n$ which (in one-line notation) is this word, qed.}, hence, $iper(\alpha, \tau) \in \mathcal{S}_n$ is defined.

(c) The interval system corresponding to $\alpha$ is an $\ell$-tuple of intervals (since $\ell(\alpha) = \ell$); denote this $\ell$-tuple by $(I_1, I_2, \ldots, I_\ell)$.

A permutation $\sigma \in \mathcal{S}_n$ is said to be $\alpha$-clumping if every $i \in \{1, 2, \ldots, \ell\}$ has the two properties that:

- the set $\sigma^{-1}(I_i)$ is an interval;
- the restriction of the map $\sigma^{-1}$ to the interval $I_i$ is increasing.

Example 6.55. For this example, let $n = 7$ and $\alpha = (2, 1, 3, 1)$. Then, $\ell = \ell(\alpha) = 4$ and $(I_1, I_2, I_3, I_4) = (\{1, 2\}, \{3\}, \{4, 5, 6\}, \{7\})$ (where we are using the notations of Definition 6.54). Hence, $I_1 = 12$, $I_2 = 3$, $I_3 = 456$ and $I_4 = 7$.

(a) If $\tau \in \mathcal{S}_4 = \mathcal{S}_\ell$ is the permutation $(2, 3, 1, 4)$, then $iper(\alpha, \tau)$ is the permutation in $\mathcal{S}_7$ (which (in one-line notation) is the word $I_{\tau(1)} I_{\tau(2)} I_{\tau(3)} I_{\tau(4)} = I_2 I_3 I_1 I_4 = 3456127$.

If $\tau \in \mathcal{S}_4 = \mathcal{S}_\ell$ is the permutation $(3, 1, 4, 2)$, then $iper(\alpha, \tau)$ is the permutation in $\mathcal{S}_7$ (which (in one-line notation) is the word $I_{\tau(1)} I_{\tau(2)} I_{\tau(3)} I_{\tau(4)} = I_3 I_1 I_4 I_2 = 4561273$.

(b) The permutation $\sigma = (3, 7, 4, 5, 6, 1, 2) \in \mathcal{S}_\ell$ (given here in one-line notation) is $\alpha$-clumping, because:

- every $i \in \{1, 2, \ldots, \ell\} = \{1, 2, 3, 4\}$ has the property that $\sigma^{-1}(I_i)$ is an interval (namely, $\sigma^{-1}(I_1) = \sigma^{-1}(\{1, 2\}) = \{6, 7\}$, $\sigma^{-1}(I_2) = \sigma^{-1}(\{3\}) = \{1\}$, $\sigma^{-1}(I_3) = \sigma^{-1}(\{4, 5, 6\}) = \{3, 4, 5\}$ and $\sigma^{-1}(I_4) = \sigma^{-1}(\{7\}) = \{2\}$), and
- the restrictions of the map $\sigma^{-1}$ to the intervals $I_i$ are increasing (this means that $\sigma^{-1}(1) < \sigma^{-1}(2)$ and $\sigma^{-1}(4) < \sigma^{-1}(5) < \sigma^{-1}(6)$, since the one-element intervals $I_2$ and $I_4$ do not contribute anything to this condition).

Here is a more or less trivial observation:

Proposition 6.56. Let $n \in \mathbb{N}$. Let $\alpha$ be a composition of $n$. Let $\ell = \ell(\alpha)$. Write $\alpha$ in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$. The interval system corresponding to $\alpha$ is an $\ell$-tuple of intervals (since $\ell(\alpha) = \ell$); denote this $\ell$-tuple by $(I_1, I_2, \ldots, I_\ell)$. Let $\tau \in \mathcal{S}_\ell$. Set $\sigma = iper(\alpha, \tau)$.\footnote{In fact, from the properties of interval systems, we know that the intervals $I_1, I_2, \ldots, I_\ell$ form a set partition of the set $[0: n]^+$. Hence, the intervals $I_{\tau(1)}, I_{\tau(2)}, \ldots, I_{\tau(\ell)}$ form a set partition of the set $[0: n]^+$. As a consequence, the word $I_{\tau(1)} I_{\tau(2)} \cdots I_{\tau(\ell)}$ is a permutation of the word $12 \ldots n$, and so there exists a permutation in $\mathcal{S}_n$ which (in one-line notation) is this word, qed.}
(a) We have \( \sigma^{-1}(I_{\tau(j)}) = \left[ \sum_{k=1}^{j-1} \alpha_{\tau(k)} : \sum_{k=1}^{j} \alpha_{\tau(k)} \right]^+ \) for every \( j \in \{1, 2, \ldots, \ell \} \).

(b) For every \( j \in \{1, 2, \ldots, \ell \} \), the restriction of the map \( \sigma^{-1} \) to the interval \( I_{\tau(j)} \) is increasing.

(c) The permutation \( \text{iper}(\alpha, \tau) \) is \( \alpha \)-clumping.

(d) Let \( i \in \{1, 2, \ldots, \ell - 1\} \). Then, the sets \( \sigma^{-1}(I_{\tau(i)}), \sigma^{-1}(I_{\tau(i+1)}) \) and \( \sigma^{-1}(I_{\tau(i)}) \cup \sigma^{-1}(I_{\tau(i+1)}) \) are nonempty intervals. Also, \( \sigma^{-1}(I_{\tau(i)}) < \sigma^{-1}(I_{\tau(i+1)}) \).

**Exercise 6.57.** Prove Proposition 6.66.

**Proposition 6.58.** Let \( n \in \mathbb{N} \). Let \( \alpha \) be a composition of \( n \). Let \( \ell = \ell(\alpha) \).

(a) Define a map

\[
\text{iper}_\alpha : \mathfrak{S}_\ell \rightarrow \{ \omega \in \mathfrak{S}_n \mid \omega \text{ is } \alpha \text{-clumping} \},
\]

\[\tau \mapsto \text{iper}(\alpha, \tau)\]

226. This map \( \text{iper}_\alpha \) is bijective.

(b) Let \( \sigma \in \mathfrak{S}_n \) be an \( \alpha \)-clumping permutation. Then, there exists a unique \( \tau \in \mathfrak{S}_\ell \) satisfying \( \sigma = \text{iper}(\alpha, \tau) \).

**Exercise 6.59.** Prove Proposition 6.58.

Next, we recall that the concatenation \( \alpha \cdot \beta \) of two compositions \( \alpha \) and \( \beta \) is defined in the same way as the concatenation of two words; if we regard compositions as words over the alphabet \( \{1, 2, 3, \ldots\} \), then the concatenation \( \alpha \cdot \beta \) of two compositions \( \alpha \) and \( \beta \) is the concatenation \( \alpha \beta \) of the words \( \alpha \) and \( \beta \). Thus, we are going to write \( \alpha \beta \) for the concatenation \( \alpha \cdot \beta \) of two compositions \( \alpha \) and \( \beta \) from now on.

**Proposition 6.60.** Let \( n \in \mathbb{N} \) and \( m \in \mathbb{N} \). Let \( \alpha \) be a composition of \( n \), and \( \beta \) be a composition of \( m \). Let \( p = \ell(\alpha) \) and \( q = \ell(\beta) \). Let \( \tau \in \mathfrak{S}_{p+q} \). Notice that \( \text{iper}(\alpha \beta, \tau) \in \mathfrak{S}_{n+m} \) (since \( \alpha \beta \) is a composition of \( n + m \) having length \( \ell(\alpha \beta) = \ell(\alpha) + \ell(\beta) = p + q \)). Then, \( \tau \in \text{Sh}_{p,q} \) if and only if \( \text{iper}(\alpha \beta, \tau) \in \text{Sh}_{n,m} \).

**Exercise 6.61.** Prove Proposition 6.60.

Here is one more simple fact:

**Lemma 6.62.** Let \( u \) and \( v \) be two words. Let \( n = \ell(u) \) and \( m = \ell(v) \). Let \( \alpha \) be a composition of \( n \), and let \( \beta \) be a composition of \( m \). Let \( p = \ell(\alpha) \) and \( q = \ell(\beta) \). The concatenation \( \alpha \beta \) is a composition of \( n + m \) having length \( \ell(\alpha \beta) = \ell(\alpha) + \ell(\beta) = p + q \). Thus, the interval system corresponding to \( \alpha \beta \) is a \((p + q)\)-tuple of intervals which covers \([0 : n + m]^+ \). Denote this \((p + q)\)-tuple by \((I_1, I_2, \ldots, I_{p+q})\).

Let \( \tau \in \text{Sh}_{p,q} \). Set \( \sigma = \text{iper}(\alpha \beta, \tau) \). Then,

\[
\begin{align*}
\bigcup_{\sigma} u \cup v &= (uv) \left[ I_{\tau(1)} \right] \cdot (uv) \left[ I_{\tau(2)} \right] \cdots (uv) \left[ I_{\tau(p+q)} \right] .
\end{align*}
\]

**Exercise 6.63.** Prove Lemma 6.62.

Having these notations and trivialities in place, we can say a bit more about the lexicographically highest element of a shuffle product than what was said in Theorem 6.44:

**Theorem 6.64.** Let \( u \) and \( v \) be two words. Let \( n = \ell(u) \) and \( m = \ell(v) \).

Let \( (a_1, a_2, \ldots, a_p) \) be the CFL factorization of \( u \). Let \( (b_1, b_2, \ldots, b_q) \) be the CFL factorization of \( v \).

Let \( \alpha \) be the \( p \)-tuple \((\ell(a_1), \ell(a_2), \ldots, \ell(a_p))\). Then, \( \alpha \) is a composition \(^{227}\) of length \( p \) and size \( \sum_{k=1}^{p} \ell(a_k) = \ell\left(\underbrace{a_1 a_2 \cdots a_p}_{= u}\right) = \ell(u) = n \).

Let \( \beta \) be the \( q \)-tuple \((\ell(b_1), \ell(b_2), \ldots, \ell(b_q))\). Then, \( \beta \) is a composition of length \( q \) and size \( \sum_{k=1}^{q} \ell(b_k) = m \). \(^{228}\)

---

226 This map is well-defined because for every \( \tau \in \mathfrak{S}_\ell \), the permutation \( \text{iper}(\alpha, \tau) \) is \( \alpha \)-clumping (according to Proposition 6.56(c)).

227 since Lyndon words are nonempty, and thus \( \ell(\alpha_i) > 0 \) for every \( i \)

228 The proof of this is the same as the proof of the fact that \( \alpha \) is a composition of length \( p \) and size \( \sum_{k=1}^{p} \ell(\alpha_k) = n \).
Now, $\alpha$ is a composition of length $p$ and size $n$, and $\beta$ is a composition of length $q$ and size $m$. Thus, the concatenation $\alpha\beta$ of these two tuples is a composition of length $p + q$ and size $n + m$. The interval system corresponding to this composition $\alpha\beta$ is a $(p + q)$-tuple (since said composition has length $p + q$); denote this $(p + q)$-tuple by $(I_1, I_2, \ldots, I_{p+q})$.

(a) If $\tau \in S_{p,q}$ satisfies $(uv) [I_\tau(1)] \geq (uv) [I_\tau(2)] \geq \cdots \geq (uv) [I_\tau(p+q)]$, and if we set $\sigma = \text{iper}(\alpha\beta, \tau)$, then $\sigma \in S_{n,m}$, and the word $u \cup v$ is the lexicographically highest element of the multiset $u \cup v$.

(b) Let $\sigma \in S_{n,m}$ be a permutation such that $u \cup v$ is the lexicographically highest element of the multiset $u \cup v$. Then, there exists a unique permutation $\tau \in S_{p,q}$ satisfying $(uv) [I_\tau(1)] \geq (uv) [I_\tau(2)] \geq \cdots \geq (uv) [I_\tau(p+q)]$ and $\sigma = \text{iper}(\alpha\beta, \tau)$.

Proof. Before we step to the actual proof, we need to make some preparation. First of all, $(I_1, I_2, \ldots, I_{p+q})$ is the interval system corresponding to the composition $\alpha\beta$. In other words,

$$(6.5) \quad (I_1, I_2, \ldots, I_{p+q}) = \text{intsys}(\alpha\beta).$$

But since $\alpha = (\ell(a_1), \ell(a_2), \ldots, \ell(a_p))$ and $\beta = (\ell(b_1), \ell(b_2), \ldots, \ell(b_q))$, we have

$$\alpha\beta = (\ell(a_1), \ell(a_2), \ldots, \ell(a_p), \ell(b_1), \ell(b_2), \ldots, \ell(b_q)).$$

Thus, (6.5) rewrites as

$$(6.6) \quad (I_1, I_2, \ldots, I_{p+q}) = \text{intsys}(\ell(a_1), \ell(a_2), \ldots, \ell(a_p), \ell(b_1), \ell(b_2), \ldots, \ell(b_q)).$$

By the definition of intsys, $(\ell(a_1), \ell(a_2), \ldots, \ell(a_p), \ell(b_1), \ell(b_2), \ldots, \ell(b_q))$, we thus have

$$I_i = \left[\sum_{k=1}^{i-1} \ell(a_k) : \sum_{k=1}^i \ell(a_k) \right]^+$$

for every $i \in \{1, 2, \ldots, p\}$, and besides

$$I_{p+j} = \left[n + \sum_{k=1}^{j-1} \ell(b_k) : n + \sum_{k=1}^j \ell(b_k) \right]^+$$

for every $j \in \{1, 2, \ldots, q\}$ (since $\sum_{k=1}^p \ell(a_k) = n$). Moreover, Remark 6.48(c) shows that $(I_1, I_2, \ldots, I_{p+q})$ is a $(p + q)$-tuple of nonempty intervals of $Z$ and satisfies the following three properties:

- The intervals $I_1, I_2, \ldots, I_{p+q}$ form a set partition of the set $[0 : n + m]^+$.
- We have $I_1 < I_2 < \cdots < I_{p+q}$.
- We have $|I_i| = \ell(a_i)$ for every $i \in \{1, 2, \ldots, p\}$ and $|I_{p+j}| = \ell(b_j)$ for every $j \in \{1, 2, \ldots, q\}$.

Of course, every $i \in \{1, 2, \ldots, p\}$ satisfies

$$(6.6) \quad I_i \subset [0 : n]^+ \quad \text{and} \quad (uv) [I_i] = u [I_i] = a_i.$$

Meanwhile, every $i \in \{p + 1, p + 2, \ldots, p + q\}$ satisfies

$$(6.7) \quad I_i \subset [n : n + m]^+ \quad \text{and} \quad (uv) [I_i] = v [I_i - n] = b_{i-p}$$

(where $I_i - n$ denotes the interval $\{k - n \mid k \in I_i\}$). We thus see that

$$(6.8) \quad (uv) [I_i] \text{ is a Lyndon word } \quad \text{for every } i \in \{1, 2, \ldots, p + q\}$$

By the definition of a CFL factorization, we have $a_1 \geq a_2 \geq \cdots \geq a_p$ and $b_1 \geq b_2 \geq \cdots \geq b_q$. We have $\sigma \in S_{n,m}$, so that $\sigma^{-1}(1) < \sigma^{-1}(2) < \cdots < \sigma^{-1}(n)$ and $\sigma^{-1}(n + 1) < \sigma^{-1}(n + 2) < \cdots < \sigma^{-1}(n + m)$. In other words, the restriction of the map $\sigma^{-1}$ to the interval $[0 : n]^+$ is strictly increasing, and so is the restriction of the map $\sigma^{-1}$ to the interval $[n : n + m]^+$.

(b) We will first show that

$$(6.9) \quad \text{if } J \subset [0 : n]^+ \text{ is an interval such that the word } (uv) [J] \text{ is Lyndon, then } \sigma^{-1}(J) \text{ is an interval.}$$

Proof of (6.9): We will prove (6.9) by strong induction over $|J|$. 

\[^{229}\] Indeed, when $i \leq p$, this follows from (6.6) and the fact that $a_i$ is Lyndon; whereas in the other case, this follows from (6.7) and the fact that $b_{i-p}$ is Lyndon.
So, fix some \( N \in \mathbb{N} \). Assume (as the induction hypothesis) that (6.9) has been proven whenever \(|J| < N\). We now need to prove (6.9) when \(|J| = N\).

Let \( J \subset [0 : n]^+ \) be an interval such that the word \((uv)[J]\) is Lyndon and such that \(|J| = N\). We have to prove that \(\sigma^{-1}(J)\) is an interval. This is obvious if \(|J| = 1\) (because in this case, \(\sigma^{-1}(J)\) is a one-element set, thus trivially an interval). Hence, we WLOG assume that we don’t have \(|J| = 1\). We also don’t have \(|J| = 0\), because \((uv)[J]\) has to be Lyndon (and the empty word is not). So we have \(|J| > 1\). Now, \(\ell((uv)[J]) = |J| > 1\), and thus \((uv)[J]\) is a Lyndon word of length \(> 1\). Let \(v'\) be the (lexicographically) smallest nonempty proper suffix of \((uv)[J]\). Since \(v'\) is a proper suffix of \(w\), there exists a nonempty \(u' \in \mathbb{A}^*\) such that \((uv)[J] = u'v'\). Consider this \(u'\).

Now, Theorem 6.30(a) (applied to \((uv)[J]\), \(u'\) and \(v'\) instead of \(w\), \(u\) and \(v\)) yields that the words \(u'\) and \(v'\) are Lyndon. Also, Theorem 6.30(b) (applied to \((uv)[J]\), \(u'\) and \(v'\) instead of \(w\), \(u\) and \(v\)) yields that \(u' < (uv)[J] < v'\).

But from the fact that \((uv)[J] = u'v'\) with \(u'\) and \(v'\) both being nonempty, it becomes immediately clear that we can write \(J\) as a union of two disjoint nonempty intervals \(K\) and \(L\) such that \(K < L\), \(u' = (uv)[K]\) and \(v' = (uv)[L]\). Consider these \(K\) and \(L\). The intervals \(K\) and \(L\) are nonempty and have their sizes add up to \(N\) (since they are disjoint and their union is \(J\)), and hence both must have size smaller than \(|J| = N\). So \(K \subset [0 : n]^+\) is an interval of size \(|K| < N\) having the property that \((uv)[K]\) is Lyndon (since \((uv)[K] = u'\) is Lyndon). Thus, we can apply (6.9) to \(K\) instead of \(J\) (because of the induction hypothesis). As a result, we conclude that \(\sigma^{-1}(K)\) is an interval. Similarly, we can apply (6.9) to \(L\) instead of \(J\) (we know that \((uv)[L]\) is Lyndon since \((uv)[L] = v'\)), and learn that \(\sigma^{-1}(L)\) is an interval. The intervals \(\sigma^{-1}(K)\) and \(\sigma^{-1}(L)\) are both nonempty (since \(K\) and \(L\) are nonempty), and their union is \(\sigma^{-1}(J)\) (because the union of \(K\) and \(L\) is \(J\)). The nonempty intervals \(K\) and \(L\) both are subsets of \([0 : n]^+\) (since their union is \(J \subset [0 : n]^+\)), and their union \(K \cup L\) is an interval (since their union \(K \cup L = J\), and we know that \(J\) is an interval).

Now, assume (for the sake of contradiction) that \(\sigma^{-1}(J)\) is not an interval. Since \(J\) is the union of \(K\) and \(L\), we have \(J = K \cup L\) and thus \(\sigma^{-1}(J) = \sigma^{-1}(K \cup L) = \sigma^{-1}(K) \cup \sigma^{-1}(L)\) (since \(\sigma\) is a bijection). Therefore, \(\sigma^{-1}(K) \cup \sigma^{-1}(L)\) is not an interval (since \(\sigma^{-1}(J)\) is not an interval). Thus, Lemma 6.50(b) yields that there exists a nonempty interval \(P \subset [n : n + m]^+\) such that \(\sigma^{-1}(P)\), \(\sigma^{-1}(K) \cup \sigma^{-1}(P)\) and \(\sigma^{-1}(P) \cup \sigma^{-1}(L)\) are intervals and such that \(\sigma^{-1}(K) < \sigma^{-1}(P) < \sigma^{-1}(L)\). Consider this \(P\). Since \(P\) is nonempty, we have \(|P| \neq 0\).

Lemma 6.52(b) (applied to \(K\) and \(P\) instead of \(I\) and \(J\)) yields

\[(uv)[K] : (uv)[P] \geq (uv)[P] : (uv)[K].\]

Since \((uv)[K] = u'\), this rewrites as

\[(uv)[P] : (uv)[P] \geq (uv)[P] : u'.\]

But Lemma 6.52(c) (applied to \(P\) and \(L\) instead of \(I\) and \(J\)) yields

\[(uv)[P] : (uv)[L] \geq (uv)[L] : (uv)[P].\]

Since \((uv)[L] = v'\), this rewrites as

\[(uv)[P] : v' \geq v' : (uv)[P].\]

Recall also that \(u' < v'\), and that both words \(u'\) and \(v'\) are Lyndon. Now, Corollary 6.17 (applied to \(u'\), \(v'\) and \((uv)[P]\) instead of \(u\), \(v\) and \(z\)) yields that \((uv)[P]\) is the empty word (because of (6.11) and (6.13)), so that \(\ell((uv)[P]) = 0\). This contradicts \(\ell((uv)[P]) = |P| \neq 0\). This contradiction shows that our assumption (that \(\sigma^{-1}(J)\) is not an interval) was wrong. Hence, \(\sigma^{-1}(J)\) is an interval. This completes the induction step, and thus (6.9) is proven.

Similarly to (6.9), we can show that

\[(uv)[P] : v' \geq u' : (uv)[P].\]

if \(J \subset [n : n + m]^+\) is an interval such that the word \((uv)[J]\) is Lyndon, then \(\sigma^{-1}(J)\) is an interval.

Now, let \(i \in \{1, 2, \ldots, p + q\}\) be arbitrary. We are going to prove that

\(\sigma^{-1}(I_i)\) is an interval.

**Proof of (6.15):** We must be in one of the following two cases:

**Case 1:** We have \(i \in \{1, 2, \ldots, p\}\).

**Case 2:** We have \(i \in \{p + 1, p + 2, \ldots, p + q\}\).
Let us first consider Case 1. In this case, we have \( i \in \{1, 2, \ldots, p\} \). Thus, \( I_i \subset [0 : n]^+ \) (by (6.6)). Also, (6.6) yields that \((uv)[I_i] = a_i\) is a Lyndon word. Hence, (6.9) (applied to \( J = I_i \)) yields that \( \sigma^{-1}(I_i) \) is an interval. Thus, (6.15) is proven in Case 1.

Similarly, we can prove (6.15) in Case 2, using (6.7) and (6.14) instead of (6.6) and (6.9), respectively. Hence, (6.15) is proven.

We know that \( \sigma^{-1}(I_i) \) is an interval. But we also know that either \( I_i \subset [0 : n]^+ \) or \( I_i \subset [n : n + m]^+ \) (depending on whether \( i \leq p \) or \( i > p \)). As a consequence, the restriction of the map \( \sigma^{-1} \) to the interval \( I_i \) is increasing (because the restriction of the map \( \sigma^{-1} \) to the interval \([0 : n]^+\) is strictly increasing, and so is the restriction of the map \( \sigma^{-1} \) to the interval \([n : n + m]^+\).

Now, let us forget that we fixed \( i \). We thus have shown that every \( i \in \{1, 2, \ldots, p + q\} \) has the two properties that:

- the set \( \sigma^{-1}(I_i) \) is an interval;
- the restriction of the map \( \sigma^{-1} \) to the interval \( I_i \) is increasing.

In other words, the permutation \( \sigma \) is \((\alpha \beta)\)-clumping (since \( I_1, I_2, \ldots, I_{p+q} \) is the interval system corresponding to the composition \( \alpha \beta \)). Hence, Proposition 6.58(b) (applied to \( n + m, \alpha \beta \) and \( p + q \) instead of \( n, \alpha \) and \( \ell \)) shows that there exists a unique \( \tau \in \mathfrak{S}_{p+q} \) satisfying \( \sigma = \text{iper}(\alpha \beta, \tau) \). Thus, the uniqueness part of Theorem 6.64(b) (i.e., the claim that the \( \tau \) in Theorem 6.64(b) is unique if it exists) is proven.

It now remains to prove the existence part of Theorem 6.64(b), i.e., to prove that there exists at least one permutation \( \tau \in \mathfrak{S}_{p+q} \) satisfying \((uv)[I_{\tau(1)}] \geq (uv)[I_{\tau(2)}] \geq \cdots \geq (uv)[I_{\tau(p+q)}] \) and \( \sigma = \text{iper}(\alpha \beta, \tau) \). Consider this \( \tau \). We will now prove that \((uv)[I_{\tau(1)}] \geq (uv)[I_{\tau(2)}] \geq \cdots \geq (uv)[I_{\tau(p+q)}] \) and \( \tau \in \mathfrak{S}_{p+q} \). Once this is done, the existence part of Theorem 6.64(b) will be proven, and thus the proof of Theorem 6.64(b) will be complete.

Proposition 6.60 yields that \( \tau \in \mathfrak{S}_{p+q} \) if and only if is \( \text{iper}(\alpha \beta, \tau) \in \mathfrak{S}_{n,m} \). Since we know that \( \text{iper}(\alpha \beta, \tau) = \sigma \in \mathfrak{S}_{n,m} \), we thus conclude that \( \tau \in \mathfrak{S}_{p+q} \). The only thing that remains to be proven now is that

\[(6.16) \quad (uv)[I_{\tau(1)}] \geq (uv)[I_{\tau(2)}] \geq \cdots \geq (uv)[I_{\tau(p+q)}].\]

Proof of (6.16): Every \( \tau \in \mathfrak{S}_{p+q} \). In other words, \( \tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p) \) and \( \tau^{-1}(p+1) < \tau^{-1}(p+2) < \cdots < \tau^{-1}(p+q) \). In other words, the restriction of the map \( \tau^{-1} \) to the interval \([0 : p]^+\) is strictly increasing and so is the restriction of the map \( \tau^{-1} \) to the interval \([p : p+q]^+\).

Let \( i \in \{1, 2, \ldots, p+q-1\} \). We will show that

\[(6.17) \quad (uv)[I_{\tau(i)}] \geq (uv)[I_{\tau(i+1)}].\]

Clearly, both \( \tau(i) \) and \( \tau(i+1) \) belong to \( \{1, 2, p+q\} = \{1, 2, \ldots, p\} \cup \{p+1, p+2, \ldots, p+q\} \). Thus, we must be in one of the following four cases:

- **Case 1**: We have \( \tau(i) \in \{1, 2, \ldots, p\} \) and \( \tau(i+1) \in \{1, 2, \ldots, p\} \).
- **Case 2**: We have \( \tau(i) \in \{1, 2, \ldots, p\} \) and \( \tau(i+1) \in \{p+1, p+2, \ldots, p+q\} \).
- **Case 3**: We have \( \tau(i) \in \{p+1, p+2, \ldots, p+q\} \) and \( \tau(i+1) \in \{1, 2, \ldots, p\} \).
- **Case 4**: We have \( \tau(i) \in \{p+1, p+2, \ldots, p+q\} \) and \( \tau(i+1) \in \{p+1, p+2, \ldots, p+q\} \).

Let us consider Case 1 first. In this case, we have \( \tau(i) \in \{1, 2, \ldots, p\} \) and \( \tau(i+1) \in \{1, 2, \ldots, p\} \). From the fact that the restriction of the map \( \tau^{-1} \) to the interval \([0 : p]^+\) is strictly increasing, we can easily deduce \( \tau(i) < \tau(i+1) \). Therefore, \( a_{\tau(i)} \geq a_{\tau(i+1)} \) (since \( a_1 \geq a_2 \geq \cdots \geq a_p \)).

But \((uv)[I_{\tau(i)}] = a_{\tau(i)} \) (by (6.6), applied to \( \tau(i) \) instead of \( i \)) and \((uv)[I_{\tau(i+1)}] = a_{\tau(i+1)} \) (similarly). In view of these equalities, the inequality \( a_{\tau(i)} \geq a_{\tau(i+1)} \) rewrites as \((uv)[I_{\tau(i)}] \geq (uv)[I_{\tau(i+1)}]\). Thus, (6.17) is proven in Case 1.

Similarly, we can show (6.17) in Case 4 (observing that \((uv)[I_{\tau(i)}] = b_{r(i)-p} \) and \((uv)[I_{\tau(i+1)}] = b_{r(i+1)-p} \) in this case).

Let us now consider Case 2. In this case, we have \( \tau(i) \in \{1, 2, \ldots, p\} \) and \( \tau(i+1) \in \{p+1, p+2, \ldots, p+q\} \). From \( \tau(i) \in \{1, 2, \ldots, p\} \), we conclude that \( I_{\tau(i)} \subset [0 : n]^+ \). From \( \tau(i+1) \in \{p+1, p+2, \ldots, p+q\} \), we conclude that \( I_{\tau(i+1)} \subset [n : n+m]^+ \). The intervals \( I_{\tau(i)} \) and \( I_{\tau(i+1)} \) are clearly nonempty.
Proposition 6.56(d) (applied to \(n+m, \alpha \beta, p+q\) and \((I_1, I_2, \ldots, I_{p+q})\) instead of \(n, \alpha, \ell\) and \((I_1, I_2, \ldots, I_{\ell})\)) yields that the sets \(\sigma^{-1}(I_{i(i)}, I_{i(i+1)})\) and \(\sigma^{-1}(I_{i(i)}, I_{i(i+1)})\cup\sigma^{-1}(I_{i(i+1)}, I_{i(i)})\) are nonempty intervals, and that we have \(\sigma^{-1}(I_{i(i)}) < \sigma^{-1}(I_{i(i+1)})\). Hence, Lemma 6.52(b) (applied to \(I = I_{r(i)}\) and \(J = I_{r(i+1)}\)) yields
\[
(uv) [I_{r(i)}] \cdot (uv) [I_{r(i+1)}] \geq (uv) [I_{r(i)}] \cdot (uv) [I_{r(i+1)}].
\]
But \((uv) [I_{r(i)}] \cdot (uv) [I_{r(i+1)}]\) are Lyndon words (as a consequence of (6.8)). Thus, Proposition 6.18 (applied to \((uv) [I_{r(i)}] \cdot (uv) [I_{r(i+1)}]\) instead of \(u\) and \(v\)) shows that \((uv) [I_{r(i)}] \geq (uv) [I_{r(i+1)}]\) if and only if \((uv) [I_{r(i)}] \cdot (uv) [I_{r(i+1)}] \geq (uv) [I_{r(i+1)}] \cdot (uv) [I_{r(i)}]\). Since we know that \((uv) [I_{r(i)}] \cdot (uv) [I_{r(i+1)}] \geq (uv) [I_{r(i+1)}] \cdot (uv) [I_{r(i)}]\) holds, we thus conclude that \((uv) [I_{r(i)}] \geq (uv) [I_{r(i+1)}]\). Thus, (6.17) is proven in Case 2.

The proof of (6.17) in Case 3 is analogous to that in Case 2 (the main difference being that Lemma 6.52(c) is used in lieu of Lemma 6.52(b)).

Thus, (6.17) is proven in all possible cases. So we always have (6.17). In other words, \((uv) [I_{r(i)}] \geq (uv) [I_{r(i+1)}]\).

Now, forget that we fixed \(i\). We hence have shown that \((uv) [I_{r(i)}] \geq (uv) [I_{r(i+1)}]\) for all \(i \in \{1, 2, \ldots, p+q-1\}\). This proves (6.16), and thus completes our proof of Theorem 6.64(b).

(a) Let \(\sigma \in \text{Sh}_{p,q}\) be such that
\[
(uv) [I_{r(1)}] \geq (uv) [I_{r(2)}] \geq \cdots \geq (uv) [I_{r(p+q)}].
\]
Set \(\sigma = \text{iper}(\alpha \beta, \tau)\). Then, Proposition 6.60 yields that \(\tau \in \text{Sh}_{p,q}\) if and only if \(\text{iper}(\alpha \beta, \tau) \in \text{Sh}_{n,m}\). Since we know that \(\tau \in \text{Sh}_{p,q}\), we can deduce from this that \(\text{iper}(\alpha \beta, \tau) \in \text{Sh}_{p,q}\), so that \(\sigma = \text{iper}(\alpha \beta, \tau) \in \text{Sh}_{n,m}\).

It remains to prove that the word \(u \uplus v\) is the lexicographically highest element of the multiset \(u \uplus v\).

It is clear that the multiset \(u \uplus v\) has some lexicographically highest element. This element has the form \(u \uplus \sigma v\) for some \(\sigma \in \text{Sh}_{n,m}\) (because any element of this multisets has such a form). Consider this \(\sigma\).

Theorem 6.64(b) (applied to \(\sigma\) instead of \(\sigma\)) yields that there exists a unique permutation \(\bar{\tau} \in \text{Sh}_{p,q}\) satisfying \((uv) [I_{r(1)}] \geq (uv) [I_{r(2)}] \geq \cdots \geq (uv) [I_{r(p+q)}]\) and \(\bar{\tau} = \text{iper}(\alpha \beta, \bar{\tau})\). (What we call \(\bar{\tau}\) here is what has been called \(\tau\) in Theorem 6.64(b)).

Now, the chain of inequalities \((uv) [I_{r(1)}] \geq (uv) [I_{r(2)}] \geq \cdots \geq (uv) [I_{r(p+q)}]\) shows that the list \(\langle (uv) [I_{r(1)}], (uv) [I_{r(2)}], \ldots, (uv) [I_{r(p+q)}]\rangle\) is the result of sorting the list \(\langle (uv) [I_{1}], (uv) [I_{2}], \ldots, (uv) [I_{p+q}]\rangle\) in increasing order. But the chain of inequalities (6.18) shows that the list \(\langle (uv) [I_{r(1)}], (uv) [I_{r(2)}], \ldots, (uv) [I_{r(p+q)}]\rangle\) is the result of sorting the same list \(\langle (uv) [I_{1}], (uv) [I_{2}], \ldots, (uv) [I_{p+q}]\rangle\) in increasing order. So each of the two lists \(\langle (uv) [I_{r(1)}], (uv) [I_{r(2)}], \ldots, (uv) [I_{r(p+q)}]\rangle\) and \(\langle (uv) [I_{r(1)}], (uv) [I_{r(2)}], \ldots, (uv) [I_{r(p+q)}]\rangle\) is the result of sorting one and the same list \(\langle (uv) [I_{1}], (uv) [I_{2}], \ldots, (uv) [I_{p+q}]\rangle\) in decreasing order. Since the result of sorting a given list in decreasing order is unique, this yields
\[
\langle (uv) [I_{r(1)}], (uv) [I_{r(2)}], \ldots, (uv) [I_{r(p+q)}]\rangle = \langle (uv) [I_{r(1)}], (uv) [I_{r(2)}], \ldots, (uv) [I_{r(p+q)}]\rangle.
\]
Hence,
\[
(uv) [I_{r(1)}] \cdot (uv) [I_{r(2)}] \cdots \cdot (uv) [I_{r(p+q)}] = (uv) [I_{r(1)}] \cdot (uv) [I_{r(2)}] \cdots \cdot (uv) [I_{r(p+q)}].
\]
But Lemma 6.62 yields
\[
u \uplus v = (uv) [I_{r(1)}] \cdot (uv) [I_{r(2)}] \cdots \cdot (uv) [I_{r(p+q)}].
\]
Meanwhile, Lemma 6.62 (applied to \(\bar{\tau}\) and \(\bar{\sigma}\) instead of \(\tau\) and \(\sigma\)) yields
\[
\nu \uplus v = (uv) [I_{r(1)}] \cdot (uv) [I_{r(2)}] \cdots \cdot (uv) [I_{r(p+q)}]
\]
\[
= (uv) [I_{r(1)}] \cdot (uv) [I_{r(2)}] \cdots \cdot (uv) [I_{r(p+q)}] \quad \text{(by (6.19))}
\]
\[
= u \uplus v \quad \text{(by (6.20))}.
\]
Thus, \(u \uplus v\) is the lexicographically highest element of the multiset \(u \uplus v\) (since we know that \(u \uplus v\) is the lexicographically highest element of the multiset \(u \uplus v\)). This proves Theorem 6.64(a).
Proposition 6.65. Let $p \in \mathbb{N}$ and $q \in \mathbb{N}$. Let $\mathcal{W}$ be a totally ordered set, and let $h : \{1, 2, \ldots, p + q\} \to \mathcal{W}$ be a map. Assume that $h(1) \geq h(2) \geq \cdots \geq h(p)$ and $h(p + 1) \geq h(p + 2) \geq \cdots \geq h(p + q)$.

For every $w \in \mathcal{W}$, let $a(w)$ denote the number of all $i \in \{1, 2, \ldots, p\}$ satisfying $h(i) = w$, and let $b(w)$ denote the number of all $i \in \{p + 1, p + 2, \ldots, p + q\}$ satisfying $h(i) = w$.

Then, the number of $\tau \in \mathcal{S}_{p,q}$ satisfying $h(\tau(1)) \geq \cdots \geq h(\tau(p + q))$ is $\prod_{w \in \mathcal{W}} \left( a(w) + b(w) \right)^{a(w)}$.

(Of course, all but finitely many factors of this product are 1.)


Proof of Theorem 6.64. Let $n = \ell(u)$ and $m = \ell(v)$. Define $\alpha, \beta$ and $(I_1, I_2, \ldots, I_{p+q})$ as in Theorem 6.64.

Since $(a_1, a_2, \ldots, a_p)$ is the CFL factorization of $u$, we have $a_1 \geq a_2 \geq \cdots \geq a_p$ and $a_1 a_2 \cdots a_p = u$. Similarly, $b_1 \geq b_2 \geq \cdots \geq b_q$ and $b_1 b_2 \cdots b_q = v$.

From (6.6), we see that $(uv)[I_i] = a_i$ for every $i \in \{1, 2, \ldots, p\}$. From (6.7), we see that $(uv)[I_i] = b_{i-p}$ for every $i \in \{p + 1, p + 2, \ldots, p + q\}$. Combining these two equalities, we obtain

$$(uv)[I_i] = \begin{cases} a_i, & \text{if } i \leq p; \\ b_{i-p}, & \text{if } i > p \end{cases} \text{ for every } i \in \{1, 2, \ldots, p + q\}. \tag{6.21}$$

In other words,

$$(uv)[I_1], (uv)[I_2], \ldots, (uv)[I_{p+q}] = (a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q). \tag{6.22}$$

(a) Let $z$ be the lexicographically highest element of the multiset $u \uplus v$. We must prove that $z = c_1 c_2 \cdots c_{p+q}$.

Since $z \in u \uplus v$, we can write $z$ in the form $u_{\sigma} v$ for some $\sigma \in \mathcal{S}_{n,m}$ (since we can write any element of $u \uplus v$ in this form). Consider this $\sigma$. Then, $u \uplus v = z$ is the lexicographically highest element of the multiset $u \uplus v$. Hence, Theorem 6.64(b) yields that there exists a unique permutation $\tau \in \mathcal{S}_{p,q}$ satisfying $(uv)[I_{\tau(1)}] \geq (uv)[I_{\tau(2)}] \geq \cdots \geq (uv)[I_{\tau(p+q)}]$ and $\sigma = \mathrm{iper}(\alpha, \beta, \tau)$. Consider this $\tau$.

Now, $\tau \in \mathcal{S}_{p,q} \subset \mathcal{S}_{n,m}$ is a permutation, and thus the list $((uv)[I_{\tau(1)}], (uv)[I_{\tau(2)}], \ldots, (uv)[I_{\tau(p+q)}])$ is a rearrangement of the list $((uv)[I_1], (uv)[I_2], \ldots, (uv)[I_{p+q}])$. Due to (6.22), this rewrites as follows: The list $((uv)[I_{\tau(1)}], (uv)[I_{\tau(2)}], \ldots, (uv)[I_{\tau(p+q)}])$ is a rearrangement of the list $(a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)$. Hence, $((uv)[I_{\tau(1)}], (uv)[I_{\tau(2)}], \ldots, (uv)[I_{\tau(p+q)}])$ is the result of sorting the list $(a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)$ in decreasing order (since $(uv)[I_{\tau(1)}] \geq (uv)[I_{\tau(2)}] \geq \cdots \geq (uv)[I_{\tau(p+q)}]$). But since the result of sorting the list $(a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)$ in decreasing order is $(c_1, c_2, \ldots, c_{p+q}),$ this becomes

$$((uv)[I_{\tau(1)}], (uv)[I_{\tau(2)}], \ldots, (uv)[I_{\tau(p+q)}]) = (c_1, c_2, \ldots, c_{p+q}).$$

Hence,

$$(uv)[I_{\tau(1)}] \cdot (uv)[I_{\tau(2)}] \cdots (uv)[I_{\tau(p+q)}] = c_1 c_2 \cdots c_{p+q}.$$ But Lemma 6.62 yields

$$u \uplus v = (uv)[I_{\tau(1)}] \cdot (uv)[I_{\tau(2)}] \cdots (uv)[I_{\tau(p+q)}].$$

Altogether, we have

$$z = u \uplus v = (uv)[I_{\tau(1)}] \cdot (uv)[I_{\tau(2)}] \cdots (uv)[I_{\tau(p+q)}] = c_1 c_2 \cdots c_{p+q} = c_1 c_2 \cdots c_{p+q}.$$ This proves Theorem 6.64(a).

(b) Recall that $u \uplus v = \left\{ u \uplus v : \sigma \in \mathcal{S}_{n,m} \right\}_{\mathrm{multiset}}$. Hence,

$$\begin{align*}
\text{(the multiplicity with which the lexicographically highest element of the multiset} \\
\text{u \uplus v appears in the multiset u \uplus v}) \\
\text{= \left( the number of all } \sigma \in \mathcal{S}_{n,m} \text{ such that u \uplus v is the} \\
\text{lexicographically highest element of the multiset u \uplus v).}
\end{align*}$$
But for a given $\sigma \in \text{Sh}_{n,m}$, we know that $u \uplus v$ is the lexicographically highest element of the multiset $u \uplus v$ if and only if $\sigma$ can be written in the form $\sigma = \text{iper} (a,\beta,\tau)$ for some $\tau \in \text{Sh}_{p,q}$ satisfying $(uv) \left[ I_{(1)} \right] \geq (uv) \left[ I_{(2)} \right] \geq \cdots \geq (uv) \left[ I_{(p+q)} \right]$. Hence,

\[
\left( \text{the number of all } \sigma \in \text{Sh}_{n,m} \text{ such that } u \uplus v \text{ is the lexicographically highest element of the multiset } u \uplus v \right)
= \left( \text{the number of all } \sigma \in \text{Sh}_{n,m} \text{ which can be written in the form } \sigma = \text{iper} (a,\beta,\tau) \right.
\]

for some $\tau \in \text{Sh}_{p,q}$ satisfying $(uv) \left[ I_{(1)} \right] \geq (uv) \left[ I_{(2)} \right] \geq \cdots \geq (uv) \left[ I_{(p+q)} \right])
= \left( \text{the number of all } \tau \in \text{Sh}_{p,q} \text{ satisfying } (uv) \left[ I_{(1)} \right] \geq (uv) \left[ I_{(2)} \right] \geq \cdots \geq (uv) \left[ I_{(p+q)} \right] \right)
\]

(because if $\sigma \in \text{Sh}_{n,m}$ can be written in the form $\sigma = \text{iper} (a,\beta,\tau)$ for some $\tau \in \text{Sh}_{p,q}$ satisfying $(uv) \left[ I_{(1)} \right] \geq (uv) \left[ I_{(2)} \right] \geq \cdots \geq (uv) \left[ I_{(p+q)} \right]$), then $\sigma$ can be written uniquely in this form. Thus,

\[
\begin{equation}
\left( \text{the multiplicity with which the lexicographically highest element of the multiset } u \uplus v \text{ appears in the multiset } u \uplus v \right)
= \left( \text{the number of all } \sigma \in \text{Sh}_{n,m} \text{ such that } u \uplus v \text{ is the lexicographically highest element of the multiset } u \uplus v \right)
= \left( \text{the number of all } \tau \in \text{Sh}_{p,q} \text{ satisfying } (uv) \left[ I_{(1)} \right] \geq (uv) \left[ I_{(2)} \right] \geq \cdots \geq (uv) \left[ I_{(p+q)} \right] \right).
\end{equation}
\]

Now, define a map $h: \{1, 2, \ldots, p + q\} \rightarrow \mathcal{L}$ by

\[
h (i) = \begin{cases} 
a_i, & \text{if } i \leq p; \\
b_{i-p}, & \text{if } i > p \end{cases}
\text{ for every } i \in \{1, 2, \ldots, p + q\}.
\]

Then, $h (1) \geq h (2) \geq \cdots \geq h (p)$ (because this is just a rewriting of $a_1 \geq a_2 \geq \cdots \geq a_p$) and $h (p + 1) \geq h (p + 2) \geq \cdots \geq h (p + q)$ (since this is just a rewriting of $b_1 \geq b_2 \geq \cdots \geq b_q$). For every $w \in \mathcal{L}$, the number of all $i \in \{1, 2, \ldots, p\}$ satisfying $h (i) = w$ is

\[
\left| \left\{ i \in \{1, 2, \ldots, p\} \mid h (i) = w \right\} \right|
= \left| \left\{ i \in \{1, 2, \ldots, p\} \mid a_i = w \right\} \right|
= \left( \text{the number of terms in the list } (a_1, a_2, \ldots, a_p) \text{ which are equal to } w \right)
= \left( \text{the number of terms in the CFL factorization of } u \text{ which are equal to } w \right)
= \left( \text{since the list } (a_1, a_2, \ldots, a_p) \text{ is the CFL factorization of } u \right)
= \text{mult}_w u
\]

(because $\text{mult}_w u$ is defined as the number of terms in the CFL factorization of $u$ which are equal to $w$). Similarly, for every $w \in \mathcal{L}$, the number of all $i \in \{p + 1, p + 2, \ldots, p + q\}$ satisfying $h (i) = w$ equals $\text{mult}_w v$. Thus, we can apply Proposition 6.65 to $\mathcal{W} = \mathcal{L}$, $a (w) = \text{mult}_w u$ and $b (w) = \text{mult}_w v$. As a result, we see that the number of $\tau \in \text{Sh}_{p,q}$ satisfying $h (\tau (1)) \geq h (\tau (2)) \geq \cdots \geq h (\tau (p + q))$ is $\prod_{w \in \mathcal{L}} \left( \text{mult}_w u + \text{mult}_w v \right)$.
In other words,

\[(6.24) \quad \prod_{w \in \mathcal{L}} \left( \frac{\text{mult}_w u + \text{mult}_w v}{\text{mult}_w u} \right).\]

But for every \(i \in \{1, 2, \ldots, p + q\}\), we have

\[h(i) = \begin{cases} a_i, & \text{if } i \leq p; \\ b_i - p, & \text{if } i > p \end{cases} = (uv)[I_i] \quad \text{(by (6.21))}.\]

Hence, for any \(\tau \in \text{Sh}_{p,q}\), the condition \(h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p + q))\) is equivalent to \((uv)[I_{\tau(1)}] \geq (uv)[I_{\tau(2)}] \geq \cdots \geq (uv)[I_{\tau(p+q)}]\). Thus,

\[
\left(\begin{array}{c}
\text{the number of all } \tau \in \text{Sh}_{p,q} \text{ satisfying } h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p + q))
\end{array}\right)
\]

\[= (\text{the number of all } \tau \in \text{Sh}_{p,q} \text{ satisfying } (uv)[I_{\tau(1)}] \geq (uv)[I_{\tau(2)}] \geq \cdots \geq (uv)[I_{\tau(p+q)}]).\]

\[= (\text{the multiplicity with which the lexicographically highest element of the multiset } u \sqcup v \text{ appears in the multiset } u \sqcup v)\]

(by (6.23)). Compared with (6.24), this yields

\[= \prod_{w \in \mathcal{L}} \left( \frac{\text{mult}_w u + \text{mult}_w v}{\text{mult}_w u} \right).\]

This proves Theorem 6.44(b).

(c) We shall use the notations of Theorem 6.44(a) and Theorem 6.44(b).

Assume that \(a_i \geq b_j\) for every \(i \in \{1, 2, \ldots, p\}\) and \(j \in \{1, 2, \ldots, q\}\). This, combined with \(a_1 \geq a_2 \geq \cdots \geq a_p\) and \(b_1 \geq b_2 \geq \cdots \geq b_q\), yields that \(a_1 \geq a_2 \geq \cdots \geq a_p \geq b_1 \geq b_2 \geq \cdots \geq b_q\). Thus, the list \((a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)\) is weakly decreasing. Thus, the result of sorting the list \((a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)\) in decreasing order is the list \((a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)\) itself. But since this result is \((c_1, c_2, \ldots, c_{p+q})\), this shows that \((c_1, c_2, \ldots, c_{p+q}) = (a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)\). Hence, \(c_1 c_2 \cdots c_{p+q} = a_1 a_2 \cdots a_p b_1 b_2 \cdots b_q = u v\). Now, Theorem 6.44(a) yields that the lexicographically highest element of the multiset \(u \sqcup v\) is \(c_1 c_2 \cdots c_{p+q} = u v\). This proves Theorem 6.44(c).

(d) We shall use the notations of Theorem 6.44(a) and Theorem 6.44(b).

Assume that \(a_i > b_j\) for every \(i \in \{1, 2, \ldots, p\}\) and \(j \in \{1, 2, \ldots, q\}\). Thus, \(a_i \geq b_j\) for every \(i \in \{1, 2, \ldots, p\}\) and \(j \in \{1, 2, \ldots, q\}\). Hence, Theorem 6.44(c) yields that the lexicographically highest element of the multiset \(u \sqcup v\) is \(u v\). Therefore, Theorem 6.44(b) shows that the multiplicity with which this word \(u v\) appears in the multiset \(u \sqcup v\) is \(\prod_{w \in \mathcal{L}} \left( \frac{\text{mult}_w u + \text{mult}_w v}{\text{mult}_w u} \right).\)
Now, every \( w \in \mathcal{L} \) satisfies \[
\left( \frac{\text{mult}_w u + \text{mult}_w v}{\text{mult}_w u} \right) = 1 \]
Thus, as we know, the multiplicity with which this word \( uv \) appears in the multiset \( u \cup v \) is
\[
\prod_{w \in \mathcal{L}} \left( \frac{\text{mult}_w u + \text{mult}_w v}{\text{mult}_w u} \right) = 1.
\]
This proves Theorem 6.44(d).

(c) We shall use the notations of Theorem 6.44(a) and Theorem 6.44(b).
Since \( u \) is a Lyndon word, the 1-tuple \( (u) \) is the CFL factorization of \( u \). Hence, we can apply Theorem 6.44(c) to 1 and \( (u) \) instead of \( p \) and \( (a_1, a_2, \ldots, a_p) \). As a result, we conclude that the lexicographically highest element of the multiset \( u \cup v \) is \( uv \). It remains to prove that the multiplicity with which this word \( uv \) appears in the multiset \( u \cup v \) is \( \text{mult}_u v + 1 \).

For every \( w \in \mathcal{L} \) satisfying \( w \neq u \), we have
\[
\text{mult}_w u = 0
\]
Also, \( \text{mult}_u u = 1 \) (for a similar reason). But \( uv \) is the lexicographically highest element of the multiset \( u \cup v \). Hence, the multiplicity with which the word \( uv \) appears in the multiset \( u \cup v \) is the multiplicity with which the lexicographically highest element of the multiset \( u \cup v \) appears in the multiset \( u \cup v \). According to Theorem 6.44(b), the latter multiplicity is
\[
\prod_{w \in \mathcal{L}} \left( \frac{\text{mult}_w u + \text{mult}_w v}{\text{mult}_w u} \right) = \frac{1}{\text{mult}_u u} \prod_{w \in \mathcal{L} ; \ w \neq u} \left( \frac{\text{mult}_w u + \text{mult}_w v}{\text{mult}_w u} \right) = \frac{1 + \text{mult}_u v}{1} \cdot \prod_{w \in \mathcal{L} ; \ w \neq u} \left( \frac{0 + \text{mult}_w v}{0} \right) = (\text{mult}_u u + 1) \cdot \prod_{w \in \mathcal{L} ; \ w \neq u} 1 = \text{mult}_u v + 1.
\]
This proves Theorem 6.44(c).

As an application of our preceding results, we can prove a further necessary and sufficient criterion for a word to be Lyndon; this criterion is due to Chen/Fox/Lyndon [31, \( \mathcal{A}' = \mathcal{A}'' \)]:

**Exercise 6.67.** Let \( w \in \mathcal{A}^* \) be a nonempty word. Prove that \( w \) is Lyndon if and only if for any two nonempty words \( u \in \mathcal{A}^* \) and \( v \in \mathcal{A}^* \) satisfying \( w = uv \), there exists at least one \( s \in u \cup v \) satisfying \( s > w \).

### 6.3. Radford’s theorem on the shuffle algebra.
We recall that our goal in Chapter 6 is to exhibit an algebraically independent generating set of the \( k \)-algebra \( \text{QSym} \). Having the notion of Lyndon words — which will, to some extent, but not literally, parametrize this generating set — in place, we could start the construction of this generating set immediately. However, it might come off as rather unmotivated this way, and so we begin with some warmups. First, we shall prove Radford’s theorem on the shuffle algebra.

### Definition 6.68. A polynomial algebra
will mean a \( k \)-algebra which is isomorphic to the polynomial ring \( k \{ x_i \mid \ i \in I \} \) as a \( k \)-algebra (for some indexing set \( I \)). Note that \( I \) need not be finite.

---

\[233\text{Proof.} \text{ Assume the contrary. Then, there exists at least one } w \in \mathcal{L} \text{ such that } \left( \frac{\text{mult}_w u + \text{mult}_w v}{\text{mult}_w u} \right) \neq 1. \text{ Consider this } w. \text{ Both } \text{mult}_w u \text{ and } \text{mult}_w v \text{ must be positive (since } \left( \frac{\text{mult}_w u + \text{mult}_w v}{\text{mult}_w u} \right) \neq 1). \text{ Since } \text{mult}_w u \text{ is positive, there must be at least one term in the CFL factorization of } u \text{ which is equal to } w. \text{ In other words, there is at least one } i \in \{1, 2, \ldots, p\} \text{ satisfying } a_i = w \text{ (since } (a_1, a_2, \ldots, a_p) \text{ is the CFL factorization of } u). \text{ Similarly, there is at least one } j \in \{1, 2, \ldots, q\} \text{ satisfying } b_j = w. \text{ These } i \text{ and } j \text{ satisfy } a_i = b_j \text{, which contradicts } a_i > b_j. \text{ This contradiction shows that our assumption was false, qed.}

\[234\text{Proof of (6.25):} \text{ Let } w \in \mathcal{L} \text{ be such that } w \neq u. \text{ Then, the number of terms in the list } (u) \text{ which are equal to } w \text{ is } 0. \text{ Since } (u) \text{ is the CFL factorization of } u, \text{ this rewrites as follows: The number of terms in the CFL factorization of } u \text{ which are equal to } w \text{ is } 0. \text{ In other words, } \text{mult}_w u = 0. \text{ This proves (6.25).} \]
Equivalently, a polynomial algebra can be defined as a $k$-algebra which has an algebraically independent (over $k$) generating set. Yet equivalently, a polynomial algebra can be defined as a $k$-algebra which is isomorphic to the symmetric algebra of a free $k$-module.

Keep in mind that when we say that a certain bialgebra $A$ is a polynomial algebra, we are making no statement about the coalgebra structure on $A$. The isomorphism from $A$ to the symmetric algebra of a free $k$-module need not be a coalgebra isomorphism, and the algebraically independent generating set of $A$ need not consist of primitives. Thus, showing that a bialgebra $A$ is a polynomial algebra does not trivialize the study of its bialgebraic structure.

**Remark 6.69.** Let $V$ be a $k$-module, and let $\mathfrak{A}$ be a totally ordered set. Let $b_a$ be an element of $V$ for every $a \in \mathfrak{A}$. Consider the shuffle algebra $Sh(V)$ (defined in Definition 1.73).

For every word $w \in \mathfrak{A}^*$ over the alphabet $\mathfrak{A}$, let us define an element $b_w$ of $Sh(V)$ by $b_w = b_{w_1} b_{w_2} \cdots b_{w_\ell}$, where $\ell$ is the length of $w$. (The multiplication used here is that of $T(V)$, not that of $Sh(V)$; the latter is denoted by $\shuffle$.)

Let $u \in \mathfrak{A}^*$ and $v \in \mathfrak{A}^*$ be two words over the alphabet $\mathfrak{A}$. Let $n = \ell(u)$ and $m = \ell(v)$. Then,

$$b_u \shuffle b_v = \sum_{\sigma \in Sh_{n,m}} b_{\sigma w}.$$  

**Exercise 6.70.** Prove Remark 6.69.

**Hint:** This follows from the definition of $\shuffle$.

We can now state Radford’s theorem [134, Theorem 3.1.1(e)]:

**Theorem 6.71.** Assume that $\mathbb{Q}$ is a subring of $k$. Let $V$ be a free $k$-module with a basis $(b_a)_{a \in \mathfrak{A}}$, where $\mathfrak{A}$ is a totally ordered set. Then, the shuffle algebra $Sh(V)$ (defined in Definition 1.73) is a polynomial $k$-algebra.

An algebraically independent generating set of $Sh(V)$ can be constructed as follows:

For every word $w \in \mathfrak{A}^*$ over the alphabet $\mathfrak{A}$, let us define an element $b_w$ of $Sh(V)$ by $b_w = b_{w_1} b_{w_2} \cdots b_{w_\ell}$, where $\ell$ is the length of $w$. (The multiplication used here is that of $T(V)$, not that of $Sh(V)$; the latter is denoted by $\shuffle$.) Let $L$ denote the set of all Lyndon words over the alphabet $\mathfrak{A}$. Then, $(b_w)_{w \in L}$ is an algebraically independent generating set of the $k$-algebra $Sh(V)$.

**Example 6.72.** For this example, let $\mathfrak{A}$ be the alphabet $\{1, 2, 3, \ldots\}$ with total order given by $1 < 2 < 3 < \cdots$, and assume that $\mathbb{Q}$ is a subring of $k$. Let $V$ be the free $k$-module with basis $(b_a)_{a \in \mathfrak{A}}$. We use the notations of Theorem 6.71. Then, Theorem 6.71 yields that $(b_w)_{w \in L}$ is an algebraically independent generating set of the $k$-algebra $Sh(V)$. Here are some examples of elements of $Sh(V)$ written as polynomials in this generating set:

- $b_{12} = b_{12}$ (the word 12 itself is Lyndon);
- $b_{21} = b_{1} \shuffle b_{2} - b_{12}$;
- $b_{11} = \frac{1}{2} b_{1} \shuffle b_{1}$;
- $b_{123} = b_{123}$ (the word 123 itself is Lyndon);
- $b_{132} = b_{132}$ (the word 132 itself is Lyndon);
- $b_{213} = b_{2} \shuffle b_{1} - b_{12} - b_{132}$;
- $b_{231} = b_{2} \shuffle b_{1} - b_{2} \shuffle b_{13} + b_{132}$;
- $b_{312} = b_{3} \shuffle b_{12} - b_{13} - b_{132}$;
- $b_{321} = b_{1} \shuffle b_{2} \shuffle b_{1} - b_{2} \shuffle b_{3} \shuffle b_{1} - b_{312} + b_{123}$;
- $b_{112} = b_{112}$ (the word 112 itself is Lyndon);
- $b_{121} = b_{12} \shuffle b_{1} - 2 b_{112}$;
- $b_{1212} = \frac{1}{2} b_{12} \shuffle b_{12} - 2 b_{1122}$;
- $b_{4321} = b_{1} \shuffle b_{2} \shuffle b_{1} \shuffle b_{1} - b_{1} \shuffle b_{2} \shuffle b_{13} - b_{1321} - b_{1231} - b_{1232} + b_{1234} + b_{12} \shuffle b_{3} \shuffle b_{1} + b_{12} \shuffle b_{3} \shuffle b_{1} - b_{1324}.$
Note that Theorem 6.71 cannot survive without the condition that $\mathbb{Q}$ be a subring of $\mathbb{k}$. For instance, for any $v \in V$, we have $\mathbb{V} \cdot v = 2 \mathbb{V} v$ in $\mathbb{Sh}(V)$, which vanishes if $2 = 0$ in $\mathbb{k}$; this stands in contrast to the fact that polynomial $\mathbb{k}$-algebras are integral domains when $\mathbb{k}$ itself is one. We will see that $\text{QSym}$ is less sensitive towards the base ring in this regard (although proving that $\text{QSym}$ is a polynomial algebra is much easier when $\mathbb{Q}$ is a subring of $\mathbb{k}$).

In our proof of Theorem 6.71 (but not only there), we will use part (a) of the following lemma\textsuperscript{236}, which makes proving that certain families indexed by Lyndon words generate certain $\mathbb{k}$-algebras more comfortable:

**Lemma 6.73.** Let $A$ be a commutative $\mathbb{k}$-algebra. Let $\mathfrak{A}$ be a totally ordered set. Let $\mathfrak{L}$ be the set of all Lyndon words over the alphabet $\mathfrak{A}$. Let $b_w$ be an element of $A$ for every $w \in \mathfrak{L}$. For every word $u \in \mathfrak{A}^*$, define an element $b_u$ of $A$ by $b_u = b_{a_1} b_{a_2} \cdots b_{a_p}$, where $(a_1, a_2, \ldots, a_p)$ is the CFL factorization of $u$.

(a) The family $(b_w)_{w \in \mathfrak{L}}$ is an algebraically independent generating set of the $\mathbb{k}$-algebra $A$ if and only if the family $(b_u)_{u \in \mathfrak{A}^*}$ is a basis of the $\mathbb{k}$-module $A$.

(b) The family $(b_w)_{w \in \mathfrak{L}}$ generates the $\mathbb{k}$-algebra $A$ if and only if the family $(b_u)_{u \in \mathfrak{A}^*}$ spans the $\mathbb{k}$-module $A$.

(c) Assume that the $\mathbb{k}$-algebra $A$ is graded. Let $\omega : \mathfrak{A} \rightarrow \{1, 2, 3, \ldots \}$ be any map such that for every $N \in \{1, 2, 3, \ldots \}$, the set $\omega^{-1}(N)$ is finite.

For every word $w \in \mathfrak{A}^*$, define an element $W_t(w) \in \mathbb{N}$ by $W_t(w) = \omega(w_1) + \omega(w_2) + \cdots + \omega(w_k)$, where $k$ is the length of $w$.

Assume that for every $w \in \mathfrak{L}$, the element $b_w$ of $A$ is homogeneous of degree $W_t(w)$.

Assume further that the $\mathbb{k}$-module $A$ has a basis $(g_u)_{u \in \mathfrak{A}^*}$ having the property that for every $u \in \mathfrak{A}^*$, the element $g_u$ of $A$ is homogeneous of degree $W_t(u)$.

Assume also that the family $(b_w)_{w \in \mathfrak{L}}$ generates the $\mathbb{k}$-algebra $A$. Then, this family $(b_w)_{w \in \mathfrak{L}}$ is an algebraically independent generating set of the $\mathbb{k}$-algebra $A$.

**Exercise 6.74.** Prove Lemma 6.73.

[Hint: For (a) and (b), notice that the $b_u$ are the “monomials” in the $b_w$. For (c), use Exercise 2.46(b) in every homogeneous component of $A$.]

The main workhorse of our proof of Theorem 6.71 will be the following consequence of Theorem 6.44(c):

**Proposition 6.75.** Let $V$ be a free $\mathbb{k}$-module with a basis $(b_a)_{a \in \mathfrak{A}}$, where $\mathfrak{A}$ is a totally ordered set.

For every word $w \in \mathfrak{A}^*$ over the alphabet $\mathfrak{A}$, let us define an element $b_w$ of $\mathbb{Sh}(V)$ by $b_w = b_{w_1} b_{w_2} \cdots b_{w_{\ell}}$, where $\ell$ is the length of $w$. (The multiplication used here is that of $T(V)$, not that of $\mathbb{Sh}(V)$; the latter is denoted by $\mathbb{W}$.)

For every word $u \in \mathfrak{A}^*$, define an element $b_u$ by $b_u = b_u \mathbb{W} b_{a_1} \mathbb{W} \cdots \mathbb{W} b_{a_p}$, where $(a_1, a_2, \ldots, a_p)$ is the CFL factorization of $u$.

If $\ell \in \mathbb{N}$ and if $x \in \mathfrak{A}^\ell$ is a word, then there is a family $(\eta_{x,y})_{y \in \mathfrak{A}^\ell} \in \mathbb{N}^{\mathfrak{A}^\ell}$ of elements of $\mathbb{N}$ satisfying

$$
\eta_{x,y} b_y = \sum_{y \leq x} \eta_{x,y} b_y
$$

and $\eta_{x,x} \neq 0$ (in $\mathbb{N}$).

Before we prove this, let us show a very simple lemma:

\textsuperscript{235}A pattern emerges in the formulas for $b_{21}$, $b_{321}$ and $b_{1321}$: for every $n \in \mathbb{N}$, we have

$$
b_{(n,n-1,\ldots,1)} = \sum_{\alpha \in \text{Comp}_n} (-1)^{n-\ell(\alpha)} b_{d_{1}(\alpha)} \mathbb{W} b_{d_{2}(\alpha)} \mathbb{W} \cdots \mathbb{W} b_{d_{\ell}(\alpha)}.
$$

where $(d_1(\alpha)) \cdot (d_2(\alpha)) \cdots (d_{\ell}(\alpha))$ is the factorization of the word $(1, 2, \ldots, n)$ into factors of length $\alpha_1$, $\alpha_2$, $\ldots$, $\alpha_{\ell}$ (where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{\ell})$). This can be proved by an application of the principle of inclusion and exclusion (as it is easy to see that for any composition $\alpha$ of $n$, the product $b_{d_{1}(\alpha)} \mathbb{W} b_{d_{2}(\alpha)} \mathbb{W} \cdots \mathbb{W} b_{d_{\ell}(\alpha)}$ equals the sum of $b_\beta$ for all words $\beta$ of $\mathbb{A}_n$ satisfying $\text{Des}(\beta^{-1}) \subset D(\alpha)$).

\textsuperscript{236}And in a later proof, we will also use its part (c) (which is tailored for application to $\text{QSym}$).
Lemma 6.76. Let \( \mathfrak{A} \) be a totally ordered set. Let \( n \in \mathbb{N} \) and \( m \in \mathbb{N} \). Let \( \sigma \in \mathrm{Sh}_{n,m} \).

(a) If \( u, v \) and \( v' \) are words satisfying \( \ell(u) = n, \ell(v) = m, \ell(v') = m \) and \( v' < v \), then \( u \sqcup v' < u \sqcup v \).

(b) If \( u, u' \) and \( v \) are three words satisfying \( \ell(u) = n, \ell(u') = n, \ell(v) = m \) and \( u' < u \), then \( u' \sqcup v < u \sqcup v \).

(c) If \( u, v \) and \( v' \) are three words satisfying \( \ell(u) = n, \ell(v) = m, \ell(v') = m \) and \( v' \leq v \), then \( u \sqcup v' \leq u \sqcup v \).

Exercise 6.77. Prove Lemma 6.76.

Exercise 6.78. Prove Proposition 6.75.

[Hint: Proceed by induction over \( \ell \). In the induction step, apply Theorem 6.44(c)\(^{237}\) to \( u = a_1 \) and \( v = a_2 a_3 \cdots a_p \), where \((a_1, a_2, \ldots, a_p)\) is the CFL factorization of \( x \). Use Lemma 6.76 to get rid of smaller terms.]


[Hint: According to Lemma 6.73(a), it suffices to show that the family \((b_w)_{w \in \mathfrak{A}}\) defined in Proposition 6.75 is a basis of the \( k \)-module \( \mathrm{Sh}(V) \). When \( \mathfrak{A} \) is finite, the latter can be proven by triangularity using Proposition 6.75. Reduce the general case to that of finite \( \mathfrak{A} \).]

6.4. Polynomial freeness of \( \mathrm{QSym} \): statement and easy parts.

Definition 6.80. For the rest of Chapter 6, we introduce the following notations: We let \( \mathfrak{A} \) be the totally ordered set \( \{1, 2, 3, \ldots\} \) with its natural order (that is, \( 1 < 2 < 3 < \cdots \)). Thus, the words over \( \mathfrak{A} \) are precisely the compositions. That is, \( \mathfrak{A}^* = \text{Comp} \). We let \( \mathfrak{L} \) denote the set of all Lyndon words over \( \mathfrak{A} \). These Lyndon words are also called Lyndon compositions.

A natural question is how many Lyndon compositions of a given size exist. While we will not use the answer, we nevertheless record it:

Exercise 6.81. Show that the number of Lyndon compositions of size \( n \) equals

\[
\frac{1}{n} \sum_{d|n} \mu(d) \left( 2^{n/d} - 1 \right) = \frac{1}{n} \sum_{d|n} \mu(d) 2^{n/d} - \delta_{n,1}
\]

for every positive integer \( n \) (where "\( \sum \)" means a sum over all positive divisors of \( n \), and where \( \mu \) is the number-theoretic Möbius function).

[Hint: One solution is similar to the solution of Exercise 6.29 using CFL factorization. Another proceeds by defining a bijection between Lyndon compositions and Lyndon words over a two-letter alphabet \( \{0, 1\} \) (with \( 0 < 1 \)) which are \( \neq 1 \). \(^{238}\)]

Let us now state Hazewinkel’s result ([65, Theorem 8.1], [69, §6.7]) which is the main goal of Chapter 6:

Theorem 6.82. The \( k \)-algebra \( \mathrm{QSym} \) is a polynomial algebra. It is isomorphic, as a graded \( k \)-algebra, to the \( k \)-algebra \( k[x_w \mid w \in \mathfrak{L}] \). Here, the grading on \( k[x_w \mid w \in \mathfrak{L}] \) is defined by setting \( \deg(x_w) = \sum_{i=1}^{\ell(w)} w_i \) for every \( w \in \mathfrak{L} \).

We shall prove Theorem 6.82 in the next section (Section 6.5). But the particular case of Theorem 6.82 when \( \mathbb{Q} \) is a subring of \( k \) can be proven more easily; we state it as a proposition:

Proposition 6.83. Assume that \( \mathbb{Q} \) is a subring of \( k \). Then, Theorem 6.82 holds.

We will give two proofs of Proposition 6.83 in this Section 6.4; a third proof of Proposition 6.83 will immediately result from the proof of Theorem 6.82 in Section 6.5. (There is virtue in giving three different proofs, as they all construct different isomorphisms \( k[x_w \mid w \in \mathfrak{L}] \to \mathrm{QSym} \).

Our first proof – originating in Malvenuto’s [112, Corollaire 4.20] – can be given right away; it relies on Exercise 5.44:

\(^{237}\)Or Theorem 6.44(e), if you prefer.

\(^{238}\)This bijection is obtained by restricting the bijection

\[
\text{Comp} \to \{w \in \{0, 1\}^* \mid w \text{ does not start with } 1\},
\]

\((a_1, a_2, \ldots, a_\ell) \to 01^{a_1-1}01^{a_2-1} \cdots 01^{a_\ell-1}
\]

(where \(01^k\) is to be read as \(0(1^k)\), not as \((01)^k\)) to the set of Lyndon compositions. The idea behind this bijection is well-known in the Grothendieck-Teichmüller community: see, e.g., [70, §3.1] (and see [58, Note 5.16] for a different appearance of this idea).
First proof of Proposition 6.83. Let $V$ be the free $k$-module with basis $(b_n)_{n \in \{1, 2, 3, \ldots\}}$. Endow the $k$-module $V$ with a grading by assigning to each basis vector $b_n$ the degree $n$. Exercise 5.44(k) shows that QSym is isomorphic to the shuffle algebra $Sh(V)$ (defined as in Proposition 1.73) as Hopf algebras. By being a bit more careful, we can obtain the slightly stronger result that QSym is isomorphic to the shuffle algebra $Sh(V)$ as graded Hopf algebras\(^{239}\). In particular, QSym $\cong Sh(V)$ as graded $k$-algebras.

Theorem 6.71 (applied to $b_n = b_n$) yields that the shuffle algebra $Sh(V)$ is a polynomial $k$-algebra, and that an algebraically independent generating set of $Sh(V)$ can be constructed as follows:

For every word $w \in \mathfrak{A}^*$ over the alphabet $\mathfrak{A}$, let us define an element $b_w$ of $Sh(V)$ by $b_w = b_{w_1}b_{w_2}\cdots b_{w_{\ell(w)}}$, where $\ell$ is the length of $w$. (The multiplication used here is that of $T(V)$, not that of $Sh(V)$; the latter is denoted by $\oplus$.) Then, $(b_w)_{w \in \mathfrak{A}^*}$ is an algebraically independent generating set of the $k$-algebra $Sh(V)$.

For every $w \in \mathfrak{A}^*$, we have $b_w = b_{w_1}b_{w_2}\cdots b_{w_{\ell(w)}}$ (by the definition of $b_w$). For every $w \in \mathfrak{A}^*$, the element $b_{w} = b_{w_1}b_{w_2}\cdots b_{w_{\ell(w)}}$ of $Sh(V)$ is homogeneous of degree $\sum_{i=1}^{\ell(w)} \deg(b_{w_i}) = \sum_{i=1}^{\ell(w)} w_i$.

Now, define a grading on the $k$-algebra $k[x_w \mid w \in \mathcal{L}]$ by setting $\deg(x_w) = \sum_{i=1}^{\ell(w)} w_i$ for every $w \in \mathcal{L}$. By the universal property of the polynomial algebra $k[x_w \mid w \in \mathcal{L}]$, we can define a $k$-algebra homomorphism $\Phi : k[x_w \mid w \in \mathcal{L}] \rightarrow Sh(V)$ by setting

$$\Phi(x_w) = b_w$$

for every $w \in \mathcal{L}$.

This homomorphism $\Phi$ is a $k$-algebra isomorphism (since $(b_w)_{w \in \mathcal{L}}$ is an algebraically independent generating set of the $k$-algebra $Sh(V)$) and is graded (because for every $w \in \mathcal{L}$, the element $b_w$ of $Sh(V)$ is homogeneous of degree $\sum_{i=1}^{\ell(w)} w_i = \deg(x_w)$). Thus, $\Phi$ is an isomorphism of graded $k$-algebras. Hence, $Sh(V) \cong k[x_w \mid w \in \mathcal{L}]$ as graded $k$-algebras. Altogether, QSym $\cong Sh(V)$ as graded $k$-algebras. Thus, QSym is a polynomial $k$-algebra. This proves Theorem 6.82 under the assumption that $Q$ be a subring of $k$. In other words, this proves Proposition 6.83.\( \square \)

Our second proof of Proposition 6.83 comes from Hazewinkel/Gubareni/Kirichenko \[69\] (where Proposition 6.83 appears as \[69, \text{Theorem 6.5.13}\]). This proof will construct an explicit algebraically independent family generating the $k$-algebra QSym. \(^{240}\) The generating set will be very unsophisticated: it will be $(M_\alpha)_{\alpha \in \mathfrak{A}^*}$, where $\mathfrak{A}$ and $\mathcal{L}$ are as in Theorem 6.82. Here, we are using the fact that words over the alphabet $\{1, 2, 3, \ldots\}$ are the same thing as compositions, so, in particular, a monomial quasisymmetric function $M_\alpha$ is defined for every such word $\alpha$.

It takes a bit of work to show that this family indeed fits the bill. We begin with a corollary of Proposition 5.3 that is essentially obtained by throwing away all non-bijective maps $f$:

**Proposition 6.84.** Let $\alpha \in \mathfrak{A}^*$ and $\beta \in \mathfrak{A}^*$. Then, $M_\alpha M_\beta = \sum_{\gamma \in \mathcal{O} \cup \beta} M_\gamma + (\text{a sum of terms of the form } M_\delta \text{ with } \delta \in \mathfrak{A}^* \text{ satisfying } \ell(\delta) < \ell(\alpha) + \ell(\beta))$.\(^{241}\)

**Exercise 6.85.** Prove Proposition 6.84.

[Hint: Recall what was said about the $p = \ell + m$ case in Example 5.4.]

**Corollary 6.86.** Let $\alpha \in \mathfrak{A}^*$ and $\beta \in \mathfrak{A}^*$. Then, $M_\alpha M_\beta$ is a sum of terms of the form $M_\delta$ with $\delta \in \mathfrak{A}^*$ satisfying $\ell(\delta) \leq \ell(\alpha) + \ell(\beta)$.\(^{242}\)

\(^{239}\)Proof. In the solution of Exercise 5.44(k), we have shown that QSym $\cong T(V')^n$ as graded Hopf algebras. But Remark 1.75(b) shows that the Hopf algebra $T(V')^n$ is naturally isomorphic to the shuffle algebra $Sh(V^n)$ as Hopf algebras; it is easy to see that the natural isomorphism $T(V')^n \rightarrow Sh(V^n)$ is graded (because it is the direct sum of the isomorphisms $(T(V')^n)^{g^n} \rightarrow (Sh(V))^g$ over all $n \in \mathbb{N}$, and each of these isomorphisms is graded). Hence, $T(V')^n \cong Sh(V^n)$ as graded Hopf algebras. But $V^n \cong V$ as graded $k$-modules (since $V$ is of finite type), and thus $Sh(V^n) \cong Sh(V)$ as graded Hopf algebras. Altogether, we obtain QSym $\cong T(V')^n \cong Sh(V^n) \cong Sh(V)$ as graded Hopf algebras, qed.

\(^{240}\)We could, of course, obtain such a family from our above proof as well (this is done by Malvenuto in \[112, \text{Corollaire 4.20}\]), but it won’t be a very simple one.

\(^{241}\)The sum $\sum_{\gamma \in \mathcal{O} \cup \beta} M_\gamma$ ranges over the multiset $\alpha \cup \beta$; if an element appears several times in $\alpha \cup \beta$, then it has accordingly many addends corresponding to it.
Exercise 6.87. Prove Corollary 6.86.

We now define a partial order on the compositions of a given nonnegative integer:

**Definition 6.88.** Let \( n \in \mathbb{N} \). We define a binary relation \( \leq \) on the set \( \text{Comp}_n \) as follows: For two compositions \( \alpha \) and \( \beta \) in \( \text{Comp}_n \), we set \( \alpha \leq \beta \) if and only if

\[
\text{either } \ell(\alpha) < \ell(\beta) \text{ or } (\ell(\alpha) = \ell(\beta) \text{ and } \alpha \leq \beta \text{ in lexicographic order}).
\]

This binary relation \( \leq \) is the smaller-or-equal relation of a total order on \( \text{Comp}_n \); we refer to said total order as the \textit{will-order} on \( \text{Comp}_n \), and we denote by \( < \) the smaller relation of this total order.

Notice that if \( \alpha \) and \( \beta \) are two compositions satisfying \( \ell(\alpha) = \ell(\beta) \), then \( \alpha \leq \beta \) in lexicographic order if and only if \( \alpha \leq \beta \) with respect to the relation \( \leq \) defined in Definition 6.1.

A remark about the name “will-order” is in order. We have taken this notation from [65, Definition 6.7.14], where it is used for an extension of this order to the whole set \( \text{Comp} \). We will never use this extension, as we will only ever compare two compositions of the same integer.\(^{242}\)

We now state a fact which is similar (and plays a similar role) to Proposition 6.75:

**Proposition 6.89.** For every composition \( u \in \text{Comp} = \mathbb{A}^* \), define an element \( M_u \in \text{QSym} \) by \( M_u = M_{a_1}M_{a_2} \cdots M_{a_p} \), where \((a_1, a_2, \ldots, a_p)\) is the CFL factorization of the word \( u \).

If \( n \in \mathbb{N} \) and if \( x \in \text{Comp}_n \), then there is a family \( \eta_{x,y} \in \mathbb{N}^{\text{Comp}_n} \) of elements of \( \mathbb{N} \) satisfying

\[
M_x = \sum_{y \in \text{Comp}_n: \ y \leq wll x \atop y \leq u} \eta_{x,y}M_y
\]

and \( \eta_{x,x} \neq 0 \) (in \( \mathbb{N} \)).

Before we prove it, let us show the following lemma:

**Lemma 6.90.** Let \( n \in \mathbb{N} \) and \( m \in \mathbb{N} \). Let \( u \in \text{Comp}_n \) and \( v \in \text{Comp}_m \). Let \( z \) be the lexicographically highest element of the multiset \( u \cup v \).

(a) We have \( z \in \text{Comp}_{n+m} \).

(b) There exists a positive integer \( h \) such that

\[
M_uM_v = hM_z + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < z \right).
\]

(c) Let \( v' \in \text{Comp}_m \) be such that \( v' < v \). Then,

\[
M_uM_{v'} = \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < z \right).
\]

**Exercise 6.91.** Prove Lemma 6.90.

[Hint: For (b), set \( h \) to be the multiplicity with which the word \( z \) appears in the multiset \( u \cup v \), then use Proposition 6.84 and notice that \( M_uM_v \) is homogeneous of degree \( n+m \). For (c), use (b) for \( v' \) instead of \( v \) and notice that Lemma 6.76(a) shows that the lexicographically highest element of the multiset \( u \cup v' \) is \( < z \).]

**Exercise 6.92.** Prove Proposition 6.89.

[Hint: Proceed by strong induction over \( n \). In the induction step, let \((a_1, a_2, \ldots, a_p)\) be the CFL factorization of \( x \), and set \( u = a_1 \) and \( v = a_2a_3 \cdots a_p \); then apply Proposition 6.89 to \( v \) instead of \( x \), and multiply the resulting equality \( M_x = \sum_{y \in \text{Comp}_{n+m}} \eta_{x,y}M_y \) with \( M_u \) to obtain an expression for \( M_uM_v = M_x \). Use

\[^{242}\text{In [65, Definition 6.7.14], the name “will-order” is introduced as an abbreviation for “weight first, then length, then lexicographic” (in the sense that two compositions are first compared by their weights, then, if the weights are equal, by their lengths, and finally, if the lengths are also equal, by the lexicographic order). For us, the alternative explanation “word length, then lexicographic” serves just as well.\]
Lemma 6.90 to show that this expression has the form \( \sum_{y \in \text{Comp}_n} \eta_{x,y} M_y \) with \( \eta_{x,x} \neq 0 \); here it helps to remember that the lexicographically highest element of the multiset \( u \cup v \) is \( uv = x \) (by Theorem 6.44(c)).

We are almost ready to give our second proof of Proposition 6.83; our last step is the following proposition:

**Proposition 6.93.** Assume that \( \mathbb{Q} \) is a subring of \( k \). Then, \((M_w)_{w \in \mathbb{L}}\) is an algebraically independent generating set of the \( k \)-algebra \( \text{QSym} \).

**Exercise 6.94.** Prove Proposition 6.93.

[Hint: Define \( M_u \) for every \( u \in \text{Comp} \) as in Proposition 6.89. Conclude from Proposition 6.89 that, for every \( n \in \mathbb{N} \), the family \((M_u)_{u \in \text{Comp}_n}\) expands upper-triangularly (with respect to the total order \( \leq \) on \( \text{Comp}_n \)) with respect to the basis \((M_u)_{u \in \text{Comp}_n}\) of \( \text{QSym}_n \), and the diagonal entries of this expansion are \( \neq 0 \). Conclude that this family \((M_u)_{u \in \text{Comp}_n}\) is a basis of \( \text{QSym}_n \) itself, and so the whole family \((M_u)_{u \in \text{Comp}}\) is a basis of \( \text{QSym} \). Conclude using Lemma 6.73(a).]

**Second proof of Proposition 6.83.** Proposition 6.93 yields that \((M_w)_{w \in \mathbb{L}}\) is an algebraically independent generating set of the \( k \)-algebra \( \text{QSym} \).

Define a grading on the \( k \)-algebra \( k[x_w \mid w \in \mathbb{L}] \) by setting \( \deg(x_w) = \sum_{i=1}^{\ell(w)} w_i \) for every \( w \in \mathbb{L} \). By the universal property of the polynomial algebra \( k[x_w \mid w \in \mathbb{L}] \), we can define a \( k \)-algebra homomorphism \( \Phi : k[x_w \mid w \in \mathbb{L}] \to \text{QSym} \) by setting

\[
\Phi(x_w) = M_w \quad \text{for every } w \in \mathbb{L}.
\]

This homomorphism \( \Phi \) is a \( k \)-algebra isomorphism (since \((M_w)_{w \in \mathbb{L}}\) is an algebraically independent generating set of the \( k \)-algebra \( \text{QSym} \)) and is graded (because for every \( w \in \mathbb{L} \), the element \( M_w \) of \( \text{QSym} \) is homogeneous of degree \( |w| = \sum_{i=1}^{\ell(w)} w_i = \deg(x_w) \)). Thus, \( \Phi \) is an isomorphism of graded \( k \)-algebras. Hence, \( \text{QSym} \cong k[x_w \mid w \in \mathbb{L}] \) as graded \( k \)-algebras. In particular, this shows that \( \text{QSym} \) is a polynomial algebra. This proves Theorem 6.82 under the assumption that \( \mathbb{Q} \) be a subring of \( k \). Proposition 6.83 is thus proven again. \( \square \)

### 6.5. Polynomial freeness of \( \text{QSym} \): the general case

We now will prepare for proving Theorem 6.82 without any assumptions on \( k \). In our proof, we follow [65] and [69, §6.7], but without using the language of plethysm and Frobenius maps. We start with the following definition:

**Definition 6.95.** Let \( \alpha \) be a composition. Write \( \alpha \) in the form \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \) with \( \ell = \ell(\alpha) \).

(a) Let \( \text{SIS}(\ell) \) denote the set of all strictly increasing \( \ell \)-tuples \((i_1, i_2, \ldots, i_{\ell}) \) of positive integers.\(^{243}\) For every \( \ell \)-tuple \( \mathbf{i} = (i_1, i_2, \ldots, i_{\ell}) \in \text{SIS}(\ell) \), we denote the monomial \( x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_{\ell}^{\alpha_{\ell}} \) by \( x_{\mathbf{i}}^{\alpha} \). This \( x_{\mathbf{i}}^{\alpha} \) is a monomial of degree \( \alpha_1 + \alpha_2 + \cdots + \alpha_\ell = |\alpha| \).

\[ M_\alpha = \sum_{\mathbf{i} \in \text{SIS}(\ell)} x_{\mathbf{i}}^{\alpha}. \quad (6.26)\]

\(^{243}\)“Strictly increasing” means that \( i_1 < i_2 < \cdots < i_{\ell} \) here. Of course, the elements of \( \text{SIS}(\ell) \) are in 1-1 correspondence with \( \ell \)-element subsets of \( \{1, 2, 3, \ldots\} \).

\(^{244}\)Proof of (6.26): By the definition of \( M_\alpha \), we have

\[
M_\alpha = \sum_{\substack{i_1 < i_2 < \cdots < i_{\ell} \in \{1, 2, 3, \ldots\} \\text{subject to } \sum (i_1, i_2, \ldots, i_{\ell}) \in \text{SIS}(\ell)}} x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_{\ell}}^{\alpha_{\ell}} = \sum_{(i_1, i_2, \ldots, i_{\ell}) \in \text{SIS}(\ell)} x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_{\ell}}^{\alpha_{\ell}} = \sum_{(i_1, i_2, \ldots, i_{\ell}) \in \text{SIS}(\ell)} x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_{\ell}}^{\alpha_{\ell}} \quad (\text{by the definition of } x_{\mathbf{i}}^{\alpha})
\]

\[\text{qed.}\]
(b) Consider the ring \( k[[x]] \) endowed with the coefficientwise topology\(^{245}\). The family \( (x^\alpha_{i})_{i \in SIS(t)} \) of elements of \( k[[x]] \) is power-summable\(^{246}\). Hence, for every \( f \in \Lambda \), there is a well-defined power series \( f \left( (x^\alpha_{i})_{i \in SIS(t)} \right) \in k[[x]] \) obtained by “evaluating” \( f \) at \( (x^\alpha_{i})_{i \in SIS(t)} \)\(^{247}\). In particular, for every \( s \in \mathbb{Z} \), we can evaluate the symmetric function \( e_s \in \Lambda \)\(^{248}\) at \( (x^\alpha_{i})_{i \in SIS(t)} \). The resulting power series

\[
M_s^{(s)} = e_s \left( (x^\alpha_{i})_{i \in SIS(t)} \right)
\]

The power series \( M_s^{(s)} \) are the power series \( e_s (\alpha) \) in [69]. We will shortly (in Corollary 6.102(a)) see that \( M_s^{(s)} \in \text{QSym} \) (although this is also easy to prove by inspection). Here are some examples of \( M_s^{(s)} \):

\(^{245}\)This topology is defined as follows:

We endow the ring \( k \) with the discrete topology. Then, we can regard the \( k \)-module \( k[[x]] \) as a direct product of infinitely many copies of \( k \) (by identifying every power series in \( k[[x]] \) with the family of its coefficients). Hence, the product topology is a well-defined topology on \( k[[x]] \); this topology is denoted as the coefficientwise topology. A sequence \((a_n)_{n \in \mathbb{N}}\) of power series converges to a power series \( a \) with respect to this topology if and only if for every monomial \( m \), all sufficiently high \( n \in \mathbb{N} \) satisfy

\[
(a_n) \rightarrow a \text{ (the coefficient of } m \text{ in } a_n) = (\text{the coefficient of } m \text{ in } a).
\]

Note that this is not the topology obtained by taking the completion of \( k[[x_1, x_2, x_3, \ldots]] \) with respect to the standard grading (in which all \( x_i \) have degree 1). (The latter completion is actually a smaller ring than \( k[[x]] \).)

\(^{246}\)Let us define what “power-summable” means for us:

A family \((n_i)_{i \in I} \in \mathbb{N}^{I}\) (where \( I \) is some set) is said to be \textit{finitely supported} if all but finitely many \( i \in I \) satisfy \( n_i = 0 \). If \((n_i)_{i \in I} \in \mathbb{N}^{I}\) is a finitely supported family, then \( \sum_{i \in I} n_i \) is a well-defined element of \( \mathbb{N} \). If \( N \in \mathbb{N} \), then a family \((n_i)_{i \in I} \in \mathbb{N}^{I}\) will be called \( (\leq N) \)-\textit{supported} if it is finitely supported and satisfies \( \sum_{i \in I} n_i \leq N \).

We say that a family \((s_i)_{i \in I} \in \mathbb{R}^{I}\) of elements of a topological commutative \( k \)-algebra \( R \) is \textit{power-summable} if it satisfies the following property: For every \( N \in \mathbb{N} \), the sum

\[
\sum_{(n_i)_{i \in I} \in (\leq N)\text{-supported}} \alpha_{(n_i)_{i \in I}} \prod_{i \in I} s_i^{n_i}
\]

converges in the topology on \( R \) for every choice of scalars \( \alpha_{(n_i)_{i \in I}} \in k \) corresponding to all \((\leq N)\)-supported \((n_i)_{i \in I} \in \mathbb{N}^{I}\). In our specific case, we consider \( k[[x]] \) as a topological commutative \( k \)-algebra, where the topology is the coefficientwise topology. The fact that the family \((x^\alpha_{i})_{i \in SIS(t)} \) is power-summable then can be proven as follows:

- If \( \alpha \neq \emptyset \), then this fact follows from the (easily-verified) observation that every given monomial in the variables \( x_1, x_2, x_3, \ldots \) can be written as a product of monomials of the form \( x^\alpha_{i} \) (with \( i \in SIS(t) \)) in only finitely many ways.
- If \( \alpha = \emptyset \), then this fact follows by noticing that \((x^\alpha_{i})_{i \in SIS(t)} \) is a finite family (indeed, \( SIS(t) = SIS(0) = \{()\} \)), and every finite family is power-summable.

\(^{247}\)Here is how this power series \( f \left( (x^\alpha_{i})_{i \in SIS(t)} \right) \) is formally defined:

Let \( R \) be any topological commutative \( k \)-algebra, and let \((s_i)_{i \in I} \in \mathbb{R}^{I}\) be any power-summable family of elements of \( R \). Assume that the indexing set \( I \) is countably infinite, and fix a bijection \( j : \{1, 2, 3, \ldots\} \rightarrow I \). Let \( g \in R(x) \) be arbitrary. Then, we can substitute \( s_{j(1)}, s_{j(2)}, s_{j(3)}, \ldots \) for the variables \( x_1, x_2, x_3, \ldots \) in \( g \), thus obtaining an infinite sum which converges in \( R \) (in fact, its convergence follows from the fact that the family \((s_i)_{i \in I} \in \mathbb{R}^{I}\) is power-summable). The value of this sum will be denoted by \( g ((s_i)_{i \in I}) \). In general, this value depends on the choice of the bijection \( j \), so the notation \( g ((s_i)_{i \in I}) \) is unambiguous only if this bijection \( j \) is chosen once and for all. However, when \( g \in \Lambda \), one can easily see that the choice of \( j \) has no effect on \( g ((s_i)_{i \in I}) \).

We can still define \( g ((s_i)_{i \in I}) \) when the set \( I \) is finite instead of being countably infinite. In this case, we only need to modify our above definition as follows: Instead of fixing a bijection \( j : \{1, 2, 3, \ldots\} \rightarrow I \), we now fix a bijection \( j : \{1, 2, \ldots, |I|\} \rightarrow I \), and instead of substituting \( s_{j(1)}, s_{j(2)}, s_{j(3)}, \ldots \) for the variables \( x_1, x_2, x_3, \ldots \) in \( g \), we now substitute \( s_{j(1)}, s_{j(2)}, \ldots, s_{j(|I|)} \) for the variables \( x_1, x_2, x_3, \ldots \) in \( g \). Again, the same observations hold as before: \( g ((s_i)_{i \in I}) \) is independent of \( j \) if \( g \in \Lambda \).

Hence, \( g ((s_i)_{i \in I}) \) is well-defined for every \( g \in R(x) \), every countable (i.e., finite or countably infinite) set \( I \), every topological commutative \( k \)-algebra \( R \) and every power-summable family \((s_i)_{i \in I} \in \mathbb{R}^{I}\) of elements of \( R \), as long as a bijection \( j \) is chosen. In particular, we can apply this to \( g = f, I = SIS(t), R = k[[x]] \) and \((s_i)_{i \in I} = (x^\alpha_{i})_{i \in SIS(t)} \), meaning \( j \) to be the bijection which sends every positive integer \( k \) to the \( k \)-th smallest element of \( SIS(t) \) in the lexicographic order. (Of course, since \( f \in \Lambda \), the choice of \( j \) is irrelevant.)

\(^{248}\)Recall that \( e_0 = 1 \), and that \( e_s = 0 \) for \( s < 0 \).
Example 6.96. If \( \alpha \) is a composition and \( \ell \) denotes its length \( \ell (\alpha) \), then

\[
M^{(0)}_\alpha = e_0 \left( \left( x_i^\alpha \right)_{i \in \text{SIS}(\ell)} \right) = 1 \left( \left( x_i^\alpha \right)_{i \in \text{SIS}(\ell)} \right) = 1
\]

and

\[
M^{(1)}_\alpha = e_1 \left( \left( x_i^\alpha \right)_{i \in \text{SIS}(\ell)} \right) = \sum_{i \in \text{SIS}(\ell)} x_i^{\alpha} = M_\alpha \quad \text{(by (6.26))}
\]

and

\[
M^{(2)}_\alpha = e_2 \left( \left( x_i^\alpha \right)_{i \in \text{SIS}(\ell)} \right) = \sum_{i < j} x_i^{\alpha} x_j^{\alpha}
\]

(where the notation “\( i < j \)” should be interpreted with respect to an arbitrary but fixed total order on the set \( \text{SIS}(\ell) \) – for example, the lexicographic order). Applying the last of these three equalities to \( \alpha = (2, 1) \), we obtain

\[
M^{(2)}_{(2, 1)} = \sum_{\substack{i < j \in \text{SIS}(2), \ j \in \text{SIS}(2); \ (i, i+2) \in \text{SIS}(2); \ (i+2, j) \in \text{SIS}(2); \ (i+2, j+2) \in \text{SIS}(2); \ (i+2, j+2) \in \text{SIS}(2); \ i+2 < j; \ i+2 < j+1}} x_{i+1}^{\alpha} x_j^{\alpha}
\]

\[
= \sum_{i+2 < j; i+2 < j+2} x_i^{\alpha} x_j^{\alpha}
\]

\[
= M_{(2, 1, 2, 1)} + M_{(2, 3, 1)} + 2M_{(2, 2, 1, 1)} + M_{(2, 2, 2)} = M_{(4, 1, 1)}
\]

(here, we have WLOG assumed that the order on \( \text{SIS}(2) \) is lexicographic)

\[
= M_{(2, 1, 2, 1)} + M_{(2, 3, 1)} + 2M_{(2, 2, 1, 1)} + M_{(2, 2, 2)} + M_{(4, 1, 1)}.
\]

Of course, every negative integer \( s \) satisfies

\[
M^{(s)}_\alpha = e_s \left( \left( x_i^\alpha \right)_{i \in \text{SIS}(\ell)} \right) = 0.
\]

There is a determinantal formula for the \( s!M^{(s)}_\alpha \) (and thus also for \( M^{(s)}_\alpha \) when \( s! \) is invertible in \( k \)), but in order to state it, we need to introduce one more notation:

Definition 6.97. Let \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \) be a composition, and let \( k \) be a positive integer. Then, \( \alpha \{ k \} \) will denote the composition \( (k\alpha_1, k\alpha_2, \ldots, k\alpha_\ell) \). Clearly, \( \ell (\alpha \{ k \}) = \ell (\alpha) \) and \( |\alpha \{ k \}| = k |\alpha| \).

Exercise 6.98. Let \( \alpha \) be a composition. Write the composition \( \alpha \) in the form \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \) with \( \ell = \ell (\alpha) \).

(a) Show that the \( s \)-th power-sum symmetric function \( p_s \in \Lambda \) satisfies

\[
p_s \left( \left( x_i^\alpha \right)_{i \in \text{SIS}(\ell)} \right) = M^{(s)}_{\alpha (s)}
\]

for every positive integer \( s \).

(b) Let us fix a total order on the set \( \text{SIS}(\ell) \) (for example, the lexicographic order). Show that the \( s \)-th elementary symmetric function \( e_s \in \Lambda \) satisfies

\[
M^{(s)}_{\alpha} = e_s \left( \left( x_i^\alpha \right)_{i \in \text{SIS}(\ell)} \right) = \sum_{(i_1, i_2, \ldots, i_s) \in \text{SIS}(\ell)^s} x_{i_1}^{\alpha} x_{i_2}^{\alpha} \cdots x_{i_s}^{\alpha}
\]

for every \( s \in \mathbb{N} \).

(c) Let \( s \in \mathbb{N} \), and let \( n \) be a positive integer. Let \( e_n^{(s)} \) be the symmetric function \( \sum_{i_1 < i_2 < \cdots < i_s} x_{i_1}^{n} x_{i_2}^{n} \cdots x_{i_s}^{n} \in \Lambda \). Then, show that

\[
M^{(s)}_{\alpha_n} = e_n^{(s)} \left( \left( x_i^\alpha \right)_{i \in \text{SIS}(\ell)} \right).
\]

\[\text{This is not completely obvious, but easy to check (see Exercise 6.98(b)).}\]
**Proposition 6.100.** Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ be a composition.

(a) Let $n \in \mathbb{N}$. Define a matrix $A_n^{(\alpha)} = \left( a_{i,j}^{(\alpha)} \right)_{i,j=1,2,\ldots,n}$ by

$$a_{i,j}^{(\alpha)} = \begin{cases} M_{\alpha(i-j+1)}, & \text{if } i \geq j; \\ i, & \text{if } i = j - 1; \\ 0, & \text{if } i < j - 1 \end{cases} \quad \text{for all } (i,j) \in \{1,2,\ldots,n\}^2.$$ 

This matrix $A_n^{(\alpha)}$ looks as follows:

$$A_n^{(\alpha)} = \begin{pmatrix} M_{\alpha(1)} & 1 & 0 & \cdots & 0 & 0 \\ M_{\alpha(2)} & M_{\alpha(1)} & 2 & \cdots & 0 & 0 \\ M_{\alpha(3)} & M_{\alpha(2)} & M_{\alpha(1)} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ M_{\alpha(n-1)} & M_{\alpha(n-2)} & M_{\alpha(n-3)} & \cdots & M_{\alpha(1)} & n - 1 \\ M_{\alpha(n)} & M_{\alpha(n-1)} & M_{\alpha(n-2)} & \cdots & M_{\alpha(2)} & M_{\alpha(1)} \end{pmatrix}.$$ 

Then, $\det \left( A_n^{(\alpha)} \right) = n! M_n^{(\alpha)}$.

(b) Let $n$ be a positive integer. Define a matrix $B_n^{(\alpha)} = \left( b_{i,j}^{(\alpha)} \right)_{i,j=1,2,\ldots,n}$ by

$$b_{i,j}^{(\alpha)} = \begin{cases} i M_{\alpha(i)}, & \text{if } j = 1; \\ M_{\alpha(i-j+1)}, & \text{if } j > 1 \end{cases} \quad \text{for all } (i,j) \in \{1,2,\ldots,n\}^2.$$ 

---

250There are two subtleties that need to be addressed:

- the fact that the definition of $f \left( (x_i^\alpha)_{i \in \text{SIS}(\ell)} \right)$ distinguishes between two cases depending on whether or not SIS(\ell) is finite;
- the fact that the total order on the set \{1,2,3,\ldots\} (which appears in the summation subscript in the equality $e_s = \sum_{(i_1,i_2,\ldots,i_\ell) \in \{1,2,3,\ldots\}^\ell} x_{i_1} x_{i_2} \cdots x_{i_\ell}$) has nothing to do with the total order on the set SIS(\ell) (which appears in the summation subscript in $\sum_{(i_1,i_2,\ldots,i_\ell) \in \text{SIS}(\ell)^\ell} x_{i_1}^\alpha x_{i_2}^\alpha \cdots x_{i_\ell}^\alpha$). For instance, the former total order is well-founded, whereas the latter may and may not be. So there is (generally) no bijection between \{1,2,3,\ldots\} and SIS(\ell) preserving these orders (even if SIS(\ell) is infinite). Fortunately, this does not matter much, because the total order is only being used to ensure that every product of $s$ pairwise distinct elements appears exactly once in the sum.
The matrix $B_{n}^{(\alpha)}$ looks as follows:

$$B_{n}^{(\alpha)} = \begin{pmatrix}
M_{\alpha}^{(1)} & M_{\alpha}^{(0)} & M_{\alpha}^{(-1)} & \cdots & M_{\alpha}^{(-n+3)} & M_{\alpha}^{(-n+2)} \\
2M_{\alpha}^{(2)} & M_{\alpha}^{(1)} & M_{\alpha}^{(0)} & \cdots & M_{\alpha}^{(-n+4)} & M_{\alpha}^{(-n+3)} \\
3M_{\alpha}^{(3)} & M_{\alpha}^{(2)} & M_{\alpha}^{(1)} & \cdots & M_{\alpha}^{(-n+5)} & M_{\alpha}^{(-n+4)} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
(n-1)M_{\alpha}^{(n-1)} & M_{\alpha}^{(n-2)} & M_{\alpha}^{(n-3)} & \cdots & M_{\alpha}^{(1)} & M_{\alpha}^{(0)} \\
nM_{\alpha}^{(n)} & M_{\alpha}^{(n-1)} & M_{\alpha}^{(n-2)} & \cdots & M_{\alpha}^{(2)} & M_{\alpha}^{(1)}
\end{pmatrix}$$

Then, $\det(B_{n}^{(\alpha)}) = M_{\alpha \{n\}}$.

**Exercise 6.101.** Prove Proposition 6.100.

[Hint: Substitute $(x_{i})_{i \in \text{SIS} (t)}$ for the variable set in Exercise 2.89, and recall Exercise 6.98(a).]

**Corollary 6.102.** Let $\alpha$ be a composition. Let $s \in \mathbb{Z}$.

(a) We have $M_{\alpha}^{(s)} \in \text{QSym}$.  
(b) We have $M_{\alpha}^{(s)} \in \text{QSym}_{s[\alpha]}$.

**Exercise 6.103.** Prove Corollary 6.102.

We make one further definition:

**Definition 6.104.** Let $\alpha$ be a nonempty composition. Then, we denote by $\gcd \alpha$ the greatest common divisor of the parts of $\alpha$. (For instance, $\gcd(8, 6, 4) = 2$.) We also define $\text{red} \alpha$ to be the composition

$$\left(\frac{\alpha_1}{\gcd \alpha}, \frac{\alpha_2}{\gcd \alpha}, \ldots, \frac{\alpha_\ell}{\gcd \alpha}\right),$$

where $\alpha$ is written in the form $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$.

We say that a nonempty composition $\alpha$ is *reduced* if $\gcd \alpha = 1$.

We define $\mathcal{RL}$ to be the set of all reduced Lyndon compositions. In other words, $\mathcal{RL} = \{w \in \mathcal{L} \mid w \text{ is reduced}\}$ (since $\mathcal{L}$ is the set of all Lyndon compositions).

Hazewinkel, in [69, proof of Thm. 6.7.5], denotes $\mathcal{RL}$ by $eLYN$, calling reduced Lyndon compositions “elementary Lyndon words”.

**Remark 6.105.** Let $\alpha$ be a nonempty composition.

(a) We have $\alpha = (\text{red} \alpha) \{\gcd \alpha\}$.

(b) The composition $\alpha$ is Lyndon if and only if the composition $\text{red} \alpha$ is Lyndon.

(c) The composition $\text{red} \alpha$ is reduced.

(d) If $\alpha$ is reduced, then $\text{red} \alpha = \alpha$.

(e) If $s \in \{1, 2, 3, \ldots\}$, then the composition $\alpha \{s\}$ is nonempty and satisfies $\text{red} (\alpha \{s\}) = \text{red} \alpha$ and $\gcd (\alpha \{s\}) = s \gcd \alpha$.

(f) We have $(\gcd \alpha) | \text{red} \alpha | |\alpha|$.


Our goal in this section is now to prove the following result of Hazewinkel:

**Theorem 6.107.** The family $\left(M_{w}^{(s)}\right)_{(w, s) \in \mathcal{RL} \times \{1, 2, 3, \ldots\}}$ is an algebraically independent generating set of the $k$-algebra QSym.
This will (almost) immediately yield Theorem 6.82.

Our first step towards proving Theorem 6.107 is the following observation:

**Lemma 6.108.** The family \( \left( M_{\alpha}^{(s)} \right)_{(w,s) \in \mathbb{N} \times \{1,2,3,\ldots\}} \) is a reindexing of the family \( \left( M_{\alpha}^{(\gcd(\alpha))} \right)_{\alpha \in \mathbb{Z}} \).


Next, we show a lemma:

**Lemma 6.110.** Let \( \alpha \) be a nonempty composition. Let \( s \in \mathbb{N} \). Then,

\[
(6.27) \quad s!M_{\alpha}^{(s)} - M_{\alpha}^s = \sum_{\beta \in \text{Comp}_{s}|\alpha|: \ell(\beta) \leq (s-1)\ell(\alpha)} k_{M_{\beta}}.
\]

(That is, \( s!M_{\alpha}^{(s)} - M_{\alpha}^s \) is a \( k \)-linear combination of terms of the form \( M_{\beta} \) with \( \beta \) ranging over the compositions of \( s|\alpha| \) satisfying \( \ell (\beta) \leq (s-1) \ell(\alpha) \).


[**Hint:** There are two approaches: One is to apply Proposition 6.100(a) and expand the determinant; the other is to argue which monomials can appear in \( s!M_{\alpha}^{(s)} - M_{\alpha}^s \).

We now return to studying products of monomial quasisymmetric functions:

**Lemma 6.112.** Let \( n \in \mathbb{N} \) and \( m \in \mathbb{N} \). Let \( u \in \text{Comp}_n \) and \( v \in \text{Comp}_m \). Let \( z \) be the lexicographically highest element of the multiset \( u \cup v \). Let \( h \) be the multiplicity with which the word \( z \) appears in the multiset \( u \cup v \). Then,

\[
M_u M_v = h M_z + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < z \right).
\]

**Proof of Lemma 6.112.** Lemma 6.112 was shown during the proof of Lemma 6.90(b). \( \square \)

**Corollary 6.113.** Let \( n \in \mathbb{N} \) and \( m \in \mathbb{N} \). Let \( u \in \text{Comp}_n \) and \( v \in \text{Comp}_m \). Regard \( u \) and \( v \) as words in \( \mathbb{N}^* \). Assume that \( u \) is a Lyndon word. Let \( (b_1,b_2,\ldots,b_q) \) be the CFL factorization of the word \( v \).

Assume that \( u \geq b_j \) for every \( j \in \{1,2,\ldots,q\} \). Let \( h = 1 + |\{j \in \{1,2,\ldots,q\} \mid b_j = u\}|. \)

Then,

\[
M_u M_v = h M_{uv} + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < uv \right).
\]


[**Hint:** Apply Lemma 6.112, and notice that \( uv \) is the lexicographically highest element of the multiset \( u \cup v \) (by Theorem 6.44(e)), and that \( h \) is the multiplicity with which this word \( uv \) appears in the multiset \( u \cup v \) (this is a rewriting of Theorem 6.44(e)).]

**Corollary 6.115.** Let \( k \in \mathbb{N} \) and \( s \in \mathbb{N} \). Let \( x \in \text{Comp}_k \) be such that \( x \) is a Lyndon word. Then:

(a) The lexicographically highest element of the multiset \( x \cup x^* \) is \( x^{s+1} \).

(b) We have

\[
M_x M_{x^*} = (s+1)M_{x^{s+1}} + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{(s+1)k} \text{ satisfying } w < x^{s+1} \right).
\]

(c) Let \( t \in \text{Comp}_{sk} \) be such that \( t < x^s \). Then,

\[
M_x M_t = \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{(s+1)k} \text{ satisfying } w < x^{s+1} \right).
\]

\[251] The following equality makes sense because we have \( z \in \text{Comp}_{n+m} \) (by Lemma 6.90(a)).

[Hint: Notice that \((x, x, \ldots, x)\) is the CFL factorization of the word \(x^s\). Now, part (a) of Corollary 6.115 follows from Theorem 6.44(c), part (b) follows from Corollary 6.113, and part (c) from Lemma 6.90(c) (using part (a)).]

Corollary 6.117. Let \(n \in \mathbb{N}\) and \(m \in \mathbb{N}\). Let \(u \in \text{Comp}_n\) and \(v \in \text{Comp}_m\). Regard \(u\) and \(v\) as words in \(A^*\). Let \((a_1, a_2, \ldots, a_p)\) be the CFL factorization of \(u\). Let \((b_1, b_2, \ldots, b_q)\) be the CFL factorization of the word \(v\). Assume that \(a_i > b_j\) for every \(i \in \{1, 2, \ldots, p\}\) and \(j \in \{1, 2, \ldots, q\}\). Then,

\[ M_u M_v = M_{uv} + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < uv \right). \]


[Hint: Combine Lemma 6.112 with the parts (c) and (d) of Theorem 6.44.]

Corollary 6.119. Let \(n \in \mathbb{N}\). Let \(u \in \text{Comp}_n\) be a nonempty composition. Regard \(u\) as a word in \(A^*\). Let \((a_1, a_2, \ldots, a_p)\) be the CFL factorization of \(u\). Let \(k \in \{1, 2, \ldots, p-1\}\) be such that \(a_k > a_{k+1}\). Let \(x\) be the word \(a_1 a_2 \cdots a_k\), and let \(y\) be the word \(a_{k+1} a_{k+2} \cdots a_p\). Then,

\[ M_u = M_x M_y - \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_n \text{ satisfying } w < u \right). \]

Exercise 6.120. Prove Corollary 6.119.

[Hint: Apply Corollary 6.117 to \(x, y, |x|, |y|, k, p - k, (a_1, a_2, \ldots, a_k)\) and \((a_k+1, a_k+2, \ldots, a_p)\) instead of \(u, v, n, m, p, q, (a_1, a_2, \ldots, a_p)\) and \((b_1, b_2, \ldots, b_q)\); then, notice that \(xy = u\) and \(|x| + |y| = n\).]

Corollary 6.121. Let \(k \in \mathbb{N}\). Let \(x \in \text{Comp}_k\) be a composition. Assume that \(x\) is a Lyndon word. Let \(s \in \mathbb{N}\). Then,

\[ M_x^s - s! M_{x^s} \in \sum_{w \in \text{Comp}_{sk}: \text{will} x^s} k M_w. \]

(Recall that \(x^s\) is defined to be the word \(x x \cdots x\).)


[Hint: Rewrite the claim of Corollary 6.121 in the form \(M_x^s \in s! M_{x^s} + \sum_{w \in \text{Comp}_{sk}: \text{will} x^s} k M_w\). This can be proven by induction over \(s\), where in the induction step we need the following two observations:

1. We have \(M_x M_{x^s} \in (s + 1) M_{x^{s+1}} + \sum_{w < x^s} \text{will} x^{s+1} k M_w\).

2. For every \(t \in \text{Comp}_{sk}\) satisfying \(t < x^s\), we have \(M_x M_t \in \sum_{w < x^s} \text{will} x^{s+1} k M_w\).

These two observations follow from parts (b) and (c) of Corollary 6.115.]

Corollary 6.123. Let \(k \in \mathbb{N}\). Let \(x \in \text{Comp}_k\) be a composition. Assume that \(x\) is a Lyndon word. Let \(s \in \mathbb{N}\). Then,

\[ M_x^{(s)} - M_{x^s} \in \sum_{w \in \text{Comp}_{sk}: \text{will} x^s} k M_w. \]

(Recall that \(x^s\) is defined to be the word \(x x \cdots x\).)

M = QSym. To this purpose, it is enough to prove that \( \sum_{\beta \in \text{Comp}_k; \ell(\beta) \leq (s-1)\ell(x)} M_\beta = \sum_{w \in \text{Comp}_k; \ell(w) \leq (s-1)\ell(x)} kM_w \) \( \sum_{w \in \text{Comp}_k; w < x} kM_w \).

252. Adding this to the claim of Corollary 6.121, obtain \( s!M_x^{(s)} - M_x \in \sum_{w \in \text{Comp}_k} kM_w \), that is, \( s! \left( M_x^{(s)} - M_x \right) \in \sum_{w \in \text{Comp}_k} kM_w \). It remains to get rid of the \( s! \) on the left hand side. Assume WLOG that \( k = \mathbb{Z} \), and argue that every \( f \in \text{QSym} \) satisfying \( s! \cdot f \in \sum_{w \in \text{Comp}_k} kM_w \) must itself lie in \( \sum_{w \in \text{Comp}_k} kM_w \).

We are now ready to prove Theorem 6.107:


**Hint:** Lemma 6.108 yields that the family \( \left( M_w^{(\gcd w)} \right)_{(w,s) \in \mathfrak{M}_E \times \{1,2,3,\ldots\}} \) is a reindexing of the family \( \left( M_{\text{red}_w}^{(\gcd w)} \right)_{w \in \mathfrak{L}} \). Hence, it is enough to prove that the family \( \left( M_{\text{red}_w}^{(\gcd w)} \right)_{w \in \mathfrak{L}} \) is an algebraically independent generating set of the \( k \)-algebra \( \text{QSym} \). The latter claim, in turn, will follow from Lemma 6.73(c)\(^{253}\) once it is proven that the family \( \left( M_{\text{red}_w}^{(\gcd w)} \right)_{w \in \mathfrak{L}} \) generates the \( k \)-algebra \( \text{QSym} \). So it remains to show that the family \( \left( M_{\text{red}_w}^{(\gcd w)} \right)_{w \in \mathfrak{L}} \) generates the \( k \)-algebra \( \text{QSym} \).

Let \( U \) denote the \( k \)-subalgebra of \( \text{QSym} \) generated by \( \left( M_{\text{red}_w}^{(\gcd w)} \right)_{w \in \mathfrak{L}} \). It then suffices to prove that \( U = \text{QSym} \). To this purpose, it is enough to prove that

\[
(6.28) \quad M_\beta \in U \quad \text{for every composition } \beta.
\]

For every reduced Lyndon composition \( \alpha \) and every \( j \in \{1,2,3,\ldots\} \), the quasisymmetric function \( M^{(j)}_\alpha \) is an element of the family \( \left( M_{\text{red}_w}^{(\gcd w)} \right)_{w \in \mathfrak{L}} \) and thus belongs to \( U \). Combine this with Exercise 6.98(d) to see that

\[
(6.29) \quad M^{(s)}_\beta \in U \quad \text{for every Lyndon composition } \beta \text{ and every } s \in \{1,2,3,\ldots\}
\]

(because every Lyndon composition \( \beta \) can be written as \( \alpha \{n\} \) for a reduced Lyndon composition \( \alpha \) and an \( n \in \{1,2,3,\ldots\} \)). Now, prove (6.28) by strong induction: first, induce over \( |\beta| \), and then, for fixed \( |\beta| \), induce over \( \beta \) in the wll-order. The induction step looks as follows: Fix some composition \( \alpha \), and assume (as induction hypothesis) that:

- (6.28) holds for every composition \( \beta \) satisfying \( |\beta| < |\alpha| \);
- (6.28) holds for every composition \( \beta \) satisfying \( |\beta| = |\alpha| \) and \( \beta \) wll.

It remains to prove that (6.28) holds for \( \beta = \alpha \). In other words, it remains to prove that \( M_\alpha \in U \). Let \( (a_1,a_2,\ldots,a_p) \) be the CFL factorization of the word \( \alpha \). Assume WLOG that \( p \neq 0 \) (else, all is trivial). We are in one of the following two cases:

**Case 1:** All of the words \( a_1, a_2, \ldots, a_p \) are equal.

**Case 2:** Not all of the words \( a_1, a_2, \ldots, a_p \) are equal.

---

\(^{252}\) since every \( w \in \text{Comp}_k \) with the property that \( \ell(w) \leq (s-1)\ell(x) \) must satisfy \( w < x \)\(^{253}\) applied to \( A = \text{QSym}, b_w = M_{\text{red}_w}^{(\gcd w)}, \text{wt} (N) = N \) and \( g_u = M_u \)
In Case 2, there exists a $k \in \{1, 2, \ldots, p-1\}$ satisfying $a_k > a_{k+1}$ (since $a_1 \geq a_2 \geq \cdots \geq a_p$), and thus Corollary 6.119 (applied to $u = \alpha$, $n = |\alpha|$, $x = a_1 a_2 \cdots a_k$ and $y = a_k a_{k+1} \cdots a_p$) shows that

$$M_\alpha = \sum_{\substack{M_{w_1 a_2 \cdots a_k} \in U \cap M_{a_k a_{k+1} \cdots a_p} \in U \cap \text{(by the induction hypothesis)}}} \left( \begin{array}{c} \text{a sum of terms of the form} \ M_w \ \text{with} \ w \in \text{Comp}[\alpha] \text{ satisfying} \ w < \alpha \ \text{will} \end{array} \right)$$

$$\in U \cap -U \subset U \cup -U \subset U,$$

so that $M_{\alpha} \in U \cup -U \subset U$. This rewrites as $M_{\alpha} \in U$ (since $\alpha = x^p$). So $M_{\alpha} \in U$ is proven in both Cases 1 and 2, and thus the induction proof of (6.28) is finished.

**Exercise 6.126.** Prove Theorem 6.82.

Of course, this proof of Theorem 6.82 yields a new (third) proof for Proposition 6.83.

We notice the following corollary of our approach to Theorem 6.82:

**Corollary 6.127.** The $\Lambda$-algebra $\text{QSym}$ is a polynomial algebra (over $\Lambda$).

**Exercise 6.128.** Prove Corollary 6.127.

[Hint: The algebraically independent generating set $\left( M_{w}^{(s)} \right)_{(w,s) \in \mathbb{R} \times \{1,2,3,\ldots\}}$ of $\text{QSym}$ contains the elements $M_{(1)}^{(s)} = e_s \in \Lambda$ for all $s \in \{1,2,3,\ldots\}$.]
7. Aguiar-Bergeron-Sottile character theory Part I: QSym as a terminal object

It turns out that the universal mapping property of NSym as a free associative algebra leads via duality to a universal property for its dual QSym, elegantly explaining several combinatorial invariants that take the form of quasisymmetric or symmetric functions:

- Ehrenborg’s quasisymmetric function of a ranked poset [48],
- Stanley’s chromatic symmetric function of a graph [156],
- the quasisymmetric function of a matroid considered in [19].

7.1. Characters and the universal property.

**Definition 7.1.** Given a Hopf algebra \(A\) over \(k\), a **character** is an algebra morphism \(\zeta: A \rightarrow k\), that is,

- \(\zeta(1_A) = 1_k\),
- \(\zeta\) is \(k\)-linear, and
- \(\zeta(ab) = \zeta(a)\zeta(b)\) for \(a, b\) in \(A\).

**Example 7.2.** A particularly important character for \(A = \text{QSym}\) is defined as follows:

\[
\text{QSym} \xrightarrow{\zeta_Q} k
\]

\[
f(x) \mapsto f(1, 0, 0, \ldots) = \left[ f(x) \right]_{x_1=1, x_2=x_3=\cdots=0}.
\]

Hence,

\[
\zeta_Q(M_\alpha) = \zeta_Q(L_\alpha) = \begin{cases} 1 & \text{if } \alpha = (n) \text{ for some } n, \\ 0 & \text{otherwise.} \end{cases}
\]

In other words, the restriction \(\zeta_Q|_{\text{QSym}_n}\) coincides with the functional \(H_n\) in \(\text{NSym}_n = \text{Hom}_k(\text{QSym}_n, k)\): one has for \(f\) in \(\text{QSym}_n\) that

\[
(7.1) \quad \zeta_Q(f) = (H_n, f).
\]

It is worth remarking that there is nothing special about setting \(x_1 = 1\) and \(x_2 = x_3 = \cdots = 0\): for quasisymmetric \(f\), we could have defined the same character \(\zeta_Q\) by picking any variable, say \(x_n\), and sending

\[
f(x) \mapsto \left[ f(x) \right]_{x_n=1, \text{ and } x_m=0 \text{ for } m \neq n}.
\]

This character \(\text{QSym} \xrightarrow{\zeta_Q} k\) has a certain universal property.

**Theorem 7.3.** A connected graded Hopf algebra \(A\) together with a character \(A \xrightarrow{\zeta} k\) induces a unique graded Hopf morphism \(A \xrightarrow{\Psi} \text{QSym}\) making this diagram commute:

\[
\begin{array}{ccc}
A & \xrightarrow{\Psi} & \text{QSym} \\
\downarrow{\zeta} & & \downarrow{\zeta_Q} \\
& k & \\
\end{array}
\]

Furthermore, \(\Psi\) has this formula on elements of \(A_n\):

\[
(7.3) \quad \Psi(a) = \sum_{\alpha \in \text{Comp}_n} \zeta_\alpha(a)M_\alpha
\]

where for \(\alpha = (\alpha_1, \ldots, \alpha_\ell)\), the map \(\zeta_\alpha\) is the composite

\[
A_n \xrightarrow{\Delta^{(\ell-1)}} A^\otimes \ell \xrightarrow{\pi_\alpha} A_{\alpha_1} \otimes \cdots \otimes A_{\alpha_\ell} \xrightarrow{\zeta^\otimes} k
\]

in which \(A^\otimes \ell \xrightarrow{\pi_\alpha} A_{\alpha_1} \otimes \cdots \otimes A_{\alpha_\ell}\) is the canonical projection.

---

\(^{254}\)We are using the notation of Proposition 5.9 here, and we are still identifying QSym with QSym \((x)\), where \(x\) denotes the infinite chain \((x_1 < x_2 < \cdots)\).
Proof. One argues that $\Psi$ is unique, and has formula (7.3), using only that $\zeta$ is $k$-linear and sends 1 to 1 and that $\Psi$ is a graded $k$-coalgebra map making (7.2) commute. Equivalently, consider the adjoint $k$-algebra map $^{255}$

$$\text{NSym} = \text{QSym}^* \xrightarrow{\Psi^*} A^*.$$ 

Commutativity of (7.2) implies that for $a$ in $A_n$,

$$(\Psi^*(H_n), a) = (H_n, \Psi(a)) \quad (7.1)$$

whereas gradedness of $\Psi^*$ yields that $\Psi^*(H_n, a) = 0$ whenever $a \in A_n$ and $m \neq n$. In other words, $\Psi^*(H_n)$ is the element of $A^o$ defined as the following functional on $A$:

$$\Psi^*(H_n)(a) = \begin{cases} \zeta(a) & \text{if } a \in A_n, \\ 0 & \text{if } a \in A_m \text{ for some } m \neq n. \end{cases} 

(7.4)$$

By the universal property for NSym $\cong k(H_1, H_2, \ldots)$ as free associative $k$-algebra, we see that any choice of a $k$-linear map $A \xrightarrow{\zeta} k$ uniquely produces a $k$-algebra map $\Psi^* : \text{QSym}^o \rightarrow A^o$ which satisfies (7.4) for all $n \geq 1$. It is easy to see that this $\Psi^*$ then automatically satisfies (7.4) for $n = 0$ as well if $\zeta$ sends 1 to 1 (it is here that we use $\zeta(1) = 1$ and the connectedness of $A$). Hence, any given $k$-linear map $A \xrightarrow{\zeta} k$ sending 1 to 1 uniquely produces a $k$-algebra map $\Psi^* : \text{QSym}^o \rightarrow A^o$ which satisfies (7.4) for all $n \geq 0$. Formula (7.3) follows as

$$\Psi(a) = \sum_{a \in \text{Comp}} (H_\alpha, \Psi(a)) M_\alpha$$

and for a composition $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, one has

$$(H_\alpha, \Psi(a)) = (\Psi^*(H_\alpha), a) = (\Psi^*(H_{\alpha_1}) \cdots \Psi^*(H_{\alpha_\ell}), a)$$

$$= (\Psi^*(H_{\alpha_1}) \otimes \cdots \otimes \Psi^*(H_{\alpha_\ell}), \Delta^{(\ell-1)}(a))$$

$$\overset{(7.4)}{=} (\zeta \otimes \pi_\alpha)(\Delta^{(\ell-1)}(a)) = \zeta_\alpha(a),$$

where the definition of $\zeta_\alpha$ was used in the last equality.

We wish to show that if, in addition, $A$ is a Hopf algebra and $A \xrightarrow{\zeta} k$ is a character (algebra map), then $A \xrightarrow{\Psi} \text{QSym}$ will be an algebra map, that is, the two maps $A \otimes A \rightarrow \text{QSym}$ given by $\Psi \circ m$ and $m \circ (\Psi \otimes \Psi)$ coincide. To see this, consider these two diagrams having the two maps in question as the composites of their top rows:

$$\begin{array}{cccc}
A \otimes A & \xrightarrow{m} & A & \xrightarrow{\Psi} \text{QSym} \\
\downarrow \zeta \otimes \zeta & & \downarrow \zeta & \\
\downarrow \zeta & & \zetaQ & \\
\downarrow \zeta & & \downarrow \zetaQ \otimes \zetaQ & \\
\text{NSym} & \xrightarrow{m} & \text{QSym} & \\
\end{array}$$

The fact that $\zeta, \zetaQ$ are algebra maps makes the above diagrams commute, so that applying the uniqueness in the first part of the proof to the character $A \otimes A \xrightarrow{\zeta \otimes \zeta} k$ proves the desired equality $\Psi \circ m = m \circ (\Psi \otimes \Psi)$. \qed

Remark 7.4. When one assumes in addition that $A$ is cocommutative, it follows that the image of $\Psi$ will lie in the subalgebra $\Lambda \subset \text{QSym}$, e.g. from the explicit formula (7.3) and the fact that one will have $\zeta_\alpha = \zeta_\beta$ whenever $\beta$ is a rearrangement of $\alpha$. In other words, the character $A \xrightarrow{\zeta} k$ defined by restricting $\zeta$ to $\Lambda$, or by

$$\zeta_\Lambda, (m_\lambda) = \begin{cases} 1 & \text{if } \lambda = (n) \text{ for some } n, \\ 0 & \text{otherwise}. \end{cases}$$

has a universal property as terminal object with respect to characters on cocommutative co- or Hopf algebras.

$^{255}$Here we are using the fact that there is a 1-to-1 correspondence between graded $k$-linear maps $A \rightarrow \text{QSym}$ and graded $k$-linear maps $\text{QSym}^o \rightarrow A^o$ given by $f \mapsto f^*$, and this correspondence has the property that a given graded map $f : A \rightarrow \text{QSym}$ is a $k$-coalgebra map if and only if $f^*$ is a $k$-algebra map. This is a particular case of Exercise 1.67(f).
We close this section by discussing a well-known polynomiality and reciprocity phenomenon; see, e.g., Humpert and Martin [78, Prop. 2.2], Stanley [156, §4].

**Definition 7.5.** The *binomial Hopf algebra* (over the commutative ring $k$) is the polynomial algebra $k[m]$ in a single variable $m$, with a Hopf algebra structure transported from the symmetric algebra Sym $(k^1)$ (which is a Hopf algebra by virtue of Example 1.21, applied to $V = k^1$) along the isomorphism Sym $(k^1) \to k[m]$ which sends the standard basis element of $k^1$ to $m$. Thus the element $m$ is primitive; that is, $\Delta m = 1 \otimes m + m \otimes 1$ and $S(m) = -m$. As $S$ is an algebra anti-endomorphism by Proposition 1.32 and $k[m]$ is commutative, one has $S(g)(m) = g(-m)$ for all polynomials $g(m)$ in $k[m]$.

**Definition 7.6.** For an element $f(x)$ in QSym and a nonnegative integer $m$, let $\text{ps}^1(f)(m)$ denote the element of $k$ obtained by principal specialization at $q = 1$

\[
\text{ps}^1(f)(m) = [f(x)]_{x_1 = x_2 = \cdots = x_m = 1, x_{m+1} = x_{m+2} = \cdots = 0} = f(1, 1, \ldots, 1, 0, 0, \ldots).
\]

**Proposition 7.7.** Assume that $Q$ is a subring of $k$. The map $\text{ps}^1$ has the following properties.

(i) Let $f \in \text{QSym}$. There is a unique polynomial in $k[m]$ which agrees for each nonnegative integer $m$ with $\text{ps}^1(f)(m)$, and which, by abuse of notation, we will also denote $\text{ps}^1(f)(m)$. If $f$ lies in $\text{QSym}_n$, then $\text{ps}^1(f)(m)$ is a polynomial of degree at most $n$, taking these values on $M_\alpha, L_\alpha$ for $\alpha = (\alpha_1, \ldots, \alpha_\ell)$ in Comp$_n$:

\[
\text{ps}^1(M_\alpha)(m) = {m \choose \ell},
\]

\[
\text{ps}^1(L_\alpha)(m) = {m - \ell + n \choose n}.
\]

(ii) The map $\text{QSym} \xrightarrow{\text{ps}^1} k[m]$ is a Hopf morphism into the binomial Hopf algebra.

(iii) For all $m$ in $\mathbb{Z}$ and $f$ in $\text{QSym}$ one has

\[
\zeta^m(f) = \text{ps}^1(f)(m).
\]

In particular, one also has

\[
\zeta^{(-m)}(f) = \text{ps}^1(S(f))(m) = \text{ps}^1(f)(-m).
\]

(iv) For a graded Hopf algebra $A$ with a character $A \xrightarrow{\zeta} k$, and any element $a$ in $A_n$, the polynomial $\text{ps}^1(\Phi(a))(m)$ in $k[m]$ has degree at most $n$, and when specialized to $m$ in $\mathbb{Z}$ satisfies

\[
\zeta^m(a) = \text{ps}^1(\Phi(a))(m).
\]

**Proof.** To prove assertion (i), note that one has

\[
\text{ps}^1(M_\alpha)(m) = M_\alpha(1, 1, \ldots, 1, 0, 0, \ldots) = \sum_{1 \leq i_1 < \cdots < i_\ell \leq m} [x_1^{\alpha_1} \cdots x_\ell^{\alpha_\ell}]_{x_j = 1} = {m \choose \ell},
\]

\[
\text{ps}^1(L_\alpha)(m) = L_\alpha(1, 1, \ldots, 1, 0, 0, \ldots) = \sum_{1 \leq i_1 < \cdots < i_n \leq m; i_k < h_{k+1} \text{ if } k \in D(\alpha)} [x_1 \cdots x_m]_{x_j = 1} = \left\{|1 \leq j_1 \leq j_2 \leq \cdots \leq j_n \leq m - \ell + 1|\right\} = {m - \ell + n \choose n}.
\]

As $\{M_\alpha\}_{\alpha \in \text{Comp}_n}$ form a basis for $\text{QSym}_n$, and $\binom{m}{n}$ is a polynomial function in $m$ of degree $\ell(\leq n)$, one concludes that for $f$ in $\text{QSym}_n$ one has that $\text{ps}^1(f)(m)$ is a polynomial function in $m$ of degree at most $n$. The polynomial giving rise to this function is unique, since infinitely many of its values are fixed.

To prove assertion (ii), note that $\text{ps}^1$ is an algebra morphism because it is an evaluation homomorphism. To check that it is a coalgebra morphism, it suffices to check $\Delta \circ \text{ps}^1 = (\text{ps}^1 \otimes \text{ps}^1) \circ \Delta$ on each $M_\alpha$ for
\(\alpha = (\alpha_1, \ldots, \alpha_\ell)\) in \(\text{Comp}_n\). Using the Vandermonde summation \((A + B)_{\ell} = \sum_k (A_k) (B_{\ell-k})\), one has
\[
\Delta \circ \psi^1(M_\alpha) = \Delta(m_{\ell}) = \left( \frac{m \otimes 1 + 1 \otimes m}{\ell} \right) = \sum_{k=0}^{\ell} \left( \frac{m \otimes 1}{k} \right) \left( 1 \otimes m \right)_{\ell-k} = \sum_{k=0}^{\ell} \left( \frac{m}{k} \right) \otimes \left( \frac{m)}{\ell-k} \right).
\]

while at the same time
\[
((\psi^1 \otimes \psi^1) \circ \Delta)(M_\alpha) = \sum_{k=0}^{\ell} \psi^1(M_{(\alpha_1, \ldots, \alpha_k)}) \otimes \psi^1(M_{(\alpha_{k+1}, \ldots, \alpha_\ell)}) = \sum_{k=0}^{\ell} \left( \frac{m}{k} \right) \otimes \left( \frac{m}{\ell-k} \right).
\]

Thus \(\psi^1\) is a bialgebra map, and hence also a Hopf map, by Proposition 1.46(c).

For assertion (iii), first assume \(m\) lies in \(\{0, 1, 2, \ldots\}\). Since \(\zeta_Q(f) = f(1, 0, 0, \ldots)\), one has
\[
\zeta_{Q}^{*m}(f) = \zeta_Q^{\otimes m} \circ \Delta^{(m-1)} f(x) = \zeta_Q^{\otimes m} \left( f(x^{(1)}, x^{(2)}, \ldots, x^{(m)}) \right)
= \left[ f(x^{(1)}, x^{(2)}, \ldots, x^{(m)}) \right]_{x^{(1)} = x^{(2)} = \ldots = x^{(m)} = 1, x^{(j)}_j = x^{(j)}_j = \ldots = 0 \text{ for all } j}
= f(1, 0, 0, \ldots, 1, 0, 0, \ldots, 1, 0, 0, \ldots) = f(1, 1, 1, 1, 0, 0, \ldots) = \psi^1(f)(m).
\]

But then Proposition 1.46(a) also implies
\[
\zeta_{Q}^{*(-m)}(f) = \left( \zeta_Q^{*(-1)} \right)^* m(f) = (\zeta_Q \circ S)^* m(f) = \zeta_{Q}^{*-m}(S(f))
= \psi^1(S(f))(m) = S(\psi^1(f)(m)) = \psi^1(f)(-m).
\]

For assertion (iv), note that
\[
\zeta_{Q}^{*m}(a) = (\zeta_Q \circ \Psi)^* m(a) = (\zeta_{Q}^{*m})(\Psi(a)) = \psi^1(\Psi(a))(m).
\]

where the three equalities come from (7.2), Proposition 1.46(a), and assertion (iii) above, respectively. \(\square\)

Remark 7.8. Aguiar, Bergeron and Sottile give a very cute (third) proof of the QSym antipode formula Theorem 5.11, via Theorem 7.3, in [4, Example 4.8]. They apply Theorem 7.3 to the coopposite coalgebra QSym\(\cop\) and its character \(\zeta_{Q}^{*(-1)}\). One can show that the map QSym\(\cop \rightarrow QSym\) induced by \(\zeta_{Q}^{*(-1)}\) is \(\Psi = S\), the antipode of QSym, because \(S: \text{QSym} \rightarrow \text{QSym}\) is a coalgebra anti-endomorphism (by Exercise 1.47) satisfying \(\zeta_{Q}^{*(-1)} = \zeta_Q \circ S\). Then they use the formula (7.3) for \(\Psi = S\) (together with the polynomiality Proposition 7.7) to derive Theorem 5.11.

Exercise 7.9. Show that \(\zeta_{Q}^{*m}(f) = \psi^1(f)(m)\) for all \(f \in \text{QSym}\) and \(m \in \{0, 1, 2, \ldots\}\). (This was already proven in Proposition 7.7(iii); give an alternative proof using Proposition 5.7.)

7.2. Example: Ehrenborg’s quasisymmetric function of a ranked poset. Here we consider incidence algebras, coalgebras and Hopf algebras generally, and then particularize to the case of graded posets, to recover Ehrenborg’s interesting quasisymmetric function invariant via Theorem 7.3.

7.2.1. Incidence algebras, coalgebras, Hopf algebras.

Definition 7.10. Given a family \(\mathcal{P}\) of finite partially ordered sets \(P\), let \(k[P]\) denote the free \(k\)-module whose basis consists of symbols \([P]\) corresponding to isomorphism classes of posets \(P\) in \(\mathcal{P}\).

We will assume throughout that each \(P\) in \(\mathcal{P}\) is bounded, that is, it has a unique minimal element \(\hat{0} := 0_P\) and a unique maximal element \(\hat{1} := 1_P\). In particular, \(\mathcal{P} \neq \emptyset\), although it is allowed that \(|\mathcal{P}| = 1\), so that \(\hat{0} = 1\); denote this isomorphism class of posets with one element by \([\emptyset]\).

If \(\mathcal{P}\) is closed under taking intervals
\[
[x, y] := [x, y]_\mathcal{P} := \{z \in P : x \leq z \leq y \}
\]
\[\text{See Exercise 7.9 for an alternative way to prove this, requiring less thought to verify its soundness.} \]
then one can easily that the following coproduct and counit endow $k[P]$ with the structure of a coalgebra, called the \textit{(reduced) incidence coalgebra}:

$$\Delta[P] := \sum_{x \in P} [0, x] \otimes [x, 1],$$

$$\epsilon[P] := \begin{cases} 
1 & \text{if } |P| = 1 \\
0 & \text{otherwise.}
\end{cases}$$

The dual algebra $k[P]^*$ is generally called the \textit{reduced incidence algebra (modulo isomorphism)} for the family $P$ (see, e.g., [145]). It contains the important element $k[P] \xrightarrow{\zeta} k$, called the \textit{$\zeta$-function} that takes the value $\zeta[P] = 1$ for all $P$.

If $P$ (is not empty and) satisfies the further property of being \textit{hereditary} in the sense that for every $P_1, P_2$ in $P$, the \textit{Cartesian product poset} $P_1 \times P_2$ with componentwise partial order is also in $P$, then one can check that the following product and unit endow $k[P]$ with the structure of a (commutative) algebra:

$$[P_1] \cdot [P_2] := m([P_1] \otimes [P_2]) := [P_1 \times P_2],$$

$$1_{k[P]} := [0].$$

**Proposition 7.11.** For any hereditary family $P$ of finite posets, $k[P]$ is a bialgebra, and even a Hopf algebra with antipode $S$ given as in (1.16) (Takeuchi’s formula):

$$S[P] = \sum_{k \geq 0} (-1)^k \sum_{\hat{0} = x_0 < \cdots < x_k = 1} [x_0, x_1] \cdots [x_{k-1}, x_k].$$

**Proof.** Checking the commutativity of the pentagonal diagram in (1.8) amounts to the fact that, for any $(x_1, x_2) <_{P_1 \times P_2} (y_1, y_2)$, one has a poset isomorphism

$$[(x_1, x_2), (y_1, y_2)]_{P_1 \times P_2} \cong [x_1, y_1]_{P_1} \times [x_2, y_2]_{P_2}.$$

Commutativity of the remaining diagrams in (1.8) is straightforward, and so $k[P]$ is a bialgebra. But then Remark 1.45 implies that it is a Hopf algebra, with antipode $S$ as in (1.16), because the map $f := \text{id}_{k[P]} - u\epsilon$ (sending the class $[0]$ to 0, and fixing all other $[P]$) is locally $*$-nilpotent:

$$f^k[P] = \sum_{\hat{0} = x_0 < \cdots < x_k = 1} [x_0, x_1] \cdots [x_{k-1}, x_k]$$

will vanish due to an empty sum whenever $k$ exceeds the maximum length of a chain in the finite poset $P$. \hfill \Box

It is perhaps worth remarking how this generalizes the Möbius function formula of P. Hall. Note that the zeta function $k[P] \xrightarrow{\zeta} k$ is a \textit{character}, that is, an algebra morphism. Proposition 1.46(a) then tells us that $\zeta$ should have a convolutional inverse $k[P] \xrightarrow{\mu = \zeta^{-1}} k$, traditionally called the \textit{Möbius function}, with the formula $\mu = \zeta^{-1} = \zeta \circ S$. Rewriting this via the antipode formula for $S$ given in Proposition 7.11 yields P. Hall’s formula.

**Corollary 7.12.** For a finite bounded poset $P$, one has

$$\mu[P] = \sum_{k \geq 0} (-1)^k |\{\text{chains } \hat{0} = x_0 < \cdots < x_k = \hat{1} \text{ in } P\}|.$$

We can also notice that $S$ is an algebra anti-endomorphism (by Proposition 1.32), thus an algebra endomorphism (since $k[P]$ is commutative). Hence, $\mu = \zeta \circ S$ is a composition of two algebra homomorphisms, thus an algebra homomorphism itself. We therefore obtain the following classical fact:

**Corollary 7.13.** For two finite bounded posets $P$ and $Q$, we have $\mu[P \times Q] = \mu[P] \cdot \mu[Q]$. 
7.2.2. The incidence Hopf algebras for ranked posets and Ehrenborg’s function.

**Definition 7.14.** Take \( P \) to be the class of bounded ranked finite posets, that is, those for which all maximal chains from \( \hat{0} \) to \( \hat{1} \) have the same length \( r(P) \). This is a hereditary class, as it implies that any interval is \([x, y]_P\) is also ranked, and the product of two bounded ranked posets is also bounded and ranked. It also uniquely defines a rank function \( P \rightarrow \mathbb{N} \) in which \( r(\hat{0}) = 0 \) and \( r(x) \) is the length of any maximal chain from \( \hat{0} \) to \( x \).

**Example 7.15.** Consider a pyramid with apex vertex \( a \) over a square base with vertices \( b, c, d, e \):

![Diagram of a pyramid with vertices a, b, c, d, e, and their connections]

Ordering its faces by inclusion gives a bounded ranked poset \( P \), where the rank of an element is one more than the dimension of the face it represents:

\[
\text{rank:}
\begin{align*}
4 & : \quad abcd \\
3 & : \quad abc, acd, ade, abe \\
2 & : \quad ab, ac, ad, ae, bc, bd, cd, de \\
1 & : \quad a, b, c, d, e \\
0 & : \quad \emptyset
\end{align*}
\]

**Definition 7.16.** Ehrenborg’s quasisymmetric function \( \Psi[P] \) for a bounded ranked poset \( P \) is the image of \([P]\) under the map \( k[P] \xrightarrow{\Psi} \text{QSym} \) induced by the zeta function \( k[P] \xrightarrow{\zeta} k \) as a character, via Theorem 7.3.

The quasisymmetric function \( \Psi[P] \) captures several interesting combinatorial invariants of \( P \); see Stanley [157, Chap. 3] for more background on these notions.

**Definition 7.17.** Let \( P \) be a bounded ranked poset of rank \( r(P) := r(\hat{1}) \). Define its rank-generating function

\[
\text{RGF}(P, q) := \sum_{p \in P} q^{r(p)},
\]

its characteristic polynomial

\[
\chi(P, q) := \sum_{p \in P} \mu(\hat{0}, p)q^{r(p)}
\]

(where \( \mu(p, q) \) is shorthand for \( \mu([p, q]) \)), its zeta polynomial

\[
Z(P, m) = |\{\text{multichains } \hat{0} \leq p_1 \leq p \cdots \leq p_{m-1} \leq p \hat{1}\}|
\]

\[
= \sum_{s=0}^{r(P)-1} \binom{m}{s+1} |\{\text{chains } \hat{0} < p_1 < \cdots < p_s < \hat{1}\}|
\]
and for a subset \( S \subset \{1, 2, \ldots, r(P)−1\} \), its flag number \( f_S \), as a component of its flag \( f \)-vector \((f_S)_{S \subset [r−1]}\)
\[
  f_S = |\{\text{chains } 0 < p_1 < \cdots < p_s < r \mid \{r(p_1), \ldots, r(p_s)\} = S\}|
\]
as well as the flag \( h \)-vector \( h_T \) given by \( f_S = \sum_{T \subset S} h_T \), or by inclusion-exclusion, \( h_S = \sum_{T \subset S} (-1)^{|S\setminus T|} h_T \).

**Example 7.18.** For the poset \( P \) in Example 7.15, one has \( RGF(P,q) = 1 + 5q + 8q^2 + 5q^3 + q^4 \). Since \( P \) is the poset of faces of a polytope, the Möbius function values for its intervals are easily predicted: \( \mu(x, y) = (-1)^{r(x, y)} \), that is, \( P \) is an Eulerian ranked poset; see Stanley [157, §3.16]. Hence its characteristic polynomial is trivially related to the rank generating function, sending \( q \mapsto −q \), that is,
\[
  \chi(P,q) = RGF(P,−q) = 1 − 5q + 8q^2 − 5q^3 + q^4.
\]

Its flag \( f \)-vector and \( h \)-vector entries are given in the following table.

<table>
<thead>
<tr>
<th>( S )</th>
<th>( f_S )</th>
<th>( h_S )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \emptyset )</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>{1}</td>
<td>5</td>
<td>5 − 1 = 4</td>
</tr>
<tr>
<td>{2}</td>
<td>8</td>
<td>8 − 1 = 7</td>
</tr>
<tr>
<td>{3}</td>
<td>5</td>
<td>5 − 1 = 4</td>
</tr>
<tr>
<td>{1, 2}</td>
<td>16</td>
<td>16 − (5 + 8) + 1 = 4</td>
</tr>
<tr>
<td>{1, 3}</td>
<td>16</td>
<td>16 − (5 + 5) + 1 = 7</td>
</tr>
<tr>
<td>{2, 3}</td>
<td>16</td>
<td>16 − (5 + 8) + 1 = 4</td>
</tr>
<tr>
<td>{1, 2, 3}</td>
<td>32</td>
<td>32 − (16 + 16 + 16) + (5 + 8 + 5) − 1 = 1</td>
</tr>
</tbody>
</table>

and using (7.7), its zeta polynomial is
\[
  Z(P,m) = 1 \binom{m}{1} + (5 + 8 + 5) \binom{m}{2} + (16 + 16 + 16) \binom{m}{3} + 32 \binom{m}{4} = \frac{m^2 (2m-1)(2m+1)}{3}.
\]

**Theorem 7.19.** Assume that \( Q \) is a subring of \( k \). Ehrenborg’s quasisymmetric function \( \Psi[P] \) for a bounded ranked poset \( P \) encodes

(i) the flag \( f \)-vector entries \( f_S \) and flag \( h \)-vector entries \( h_S \) as its \( M_\alpha \) and \( L_\alpha \) expansion coefficients:
\[
  \Psi[P] = \sum_\alpha f_{D(\alpha)}(P) \ M_\alpha = \sum_\alpha h_{D(\alpha)}(P) \ L_\alpha.
\]

(ii) the zeta polynomial as the specialization from Definition 7.6
\[
  Z(P,m) = p^1(\Psi[P])(m) = [\Psi[P]]_{x_1=x_2=\cdots=x_m=1}, \text{ and}
\]

(iii) the rank-generating function as the specialization
\[
  RGF(P,q) = [\Psi[P]]_{x_1=q, x_2=1, x_3=x_4=\cdots=0},
\]

(iv) the characteristic polynomial as the convolution
\[
  \chi(P,q) = ((\psi_q \circ S) \ast \zeta_Q) \circ \Psi[P]
\]

where \( QSym \overset{\psi_q}{\longrightarrow} k[q] \) maps \( f(x) \mapsto f(q,0,0,\ldots). \)

**Proof.** In assertion (i), the expansion \( \Psi[P] = \sum_\alpha f_{D(\alpha)}(P) \ M_\alpha \) is (7.3), since \( \zeta_\alpha[P] = f_{D(\alpha)}(P) \). The \( L_\alpha \) expansion follows by inclusion-exclusion, as \( L_\alpha = \sum_{\beta \supset D(\beta) \supset D(\alpha)} M_\beta \) and \( f_S(P) = \sum_{T \subset S} h_T \).

Assertion (ii) is immediate from Proposition 7.7(iv), since \( Z(P,m) = \zeta^m[P] \).

Assertion (iii) can be deduced from assertion (i), but it is perhaps more fun and in the spirit of things to proceed as follows. Note that \( \psi_q(M_\alpha) = q^\alpha \) for \( \alpha = (n) \), and \( \psi_q(M_\alpha) \) vanishes for all other \( \alpha \neq (n) \) in \( \text{Comp}_n \). Hence for a bounded ranked poset \( P \) one has
\[
  (\psi_q \circ \Psi)[P] = q^r(P).
\]

---

257 Actually, (7.7) is false if \(|P| = 1 \) (but only then). We use (7.6) to define \( Z(P,m) \) in this case.

258 In fact, Ehrenborg defined \( \Psi[P] \) in [38, Defn. 4.1] via this \( M_\alpha \) expansion, and then showed that it gave a Hopf morphism.
Consequently, using (1.17) one can compute

\[
RGF(P, q) = \sum_{p \in P} q^r(p) \cdot 1 = \sum_{p \in P} q^r([0, p]) \cdot \zeta[p, \hat{1}] = (7.2, \ (7.8), \ (1.17), \ (7.7)) \sum_{p \in P} (\psi_q \circ \Psi)[\hat{0}, p] \cdot (\zeta_Q \circ \Psi)[p, \hat{1}]
\]

Similarly, for assertion (iv) first note that Proposition 7.11 and Corollary 7.12 let one calculate that

\[
(\psi_q \circ \Psi)(P) = \sum_k (-1)^k \sum_{\hat{0} = x_0 < \cdots < x_k = \hat{1}} (\psi_q \circ \Psi)([x_0, x_1]) \cdot \cdots \cdot (\psi_q \circ \Psi)([x_k-1, x_k]) = \sum_k (-1)^k \sum_{\hat{0} = x_0 < \cdots < x_k = \hat{1}} q^{r(P)} = \mu(\hat{0}, \hat{1})q^{r(P)}.
\]

This is used in the penultimate equality here:

\[
((\psi_q \circ S) \circ \Psi)[P] \stackrel{(1.17)}{=} (((\psi_q \circ S) \circ \Psi) \circ (\zeta_Q \circ \Psi))[P] = ((\psi_q \circ \Psi \circ S) \circ \zeta)[P]
\]

\[
= \sum_{p \in P} (\psi_q \circ \Psi \circ S)[\hat{0}, p] \cdot \zeta[p, \hat{1}] = \sum_{p \in P} \mu(\hat{0}, p)q^{r(P)} = \chi(P, q).
\]

7.3. Example: Stanley’s chromatic symmetric function of a graph. We introduce the chromatic Hopf algebra of graphs and an associated character \( \zeta \) so that the map \( \Psi \) from Theorem 7.3 sends a graph \( G \) to Stanley’s chromatic symmetric function of \( G \). Then principal specialization \( ps^1 \) sends this to the chromatic polynomial of the graph.

7.3.1. The chromatic Hopf algebra of graphs.

Definition 7.20. The chromatic Hopf algebra (see Schmitt [147, §3.2]) \( \mathcal{G} \) is a free \( k \)-module whose \( k \)-basis elements \([G]\) are indexed by isomorphism classes of (finite) simple graphs \( G = (V, E) \). Define for \( G_1 = (V_1, E_1), G_2 = (V_2, E_2) \) the multiplication

\[
[G_1] \cdot [G_2] := [G_1 \sqcup G_2]
\]

where \([G_1 \sqcup G_2]\) denote the isomorphism class of the disjoint union, on vertex set \( V = V_1 \sqcup V_2 \) which is a disjoint union of copies of their vertex sets \( V_1, V_2 \), with edge set \( E = E_1 \sqcup E_2 \). For example,

\[
\begin{bmatrix}
\bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet \\
\end{bmatrix}
\cdot
\begin{bmatrix}
\bullet \\
\bullet \\
\end{bmatrix}
= \begin{bmatrix}
\bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet \\
\end{bmatrix}
\]

Thus the class \([\varnothing]\) of the empty graph \( \varnothing \) having \( V = \varnothing, E = \varnothing \) is a unit element.

Given a subset \( V' \subset V \), the subgraph induced on vertex set \( V' \) is defined as the graph \( G|_{V'} := (V', E') \) with edge set \( E' = \{ e \in E : e = \{ v_1, v_2 \} \subset V' \} \). This lets one define a comultiplication

\[
\Delta[G] := \sum_{(V_1, V_2) : V_1 \cup V_2 = V} [G|_{V_1}] \otimes [G|_{V_2}].
\]

Define a counit

\[
\epsilon[G] := \begin{cases} 
1 & \text{if } G = \varnothing \\
0 & \text{otherwise.}
\end{cases}
\]

Proposition 7.21. The above maps endow \( \mathcal{G} \) with the structure of a graded connected finite type Hopf algebra over \( k \), which is both commutative and cocommutative.
Example 7.22. Here are some examples of these structure maps:

\[
\Delta \left[ \begin{array}{ccc}
\bullet & \bullet & \bullet \\
\bullet & & \\
& & \\
\end{array} \right] = 1 \otimes \left[ \begin{array}{ccc}
\bullet & \bullet & \\
\bullet & & \\
& & \\
\end{array} \right] + 2 \left[ \begin{array}{ccc}
\bullet & & \\
\bullet & & \\
& & \\
\end{array} \right] \otimes 1
\]

Proof of Proposition 7.21. The associativity of the multiplication and comultiplication should be clear as

\[
m^{(2)}([G_1] \otimes [G_2] \otimes [G_3]) = [G_1 \sqcup G_2 \sqcup G_3]
\]

\[
\Delta^{(2)}[G] = \sum_{(V_1, V_2, V_3): V = V_1 \sqcup V_2 \sqcup V_3} [G|_{V_1}] \otimes [G|_{V_2}] \otimes [G|_{V_3}].
\]

Checking the unit and counit conditions are straightforward. Commutativity of the pentagonal bialgebra diagram in (1.8) comes down to check that, given graphs \(G_1, G_2\) on disjoint vertex sets \(V_1, V_2\), when one applies to \([G_1] \otimes [G_2]\) either the composite \(\Delta \circ m\) or the composite \((m \otimes m) \circ (\text{id} \otimes T \otimes \text{id}) \circ (\Delta \otimes \Delta)\), the result is the same:

\[
\sum_{(V_1, V_2, V_3): V_1 = V_1 \sqcup V_2 \sqcup V_3} [G_1|_{V_1}] \sqcup [G_2|_{V_2}] \otimes [G_1|_{V_2}] \otimes [G_2|_{V_3}].
\]

Letting \(\mathcal{G}_n\) be the \(k\)-span of \([G]\) having \(n\) vertices makes \(\mathcal{G}\) a bialgebra which is graded and connected, and hence also a Hopf algebra by Proposition 1.36. Cocommutativity should be clear, and commutativity follows from the graph isomorphism \(G_1 \sqcup G_2 \cong G_2 \sqcup G_1\). Finally, \(\mathcal{G}\) is of finite type since there are only finitely many isomorphism classes of simple graphs on \(n\) vertices for every given \(n\). \qed

Remark 7.23. Humpert and Martin [78, Theorem 3.1] gave the following expansion for the antipode in the chromatic Hopf algebra, containing fewer terms than Takeuchi’s general formula (1.16): given a graph \(G = (V, E)\), one has

\[
S[G] = \sum_{F} (-1)^{|V| - \text{rank}(F)} \text{acyc}(G/F)[G_{V,F}].
\]

Here \(F\) runs over all subsets of edges that form flats in the graphic matroid for \(G\), meaning that if \(e = \{v, v'\}\) is an edge in \(E\) for which one has a path of edges in \(F\) connecting \(v\) to \(v'\), then \(e\) also lies in \(F\). Here \(G/F\) denotes the quotient graph in which all of the edges of \(F\) have been contracted, while \(\text{acyc}(G/F)\) denotes its number of acyclic orientations, and \(G_{V,F} := (V, F)\) as a simple graph.\(^{259}\)

Remark 7.24. In [13], Benedetti, Hallam and Machacek define a Hopf algebra of simplicial complexes, which contains \(\mathcal{G}\) as a Hopf subalgebra (and also has \(\mathcal{G}\) as a quotient Hopf algebra). They compute a formula for its antipode similar to (and generalizing) (7.9).

Remark 7.25. The chromatic Hopf algebra \(\mathcal{G}\) is used in [92] and [32, §14.4] to study Vassiliev invariants of knots. In fact, a certain quotient of \(\mathcal{G}\) (named \(\mathcal{F}\) in [92] and \(\mathcal{L}\) in [32, §14.4]) is shown to naturally host invariants of chord diagrams and therefore Vassiliev invariants of knots.

Remark 7.26. The \(k\)-algebra \(\mathcal{G}\) is isomorphic to a polynomial algebra (in infinitely many indeterminates) over \(k\). Indeed, every finite graph can be uniquely written as a disjoint union of finitely many connected finite graphs (up to order). Therefore, the basis elements \([G]\) of \(\mathcal{G}\) corresponding to connected finite graphs \(G\) are

\(^{259}\)The notation \(\text{rank}(F)\) denotes the rank of \(F\) in the graphic matroid of \(G\). We can define it without reference to matroid theory as the maximum cardinality of a subset \(F'\) of \(F\) such that the graph \(G_{V,F'}\) is acyclic. Equivalently, \(\text{rank}(F)\) is \(|V| - \epsilon(F)|\), where \(\epsilon(F)\) denotes the number of connected components of the graph \(G_{V,F}\). Thus, the equality (7.9) can be rewritten as \(S[G] = \sum_{F} (-1)^{\epsilon(F)} \text{acyc}(G/F)[G_{V,F}]\). In this form, this equality is also proven in [14, Thm. 3.1].
algebraically independent in $\mathcal{G}$ and generate the whole $k$-algebra $\mathcal{G}$ (indeed, the disjoint unions of connected finite graphs are precisely the monomials in these elements). Thus, $\mathcal{G}$ is isomorphic to a polynomial $k$-algebra with countably many generators (one for each isomorphism class of connected finite graphs). As a consequence, for example, we see that $\mathcal{G}$ is an integral domain if $k$ is an integral domain.

7.3.2. A “ribbon basis” for $\mathcal{G}$ and selfduality. In this subsection, we shall explore a second basis of $\mathcal{G}$ and a bilinear form on $\mathcal{G}$. This material will not be used in the rest of these notes (except in Exercise 7.44), but it is of some interest and provides an example of how a commutative cocommutative Hopf algebra can be studied.

First, let us define a second basis of $\mathcal{G}$, which is obtained by Möbius inversion (in an appropriate sense) from the standard basis ($[G]$) ($G$ is an isomorphism class of finite graphs):

**Definition 7.27.** For every finite graph $G = (V,E)$, set

$$[G] = \sum_{H = (V,E') \; E' \supseteq E^c} (-1)^{|E'\setminus E^c|} [H] \in \mathcal{G},$$


**Proposition 7.28.** (a) Every finite graph $G = (V,E)$ satisfies

$$[G] = \sum_{H = (V,E') \; E' \cap E = \emptyset} [H]^\sharp.$$  

(b) The elements $[G]^\sharp$, where $[G]$ ranges over all isomorphism classes of finite graphs, form a basis of the $k$-module $\mathcal{G}$.

(c) For any graph $H = (V,E)$, we have

$$\Delta [H]^\sharp = \sum_{V = V_1 \sqcup V_2} \sum_{H = H_{|V_1} \sqcup H_{|V_2}} [H_{|V_1}]^\sharp \otimes [H_{|V_2}]^\sharp.$$  

(d) For any two graphs $H_1 = (V_1,E_1)$ and $H_2 = (V_2,E_2)$, we have

$$[H_1]^\sharp [H_2]^\sharp = \sum_{H = (V_1 \sqcup V_2,E) \; H_{|V_1} = H_1; \; H_{|V_2} = H_2} [H]^\sharp.$$  

For example,

$$\begin{bmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \end{bmatrix}^\sharp = \begin{bmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \end{bmatrix} - \begin{bmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \end{bmatrix} - \begin{bmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \end{bmatrix} + \begin{bmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \end{bmatrix}$$

$$= \begin{bmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \end{bmatrix} - 2 \begin{bmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \end{bmatrix} + \begin{bmatrix} \circ & \circ & \circ \\ \circ & \circ & \circ \end{bmatrix}.$$  

Proving Proposition 7.28 is part of Exercise 7.33 further below.

The equalities that express the elements $[G]^\sharp$ in terms of the elements $[H]$ (as in Definition 7.27), and vice versa (Proposition 7.28(a)), are reminiscent of the relations (5.22) and (5.21) between the bases ($R_\alpha$) and ($H_\alpha$) of NSym. In this sense, we can call the basis of $\mathcal{G}$ formed by the $[G]^\sharp$ a “ribbon basis” of $\mathcal{G}$.

We now define a $k$-bilinear form on $\mathcal{G}$:
Definition 7.29. For any two graphs $G$ and $H$, let $\text{Iso}(G, H)$ denote the set of all isomorphisms from $G$ to $H$.  
Let us now define a $k$-bilinear form $\langle \cdot, \cdot \rangle : G \times G \to k$ on $G$ by setting
$$\langle \{G\}, \{H\} \rangle = |\text{Iso}(G, H)|.$$  

Proposition 7.30. The form $\langle \cdot, \cdot \rangle : G \times G \to k$ is symmetric.

Again, we refer to Exercise 7.33 for a proof of Proposition 7.30.

The basis of $G$ constructed in Proposition 7.28(b) and the bilinear form $\langle \cdot, \cdot \rangle$ defined in Definition 7.29 can be used to construct a Hopf algebra homomorphism from $G$ to its graded dual $G^\circ$.

Definition 7.31. For any finite graph $G$, let $\text{aut}(G)$ denote the number $|\text{Iso}(G, G)|$. Notice that this is a positive integer, since the set $\text{Iso}(G, G)$ is nonempty (it contains $\text{id}_G$).

Now, recall that the Hopf algebra $G$ is a connected graded Hopf algebra of finite type. The $n$-th homogeneous component is spanned by the $[G]$ where $G$ ranges over the graphs with $n$ vertices. Since $G$ is of finite type, its graded dual $G^\circ$ is defined. Let $([G]^*)_{[G]}$ be an isomorphism class of finite graphs be the basis of $G^\circ$ dual to the basis $([G])_{[G]}$ is an isomorphism class of finite graphs of $G$. Define a $k$-linear map $\psi : G \to G^\circ$ by
$$\psi([G]^*) = \text{aut}(G) \cdot [G]^*$$
for every finite graph $G$.

Proposition 7.32. Consider the map $\psi : G \to G^\circ$ defined in Definition 7.31.

(a) This map $\psi$ satisfies $\psi(ab) = (a, b)$ for all $a \in G$ and $b \in G$.

(b) The map $\psi : G \to G^\circ$ is a Hopf algebra homomorphism.

(c) If $\mathbb{Q}$ is a subring of $k$, then the map $\psi$ is a Hopf algebra isomorphism $G \to G^\circ$.

Exercise 7.33. Prove Proposition 7.28, Proposition 7.30 and Proposition 7.32.

Remark 7.34. Proposition 7.32(c) shows that the Hopf algebra $G$ is self-dual when $\mathbb{Q}$ is a subring of $k$. On the other hand, if $k$ is a field of positive characteristic, then $G$ is never self-dual. Here is a quick way to see this: The elements $[G]^*$ of $G^\circ$ defined in Definition 7.31 have the property that
$$([\circ]^*)^n = n! \cdot \sum_{[G]} [G]^*$$
for every $n \in \mathbb{N}$, where $\circ$ denotes the graph with one vertex. Thus, if $p$ is a prime and $k$ is a field of characteristic $p$, then $([\circ]^*)^p = 0$. Hence, the $k$-algebra $G^\circ$ has nilpotents in this situation. However, the $k$-algebra $G$ does not (indeed, Remark 7.26 shows that it is an integral domain whenever $k$ is an integral domain). Thus, when $k$ is a field of characteristic $p$, then $G$ and $G^\circ$ are not isomorphic as $k$-algebras (let alone as Hopf algebras).

\[260\] We recall that if $G = (V, E)$ and $H = (W, F)$ are two graphs, then an isomorphism from $G$ to $H$ means a bijection $\varphi : V \to W$ such that $\varphi_* (E) = F$. Here, $\varphi_*$ denotes the map from the powerset of $V$ to the powerset of $W$ which sends every $T \subseteq V$ to $\varphi(T) \subseteq W$.

\[261\] This is well-defined, because:
- the number $|\text{Iso}(G, H)|$ depends only on the isomorphism classes $[G]$ and $[H]$ of $G$ and $H$, but not on $G$ and $H$ themselves;
- the elements $[G]^*$, where $[G]$ ranges over all isomorphism classes of finite graphs, form a basis of the $k$-module $G$ (because of Proposition 7.28(b));

\[262\] This is well-defined, since $([G]^*)_{[G]}$ is an isomorphism class of finite graphs is a basis of the $k$-module $G$ (because of Proposition 7.28(b)).

\[263\] To see this, observe that the tensor $[\circ]^n$ appears in the iterated coproduct $\Delta^{(n-1)}([G])$ exactly $n!$ times whenever $G$ is a graph on $n$ vertices.
7.3.3. Stanley’s chromatic symmetric function of a graph.

**Definition 7.35.** Stanley’s chromatic symmetric function $\Psi[G]$ for a simple graph $G = (V, E)$ is the image of $[G]$ under the map $G \xrightarrow{\Psi} \text{QSym}$ induced via Theorem 7.3 from the edge-free character $\zeta : \mathcal{G} \to \mathbf{k}$ defined by

$$\zeta[G] = \begin{cases} 1 & \text{if } G \text{ has no edges, that is, } G \text{ is an independent/stable set of vertices}, \\ 0 & \text{otherwise}. \end{cases}$$

Note that, because $\mathcal{G}$ is cocommutative, $\Psi[G]$ is symmetric and not just quasisymmetric; see Remark 7.4.

Recall that for a graph $G = (V, E)$, a (vertex-)coloring $f : V \to \{1, 2, \ldots\}$ is called proper if no edge $e = \{v, v'\} \in E$ has $f(v) = f(v')$.

**Proposition 7.36.** For a graph $G = (V, E)$, the symmetric function $\Psi[G]$ has the expansion\(^{264}\)

$$\Psi[G] = \sum_{\text{proper colorings } f : V \to \{1, 2, \ldots\}} x_f$$

where $x_f := \prod_{v \in V} x_{f(v)}$. In particular, its specialization from Proposition 7.6 gives the chromatic polynomial of $G$:

$$\text{ps}^1\Psi[G](m) = \chi_G(m) = |\{\text{proper colorings } f : V \to \{1, 2, \ldots, m\}\}|$$

**Proof.** The iterated coproduct $\mathcal{G} \xrightarrow{\Delta^{(\ell-1)}} \mathcal{G}^{\otimes \ell}$ sends

$$[G] \mapsto \sum_{\{V_1, \ldots, V_{\ell} \} : V = V_1 \sqcup \cdots \sqcup V_{\ell}} \left[ G|_{V_1} \right] \otimes \cdots \otimes \left[ G|_{V_{\ell}} \right]$$

and the map $\zeta^{\otimes \ell}$ sends each addend on the right to 1 or 0, depending upon whether each $V_i \subset V$ is a stable set or not, that is, whether the assignment of color $i$ to the vertices in $V_i$ gives a proper coloring of $G$. Thus formula (7.3) shows that the coefficient $\zeta_\alpha$ of $x_1^{\alpha_1} \cdots x_\ell^{\alpha_\ell}$ in $\Psi[G]$ counts the proper colorings $f$ in which $|f^{-1}(i)| = \alpha_i$ for each $i$. \qed

**Example 7.37.** For the complete graph $K_n$ on $n$ vertices, one has

$$\Psi[K_n] = n! e_n$$

$$\text{ps}^1(\Psi[K_n])(m) = n! e_n(1, 1, \ldots, 1) = n! \binom{m}{n}$$

$$m(m-1) \cdots (m-(n-1)) = \chi(K_n, m).$$

In particular, the single vertex graph $K_1$ has $\Psi[K_1] = e_1$, and since the Hopf morphism $\Psi$ is in particular an algebra morphism, a graph $K_{1^{\infty}}$ having $n$ isolated vertices and no edges will have $\Psi[K_{1^{\infty}}] = e_1^n$.

As a slightly more interesting example, the graph $P_3$ which is a path having three vertices and two edges will have

$$\Psi[P_3] = m(2, 1) + 6m(1, 1, 1) = e_2 e_1 + 3 e_3$$

One might wonder, based on the previous examples, when $\Psi[G]$ is e-positive, that is, when does its unique expansion in the $\{e_\lambda\}$ basis for $\Lambda$ have nonnegative coefficients? This is an even stronger assertion than s-positivity, that is, having nonnegative coefficients for the expansion in terms of Schur functions $\{s_\lambda\}$, since each $e_\lambda$ is s-positive. This weaker property fails, starting with the claw graph $K_{3,1}$, which has

$$\Psi[K_{3,1}] = s(3,1) - s(2,2) + 5 s(2,1,1) + 8 s(1,1,1,1).$$

On the other hand, a result of Gasharov [56, Theorem 2] shows that one at least has s-positivity for $\Psi[\text{inc}(P)]$ where $\text{inc}(P)$ is the incomparability graph of a poset which is $(3 + 1)$-free; we refer the reader to Stanley [156, §5] for a discussion of the following conjecture, which remains open\(^{265}\):

\(^{264}\)In fact, Stanley defined $\Psi[G]$ in [156, Defn. 2.1] via this expansion.

\(^{265}\)A recent refinement for incomparability graphs of posets which are both $(3 + 1)$- and $(2 + 2)$-free, also known as unit interval orders is discussed by Shareshian and Wachs [156].
Conjecture 7.38. For any \((3+1)\)-free poset \(P\), the incomparability graph \(\text{inc}(P)\) has \(\Psi[\text{inc}(P)]\) an \(e\)-positive symmetric function.

Here is another question about \(\Psi[G]\): how well does it distinguish nonisomorphic graphs? Stanley gave this example of two graphs \(G_1, G_2\) having \(\Psi[G_1] = \Psi[G_2]:\)

\[
G_1 = \begin{array}{c}
\bullet \\
\bullet \\
\bullet
\end{array} \quad G_2 = \begin{array}{c}
\bullet \\
\bullet \\
\bullet
\end{array}
\]

At least \(\Psi[G]\) appears to do better at distinguishing trees, much better than its specialization, the chromatic polynomial \(\chi(G, m)\), which takes the same value \(m(m-1)^{n-1}\) on all trees with \(n\) vertices.

Question 7.39. Does the chromatic symmetric function (for \(k = \mathbb{Z}\)) distinguish trees?

It has been checked that the answer is affirmative for trees on 23 vertices or less. There are also interesting partial results on this question by Martin, Morin and Wagner [122].

We close this section with a few other properties of \(\Psi[G]\) proven by Stanley which follow easily from the theory we have developed. For example, his work makes no explicit mention of the chromatic Hopf algebra \(G\), and the fact that \(\Psi\) is a Hopf morphism (although he certainly notes the trivial algebra morphism property \(\Psi(G_1 \sqcup G_2) = \Psi(G_1)\Psi(G_2)\)). One property he proves is implicitly related to \(\Psi\) as a coalgebra morphism: he considers (in the case when \(\mathbb{Q}\) is a subring of \(k\)) the effect on \(\Psi\) of the operator \(\frac{\partial}{\partial p_1}: \Lambda_q \rightarrow \Lambda_q\) which acts by first expressing a symmetric function \(f = f(p_1, p_2, \ldots)\) as a polynomial in the power sums \(\{p_n\}\), and then applying \(\frac{\partial}{\partial p_1}\). It is not hard to see that \(\frac{\partial}{\partial p_1}\) is the same as the skewing operator \(s_{(1)} = p_1^\perp\): both act as derivations on \(\Lambda_q = \mathbb{Q}[p_1, p_2, \ldots]\), and agree in their effect on each \(p_n\), in that both send \(p_1 \mapsto 1\), and both annihilate \(p_2, p_3, \ldots\).

Proposition 7.40. (Stanley [156, Cor. 2.12(a)]) For any graph \(G = (V, E)\), one has

\[
\frac{\partial}{\partial p_1} \Psi[G] = \sum_{v \in V} \Psi[G|_{V \setminus v}].
\]

Proof. One first computes

\[
\Delta \Psi[G] = (\Psi \otimes \Psi) \Delta [G] = \sum_{(V_1, V_2): V_1 \sqcup V_2 = V} \Psi[G|_{V_1}] \otimes \Psi[G|_{V_2}].
\]

Since degree considerations force \((s_{(1)}, \Psi[G|_{V_1}]) = 0\) unless \(|V_1| = 1\), in which case \(\Psi[G|_{V_1}] = s_{(1)}\), one has

\[
\frac{\partial}{\partial p_1} \Psi[G] = s_{(1)} \Psi[G] = \sum_{(V_1, V_2): V_1 \sqcup V_2 = V} (s_{(1)}, \Psi[G|_{V_1}]) \cdot \Psi[G|_{V_2}] = \sum_{v \in V} \Psi[G|_{V \setminus v}].
\]

\[\square\]

Definition 7.41. Given a graph \(G = (V, E)\), an acyclic orientation \(\Omega\) of the edges \(E\) (that is, an orientation of each edge such that the resulting directed graph has no cycles), and a vertex-coloring \(f: V \rightarrow \{1, 2, \ldots\}\), say that the pair \((\Omega, f)\) are weakly compatible if whenever \(\Omega\) orients an edge \(\{v, v'\}\) in \(E\) as \(v \rightarrow v'\), one has \(f(v) \leq f(v')\). Note that a proper vertex-coloring \(f\) of a graph \(G = (V, E)\) is weakly compatible with a unique acyclic orientation \(\Omega\).

Proposition 7.42. (Stanley [156, Prop. 4.1, Thm. 4.2]) The involution \(\omega\) of \(\Lambda\) sends \(\Psi[G]\) to \(\omega(\Psi[G]) = \sum_{(\Omega, f)} x_f\) in which the sum runs over weakly compatible pairs \((\Omega, f)\) of an acyclic orientation \(\Omega\) and vertex-coloring \(f\).

Furthermore, the chromatic polynomial \(\chi_G(m)\) has the property that \((-1)^{|V|} \chi(G, -m)\) counts all such weakly compatible pairs \((\Omega, f)\) in which \(f: V \rightarrow \{1, 2, \ldots, m\}\) is a vertex-\(m\)-coloring.
Proof. As observed above, a proper coloring \( f \) is weakly compatible with a unique acyclic orientation \( \Omega \) of \( G \).

Denote by \( P_\Omega \) the poset on \( V \) which is the transitive closure of \( \Omega \), endowed with a strict labelling by integers, that is, \( i <_p j \) implies \( i >_\Omega j \). Then proper colorings \( f \) that induce \( \Omega \) are the same as \( P_\Omega \)-partitions, so that

\[
\Psi[G] = \sum_{\Omega} F_{P_\Omega}(x).
\]

Applying the antipode \( S \) and using Corollary 5.27 gives

\[
\omega(\Psi[G]) = (-1)^{|V|} S(\Psi[G]) = \sum_{\Omega} F_{P_\Omega}(x) = \sum_{(\Omega,f)} x_f
\]

where in the last line one sums over weakly compatible pairs as in the proposition. The last equality comes from the fact that since each \( P_\Omega \) has been given a strict labelling, \( P_\Omega^{\text{opp}} \) acquires a weak (or natural) labelling, that is \( i <_{P_\Omega^{\text{opp}}} j \) implies \( i <_\Omega j \).

The last assertion follows from Proposition 7.7(iii). \( \square \)

Remark 7.43. The interpretation of \( \chi(G,-m) \) in Proposition 7.42 is a much older result of Stanley [155]. The special case interpreting \( \chi(G,-1) \) as \( (-1)^{|V|} \) times the number of acyclic orientations of \( G \) has sometimes been called Stanley’s (-1)-color theorem. It also follows (via Proposition 7.7) from Humpert and Martin’s antipode formula for \( \mathcal{G} \) discussed in Remark 7.23: taking \( \zeta \) to be the character of \( \mathcal{G} \) given in (7.12),

\[
\chi(G,-1) = \zeta^{(-1)}(G) = \zeta(S[G]) = \sum_{F} (-1)^{|V|-\text{rank}(F)} \text{acyc}(G/F) \zeta[G_F] = (-1)^{|V|} \text{acyc}(G)
\]

where the last equality uses the vanishing of \( \zeta \) on graphs that have edges, so only the \( F = \emptyset \) term survives.

Exercise 7.44. If \( V \) and \( X \) are two sets, and if \( f : V \to X \) is any map, then eqs \( f \) will denote the set

\[
\{ \{u, u'\} \mid u \in V, u' \in V, u \neq u' \text{ and } f(u) = f(u') \}.
\]

This is a subset of the set of all two-element subsets of \( V \).

If \( G = (V,E) \) is a finite graph, then show that the map \( \Psi \) introduced in Definition 7.35 satisfies

\[
\Psi\left(\begin{bmatrix} G \end{bmatrix}\right) = \sum_{f : V \to \{1,2,3,...\}} x_f,
\]

where \( x_f := \prod_{v \in V} x_{f(v)} \). Here, \( [G]^f \) is defined as in Definition 7.27.

7.4. Example: The quasisymmetric function of a matroid. We introduce the matroid-minor Hopf algebra of Schmitt [144], and studied extensively by Crapo and Schmitt [34, 35, 36]. A very simple character \( \zeta \) on this Hopf algebra will then give rise, via the map \( \Psi \) from Theorem 7.3, to the quasisymmetric function invariant of matroids from the work of Billera, Jia and the second author [19].

7.4.1. The matroid-minor Hopf algebra. We begin by reviewing some notions from matroid theory; see Oxley [124] for background, undefined terms and unproven facts.

Definition 7.45. A matroid \( M \) of rank \( r \) on a (finite) ground set \( E \) is specified by a nonempty collection \( \mathcal{B}(M) \) of \( r \)-element subsets of \( E \) with the following exchange property:

For any \( B, B' \) in \( \mathcal{B}(M) \) and \( b \) in \( B \), there exists \( b' \) in \( B' \) with \( (B \setminus \{ b \}) \cup \{ b' \} \) in \( \mathcal{B}(M) \).

Example 7.46. A matroid \( M \) with ground set \( E \) is represented by a family of vectors \( S = (v_e)_{e \in E} \) in a vector space if \( \mathcal{B}(M) \) is the collection of subsets \( B \subset E \) having the property that the subfamily \( (v_e)_{e \in B} \) is a basis for the span of all of the vectors in \( S \).

For example, if \( M \) is the matroid with \( \mathcal{B}(M) = \{(a,b), (a,c), (a,d), (b,c), (b,d)\} \) on the ground set \( E = \{a,b,c,d\} \), then \( M \) is represented by the family \( S = (v_a, v_b, v_c, v_d) \) of the four vectors \( v_a = (1,0), v_b = (1,1), v_c = (0,1), v_d = (1,1) \) in \( \mathbb{R}^2 \) depicted here:

\[
\begin{array}{c}
\text{v_c, v_d} \\
\text{v_b} \\
\text{v_a}
\end{array}
\]

\[
\begin{array}{c}
\text{v_c, v_d} \\
\text{v_b} \\
\text{v_a}
\end{array}
\]
Conversely, whenever \( E \) is a finite set and \( S = (v_e)_{e \in E} \) is a family of vectors in a vector space, then the set
\[
\{ B \subset E : \text{the subfamily } (v_e)_{e \in B} \text{ is a basis for the span of all of the vectors in } S \}
\]
is a matroid on the ground set \( E \).

A matroid is said to be \textit{linear} if there exists a family of vectors in a vector space representing it. Not all matroids are linear, but many important ones are.

**Example 7.47.** A special case of matroids \( M \) represented by vectors are \textit{graphic matroids}, coming from a graph \( G = (V, E) \), with parallel edges and self-loops allowed. One represents these by vectors in \( \mathbb{R}^V \) with standard basis \( \{e_v\}_{v \in V} \) by associating the vector \( e_v - e_{v'} \) to any edge connecting a vertex \( v \) with a vertex \( v' \).

One can check (or see [124, §1.2]) that the bases \( B \) in \( \mathcal{B}(M) \) correspond to the edge sets of \textit{spanning forests} for \( G \), that is, edge sets which are acyclic and contain one spanning tree for each connected component of \( G \). For example, the matroid \( \mathcal{B}(M) \) corresponding to the graph \( G = (V, E) \) shown below:

![Graph](image)

is exactly the matroid represented by the vectors in Example 7.46; indeed, the spanning trees of this graph \( G \) are the edge sets \( \{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\} \).

To define the matroid-minor Hopf algebra one needs the basic matroid operations of \textit{deletion} and \textit{contraction}. These model the operations of deleting or contracting an edge in a graph. For configurations of vectors they model the deletion of a vector, or the passage to images in the quotient space modulo the span of a vector.

**Definition 7.48.** Given a matroid \( M \) of rank \( r \) and an element \( e \) of its ground set \( E \), say that \( e \) is loop (resp. coloop) of \( M \) if \( e \) lies in no basis (resp. every basis) \( B \) in \( \mathcal{B}(M) \). If \( e \) is not a coloop, the deletion \( M \setminus e \) is a matroid of rank \( r \) on ground set \( E \setminus \{e\} \) having bases
\[
(7.14) \quad \mathcal{B}(M \setminus e) := \{ B \in \mathcal{B}(M) : e \notin B \}.
\]

If \( e \) is not a loop, the contraction \( M/e \) is a matroid of rank \( r \) on ground set \( E \setminus \{e\} \) having bases
\[
(7.15) \quad \mathcal{B}(M/e) := \{ B \setminus \{e\} : e \in B \in \mathcal{B}(M) \}.
\]

When \( e \) is a loop of \( M \), then \( M/e \) has rank \( r \) instead of \( r - 1 \) and one defines its bases as in (7.14) rather than (7.15); similarly, if \( e \) is a coloop of \( M \) then \( M \setminus e \) has rank \( r - 1 \) instead of \( r \) and one defines its bases as in (7.15) rather than (7.14).

**Example 7.49.** Starting with the graph \( G \) and its graphic matroid \( M \) from Example 7.47, the deletion \( M \setminus a \) and contraction \( M/c \) correspond to the graphs \( G \setminus a \) and \( G/c \) shown here:

\[
G \setminus a = \quad G/c =
\]

One has
- \( \mathcal{B}(M \setminus a) = \{\{b, c\}, \{b, d\}\} \), so that \( b \) has become a coloop in \( M \setminus a \), and
- \( \mathcal{B}(M/c) = \{\{a\}, \{b\}\} \), so that \( d \) has become a loop in \( M/c \).

**Definition 7.50.** Deletions and contractions commute with each other, leading to well-defined operations for subsets \( A \subset E \) of the
- restriction \( M|_A \) on ground set \( A \), obtained by deleting all \( e \) of \( E \setminus A \) in any order, and
- quotient/contraction \( M/A \) on ground set \( E \setminus A \), obtained by contracting all \( e \) in \( A \) in any order.
We will also need the direct sum \( M_1 \oplus M_2 \), whose ground set \( E = E_1 \sqcup E_2 \) is the disjoint union of a copy of the ground sets \( E_1, E_2 \) for \( M_1, M_2 \), and having bases
\[
\mathcal{B}(M_1 \oplus M_2) := \{B_1 \cup B_2 : B_i \in \mathcal{B}(M_i) \text{ for } i = 1, 2\}.
\]

Lastly, say that two matroids \( M_1, M_2 \) are isomorphic if there is a bijection of their ground sets \( E_1 \xrightarrow{\varphi} E_2 \)
having the property that \( \varphi \mathcal{B}(M_1) = \mathcal{B}(M_2) \).

Now one can define the matroid-minor Hopf algebra, originally introduced by Schmitt [144, §15], and studied further by Crapo and Schmitt [34, 35, 36].

**Definition 7.51.** Let \( \mathcal{M} \) have \( k \)-basis elements \([M]\) indexed by isomorphism classes of matroids. Define the multiplication via
\[
[M_1] : [M_2] := [M_1 \oplus M_2]
\]
so that the class \([\emptyset]\) of the empty matroid \( \emptyset \) having empty ground set gives a unit. Define the comultiplication for \( M \) a matroid on ground set \( E \)
via
\[
\Delta[M] := \sum_{A \subseteq E} [M|_A] \otimes [M/A],
\]
and a counit
\[
\epsilon[M] := \begin{cases} 1 & \text{if } M = \emptyset \\ 0 & \text{otherwise.} \end{cases}
\]

**Proposition 7.52.** The above maps endow \( \mathcal{M} \) with the structure of a graded connected finite type Hopf algebra over \( k \), which is commutative.

**Proof.** Checking the unit and counit conditions are straightforward. Associativity and commutativity of the multiplication follow because the direct sum operation \( \oplus \) for matroids is associative and commutative up to isomorphism. Coassociativity follows because for a matroid \( M \) on ground set \( E \), one has this equality between the two candidates for \( \Delta^{(2)}[M] \)
\[
\sum_{\emptyset \subseteq A_1 \subseteq A_2 \subseteq E} [M|_{A_1}] \otimes [(M|_{A_2})/A_1] \otimes [M/A_2]
= \sum_{\emptyset \subseteq A_1 \subseteq A_2 \subseteq E} [M|_{A_1}] \otimes [(M|_{A_1})|_{A_2 \setminus A_1}] \otimes [M/A_2]
\]
due to the matroid isomorphism \( [M|_{A_2}/A_1] \cong (M/A_1)|_{A_2 \setminus A_1} \). Commutativity of the bialgebra diagram in (1.8) amounts to the fact that for a pair of matroids \( M_1, M_2 \) and subsets \( A_1, A_2 \) of their (disjoint) ground sets \( E_1, E_2 \), one has isomorphisms
\[
M_1|_{A_1} \oplus M_2|_{A_2} \cong (M_1 \oplus M_2)|_{A_1 \sqcup A_2},
M_1/A_1 \oplus M_2/A_2 \cong (M_1 \oplus M_2)/(A_1 \sqcup A_2).
\]
Letting \( \mathcal{M}_n \) be the \( k \)-span of \([M]\) for matroids whose ground set \( E \) has cardinality \(|E| = n\), one can then easily check that \( \mathcal{M} \) becomes a bialgebra which is graded, connected, and of finite type, hence also a Hopf algebra by Proposition 1.36. \( \square \)

### 7.4.2. A quasisymmetric function for matroids.

**Definition 7.53.** Define a character \( \mathcal{M} \xrightarrow{\zeta} k \) by
\[
\zeta[M] = \begin{cases} 1 & \text{if } M \text{ has only one basis,} \\ 0 & \text{otherwise.} \end{cases}
\]
It is easily checked that this is a character, that is, an algebra map \( \mathcal{M} \xrightarrow{\zeta} k \). Note that if \( M \) has only one basis, say \( \mathcal{B}(M) = \{B\} \), then \( B := \text{coloops}(M) \) is the set of coloops of \( M \), and \( E \setminus B = \text{loops}(M) \) is the set of loops of \( M \). Equivalently, \( M = \bigoplus_{e \in E} M|_{\{e\}} \) is the direct sum of matroids each having one element, each a coloop or loop.

Define \( \Psi[M] \) for a matroid \( M \) to be the image of \([M]\) under the map \( \mathcal{M} \xrightarrow{\Psi} \text{QSym} \) induced via Theorem 7.3 from the above character \( \zeta \).
It turns out that $\Psi[M]$ is intimately related with greedy algorithms and finding minimum cost bases. A fundamental property of matroids (and one that characterizes them, in fact; see [124, §1.8]) is that no matter how one assigns costs $f : E \to \mathbb{R}$ to the elements of $E$, the following greedy algorithm (generalizing Kruskal’s algorithm for finding minimum cost spanning trees) always succeeds in finding one basis $B$ in $\mathcal{B}(M)$ achieving the minimum total cost $f(B) := \sum_{b \in B} f(b)$:

**Algorithm 7.54.** Start with the empty subset $I_0 = \emptyset$ of $E$. For $j = 1, 2, \ldots, r$, having already defined the set $I_{j-1}$, let $e$ be the element of $E \setminus I_{j-1}$ having the lowest cost $f(e)$ among all those for which $I_{j-1} \cup \{e\}$ is independent, that is, still a subset of at least one basis $B$ in $\mathcal{B}(M)$. Then define $I_j := I_{j-1} \cup \{e\}$. Repeat this until $j = r$, and $B = I_r$ will be among the bases that achieve the minimum cost.

**Definition 7.55.** Say that a cost function $f : E \to \{1, 2, \ldots\}$ is $M$-generic if there is a unique basis $B$ in $\mathcal{B}(M)$ achieving the minimum cost $f(B)$.

**Example 7.56.** For the graphic matroid $M$ of Example 7.47, this cost function $f_1 : E \to \{1, 2, \ldots\}$

$$f_1(a)=1 \quad f_1(b)=3 \quad f_1(c)=3 \quad f_1(d)=2$$

is $M$-generic, as it minimizes uniquely on the basis $\{a, d\}$, whereas this cost function $f_2 : E \to \{1, 2, \ldots\}$

$$f_2(a)=1 \quad f_2(b)=3 \quad f_2(c)=2 \quad f_2(d)=2$$

is not $M$-generic, as it achieves its minimum value on the two bases $\{a, c\}, \{a, d\}$.

**Proposition 7.57.** For a matroid $M$ on ground set $E$, one has this expansion\footnote{In fact, this expansion was the original definition of $\Psi[M]$ in [19, Defn. 1.1].}

$$\Psi[M] = \sum_{M\text{-generic} \atop f : E \to \{1, 2, \ldots\}} x_f$$

where $x_f := \prod_{e \in E} x_{f(e)}$. In particular, for $m \geq 0$, its specialization $ps^1$ from Definition 7.6 has this interpretation:

$$ps^1\Psi[M](m) = |\{M\text{-generic } f : E \to \{1, 2, \ldots, m\}\}|.$$

**Proof.** The iterated coproduct $\Delta^{(\ell-1)} \to \mathcal{M}^{\otimes \ell}$ sends

$$[M] \mapsto \sum [M|_{A_1}] \otimes [(M|_{A_2})/A_1] \otimes \cdots \otimes [(M|_{A_\ell})/A_{\ell-1}]$$

where the sum is over flags of nested subsets

\begin{equation}
\emptyset = A_0 \subseteq A_1 \subseteq \cdots \subseteq A_{\ell-1} \subseteq A_\ell = E.
\end{equation}

The map $\zeta^{\otimes \ell}$ sends each summand to 1 or 0, depending upon whether each $(M|_{A_j})/A_{j-1}$ has a unique basis or not. Thus formula (7.3) shows that the coefficient $\zeta_{\alpha}$ of $x_{i_1}^{\alpha_1} \cdots x_{i_\ell}^{\alpha_\ell}$ in $\Psi[M]$ counts the flags of subsets in (7.16) for which $|A_j \setminus A_{j-1}| = \alpha_j$ and $(M|_{A_j})/A_{j-1}$ has a unique basis, for each $j$.

Given a flag as in (7.16), associate the cost function $f : E \to \{1, 2, \ldots\}$ whose value on each element of $A_j \setminus A_{j-1}$ is $i_j$; conversely, given any cost function $f$, say whose distinct values are $i_1 < \ldots < i_\ell$, one associates the flag having $A_j \setminus A_{j-1} = f^{-1}(i_j)$ for each $j$.\hfill\square
We will prove below, using induction on \( s = 0, 1, 2, \ldots, \ell \) the following claim: After having completed \( \alpha_1 + \alpha_2 + \cdots + \alpha_s \) steps in the greedy algorithm (Algorithm 7.54), there is a unique choice for the independent set produced thus far, namely

\[
I_{\alpha_1 + \alpha_2 + \cdots + \alpha_s} = \bigcup_{j=1}^{s} \text{coloops}((M|_{A_j})/A_{j-1}),
\]

if and only if each of the matroids \((M|_{A_j})/A_{j-1}\) for \( j = 1, 2, \ldots, s \) has a unique basis.

The case \( s = \ell \) in this claim would show what we want, namely that \( f \) is \( M \)-generic, minimizing uniquely on the basis shown in (7.17) with \( s = \ell \), if and only if each \((M|_{A_j})/A_{j-1}\) has a unique basis.

The assertion of the claim is trivially true for \( s = 0 \). In the inductive step, one may assume that

- the independent set \( I_{\alpha_1 + \alpha_2 + \cdots + \alpha_{s-1}} \) takes the form in (7.17), replacing \( s \) by \( s - 1 \),
- it is the unique \( f \)-minimizing basis for \((M|_{A_{s-1}})\), and
- \((M|_{A_j})/A_{j-1}\) has a unique basis for \( j = 1, 2, \ldots, s - 1 \).

Since \( A_{s-1} \) exactly consists of all of the elements \( e \) of \( E \) whose costs \( f(e) \) lie in the range \( \{i_1, i_2, \ldots, i_{s-1}\} \), in the next \( \alpha_s \) steps the algorithm will work in the quotient matroid \((M/A_{s-1})\) and attempt to augment \( I_{\alpha_1 + \alpha_2 + \cdots + \alpha_{s-1}} \) using the next-cheapest elements, namely the elements of \( A_s \setminus A_{s-1} \), which all have cost \( f \) equal to \( i_s \). Thus the algorithm will have no choices about how to do this augmentation if and only if \((M|_{A_s})/A_{s-1}\) has a unique basis, namely its set of coloops, in which case the algorithm will choose to add all of these coloops, giving \( I_{\alpha_1 + \alpha_2 + \cdots + \alpha_s} \) as described in (7.17). This completes the induction.

The last assertion follows from Proposition 7.7.

\[ \square \]

Example 7.58. If \( M \) has one basis then every function \( f : E \to \{1, 2, \ldots\} \) is \( M \)-generic, and

\[
\Psi[M] = \sum_{f:E \to \{1,2,\ldots\}} x_f = (x_1 + x_2 + \cdots)^{|E|} = M_{(1)}^{|E|}.
\]

Example 7.59. Let \( U_{r,n} \) denote the uniform matroid of rank \( r \) on \( n \) elements \( E \), having \( B(U_{r,n}) \) equal to all of the \( r \)-element subsets of \( E \).

As \( U_{1,2} \) has \( E = \{1, 2\} \) and \( B = \{\{1\}, \{2\}\} \), genericity means \( f(1) \neq f(2) \), so

\[
\Psi[U_{1,2}] = \sum_{(f(1), f(2)): f(1) \neq f(2)} x_{f(1)}x_{f(2)} = x_1x_2 + x_2x_1 + x_1x_3 + x_3x_1 + \cdots = 2M_{(1,1)}.
\]

Similarly \( U_{1,3} \) has \( E = \{1, 2, 3\} \) with \( B = \{\{1\}, \{2\}, \{3\}\} \), and genericity means either that \( f(1), f(2), f(3) \) are all distinct, or that two of them are the same and the third is smaller. This shows

\[
\Psi[U_{1,3}] = 3\sum_{i<j} x_i x_j^2 + 6\sum_{i<j<k} x_i x_j x_k
\]

\[
= 3M_{(1,2)} + 6M_{(1,1,1)}
\]

\[
\psi^3\Psi[U_{1,3}](m) = 3\binom{m}{2} + 6\binom{m}{3} = \frac{m(m-1)(2m-1)}{2}
\]

One can similarly analyze \( U_{2,3} \) and check that

\[
\Psi[U_{2,3}] = 3M_{(2,1)} + 6M_{(1,1,1)}
\]

\[
\psi^3\Psi[U_{2,3}](m) = 3\binom{m}{2} + 6\binom{m}{3} = \frac{m(m-1)(2m-1)}{2}
\]

These last examples illustrate the behavior of \( \Psi \) under the duality operation on matroids.

Definition 7.60. Given a matroid \( M \) of rank \( r \) on ground set \( E \), its dual or orthogonal matroid \( M^\perp \) is a matroid of rank \( |E| - r \) on the same ground set \( E \), having

\[
B(M^\perp) := \{E \setminus B\}_{B \in B(M)}.
\]

Here are a few examples of dual matroids.

Example 7.61. The dual of a uniform matroid is another uniform matroid:

\[
U_{r,n}^\perp = U_{n-r,n}.
\]
Example 7.62. If $M$ is matroid of rank $r$ represented by family of vectors $\{e_1, \ldots, e_n\}$ in a vector space over some field $k$, one can find a family of vectors $\{e_1^\perp, \ldots, e_n^\perp\}$ that represent $M^\perp$ in the following way. Pick a basis for the span of the vectors $\{e_i\}_{i=1}^n$, and create a matrix $A$ in $k^{r \times n}$ whose columns express the $e_i$ in terms of this basis. Then pick any matrix $A^\perp$ whose row space is the null space of $A$, and one finds that the columns $\{e_i^\perp\}_{i=1}^n$ of $A^\perp$ represent $M^\perp$. See Oxley [124, §2.2].

Example 7.63. Let $G = (V, E)$ be a graph embedded in the plane with edge set $E$, giving rise to a graphic matroid $M$ on ground set $E$. Let $G^\perp$ be a planar dual of $G$, so that, in particular, for each edge $e$ in $E$, the graph $G^\perp$ has one edge $e^\perp$, crossing $e$ transversely. Then the graphic matroid of $G^\perp$ is $M^\perp$. See Oxley [124, §2.3].

Proposition 7.64. If $\Psi[M] = \sum_\alpha c_\alpha M_\alpha$, then $\Psi[M^\perp] = \sum_\alpha c_\alpha \Psi[M^\perp] - \sum_\alpha c_\alpha M_{\alpha}^\perp$.

Proof. This amounts to showing that for any composition $\alpha = (\alpha_1, \ldots, \alpha_\ell)$, the cardinality of the set of $M$-generic $f$ having $x_f = x^\alpha$ is the same as the cardinality of the set of $M^\perp$-generic $f^\perp$ having $x_{f^\perp} = x^{\psi(\alpha)}$. We claim that the map $f \mapsto f^\perp$ in which $f^\perp(e) = \ell + 1 - f(e)$ gives a bijection between these sets. To see this, note that any basis $B$ of $M$ satisfies

$$f(B) + f(E \setminus B) = \sum_{e \in E} f(e)$$

$$f(E \setminus B) + f^\perp(E \setminus B) = (\ell + 1)(|E| - r),$$

where $r$ denotes the rank of $M$. Thus $B$ is $f$-minimizing if and only if $E \setminus B$ is $f$-maximizing (by (7.18)) if and only if $E \setminus B$ is $f^\perp$-minimizing (by (7.19)). Consequently $f$ is $M$-generic if and only if $f^\perp$ is $M^\perp$-generic.

The last assertion follows, for example, from the calculation in Proposition 7.7(i) that $\psi^1(M_\alpha)(m) = \binom{m}{\ell(\alpha)}$ together with the fact that $\ell(\psi(\alpha)) = \ell(\alpha)$.\hfill \(\square\)

Just as (7.13) showed that Stanley’s chromatic symmetric function of a graph has an expansion as a sum of $P$-partition enumerators for certain strictly labelled posets $P$, the same holds for $\Psi[M]$.

Definition 7.65. Given a matroid $M$ on ground set $E$, and a basis $B$ in $B(M)$, define the base-cobase poset $P_B$ to have $b < b'$ whenever $b$ lies in $B$ and $b'$ lies in $E \setminus B$ and $(B \setminus \{b\}) \cup \{b'\}$ is in $B(M)$.

Proposition 7.66. For any matroid $M$, one has $\Psi[M] = \sum_{B \in B(M)} F_{(P_B, \text{strict})}(\mathbf{x})$ where $F_{(P, \text{strict})}(\mathbf{x})$ for a poset $P$ means the $P$-partition enumerator for any strict labelling of $P$, i.e. a labelling such that the $P$-partitions satisfy $f(i) < f(j)$ whenever $i <_P j$.

In particular, $\Psi[M]$ expands nonnegatively in the $\{L_\alpha\}$ basis.

Proof. A basic result about matroids, due to Edmonds [47], describes the edges in the matroid base polytope which is the convex hull of all vectors $\{\sum_{b \in B} e_b\}_{B \in B(M)}$ inside $\mathbb{R}^E$ with standard basis $\{e_e\}_{e \in E}$. He shows that all such edges connect two bases $B, B'$ that differ by a single basis exchange, that is, $B' = (B \setminus \{b\}) \cup \{b'\}$ for some $b$ in $B$ and $b'$ in $E \setminus B$.

Polyhedral theory then says that a cost function $f$ on $E$ will minimize uniquely at $B$ if and only if one has a strict increase $f(B) < f(B')$ along each such edge $B \to B'$ emanating from $B$, that is, if and only if $f(b) < f(b')$ whenever $b <_P b'$ in the base-cobase poset $P_B$, that is, $f$ lies in $A(P_B, \text{strict})$.\hfill \(\square\)

Example 7.67. The graphic matroid from Example 7.47 has this matroid base polytope, with the bases $B$ in $B(M)$ labelling the vertices:

![Diagram](image-url)
The base-cobase posets $P_B$ for its five vertices $B$ are as follows:

$$
\begin{array}{c|c|c|c|c}
 a & b & c & d \\
\hline
 b & d & a & c \\
 a & d & b & c \\
 c & d & a & b \\
\end{array}
$$

One can label the first of these five strictly as $1 \ 2 \ 3 \ 4$ and compute its strict $P$-partition enumerator from the linear extensions \{3412, 3421, 4312, 4321\} as

$$
L(2, 2) + L(2, 1, 1) + L(1, 1, 2) + L(1, 1, 1, 1)
$$

while any of the last four can be labelled strictly as $1 \ 2 \ 3 \ 4$ and they each have an extra linear extension 3142 giving their strict $P$-partition enumerators as

$$
L(2, 2) + L(2, 1, 1) + L(1, 1, 2) + L(1, 1, 1, 1) + L(1, 2, 1)
$$

Hence one has

$$
\Psi[M] = 5L(2, 2) + 5L(1, 1, 2) + 4L(1, 2, 1) + 5L(2, 1, 1) + 5L(1, 1, 1, 1)
$$

As $M$ is a graphic matroid for a self-dual planar graph, one has a matroid isomorphism $M \cong M^\perp$ (see Example 7.63), reflected in the fact that $\Psi[M]$ is invariant under the symmetry swapping $M_\alpha \leftrightarrow M_{\text{rev}(\alpha)}$ (and simultaneously swapping $L_\alpha \leftrightarrow L_{\text{rev}(\alpha)}$).

This $P$-partition expansion for $\Psi[M]$ also allows us to identify its image under the antipode of QSym.

**Proposition 7.68.** For a matroid $M$ on ground set $E$, one has

$$
S(\Psi[M]) = (-1)^{|E|} \sum_{f:E \rightarrow \{1, 2, \ldots\}} |\{f\text{-maximizing bases } B\}| \cdot x_f
$$

and

$$
\text{ps}^1 \Psi[M](-m) = (-1)^{|E|} \sum_{f:E \rightarrow \{1, 2, \ldots, m\}} |\{f\text{-maximizing bases } B\}|.
$$

In particular, the expected number of $f$-maximizing bases among all cost functions $f : E \rightarrow \{1, 2, \ldots, m\}$ is $(-m)^{-|E|} \text{ps}^1 \Psi[M](-m)$.

**Proof.** Corollary 5.27 implies

$$
S(\Psi[M]) = \sum_{B \in B(M)} S(F(P_B, \text{strict})(x)) = (-1)^{|E|} \sum_{B \in B(M)} F(P_B^{\text{natural}})(x)
$$

where $F(P, \text{natural})(x)$ is the enumerator for $P$-partitions in which $P$ has been naturally labelled, so that they satisfy $f(i) \leq f(j)$ whenever $i <_P j$. When $P = P_B^{\text{natural}}$, this is exactly the condition for $f$ to achieve its maximum value at $f(B)$ (possibly not uniquely), that is, for $f$ to lie in the closed normal cone to the vertex indexed by $B$ in the matroid base polytope; compare this with the discussion in the proof of Proposition 7.66. Thus one has

$$
S(\Psi[M]) = (-1)^{|E|} \sum_{\substack{(B, f): B \in B(M) \\text{maximizing at } B \text{ maximizing at } B}} x_f
$$

which agrees with the statement of the proposition, after reversing the order of the summation.

The rest follows from Proposition 7.7. \qed
Example 7.69. We saw in Example 7.67 that the matroid $M$ from Example 7.47 has
$$\Psi[M] = 5L_{(2,2)} + 5L_{(1,1,2)} + 4L_{(1,2,1)} + 5L_{(2,1,1)} + 5L_{(1,1,1,1)},$$
and therefore will have
$$\text{ps}^1(\Psi[M])(m) = 5 \binom{m-2+4}{4} + (5+4+5) \binom{m-3+4}{4} + 5 \binom{m-4+4}{4} = \frac{m(m-1)(2m^2-2m+1)}{2},$$
using $\text{ps}^1(L_\alpha)(m) = \binom{m-|\alpha|}{|\alpha|}$ from Proposition 7.7 (i). Let us first do a reality-check on a few of its values with $m \geq 0$ using Proposition 7.57, and for negative $m$ using Proposition 7.68:

<table>
<thead>
<tr>
<th>$m$</th>
<th>$-1$</th>
<th>$0$</th>
<th>$1$</th>
<th>$2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{ps}^1(\Psi[M])(m)$</td>
<td>$5$</td>
<td>$0$</td>
<td>$0$</td>
<td>$5$</td>
</tr>
</tbody>
</table>

When $m = 0$, interpreting the set of cost functions $f : E \to \{1, 2, \ldots, m\}$ as being empty explains why the value shown is 0. When $m = 1$, there is only one function $f : E \to \{1\}$, and it is not $M$-generic; any of the 5 bases in $B(M)$ will minimize $f(B)$, explaining both why the value for $m = 1$ is 0, but also explaining the value of 5 for $m = -1$. The value of 5 for $m = 2$ counts these $M$-generic cost functions $f : E \to \{1, 2\}$:

Lastly, Proposition 7.68 predicts the expected number of $f$-minimizing bases for $f : E \to \{1, 2, \ldots, m\}$ as
$$(-m)^{-|E|} \text{ps}^1(\Psi[M])(-m) = (-m)^{-4} \frac{m(m+1)(2m^2+2m+1)}{2} = \frac{(m+1)(2m^2+2m+1)}{2m^3},$$
whose limit as $m \to \infty$ is 1, consistent with the notion that “most” cost functions should be generic with respect to the bases of $M$, and maximize/minimize on a unique basis.

Remark 7.70. It is not coincidental that there is a similarity of results for Stanley’s chromatic symmetric function of a graph $\Psi[G]$ and for the matroid quasisymmetric function $\Psi[M]$, such as the $P$-partition expansions (7.13) versus Proposition 7.66, and the reciprocity results Proposition 7.42 versus Proposition 7.68. It was noted in [19, §9] that one can associate a similar quasisymmetric function invariant to any generalized permutohedra in the sense of Postnikov [132]. Furthermore, recent work of Ardila and Aguiar [3] has shown that there is a Hopf algebra of such generalized permutohedra, arising from a Hopf monoid in the sense of Aguiar and Mahajan [6]. This Hopf algebra generalizes the chromatic Hopf algebra of graphs and the matroid-minor Hopf algebra, and its quasisymmetric function invariant derives as usual from Theorem 7.3. Their work [3] also provides a generalization of the chromatic Hopf algebra antipode formula of Humpert and Martin [78] discussed in Remark 7.23 above.
8. The Malvenuto-Reutenauer Hopf algebra of permutations

Like so many Hopf algebras we have seen, the Malvenuto-Reutenauer Hopf algebra $\text{FQSym}$ can be thought of fruitfully in more than one way. One is that it gives a natural noncommutative lift of the quasisymmetric $P$-partition enumerators and the fundamental basis $\{L_\alpha\}$ of QSym, rendering their product and coproduct formulas even more natural.

8.1. Definition and Hopf structure.

**Definition 8.1.** Define $\text{FQSym} = \bigoplus_{n \geq 0} \text{FQSym}_n$ to be a graded $k$-module in which $\text{FQSym}_n$ has $k$-basis $\{F_w\}_{w \in \mathfrak{S}_n}$ indexed by the permutations $w = (w_1, \ldots, w_n)$ in $\mathfrak{S}_n$.

We first attempt to lift the product and coproduct formulas (5.8), (5.7) in the $\{L_\alpha\}$ basis of QSym. We attempt to define a product for $u \in \mathfrak{S}_k, v \in \mathfrak{S}_\ell$ as follows:

\begin{equation}
F_u F_v := \sum_{w \in u \cup v[k]} F_w
\end{equation}

where we regard permutations as words (namely, every $\pi \in \mathfrak{S}_n$ is identified with the word $((\pi_1, \pi_2, \ldots, \pi_n))$, and where for $v = (v_1, \ldots, v_\ell)$ one sets $v[k] := (k + v_1, \ldots, k + v_\ell)$. Note that the multiset $u \cup v[k]$ is an actual set in this situation (i.e., has each element appear only once) and is a subset of $S_{k+\ell}$.

The coproduct will be defined using the notation of standardization $\text{std}(i)$ of a word $i = (i_1, \ldots, i_n)$ in some linearly ordered alphabet. This standardization is the permutation in $\mathfrak{S}_n$ obtained by replacing all the occurrences of the smallest letter in $i$ by the numbers $1, 2, \ldots, m_1$ from left to right (where $m_1$ is the number of times this smallest letter occurs in $i$), then replacing all occurrences of the next smallest letter by the numbers $m_1 + 1, m_1 + 2, \ldots, m_1 + m_2$, from left to right (where $m_2$ is how often this letter occurs in $i$), etc.\footnote{Alternatively, one can define $\text{std}(i)$ as the unique permutation $\sigma \in \mathfrak{S}_n$ such that for every $1 \leq a < b \leq n$, we have $\sigma(a) < \sigma(b)$ if and only if $i_a \leq i_b$.}

**Example 8.2.** Considering words in the Roman alphabet $a < b < c < \cdots$

\[
\text{std}(b\ a\ c\ c\ b\ a\ a\ b\ a\ c) = (5\ 1\ 9\ 10\ 6\ 2\ 3\ 7\ 4\ 11\ 8).
\]

Using this, define for $w = (w_1, \ldots, w_n)$ in $\mathfrak{S}_n$

\begin{equation}
\Delta F_w := \sum_{k=0}^n F_{\text{std}(w_1, w_2, \ldots, w_k)} \otimes F_{\text{std}(w_{k+1}, w_{k+2}, \ldots, w_n)}.
\end{equation}

It is possible to check directly that the maps defined in (8.1) and (8.2) endow $\text{FQSym}$ with the structure of a graded connected finite type Hopf algebra; see Hazewinkel, Gubaren, Kirichenko [69, Thm. 7.1.8]. However in justifying this here, we will follow the approach of Duchamp, Hivert and Thibon [44, §3], which exhibits $\text{FQSym}$ as a subalgebra of a larger ring of (noncommutative) power series of bounded degree in a totally ordered alphabet.
Definition 8.3. Given a totally ordered set \( I \), create a totally ordered variable set \( \{x_i\}_{i \in I} \), and the ring \( R(\{x_i\}_{i \in I}) \) of noncommutative power series of bounded degree in this alphabet.\(^{268}\) Many times, we will use a variable set \( X := (x_1 < x_2 < \cdots) \), and call the ring \( R(X) \).

We first identify the algebra structure for \( FQSym \) as the subalgebra of finite type within \( R(\{x_i\}_{i \in I}) \) spanned by the elements

\[
F_w = F_w(\{x_i\}_{i \in I}) := \sum_{i=(i_1, \ldots, i_n): \text{std}(i) = w^{-1}} x_i
\]

where \( x_i := x_{i_1} \cdots x_{i_n} \), as \( w \) ranges over \( \bigcup_{n \geq 0} \mathcal{S}_n \).

Example 8.4. For the alphabet \( X = (x_1 < x_2 < \cdots) \), in \( R(X) \) one has

\[
F_1 = \sum_{1 \leq i} x_i = x_1 + x_2 + \cdots
\]

\[
F_{12} = \sum_{1 \leq i < j} x_i x_j = x_1^2 + x_2^2 + \cdots + x_1 x_2 + x_1 x_3 + x_2 x_3 + x_1 x_4 + \cdots
\]

\[
F_{21} = \sum_{1 \leq i < j} x_j x_i = x_2 x_1 + x_3 x_1 + x_3 x_2 + x_4 x_1 + \cdots
\]

\[
F_{312} = \sum_{\text{std}(i) = 231} x_i = \sum_{1 \leq i < j < k} x_j x_k x_i
\]

\[
= x_2^2 x_1 + x_3^2 x_1 + x_3^2 x_2 + \cdots + x_2 x_3 x_1 + x_2 x_4 x_1 + \cdots
\]

Proposition 8.5. For any totally ordered infinite set \( I \), the elements \( \{F_w\} \) as \( w \) ranges over \( \bigcup_{n \geq 0} \mathcal{S}_n \) form the \( k \)-basis for a subalgebra \( FQSym(\{x_i\}_{i \in I}) \) of \( R(X) \), which is graded connected and of finite type, having multiplication defined \( k \)-linearly by (8.1).

Consequently all such algebras are isomorphic to a single algebra \( FQSym \), having basis \( \{F_w\} \) and multiplication given by the rule (8.1), with the isomorphism mapping \( F_w \mapsto F_w(\{x_i\}_{i \in I}) \).

For example,

\[
F_1 F_{21} = (x_1 + x_2 + x_3 + \cdots)(x_2 x_1 + x_3 x_1 + x_3 x_2 + x_4 x_1 + \cdots)
\]

\[
= x_1 x_3 x_2 + x_1 x_3 x_2 + \cdots + x_1 x_2 x_1 + x_2 x_3 x_2 + x_2 x_4 x_1 + \cdots
\]

\[
= x_2 x_3 x_1 + x_2 x_4 x_1 + \cdots
\]

\[
= \sum_{\text{std}(i) = 132} x_i + \sum_{\text{std}(i) = 231} x_i + \sum_{\text{std}(i) = 321} x_i = F_{132} + F_{312} + F_{321} = \sum_{w \in 1 \cup 32} F_w
\]

Proof of Proposition 8.5. The elements \( \{F_w(\{x_i\}_{i \in I})\} \) are linearly independent as they are supported on disjoint monomials, and so form a \( k \)-basis for their span. The fact that they multiply via rule (8.1) is the equivalence of conditions (i) and (iii) in the following Lemma 8.6, from which all the remaining assertions follow.

\(^{268}\)Let us recall the definition of \( R(\{x_i\}_{i \in I}) \).

Let \( N \) denote the free monoid on the alphabet \( \{x_i\}_{i \in I} \); it consists of words \( x_{i_1} x_{i_2} \cdots x_{i_k} \). We define a topological \( k \)-module \( k \langle\langle \{x_i\}_{i \in I}\rangle\rangle \) to be the Cartesian product \( k^N \) (equipped with the product topology), but we identify its element \( (\delta_w, u)_{w \in N} \in k^N = k \langle\langle \{x_i\}_{i \in I}\rangle\rangle \) can be rewritten as the convergent sum \( \sum_{w \in N} \lambda_w w \). We call \( \lambda_w \) the coefficient of \( w \) in this element (or the coefficient of this element before \( w \)). The elements of \( k \langle\langle \{x_i\}_{i \in I}\rangle\rangle \) will be referred to as noncommutative power series. We define a multiplication on \( k \langle\langle \{x_i\}_{i \in I}\rangle\rangle \) by the formula

\[
\left( \sum_{w \in N} \lambda_w w \right) \left( \sum_{w \in N} \mu_w w \right) = \sum_{w \in N} \sum_{(u,v) \in N^2, uv = w} \lambda_u \mu_v \ w.
\]

(This is well-defined thanks to the fact that, for each \( w \in N \), there are only finitely many \( (u,v) \in N^2 \) satisfying \( w = uv \).) Thus, \( k \langle\langle \{x_i\}_{i \in I}\rangle\rangle \) becomes a \( k \)-algebra with unity 1 (the empty word). (It is similar to the monoid algebra \( kN \) of \( N \) over \( k \), with the only difference that infinite sums are allowed.)

Now, we define \( R(\{x_i\}_{i \in I}) \) to be the \( k \)-subalgebra of \( k \langle\langle \{x_i\}_{i \in I}\rangle\rangle \) consisting of all noncommutative power series \( \sum_{w \in N} \lambda_w w \in k \langle\langle \{x_i\}_{i \in I}\rangle\rangle \) of bounded degree (i.e., such that all words \( w \in N \) of sufficiently high length satisfy \( \lambda_w = 0 \)).
Lemma 8.6. For a triple of permutations 
\[ u = (u_1, \ldots, u_k) \text{ in } \mathfrak{S}_k, \]
\[ v = (v_1, \ldots, v_{n-k}) \text{ in } \mathfrak{S}_{n-k}, \]
\[ w = (w_1, \ldots, w_n) \text{ in } \mathfrak{S}_n, \]
the following conditions are equivalent:

(i) \( w^{-1} \) lies in the set \( u^{-1} \sqcup v^{-1}[k] \).

(ii) \( u = \text{std}(w_1, \ldots, w_k) \) and \( v = \text{std}(w_{k+1}, \ldots, w_n) \).

(iii) for some word \( i = (i_1, \ldots, i_n) \) with \( w = \text{std}(i) = w \) one has \( u = \text{std}(i_1, \ldots, i_k) \) and \( v = \text{std}(i_{k+1}, \ldots, i_n) \).

Proof of Lemma 8.6. The implication (ii) \( \Rightarrow \) (iii) is clear since \( \text{std}(w) = w \). The reverse implication (iii) \( \Rightarrow \) (ii) is best illustrated by example, e.g., considering Example 8.2 as concatenated, with \( n = 11, k = 6, n-k = 5 \):

\[
\begin{array}{c|c}
\text{w} & \text{v} \\
\hline
(5 & 1 & 9 & 10 & 6 & 2) & (3 & 7 & 4 & 11 & 8) \\
\end{array}
\]

The equivalence of (i) and (ii) is a fairly standard consequence of unique parabolic factorization \( W = W^J W_J \) where \( W = \mathfrak{S}_n \) and \( W_J = \mathfrak{S}_k \times \mathfrak{S}_{n-k} \), so that \( W^J \) are the minimum-length coset representatives for cosets \( xW_J \) (that is, the permutations \( x \in \mathfrak{S}_n \) satisfying \( x_1 < \cdots < x_k \) and \( x_{k+1} < \cdots < x_n \)). One can uniquely express any \( w \) in \( W \) as \( w = xy \) with \( x \) in \( W^J \) and \( y \in W_J \), which here means that \( y = u \cdot v[k] = v[k] \cdot u \) for some \( u \) in \( \mathfrak{S}_k \) and \( v \) in \( \mathfrak{S}_{n-k} \). Therefore \( w = xw[k] \), if and only if \( w^{-1} = u^{-1}v^{-1}[k]x^{-1} \), which means that \( w^{-1} \) is the shuffle of the sequences \( u^{-1} \) in positions \( \{x_1, \ldots, x_k\} \) and \( v^{-1}[k] \) in positions \( \{x_{k+1}, \ldots, x_n\} \). \( \square \)

Example 8.7. To illustrate the equivalence of (i) and (ii) and the parabolic factorization in the preceding proof, let \( n = 9 \) and \( k = 5 \) with

\[
\begin{align*}
w &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 9 & 6 & 1 & 5 & 8 & 2 & 3 & 7 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 4 & 5 & 6 & 9 & 2 & 3 & 7 & 8 \end{pmatrix} \cdot \begin{pmatrix} 2 & 5 & 4 & 1 & 3 \\ 9 & 6 & 7 & 8 \end{pmatrix} \\
&= x \cdot u \cdot v[k]
\end{align*}
\]

\[
w^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 9 & 6 & 1 & 5 & 8 & 2 & 3 & 7 \end{pmatrix}
\]

\[
= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 1 & 5 & 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 7 & 8 & 9 & 6 \\ 1 & 6 & 7 & 2 & 3 & 4 & 8 & 9 & 5 \end{pmatrix} \\
= u^{-1} \cdot v^{-1}[k] \cdot x^{-1}
\]

Proposition 8.5 yields that \( \text{FQSym} \) is isomorphic to the \( \mathfrak{k} \)-subalgebra \( \text{FQSym}(\mathbf{X}) \) of the \( \mathfrak{k} \)-algebra \( \mathfrak{R} \langle \mathbf{X} \rangle \) when \( \mathbf{X} \) is the variable set \( \{X_1 < X_2 < \cdots \} \). We identify \( \text{FQSym} \) with \( \text{FQSym}(\mathbf{X}) \) along this isomorphism. For any infinite alphabet \( \{X_i\}_{i \in I} \) and any \( f \in \text{FQSym} \), we denote by \( f(\{X_i\}_{i \in I}) \) the image of \( f \) under the algebra isomorphism \( \text{FQSym} \to \text{FQSym}(\{X_i\}_{i \in I}) \) defined in Proposition 8.5.

One can now use this to define a coalgebra structure on \( \text{FQSym} \). Roughly speaking, one wants to first evaluate an element \( f \) in \( \text{FQSym} \cong \text{FQSym}(\mathbf{X}) \cong \text{FQSym}(\mathbf{X}, \mathbf{Y}) \) as \( f(\mathbf{X}, \mathbf{Y}) \), using the linearly ordered variable set \( (\mathbf{X}, \mathbf{Y}) := (X_1 < X_2 < \cdots Y_1 < Y_2 < \cdots) \). Then one should take the image of \( f(\mathbf{X}, \mathbf{Y}) \) after imposing the partial commutativity relations

\[
(8.3) \quad X_i Y_j = Y_j X_i \quad \text{for every pair } (X_i, Y_j) \in \mathbf{X} \times \mathbf{Y},
\]

and hope that this image lies in a subalgebra isomorphic to

\[
\text{FQSym}(\mathbf{X}) \otimes \text{FQSym}(\mathbf{Y}) \cong \text{FQSym} \otimes \text{FQSym}.
\]
We argue this somewhat carefully. Start by considering the canonical monoid epimorphism

(8.4) \[ F(X, Y) \xrightarrow{\rho} M, \]

where \( F(X, Y) \) denotes the free monoid on the alphabet \((X, Y)\) and \( M \) denotes its quotient monoid imposing the partial commutativity relations (8.3). Let \( k^M \) denote the \( k \)-module of all functions \( f : M \to k \), with pointwise addition and scalar multiplication; similarly define \( k^{F(X, Y)} \). As both monoids \( F(X, Y) \) and \( M \) enjoy the property that an element \( m \) has only finitely many factorizations as \( m = m_1m_2 \), one can define a convolution algebra structure on both \( k^{F(X, Y)} \) and \( k^M \) via

\[
(f_1 * f_2)(m) = \sum_{(m_1, m_2) \in N \times N} f_1(m_1)f_2(m_2),
\]

where \( N \) is respectively \( F(X, Y) \) or \( M \). As fibers of the map \( \rho \) in (8.4) are finite, it induces a map of convolution algebras, which we also call \( \rho \):

(8.5) \[ k^{F(X, Y)} \xrightarrow{\rho} k^M. \]

Now recall that \( R(X) \) denotes the algebra of noncommutative formal power series in the variable set \( X \), of bounded degree, with coefficients in \( k \). One similarly has the ring \( R(X, Y) \), which can be identified with the subalgebra of \( k^{F(X, Y)} \) consisting of the functions \( f : F(X, Y) \to k \) having a bound on the length of the words in their support (the value of \( f \) on a word in \((X, Y)\) gives its power series coefficient corresponding to said word). We let \( R(M) \) denote the analogous subalgebra of \( k^M \); this can be thought of as the algebra of bounded degree “partially commutative power series” in the variable sets \( X \) and \( Y \). Note that \( \rho \) restricts to a map

(8.6) \[ R(X, Y) \xrightarrow{\rho} R(M). \]

Finally, we claim (and see Proposition 8.9 below for a proof) that this further restricts to a map

(8.7) \[ FQSym(X, Y) \xrightarrow{\rho} FQSym(X) \otimes FQSym(Y) \]

in which the target is identified with its image under the (injective\(^2\)) multiplication map

\[
FQSym(X) \otimes FQSym(Y) \hookrightarrow R(M)
\]

\[
f(X) \otimes g(Y) \to f(X)g(Y).
\]

Using the identification of \( FQSym \) with all three of \( FQSym(X) \), \( FQSym(Y) \), \( FQSym(X, Y) \), the map \( \rho \) in (8.7) will then define a coproduct structure on \( FQSym \). Abusing notation, for \( f \) in \( FQSym \), we will simply write \( \Delta(f) = f(X, Y) \) instead of \( \rho(f(X, Y)) \).

**Example 8.8.** Recall from Example 8.4 that one has

\[
F_{312} = \sum_{i: \text{std}(i) = 231} X_i = \sum_{1 \leq i < j \leq k} X_jX_k X_i
\]

and therefore its coproduct is

\[
\Delta F_{312} = F_{312}(X_1, X_2, \ldots, Y_1, Y_2, \ldots)
\]

(by our abuse of notation)

\[
= \sum_{i < j \leq k} X_jX_k X_i + \sum_{i < j \leq k} X_jY_k X_i + \sum_{i < j \leq k} Y_jY_k X_i
\]

\[
= \sum_{i < j \leq k} X_jX_k X_i \cdot 1 + \sum_{i < j \leq k} X_jX_i \cdot Y_k + \sum_{i < j \leq k} X_i \cdot Y_jY_k + \sum_{i < j \leq k} 1 \cdot Y_jY_k X_i
\]

\[
= F_{312}(X) \cdot 1 + F_{21}(X) \cdot F_1(Y) + F_1(X) \cdot F_{12}(Y) + 1 \cdot F_{312}(Y)
\]

\[
= F_{312} \otimes 1 + F_{21} \otimes F_1 + F_1 \otimes F_{12} + 1 \otimes F_{312}.
\]

**Proposition 8.9.** The map \( \rho \) in (8.6) does restrict as claimed to a map as in (8.7), and hence defines a coproduct on \( FQSym \), acting on the \( \{F_w\} \) basis by the rule (8.2). This endows \( FQSym \) with the structure of a graded connected finite type Hopf algebra.

\(^2\)as images of the basis \( F_u(X) \otimes F_v(Y) \) of \( FQSym(X) \otimes FQSym(Y) \) are supported on disjoint monomials in \( R(M) \), so linearly independent.
Proof. Let $I$ be the totally ordered set $\{1 < 2 < 3 < \cdots \}$. Let $J$ be the totally ordered set $\{1 < 2 < 3 < \cdots < i < 2 < 3 < \cdots \}$. We set $X_i = Y_i$ for every integer $i$. Then, the alphabet $(X, Y)$ can be written as $\{X_i\}_{i \in J}$.

If $i$ is a word over the alphabet $I = \{1 < 2 < 3 < \cdots \}$, then we denote by $\tilde{i}$ the word over $J$ obtained from $i$ by replacing every letter $i$ by $\tilde{i}$.

For the first assertion of Proposition 8.9, it suffices to check that $F_w$ indeed has the image under $\Delta$ claimed in (8.2). Let $n \in \mathbb{N}$ and $w \in \mathbb{S}_n$. Then,

$$
\Delta F_w = F_w (X, Y) = \sum_{i \in J^n : \text{std}(i) = w^{-1}} (X, Y)_i = \sum_{t \in J^n : \text{std}(t) = w^{-1}} (X, Y)_t
$$

(8.8)

(since for every $t \in J^n$, there exists exactly one choice of $k \in \{0, 1, \ldots, n\}$ and $(i, j) \in I^k \times I^{n-k}$ satisfying $t \in i \uplus j$; namely, $i$ is the restriction of $t$ to the subalphabet $I$ of $J$, whereas $j$ is the restriction of $t$ to $J \setminus I$, and $k$ is the length of $i$).

We now fix $k$ and $(i, j)$, and try to simplify the inner sum

$$
\sum_{t \in J^n : \text{std}(t) = w^{-1}} (X, Y)_t
$$

(8.8). First we notice that this sum is nonempty if and only if there exists some $t \in i \uplus j$ satisfying $\text{std}(t) = w^{-1}$. This existence is easily seen to be equivalent to $w^{-1} \in \text{std}(i) \uplus \text{std}(j)[k]$ (since the standardization of any shuffle in $i \uplus j$ is the corresponding shuffle in $\text{std}(i) \uplus \text{std}(j)[k]$). This, in turn, is equivalent to $\text{std}(i) = (\text{std}(w_1, \ldots, w_k))^{-1}$ and $\text{std}(j) = (\text{std}(w_{k+1}, \ldots, w_n))^{-1}$ (according to the equivalence (i) $\iff$ (ii) in Lemma 8.6).

Hence, the inner sum on the right hand side of (8.8) is nonempty if and only if $\text{std}(i) = (\text{std}(w_1, \ldots, w_k))^{-1}$ and $\text{std}(j) = (\text{std}(w_{k+1}, \ldots, w_n))^{-1}$. When it is nonempty, it has only one addend270, and this addend is $(X, Y)_t = X_i Y_j$ (since $t \in i \uplus j$). Summarizing, we see that the inner sum on the right hand side of (8.8) equals $X_i Y_j$ when $\text{std}(i) = (\text{std}(w_1, \ldots, w_k))^{-1}$ and $\text{std}(j) = (\text{std}(w_{k+1}, \ldots, w_n))^{-1}$, and is empty otherwise. Thus, (8.8) simplifies to

$$
\Delta F_w = \sum_{k=0}^n \sum_{(i, j) \in I^k \times I^{n-k} : \text{std}(i) = (\text{std}(w_1, \ldots, w_k))^{-1}, \text{std}(j) = (\text{std}(w_{k+1}, \ldots, w_n))^{-1}} X_i Y_j
$$

$$
= \sum_{k=0}^n F_{\text{std}(w_1, \ldots, w_k)}(X) F_{\text{std}(w_{k+1}, \ldots, w_n)}(Y)
$$

$$
= \sum_{k=0}^n F_{\text{std}(w_1, \ldots, w_k)} \otimes F_{\text{std}(w_{k+1}, \ldots, w_n)} \in \text{FQSym} \otimes \text{FQSym}.
$$

This proves (8.2), and thus the first assertion of Proposition 8.9.

From this, it is easy to derive that $\Delta$ satisfies coassociativity (i.e., the diagram (1.3) holds for $C = \text{FQSym}$). (Alternatively, one can obtain this from the associativity of multiplication using Corollary 8.11.) We have already verified the rule (8.2). The connected graded structure on \text{FQSym} gives a counit and an antipode for free. \hfill \Box

Exercise 8.10. We say that a permutation $w \in \mathbb{S}_n$ is connected if $n$ is a positive integer and if there exists no $i \in \{1, 2, \ldots, n-1\}$ satisfying $\{1, 2, \ldots, i\} = \{1, 2, \ldots, i\}$. Let $\mathcal{C}$ denote the set of all connected permutations of all $n \in \mathbb{N}$. Show that $\text{FQSym}$ is a free (noncommutative) $k$-algebra with generators $(F_w)_{w \in \mathcal{C}}$. (This statement means that $(F_w, F_{w_2} \cdot \cdots \cdot F_{w_k})_{k \in \mathbb{N}; (w_1, w_2, \ldots, w_k) \in \mathcal{C}^k}$ is a basis of the $k$-module $\text{FQSym}$.)

270In fact, the elements std($t$) for $t \in i \uplus j$ are pairwise distinct, and thus only one of them can equal $w^{-1}$. 

HOPF ALGEBRAS IN COMBINATORICS (VERSION CONTAINING SOLUTIONS) 199
The Hopf algebra FQSym is self-dual: Let \( \{ G_w \} \) be the dual \( k \)-basis to the \( k \)-basis \( \{ F_w \} \) for FQSym. Then, the \( k \)-linear map sending \( G_w \mapsto F_{w^{-1}} \) is a Hopf algebra isomorphism \( FQSym^\alpha \rightarrow FQSym \).

**Proof.** For any \( 0 \leq k \leq n \), any \( u \in \mathfrak{S}_k \) and any \( v \in \mathfrak{S}_{n-k} \), one has

\[
F_{u^{-1}}F_{v^{-1}} = \sum_{w^{-1} \in u^{-1} v^{-1}[k]} F_{w^{-1}} = \sum_{\text{std}(u_1,\ldots,u_k) = u} \sum_{\text{std}(w_{k+1},\ldots,w_n) = v} F_{w^{-1}}
\]

via the equivalence of (i) and (ii) in Lemma 8.6. On the other hand, in FQSym, the dual \( k \)-basis \( \{ G_w \} \) to the \( k \)-basis \( \{ F_w \} \) for FQSym should have product formula

\[
G_u G_v = \sum_{\text{std}(u_1,\ldots,u_k) = u} \sum_{\text{std}(w_k+1,\ldots,w_n) = v} G_w
\]

coming from the coproduct formula (8.2) for FQSym in the \( \{ F_w \} \)-basis. Comparing these equalities, we see that the \( k \)-linear map \( \tau \) sending \( G_w \mapsto F_{w^{-1}} \) is an isomorphism \( FQSym^\alpha \rightarrow FQSym \) of \( k \)-algebras. Hence, the adjoint \( \tau^* : FQSym^\alpha \rightarrow (FQSym^\alpha)^\alpha \) of this map is an isomorphism of \( k \)-coalgebras. But identifying \( (FQSym^\alpha)^\alpha \) with FQSym in the natural way (since FQSym is of finite type), we easily see that \( \tau^* = \tau \), whence \( \tau \) itself is an isomorphism of both \( k \)-algebras and \( k \)-coalgebras, hence of \( k \)-bialgebras, hence of Hopf algebras.

We can now be a bit more precise about the relations between the various algebras

\[ \Lambda, \text{QSym, NSym, FQSym, R(X), R(x)}. \]

Not only does FQSym allow one to lift the Hopf structure of QSym, it dually allows one to extend the Hopf structure of NSym. To set up this duality, note that Corollary 8.11 motivates the choice of an inner product on FQSym in which

\[
(F_u, F_v) := \delta_{u^{-1}, v}.
\]

We wish to identify the images of the ribbon basis \( \{ R_\alpha \} \) of NSym when included in FQSym.

**Definition 8.12.** For any composition \( \alpha \), define an element \( R_\alpha \) of FQSym by

\[
R_\alpha := \sum_{w : \text{Des}(w) = D(\alpha)} F_{w^{-1}} = \sum_{(w,i) : \text{Des}(w) = D(\alpha)} \sum_{\text{std}(i) = w} X_i
\]

where the \( w \) in the sums are supposed to belong to \( \mathfrak{S}_{|\alpha|} \), and where the descent set of a sequence \( i = (i_1,\ldots,i_n) \) is defined by

\[
\text{Des}(i) := \{ j \in \{ 1,2,\ldots,n-1 \} : i_j > i_{j+1} \} = \text{Des}(|\alpha|).
\]

Alternatively, \( R_\alpha = \sum_T X_T \) in which the sum is over column-strict tableaux of the ribbon skew shape \( \alpha \), and \( X_T = X_i \) in which \( i \) is the sequence of entries of \( T \) read in order from the southwest toward the northeast.

**Example 8.13.** Taking \( \alpha = (1,3,2) \), with ribbon shape and column-strict fillings \( T \) as shown

\[
\begin{array}{cccccc}
\emptyset & \emptyset & \square & \square & \square & \square \\
\square & \square & \square & \square & \square & \square \\
\end{array}
\]

one has that

\[
R_{(1,3,2)} = \sum_{i=(i_1,i_2,i_3,i_4,i_5,i_6) : \text{Des}(i) = D(\alpha) = \{ 1,4 \}} X_i = \sum_{i_1 > i_2 \leq i_3 \leq i_4 \leq i_5 \leq i_6} X_{i_1}X_{i_2}X_{i_3}X_{i_4}X_{i_5}X_{i_6} = \sum_T X_T
\]
Corollary 8.14. For every \( n \in \mathbb{N} \) and \( w \in \mathfrak{S}_n \), we let \( \gamma(w) \) denote the unique composition \( \alpha \) of \( n \) satisfying \( D(\alpha) = \text{Des}(w) \). The following \( \mathbb{Z} \)-linear surjection and injection

\[
\begin{align*}
\text{FQSym} & \xrightarrow{\pi} \text{QSym} \\
F_w & \mapsto L_{\gamma(w)} \\
\text{NSym} & \xleftarrow{\iota} \text{FQSym} \\
R_\alpha & \mapsto R_\alpha
\end{align*}
\]

are Hopf morphisms, and are also adjoint maps with respect to the above choice of inner product on FQSym and the usual dual pairing between NSym and QSym. Furthermore, the second map \( \iota \) lets one factor the surjection NSym \( \xrightarrow{} \Lambda \) as follows

\[
\begin{align*}
\text{NSym} & \rightarrow \text{FQSym} \hookrightarrow \text{R}\langle X \rangle \rightarrow \text{R}(x) \\
R_\alpha & \mapsto R_\alpha \mapsto s_\alpha(x)
\end{align*}
\]

through the abelianization \( \text{R}(X) \rightarrow \text{R}(x) \) sending the noncommutative variable \( X_i \) to the commutative \( x_i \).

Proof. The fact that \( \text{FQSym} \xrightarrow{\pi} \text{QSym} \) is a Hopf map comes from checking that it respects the product (compare (5.8) and (8.1)) and the coproducts (compare (5.7) and (8.2)) then applying Proposition 1.46(c). It follows that \( \text{NSym} \xleftarrow{\iota} \text{FQSym} \) is a Hopf map once we check the adjointness assertion, via the following calculation:

\[
(\iota(R_\alpha), F_w) = (R_\alpha, F_w) = \sum_{u: \text{Des}(u) = D(\alpha)} (F_{w^{-1}}, F_u) = \begin{cases} 
1 & \text{if } \text{Des}(w) = D(\alpha) \\
0 & \text{otherwise}
\end{cases}
= (R_\alpha, L_{\gamma(w)}) = (R_\alpha, \pi(F_w)).
\]

The last assertion is clear: the abelianization map sends the noncommutative tableau monomial \( X_T \) to the commutative tableau monomial \( x_T \). \( \square \)

We summarize some of this picture as follows:

Furthermore, if we denote by \( \iota \) the canonical inclusion \( \Lambda \rightarrow \text{QSym} \) as well, then the diagram

\[
\begin{align*}
\text{FQSym} & \xrightarrow{} \text{QSym} \\
\text{NSym} & \xleftarrow{} \text{FQSym} \\
\Lambda & \xleftarrow{} \text{NSym} \hookrightarrow \text{QSym}
\end{align*}
\]

is commutative (according to the last sentence of Corollary 8.14).

9. Further topics

Some of these we may touch on in class, others are appropriate for student talks.
9.0.1. 0-Hecke algebras.

  Review the notions of indecomposables, simples, projectives, along with the theorems of Krull-Remak-Schmidt, of Jordan-Hölder, and the two kinds of Grothendieck groups dual to each other.

- 0-Hecke algebra representation theory.
  Describe the simples and projectives, following Denton, Hivert, Schilling, Thiery on \( J \)-trivial monoids.

- \( \text{Nsym} \) and \( \text{Qsym} \) as Grothendieck groups.
  Give Krob and Thibon’s interpretation of
  - \( \text{QSym} \) and the Grothendieck group of composition series, and
  - \( \text{NSym} \) and the Grothendieck group of projectives.

Remark 9.1. Mention P. McNamara’s interpretation, in the case of supersolvable lattices, of the Ehrenborg quasisymmetric function as the composition series enumerator for an \( H_n(0) \)-action on the maximal chains.

9.0.2. Aguiar-Bergeron-Sottile character theory Part II: Odd and even characters, subalgebras.

9.0.3. Face enumeration, Eulerian posets, and \( cd \)-indices. Borrowing from Billera’s ICM notes.
- f-vectors, h-vectors
- flag f-vectors, flag h-vectors
- ab-indices and \( cd \)-indices

9.0.4. Other topics.
- Loday-Ronco Hopf algebra of planar binary trees
- Poirier-Reutenauer Hopf algebra of tableaux
- Reading Hopf algebra of Baxter permutations
- Hopf monoids, e.g. of Hopf algebra of generalized permutohedra, of matroids, of graphs, Stanley chromatic symmetric functions and Tutte polynomials
- Lam-Pylyavskyy Hopf algebra of set-valued tableaux
- Connes-Kreimer Hopf algebra and renormalization
- Noncommutative symmetric functions and \( \Omega \Sigma \mathbb{C} P^\infty \)
- Maschke’s theorem and “integrals” for Hopf algebras
- Nichols-Zoeller structure theorem and group-like elements
- Cartier-Milnor-Moore structure theorem and primitive elements
- Quasi-triangular Hopf algebras and quantum groups
- The Steenrod algebra, its dual, and tree Hopf algebras
- Ringel-Hall algebras of quivers
- Ellis-Khovanov odd symmetric function Hopf algebras (see also Lauda-Russell)

Student talks given in class were:
(1) Al Garver, on Maschke’s theorem for finite-dimensional Hopf algebras
(2) Jonathan Hahn, on the paper by Humpert and Martin.
(3) Emily Gunawan, on the paper by Lam, Lauve and Sottile.
(4) Jonas Karlsson, on the paper by Connes and Kreimer
(5) Thomas McConville, on Butcher’s group and generalized Runge-Kutta methods.
(6) Cihan Bahran, on universal enveloping algebras and the Poincaré-Birkhoff-Witt theorem.
(7) Theodoreos Douvropoulos, on the Cartier-Milnor-Moore theorem.
(8) Alex Csar, on the Loday-Ronco Hopf algebra of binary trees
(9) Kevin Dilks, on Reading’s Hopf algebra of (twisted) Baxter permutations
(10) Becky Patrias, on the paper by Lam and Pylyavskyy
(11) Meng Wu, on multiple zeta values and Hoffman’s homomorphism from \( \text{QSym} \)

10. Some open problems and conjectures

- Is there a proof of the Assaf-McNamara skew Pieri rule that gives a resolution of Specht or Schur/Weyl modules whose character corresponds to \( s_{\lambda/\mu} h_n \), whose terms model their alternating sum?
Explicit antipodes in the Lam-Pylyavskyy Hopf algebras?

P. McNamara’s question: are $P$-partition enumerators irreducible for connected posets $P$?

Stanley’s question: are the only $P$-partition enumerators which are symmetric (not just quasisymmetric) those for which $P$ is a skew-shape with a column-strict labelling?

Does Stanley’s chromatic symmetric function distinguish trees?

Hoffman’s stuffle conjecture

Billera-Brenti’s nonnegativity conjecture for the total cd-index of Bruhat intervals

11. Solutions to the exercises

This chapter contains solutions to the exercises scattered throughout the text. These solutions vary in level of detail (some of them are detailed, some only outline the most important steps, and many lie inbetween these two extremes), and sometimes in notation (as they have been written over a long timespan). They also have seen far less quality control than the main text, so typos and worse are to be expected. Comments and alternative solutions are welcome!

11.1. Solution to Exercise 1.8. Solution to Exercise 1.8. This is analogous to the well-known fact that any nonunital nonassociative $k$-algebra has at most one multiplicative identity. If you can prove the latter fact by pure abstract nonsense (i.e., without referring to elements), then the same proof serves as a solution to Exercise 1.8 once all arrows are reversed (and all $m$’s and $u$’s are replaced by $\Delta$’s and $\epsilon$’s). Let us see how this works.

How does one classically prove that every nonunital nonassociative $k$-algebra has at most one multiplicative identity? Let $A$ be a nonunital nonassociative $k$-algebra, and let 1 and 1’ be two elements of $A$ which could both serve as multiplicative identities. That is, every $a \in A$ satisfies both $1a = a1 = a$ and $1'a = a1' = a$.

Now, applying $1a = a$ to $a = 1'$ yields $1 \cdot 1' = 1'$. But applying $a1' = a$ to $a = 1$ yields $1 \cdot 1' = 1$. Comparing $1 \cdot 1' = 1$ with $1 \cdot 1' = 1'$ yields $1 = 1'$, and thus the multiplicative identity is unique.

This argument made use of elements, which we need to get rid of in order to be able to reverse the arrows. The idea is to replace every element $\alpha$ of $A$ by the linear map $k \to A$ which sends $1 \in k$ to $\alpha$. In terms of commutative diagrams, a nonunital nonassociative $k$-algebra is a $k$-module $A$ endowed with a $k$-linear map $m : A \otimes A \to A$ which is not a priori required to satisfy any properties. A multiplicative identity of $A$ then corresponds to a $k$-linear map $u : k \to A$ making the diagram (1.2) commute. Let $u$ and $u'$ be two such $k$-linear maps $k \to A$. Instead of applying $1a = a$ to $a = 1'$, we now need to take the commutative diagram

\[
\begin{array}{ccc}
A \otimes k & \xleftarrow{\text{id}} & A \\
\downarrow{\text{id} \otimes u} & & \downarrow{\text{id}} \\
A \otimes A & \xrightarrow{m} & A
\end{array}
\]

(which commutes because $u$ makes (1.2) commute, and corresponds to the axiom $1a = a$) and pre-compose it with the morphism $u'$ (which corresponds to the multiplicative identity $1'$), thus obtaining

\[
\begin{array}{ccc}
A \otimes k & \xleftarrow{u'} & k \\
\downarrow{\text{id} \otimes u} & & \downarrow{\text{id}} \\
A \otimes A & \xrightarrow{m} & A
\end{array}
\]

271 The word “nonassociative” does not prohibit associativity; it simply means that associativity is not required. Similarly, “nonunital” does not force the nonexistence of a multiplicative identity.
Similarly, the element-free version of applying \( a1' = a \) to \( a = 1 \) yields the commutative diagram

\[
\begin{array}{ccc}
  k & \xrightarrow{u} & \k \\
  \downarrow & & \downarrow \\
  A & \xrightarrow{id} & \k \\
  \downarrow & & \downarrow \\
  A & \xleftarrow{m} & A \\
\end{array}
\]

The path \( k \xrightarrow{u'} A \xrightarrow{\text{id} \otimes u} A \otimes A \) through the diagram (11.1) and the path \( k \xrightarrow{u} A \xrightarrow{\text{id} \otimes u'} A \otimes A \) through the diagram (11.2) give the same map, since they both correspond to the element \( 1 \otimes 1' \) of \( A \otimes A \). How can this be seen without referring to elements? Being a tautology at the level of elements, this must follow from formal properties of tensor products (without using axioms like the commutativity of (1.2)). And so it does: In fact, the three small quadrilaterals in the diagram

\[
\begin{array}{ccc}
  k & \xrightarrow{u'} & \k \\
  \downarrow & & \downarrow \\
  A & \xrightarrow{\text{id} \otimes u} & \k \\
  \downarrow & & \downarrow \\
  A & \xleftarrow{m} & A \\
\end{array}
\]

commute, and therefore so does the whole diagram. We can now piece this diagram together with the diagrams (11.1) and (11.2) (this is the diagrammatic equivalent of comparing \( 1 \cdot 1' = 1 \) with \( 1 \cdot 1' = 1' \)), and obtain

\[
\begin{array}{ccc}
  k & \xrightarrow{u'} & \k \\
  \downarrow & & \downarrow \\
  A & \xrightarrow{\text{id} \otimes u} & \k \\
  \downarrow & & \downarrow \\
  A & \xleftarrow{m} & A \\
\end{array}
\]

Following the outside quadrilateral of this commutative diagram yields \( \text{id} \circ u = \text{id} \circ u' \), so that \( u = u' \), which shows that the two maps \( u \) and \( u' \) are equal. The purely diagrammatic argument that we made can now be easily transformed into a solution of Exercise 1.8 by merely reversing arrows and replacing \( m \) and \( u \) by \( \Delta \) and \( \epsilon \).

Alternatively one can solve the exercise using Sweedler notation, which probably gives some practice in using the latter.


(a) Let \( C, C', D \) and \( D' \) be four \( k \)-coalgebras. Let \( f : C \to C' \) and \( g : D \to D' \) be two \( k \)-coalgebra homomorphisms. We need to prove that \( f \otimes g : C \otimes D \to C' \otimes D' \) is a \( k \)-coalgebra homomorphism.
Recall that (by the definition of a “$k$-coalgebra homomorphism”) the map $f \otimes g : C \otimes D \to C' \otimes D'$ is a $k$-coalgebra homomorphism if and only if the two diagrams

\[
\begin{array}{ccc}
C \otimes D & \xrightarrow{f \otimes g} & C' \otimes D' \\
\Delta_{C \otimes D} & & \Delta_{C' \otimes D'} \\
(C \otimes D) \otimes (C \otimes D) & \xrightarrow{(f \otimes g) \otimes (f \otimes g)} & (C' \otimes D') \otimes (C' \otimes D')
\end{array}
\]

and

\[
\begin{array}{ccc}
C \otimes D & \xrightarrow{f \otimes g} & C' \otimes D' \\
\epsilon_{C \otimes D} & & \epsilon_{C' \otimes D'} \\
k & & k
\end{array}
\]

commute. We shall now prove that these diagrams indeed commute.

We know that the map $f : C \to C'$ is a $k$-coalgebra homomorphism. By the definition of a $k$-coalgebra homomorphism, this means that the two diagrams

\[
\begin{array}{ccc}
C & \xrightarrow{f} & C' \\
\Delta_C & & \Delta_{C'} \\
C \otimes C & \xrightarrow{f \otimes f} & C' \otimes C'
\end{array}
\]

and

\[
\begin{array}{ccc}
C & \xrightarrow{f} & C' \\
\epsilon_C & & \epsilon_{C'} \\
k & & k
\end{array}
\]

commute.

For any two $k$-modules $U$ and $V$, let $T_{U,V} : U \otimes V \to V \otimes U$ be the twist map (i.e., the $k$-linear map $U \otimes V \to V \otimes U$ sending every $u \otimes v$ to $v \otimes u$). A simple linear-algebraic fact says that if $U, V, U'$ and $V'$ are four $k$-modules and $x : U \to U'$ and $y : V \to V'$ are two $k$-linear maps, then

\[
(y \otimes x) \circ T_{U,V} = T_{U',V'} \circ (x \otimes y).
\]

The definition of the $k$-coalgebra $C \otimes D$ yields

\[
\Delta_{C \otimes D} = (\text{id}_C \otimes T_{C,D} \otimes \text{id}_D) \circ (\Delta_C \otimes \Delta_D).
\]

Similarly,

\[
\Delta_{C' \otimes D'} = (\text{id}_{C'} \otimes T_{C',D'} \otimes \text{id}_{D'}) \circ (\Delta_{C'} \otimes \Delta_{D'}).
\]

But $(f \otimes f) \circ \Delta_C = \Delta_{C'} \circ f$ (since the diagram (11.5) commutes), and similarly $(g \otimes g) \circ \Delta_D = \Delta_{D'} \circ g$. 


Now,
\[
(f \otimes g) \otimes (f \otimes g) \circ \Delta_{C \otimes D} = (f \otimes (g \otimes f) \otimes g) \circ \Delta_{C \otimes D}
\]
\[
= (f \otimes (g \otimes f) \otimes g) \circ (id_C \otimes T_{C,D} \otimes id_D) \circ (\Delta_C \otimes \Delta_D)
\]
\[
= (f \otimes (g \otimes f) \otimes g) \circ (f \otimes (g \otimes f) \otimes g) \circ \Delta_{C \otimes D}
\]
\[
= (id_C \otimes T_{C,D} \otimes id_D) \circ (f \otimes (g \otimes f) \otimes g) \circ \Delta_{C \otimes D}
\]
\[
= (id_C \otimes T_{C,D} \otimes id_D) \circ (f \otimes (g \otimes f) \otimes g) \circ \Delta_{C \otimes D}
\]
\[
= (id_C \otimes T_{C,D} \otimes id_D) \circ (f \otimes (g \otimes f) \otimes g) \circ \Delta_{C \otimes D}
\]
In other words, the diagram (11.3) commutes.

We have \( \epsilon_C \circ f = \epsilon_C \) (since the diagram (11.6) commutes) and \( \epsilon_{D'} \circ g = \epsilon_D \) (similarly).

Now, let \( \theta \) be the canonical \( k \)-module isomorphism \( k \otimes k \to k \). Then, the definition of the \( k \)-coalgebra \( C \otimes D \) yields

\[
\epsilon_{C \otimes D} = \theta \circ (\epsilon_C \otimes \epsilon_D).
\]
Similarly,

\[
\epsilon_{C' \otimes D'} = \theta \circ (\epsilon_{C'} \otimes \epsilon_{D'}). \tag{11.11}
\]

Now,
\[
\epsilon_{C' \otimes D'} \circ (f \otimes g) = \theta \circ (\epsilon_{C'} \otimes \epsilon_{D'}) \circ (f \otimes g)
\]
\[
= \theta \circ (\epsilon_{C'} \circ f \otimes \epsilon_{D'} \circ g) \circ (\Delta_C \otimes \Delta_D)
\]

(by (11.10)). In other words, the diagram (11.4) commutes.

We now know that the two diagrams (11.3) and (11.4) commute. Thus, the map \( f \otimes g : C \otimes D \to C' \otimes D' \) is a \( k \)-coalgebra homomorphism (because we know that the map \( f \otimes g : C \otimes D \to C' \otimes D' \) is a \( k \)-coalgebra homomorphism if and only if the two diagrams (11.3) and (11.4) commute). This solves Exercise 1.13(b).

(a) In order to obtain a solution to Exercise 1.13(a), it is enough to reverse all arrows in the above solution to Exercise 1.13(b) (and, of course, replace \( C, D, \Delta_C \) etc. by \( A, B, m_A \) etc.).
11.3. Solution to Exercise 1.20. Solution to Exercise 1.20. (a) Assume that $f$ is surjective. It is a known fact from linear algebra\footnote{proven, e.g., in Keith Conrad’s [33, “Tensor Products II”, Thm. 2.19]} that if $U$, $V$, $U'$ and $V'$ are four \( k \)-modules, and $\phi : U \to U'$ and $\psi : V \to V'$ are two surjective $k$-linear maps, then the kernel of $\phi \otimes \psi : U \otimes V \to U' \otimes V'$ is

$$\ker (\phi \otimes \psi) = (\ker \phi) \otimes V + U \otimes (\ker \psi).$$

Applying this to $U = A$, $U' = C$, $V = A$, $V' = C$, $\phi = f$ and $\psi = f$, we obtain $\ker (f \otimes f) = (\ker f) \otimes A + A \otimes (\ker f)$.

But $f$ is a coalgebra homomorphism, so that $\epsilon = \epsilon \circ f$. Hence, every $x \in \ker f$ satisfies $\epsilon (x) = 0$. In other words, $\epsilon (\ker f) = 0$. Also,

$$\ker f \subset \ker (\Delta \circ f) = \ker ((f \otimes f) \circ \Delta) = \Delta^{-1} (\ker (f \otimes f)),$$

so that

$$\Delta (\ker f) \subset \ker (f \otimes f) = (\ker f) \otimes A + A \otimes (\ker f).$$

Combined with $\epsilon (\ker f) = 0$, this shows that $\ker f$ is a two-sided coideal of $A$. Thus, Exercise 1.20(a) is solved.

(b) Assume that $k$ is a field. Then, it is a known fact from linear algebra\footnote{proven, e.g., in Keith Conrad’s [33, “Tensor Products II”, Thm. 2.19]} that if $U$, $V$, $U'$ and $V'$ are four \( k \)-modules, and $\phi : U \to U'$ and $\psi : V \to V'$ are two \( k \)-linear maps, then the kernel of $\phi \otimes \psi : U \otimes V \to U' \otimes V'$ is

$$\ker (\phi \otimes \psi) = (\ker \phi) \otimes V + U \otimes (\ker \psi).$$

Starting from this point, we can continue arguing as in the solution of part (a). Thus, Exercise 1.20(b) is solved.

11.4. Solution to Exercise 1.25. Solution to Exercise 1.25.

(a) Define a map $\tilde{\Delta} : A \to A \otimes A$ by

$$\tilde{\Delta} (x) = \Delta (x) - (x \otimes 1 + 1 \otimes x) \quad \text{for all } x \in A.$$

It is easily seen that $\tilde{\Delta}$ is a homomorphism of graded $k$-modules. Hence, the kernel $\ker \tilde{\Delta}$ of $\tilde{\Delta}$ is a graded $k$-submodule of $A$ (since the kernel of a homomorphism of graded $k$-modules always is a graded $k$-submodule of the domain).

We have defined $\mathfrak{p}$ as the set of all primitive elements of $A$. In other words,

$$\mathfrak{p} = \{ x \in A \mid \Delta (x) = x \otimes 1 + 1 \otimes x \}$$

$$= \left\{ x \in A \mid \frac{\Delta (x) - (x \otimes 1 + 1 \otimes x)}{-\tilde{\Delta} (x)} = 0 \right\}$$

$$= \left\{ x \in A \mid \tilde{\Delta} (x) = 0 \right\} = \ker \tilde{\Delta}.$$

Thus, $\mathfrak{p}$ is a graded $k$-submodule of $A$ (since we know that $\ker \tilde{\Delta}$ is a graded $k$-submodule of $A$). This solves Exercise 1.25 (a).
isomorphism. In other words, the restriction of \( u \in \mathfrak{p} \) and thus an isomorphism. In other words, \( \epsilon \) is an isomorphism. Since \( \phi \) is a graded algebra, \( \epsilon \) is well-defined. This inverse map must be \( \epsilon \) is connected, we have 

\[
\Delta (x) = x \otimes 1 + 1 \otimes x
\]

(11.12) Subtracting \( x \) from this equality, we obtain \( 0 = \epsilon (x) \). This proves (11.12).

(Note: We will reprove (11.12) below in Proposition 1.37.)

Now, for every \( x \in \mathfrak{p} \), we have

\[
\Delta (x) = x \otimes 1 + 1 \otimes x \quad \text{(since } x \in \mathfrak{p}, \text{ so that } x \text{ is primitive)}
\]

\[
\in \mathfrak{p} \otimes A + A \otimes \mathfrak{p}.
\]

In other words, \( \Delta (p) \in \mathfrak{p} \otimes A + A \otimes \mathfrak{p} \). Combined with \( \epsilon (p) = 0 \) (this follows from (11.12)), this yields that \( \mathfrak{p} \) is a two-sided coideal of \( A \). This solves Exercise 1.25 (b).

11.5. Solution to Exercise 1.26. Solution to Exercise 1.26. (a) The unit map \( u : k \rightarrow A \) is graded (since \( A \) is a graded algebra). Hence, \( u(k_0) \subset A_0 \), where \( k_0 \) denotes the 0-th graded component of \( k \). But the grading on \( k \) is such that \( k_0 = k \). Thus, \( u(k_0) = u(k) \), so that \( u(k) = u(k_0) \subset A_0 \). But

\[
u(k) = \begin{cases} u(\lambda) : \lambda \in k \\ = \lambda \cdot 1_A : \lambda \in k \\ = \lambda \cdot 1_A \\ = \text{(by the definition of } u) \end{cases}
\]

Hence, \( k \cdot 1_A = u(k) \subset A_0 \), so that part (a) is solved.

(b) Since \( A \) is connected, we have \( A_0 \cong k \). In other words, there exists a \( k \)-module isomorphism \( \phi : A_0 \rightarrow k \). Consider this \( \phi \). Since \( \phi \) is a \( k \)-module isomorphism \( A_0 \rightarrow k \), the inverse \( \phi^{-1} \) of \( \phi \) is a well-defined \( k \)-module isomorphism \( k \rightarrow A_0 \).

We saw in the proof of part (a) that \( u(k) \subset A_0 \). Hence, \( u \) restricts to a map \( k \rightarrow A_0 \). Denote this map \( k \rightarrow A_0 \) by \( u' \). Then, \( u' \) is a restriction of \( u \) (more precisely, a corestriction of \( u \), because we are restricting the target rather than the domain).

Let also \( \epsilon' \) denote the restriction of \( \epsilon \) to \( A_0 \). Since \( \epsilon' \) and \( u' \) are restrictions of \( \epsilon \) and \( u \), we have \( \epsilon' \circ u' = \epsilon \circ u = \text{id}_k \) (by the axioms of a coalgebra). Now, \( (\epsilon' \circ \phi^{-1}) \circ (\phi \circ u') = \epsilon' \circ u' = \text{id}_k \). Hence, the \( k \)-linear map \( \epsilon' \circ \phi^{-1} : k \rightarrow k \) has a right inverse. Thus, the \( k \)-linear map \( \epsilon' \circ \phi^{-1} : k \rightarrow k \) is surjective. Since every surjective \( k \)-linear map \( k \rightarrow k \) is an isomorphism, this shows that the \( k \)-linear map \( \epsilon' \circ \phi^{-1} : k \rightarrow k \) is an isomorphism. Since \( \phi \) also is an isomorphism, the map \( \epsilon' \circ \phi^{-1} \circ \phi \) is a composition of two isomorphisms, and thus an isomorphism. In other words, \( \epsilon' \) is an isomorphism (since \( \epsilon' \circ \phi^{-1} \circ \phi = \epsilon' \)). Hence, the inverse map of \( \epsilon' \) is well-defined. This inverse map must be \( u' \) (since \( \epsilon' \circ u' = \text{id}_k \)), and so we conclude that \( u' \) is an isomorphism. In other words, the restriction of \( u \) to a map \( k \rightarrow A_0 \) is an isomorphism. This solves part (b).

---

274 \textit{Proof.} Let \( \alpha \) be a surjective \( k \)-linear map \( k \rightarrow k \). We need to show that \( \alpha \) is an isomorphism.

Let \( \lambda \in \ker \alpha \). Then, \( \lambda \in k \) satisfies \( \alpha (\lambda) = 0 \). But \( \alpha \) is surjective, so that \( k = \alpha (k) \). Hence, \( 1 \in k = \alpha (k) \). Thus, there exists a \( \mu \in k \) such that \( 1 = \alpha (\mu) \). Consider this \( \mu \). Then, the \( k \)-linearity of \( \alpha \) yields \( \alpha (\mu \lambda) = \lambda \alpha (\mu) = \lambda \). But the \( k \)-linearity of \( \alpha \) also shows that \( \alpha (\mu \lambda) = \mu \alpha (\lambda) = 0 \). Thus, \( 0 = \alpha (\mu \lambda) = \alpha (\lambda \mu) = \lambda \), so that \( \lambda = 0 \). We thus have shown that every \( \lambda \in \ker \alpha \) satisfies \( \lambda = 0 \). Hence, \( \ker \alpha = 0 \), so that the map \( \alpha \) is injective. Since \( \alpha \) is injective and surjective, we see that \( \alpha \) is bijective, thus an isomorphism, qed.
(d) In the proof of part (b), we showed that \( u' \) is the inverse map of \( e' \). Hence, \( e' \) is the inverse map of \( u' \).

In other words, the restriction of \( e \) to \( A_0 \) is the inverse map of the restriction of \( u \) to a map \( k \to A_0 \). This solves part (d).

(c) The counit map \( e \) is graded (since \( A \) is a graded coalgebra). Hence, every \( n \geq 0 \) satisfies \( e(A_n) \subset k_n \) (where \( k_n \) denotes the \( n \)-th graded component of \( k \)). For every positive \( n \), this shows that \( A_n \subset \ker e \) 275. Hence, \( \bigoplus_{n>0} A_n = \sum_{n>0} A_n \subset \sum_{n>0} \ker e \subset \ker e \) (since \( \ker e \) is a \( k \)-submodule of \( A \)).

Now, let \( a \in \ker e \) be arbitrary. Then, \( a \in A \) satisfies \( e(a) = 0 \). We have \( a \in A = \bigoplus_{n>0} A_n = A_0 \oplus \bigoplus_{n>0} A_n \). Hence, we can write \( a \) in the form \( a = a' + a'' \) for \( a' \in A_0 \) and \( a'' \in \bigoplus_{n>0} A_n \). Consider these \( a' \) and \( a'' \). We have \( a'' \in \bigoplus_{n>0} A_n \subset \ker e \), so that \( e(a'') = 0 \). Since \( a = a' + a'' \), we have \( e(a) = e(a' + a'') = e(a') + e(a'') = e(a') \), thus \( e(a') = 0 \). Since \( e \) restricted to \( A_0 \) is injective (in fact, part (d) of this problem shows that \( e \) restricted to \( A_0 \) is an isomorphism), this yields that \( a' = 0 \) (because \( a' \in A_0 \)). Hence, \( a = a'' \in \bigoplus_{n>0} A_n \).

Now forget that we fixed \( a \). We thus have seen that every \( a \in \ker e \) satisfies \( a \in \bigoplus_{n>0} A_n \). In other words, \( \ker e \subset \bigoplus_{n>0} A_n \). Combined with \( \bigoplus_{n>0} A_n \subset \ker e \), this yields \( \ker e = \bigoplus_{n>0} A_n \). This solves part (c).

(e) Let \( x \in A \). We have \( A = \bigoplus_{n>0} A_n = A_0 \oplus \bigoplus_{n>0} A_n = k \cdot 1_A \oplus I = k \cdot 1_A + I \) and

\[
\Delta(x) \in A \otimes I = A \otimes (k \cdot 1_A + I) = A \otimes (k \cdot 1_A) + A \otimes I.
\]

Hence, there exist \( y \in A \otimes (k \cdot 1_A) \) and \( z \in A \otimes I \) such that \( \Delta(x) = y + z \). Consider these \( y \) and \( z \). We will show that \( y = x \otimes 1_A \).

Since \( y \in A \otimes (k \cdot 1_A) = k \cdot (A \otimes 1_A) = A \otimes 1_A \), we can write \( y \) in the form \( y = y' \otimes 1_A \) for some \( y' \in A \). Consider this \( y' \).

By the commutativity of (1.4), we have \((\text{id}_A \otimes e)(\Delta(x)) = x \) (where we identify \( A \otimes k \) with \( A \)). Hence,

\[
x = (\text{id}_A \otimes e) \left( \Delta(x) \right) = (\text{id}_A \otimes e) \left( (y + z) \right) = (\text{id}_A \otimes e) \left( \begin{array}{c} y \\ = y' \otimes 1_A \end{array} \right) + (\text{id}_A \otimes e) \left( \begin{array}{c} z \\ = \epsilon(I) \end{array} \right)
\]

\[
= y' + \text{id}_A(A) 0 = y' + 0 = y',
\]

which shows that \( x = y' \). Now, \( y = \frac{y'}{x} \otimes 1_A = x \otimes 1_A \) and

\[
x = \frac{y}{x} + \frac{z}{x \otimes 1_A} \in A \otimes I + A \otimes I = x \otimes 1 + A \otimes I.
\]

This solves part (e).

(f) Let \( x \in I \). Thus, \( x \in I = \ker e \) (by part (c)), so that \( e(x) = 0 \).

Let us introduce \( y \) and \( z \) as in the solution to part (e). As we saw in that solution, we have \( y = x \otimes 1_A \). We have

\[
z \in \frac{A}{k \cdot 1_A + I} \otimes I = (k \cdot 1_A + I) \otimes I = (k \cdot 1_A) \otimes I + I \otimes I.
\]

Hence, there exist \( u \in (k \cdot 1_A) \otimes I \) and \( v \in I \otimes I \) such that \( z = u + v \). Consider these \( u \) and \( v \). We will show that \( u = 1_A \otimes x \).

---

275 Proof. Let \( n \) be a positive integer. Then, \( e(A_n) \subset k_n = 0 \) (because of how the grading on \( k \) is constructed), so that \( e(A_n) = 0 \) and thus \( A_n \subset \ker e \), qed.
We have \( v \in I \otimes I \), so that \((\epsilon \otimes \text{id}_A)(v) \in (\epsilon \otimes \text{id}_A)(I \otimes I) = \epsilon(I) \text{id}_A(I) = 0\). Thus, \((\epsilon \otimes \text{id}_A)(v) = 0\).

Since \( u \in (k \cdot 1_A) \otimes I = k \cdot (1_A \otimes I) = 1_A \otimes I \), we can write \( u \) in the form \( u = 1_A \otimes u' \) for some \( u' \in I \). Consider this \( u'\).

By the commutativity of (1.4), we have \((\epsilon \otimes \text{id}_A)(\Delta(x)) = x\) (where we identify \( k \otimes A \) with \( A \)). Hence,

\[
\Delta(x) = \begin{pmatrix} y \\ z \end{pmatrix} = e(x) \begin{pmatrix} y \\ z \end{pmatrix} = e(x) \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u' \\ 1_A \otimes u' \end{pmatrix} = \epsilon(1_A) u' = u'.
\]

Now, \( u = 1_A \otimes u' = 1_A \otimes x \). Now,

\[
\Delta(x) = \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} x \otimes 1_A \\ u+v \end{pmatrix} = \begin{pmatrix} x \otimes 1_A \\ 1 \end{pmatrix} \otimes 1 = 1 \otimes (x \otimes 1_A) \otimes 1 + 1 \otimes x \otimes 1 + 1 \otimes x \otimes 1 + 1 \otimes I \otimes I.
\]

In other words, \( \Delta(x) = 1 \otimes x + x \otimes 1 + \Delta_+(x) \) for some \( \Delta_+(x) \in I \otimes I \). This solves part (f).

11.6. Solution to Exercise 1.28. Solution to Exercise 1.28. We have to prove that the binary operation \( \star \) on \( \text{Hom}(C, A) \) is associative. In other words, we have to prove that any three elements \( f, g \) and \( h \) of \( \text{Hom}(C, A) \) satisfy \( f \star (g \star h) = (f \star g) \star h \).

So let \( f \), \( g \) and \( h \) be three elements of \( \text{Hom}(C, A) \). As usual, denote by \( m : A \otimes A \to A \) the multiplication of \( A \), and by \( \Delta : C \to C \otimes C \) the comultiplication of \( C \).

By the definition of \( g \star h \), we have \( g \star h = m \circ (g \otimes h) \circ \Delta \), so that

\[
\begin{align*}
\Delta(f \otimes (g \star h)) & = (\text{id}_A \circ f \circ \text{id}_C) \otimes (m \circ (g \otimes h) \circ \Delta) \\
& = (\text{id}_A \circ f \circ \text{id}_C) \otimes (m \circ (g \otimes h) \circ \Delta) \\
& = (\text{id}_A \otimes m) \circ (f \otimes (g \otimes h)) \circ (\text{id}_C \otimes \Delta) \\
& = (\text{id}_A \otimes m) \circ (f \otimes (g \otimes h)) \circ (\text{id}_C \otimes \Delta).
\end{align*}
\]

But the definition of \( f \star (g \star h) \) yields

\[
\begin{align*}
(11.13) \quad & f \star (g \star h) = m \circ \frac{(f \otimes (g \star h))}{=} = (\text{id}_A \circ m) \circ (f \otimes (g \otimes h)) \circ (\text{id}_C \otimes \Delta) \\
& \quad \circ \Delta = m \circ (\text{id}_A \otimes m) \circ (f \otimes (g \otimes h)) \circ (\text{id}_C \otimes \Delta) \circ \Delta.
\end{align*}
\]

On the other hand, \( f \star g = m \circ (f \otimes g) \circ \Delta \) (by the definition of \( f \star g \)), so that

\[
\begin{align*}
& (f \star g) \circ h = (m \circ (f \otimes g) \circ \Delta) \circ (\text{id}_C \circ h \circ \text{id}_C) \\
& = (m \circ (f \otimes g) \circ \Delta) \circ (\text{id}_C \circ h \circ \text{id}_C) \\
& = (m \circ \text{id}_A) \circ ((f \otimes g) \circ h) \circ (\Delta \circ \text{id}_C) \\
& = (m \circ \text{id}_A) \circ (f \otimes g \circ h) \circ (\Delta \circ \text{id}_C).
\end{align*}
\]

Now, the definition of \( (f \star g) \star h \) yields

\[
\begin{align*}
(11.14) \quad & (f \star g) \star h = m \circ \frac{(f \star g) \otimes h}{=} = (m \circ \text{id}_A) \circ ((f \star g) \otimes h) \circ (\Delta \circ \text{id}_C) \circ \Delta \\
& \quad \circ \Delta = m \circ (m \circ \text{id}_A) \circ (f \otimes g \circ h) \circ (\Delta \circ \text{id}_C) \circ \Delta.
\end{align*}
\]
Now, recall that $A$ is a $k$-algebra, and hence the diagram (1.1) commutes (by the definition of a $k$-algebra). Thus, $m \circ (id_A \otimes m) = m \circ (m \otimes id_A)$. Also, $C$ is a $k$-coalgebra, and thus the diagram (1.3) commutes (by the definition of a $k$-coalgebra). Hence, $(id_C \otimes \Delta) \circ \Delta = (\Delta \otimes id_C) \circ \Delta$. Now, (11.13) becomes

$$f \ast (g \ast h) = m \circ (id_A \otimes m) \circ (f \otimes g \otimes h) \circ (id_C \otimes \Delta) \circ \Delta$$

(by (11.14)). Thus, $f \ast (g \ast h) = (f \ast g) \ast h$ is proven, and the solution of Exercise 1.28 is complete.

11.7. Solution to Exercise 1.29. Solution to Exercise 1.29. (a) There are two ways to solve an exercise like this: either by explicitly evaluating the two sides on elements of their domain (in this case, it is of course enough to only evaluate them on pure tensors) or by diagram chasing. In the case of this particular exercise, both solutions are very easy, so let us show them both.

First solution: Here is the solution by explicit computations:

We need to prove that $(f \otimes g) \ast (f' \otimes g') = (f \ast f') \otimes (g \ast g')$. For this, it is clearly enough to show that $((f \otimes g) \ast (f' \otimes g'))(c \otimes d) = ((f \ast f') \otimes (g \ast g'))(c \otimes d)$ for every $c \in C$ and $d \in D$. But this can be done directly: Since $\Delta(c \otimes d) = \sum_{(c),(d)} (c_1 \otimes d_1) \otimes (c_2 \otimes d_2)$ (we are using Sweedler’s notation here), we have

$$((f \otimes g) \ast (f' \otimes g'))(c \otimes d) = \sum_{(c),(d)} (f \otimes g)(c_1 \otimes d_1) \cdot (f' \otimes g')(c_2 \otimes d_2)$$

$$= \sum_{(c),(d)} (f(c_1) \otimes g(d_1)) \cdot (f'(c_2) \otimes g'(d_2))$$

$$= \sum_{(c),(d)} f(c_1) \ast f'(c_2) \otimes g(d_1) \ast g'(d_2).$$

Compared to

$$((f \ast f') \otimes (g \ast g'))(c \otimes d) = \sum_{(c),(d)} (f(c_1) \ast f'(c_2)) \otimes (g(d_1) \ast g'(d_2))$$

this yields $((f \otimes g) \ast (f' \otimes g'))(c \otimes d) = ((f \ast f') \otimes (g \ast g'))(c \otimes d)$, which completes our solution of Exercise 1.29(a).

Second solution: Now comes the solution by diagram chasing. Actually, we will not see any diagrams here because in this particular case it would be a waste of space to draw them; everything can be done by a short computation. Denote by $T$ the twist map $U \otimes V \to V \otimes U$ for any two $k$-modules $U$ and $V$. (We leave $U$ and $V$ out of the notation since these will always be clear from the context.) By the definition of
convolution, we have
\[
(f \otimes g) \star (f' \otimes g') = \underbrace{m_{A \otimes B}}_{= (m_A \otimes m_B) \circ (id \otimes T \otimes id)} \circ (f \otimes g) \otimes (f' \otimes g') \circ \Delta_{C \otimes D} \\
= (m_A \otimes m_B) \circ (id \otimes T \otimes id) \circ (f \otimes g) \otimes (f' \otimes g') \circ (id \otimes T \otimes id) \circ (\Delta_C \otimes \Delta_D) \\
= (m_A \otimes m_B) \circ (f \otimes f' \otimes g \otimes g') \circ (id \otimes T \otimes id) \circ (\Delta_C \otimes \Delta_D) \\
= (m_A \otimes m_B) \circ (f \otimes f') \circ (\Delta_C) \otimes (m_B \circ (g \otimes g') \circ \Delta_D) \\
= f \otimes f' \otimes (g \otimes g').
\]

This solves Exercise 1.29(a) again.

(b) Recall that the products in the \(k\)-algebras \((\text{Hom} (C, A), \star)\), \((\text{Hom} (D, B), \star)\) and \((\text{Hom} (C \otimes D, A \otimes B), \star)\) are the convolution products \(\star\). For the sake of consistency, we will also denote by \(\star\) the product in the \(k\)-algebra \((\text{Hom} (C, A), \star) \otimes (\text{Hom} (D, B), \star)\).

We need to prove that \(R\) is a \(k\)-algebra homomorphism. For this, it is clearly enough to show that \(R\) preserves products and sends the unity \((u_A \otimes C) \otimes (u_B \otimes D)\) of the \(k\)-algebra \((\text{Hom} (C, A), \star) \otimes (\text{Hom} (D, B), \star)\) to the unity \(u_{A \otimes B} \otimes C \otimes D\) of the \(k\)-algebra \((C \otimes D, A \otimes B), \star)\).

Let us check this. First, let us verify that \(R\) preserves products. So we need to show that \(R(F \star F') = R(F) \star R(F')\) for all \(F\) and \(F'\) in \((\text{Hom} (C, A), \star) \otimes (\text{Hom} (D, B), \star)\). In proving this, we can WLOG assume that \(F\) and \(F'\) are pure tensors (since the claim \(R(F \star F') = R(F) \star R(F')\) is linear in \(F\) and \(F'\)). Assume this, and write \(F\) and \(F'\) as \(F = f \otimes g\) and \(F' = f' \otimes g'\), respectively. Then,

\[
R(F \star F') = R \left( \begin{array}{c} (f \otimes g) \star (f' \otimes g') \\ = (f \otimes g) \otimes (f' \otimes g') \end{array} \right) = R \left( (f \otimes g) \otimes (f' \otimes g') \right) = R \left( (f \otimes f') \otimes (g \otimes g') \right) \\
= R \left( f \otimes f' \otimes (g \otimes g') \right) = R \left( (f \star f') \otimes (g \star g') \right) \\
= R \left( \underbrace{R(f \otimes g)}_{= F} \right) \star \underbrace{R(f' \otimes g')}_{= F'} = R(F) \star R(F'),
\]

where the meaning of the tensor sign (each time standing either for a tensor or for a tensor product of maps) should be clear from the context. Thus, we are done checking that \(R\) preserves products.

It thus remains to verify that \(R\) sends \((u_A \otimes C) \otimes (u_B \otimes D)\) to \(u_{A \otimes B} \otimes C \otimes D\). This is straightforward: If \(s\) denotes the canonical isomorphism \(k \to k \otimes k\), then

\[
R((u_A \otimes C) \otimes (u_B \otimes D)) = (u_A \otimes C) \otimes (u_B \otimes D) = (u_A \otimes u_B) \circ (\epsilon_C \otimes \epsilon_D) = u_{A \otimes B} \circ s^{-1} \circ s \circ \epsilon_C \otimes \epsilon_D = u_{A \otimes B} \otimes C \otimes D.
\]

The solution of Exercise 1.29(b) is thus complete.
11.8. Solution to Exercise 1.39. Solution to Exercise 1.39. Let us start with an observation which is irrelevant to our solution of the exercise. Namely, let us notice that $m^{(k)}(a_1 \otimes a_2 \otimes \ldots \otimes a_{k+1}) = a_1 \otimes a_2 \otimes \ldots \otimes a_{k+1}$ for any $k \geq 0$ and any $k + 1$ elements $a_1, a_2, \ldots, a_{k+1}$ of $A$. The statements of Exercise 1.39 are nothing but different aspects of what is known as “general associativity” (although they all fall short of defining an “arbitrary bracketing” of a $(k + 1)$-fold product), written in an element-free fashion (that is, written without any reference to elements of $A$, but only in terms of maps). For instance, part (a) of the exercise says that any $k + 1$ elements $a_1, a_2, \ldots, a_{k+1}$ of $A$ satisfy

$$a_1 \otimes (a_2 \otimes (a_3 \otimes (\ldots (a_{k+1} \otimes \ldots)))) = (a_1 \otimes (a_2 \otimes (a_3 \otimes (\ldots (a_{k+1} \otimes \ldots))))).$$

However, there is virtue in solving Exercise 1.39 in an element-free way (i.e., without referring to elements, but only referring to maps), because such a solution will automatically yield a solution of Exercise 1.40 by reversing all arrows. So let us show an element-free solution of Exercise 1.39.

(a) We will solve Exercise 1.39(a) by induction over $k$.

The induction base ($k = 0$) is vacuously true, since there exists no $0 \leq i \leq k - 1$ for $k = 0$. So let us proceed to the induction step. Let $K$ be a positive integer. We want to prove that the claim of Exercise 1.39(a) holds for $k = K$, assuming (as the induction hypothesis) that the claim of Exercise 1.39(a) holds for $k = K - 1$.

We first notice that

$$m^{(K-1)} = m \circ \left( m^{(i)} \otimes m^{((K-1)-i)} \right)$$

for all $0 \leq i \leq (K - 1) - 1$. (This is merely a restatement of the induction hypothesis.)

Now fix $0 \leq i \leq K - 1$. We need to show that

$$m^{(K)} = m \circ \left( m^{(i)} \otimes m^{(K-1-i)} \right).$$

If $i = 0$, then (11.16) is obviously true. Hence, we can WLOG assume that we don’t have $i = 0$. Assume this. Then, $i > 0$. Hence, the recursive definition of $m^{(i)}$ yields $m^{(i)} = m \circ (id_A \otimes m^{(i-1)})$. Thus,

$$\frac{m^{(i)}}{=m\circ(id_A \otimes m^{(i-1)})} \otimes \frac{m^{(K-1-i)}}{=id_A \circ m^{(K-1-i)}} = \left( m \circ \left( id_A \otimes m^{(i-1)} \right) \right) \otimes \left( id_A \circ m^{(K-1-i)} \right)$$

$$= (m \otimes id_A) \circ \left( id_A \otimes m^{(i-1)} \otimes m^{(K-1-i)} \right)$$

$$= (m \otimes id_A) \circ \left( id_A \otimes m^{(i-1)} \otimes m^{(K-1-i)} \right)$$

(11.17)

On the other hand,

$$\frac{id_A \otimes m^{(K-1)}}{=m\circ(m^{(i-1)} \otimes m^{((K-1)-1-(i-1))})} = id_A \otimes \left( m \circ \left( m^{(i-1)} \otimes m^{((K-1)-1-(i-1))} \right) \right)$$

(by (11.15), applied to $i - 1$ instead of $i$)

$$= id_A \otimes \left( m \circ \left( m^{(i-1)} \otimes m^{(K-1-i)} \right) \right)$$

$$= (id_A \otimes m) \circ \left( id_A \otimes m^{(i-1)} \otimes m^{(K-1-i)} \right)$$

$$= (id_A \otimes m) \circ \left( id_A \otimes m^{(i-1)} \otimes m^{(K-1-i)} \right)$$

(11.18)

---

\(^{276}\) That is, the rule stating that the product of several elements of a $k$-algebra does not depend on the bracketing.

\(^{277}\) Because if $i = 0$, then $m \circ \left( m^{(i)} \otimes m^{(K-1-i)} \right) = m \circ \left( id_A \otimes m^{(K-1)} \right) = m^{(K)}$ (by the inductive definition of $m^{(K)}$).
Now, the upper left triangle in the diagram

(11.19)

is commutative (by (11.18)), and so is the lower left triangle (according to (11.17)). Since the square in the diagram (11.19) is also commutative (by the commutativity of (1.1)), we thus conclude that the whole diagram (11.19) is commutative. Hence, following the outermost arrows in this diagram, we obtain $m \circ (\text{id}_A \otimes m^{(k-1)}) = m \circ (m^{(i)} \otimes m^{(K-1-i)})$. Now, the recursive definition of $m^{(K)}$ yields $m^{(K)} = m \circ (\text{id}_A \otimes m^{(K-1)}) = m \circ (m^{(i)} \otimes m^{(K-1-i)})$. Hence, (11.16) is proven.

We thus have shown that Exercise 1.39(a) holds for $k = K$. This completes the induction step, and thus Exercise 1.39(a) is solved by induction.

(b) Let $k \geq 1$. Then, Exercise 1.39(a) (applied to $i = k-1$) yields $m^{(k)} = m \circ \left( m^{(k-1)} \otimes m^{(k-1-(k-1))} \right) = m \circ (m^{(k-1)} \otimes \text{id}_A)$. Thus, Exercise 1.39(b) is solved.

(c) We will solve Exercise 1.39(c) by induction over $k$.

The induction base ($k = 0$) is vacuously true, since there exists no $0 \leq i \leq k-1$ for $k = 0$. So let us proceed to the induction step. Let $K$ be a positive integer. We want to prove that the claim of Exercise 1.39(c) holds for $k = K$, assuming (as the induction hypothesis) that the claim of Exercise 1.39(c) holds for $k = K-1$.

So let $0 \leq i \leq K - 1$ be arbitrary. Thus, $K - 1 \geq 0$, so that $K \geq 1$. We are going to prove that

(11.20)

$$m^{(K)} = m^{(K-1)} \circ (\text{id}_{A^{(i)}} \otimes m \otimes \text{id}_{A^{(K-1-i)}}).$$

We must be in one of the following three cases:

Case 1: We have $i \neq 0$.
Case 2: We have $i \neq K - 1$.
Case 3: We have neither $i \neq 0$ nor $i \neq K - 1$.

Let us consider Case 1 first. In this case, we have $i \neq 0$. Hence, $i \geq 1$ (because $0 \leq i$), so that $i - 1 \geq 0$. Thus, we can apply Exercise 1.39(c) to $K - 1$ and $i - 1$ instead of $k$ and $i$ (because we have assumed that the claim of Exercise 1.39(c) holds for $k = K - 1$). As a result, we obtain

$$m^{(K-1)} = m^{((K-1)-1)} \circ (\text{id}_{A^{(i-1)}} \otimes m \otimes \text{id}_{A^{(K-1-(i-1))}})$$

$$= m^{(K-2)} \circ (\text{id}_{A^{(i-1)}} \otimes m \otimes \text{id}_{A^{(K-1-i)}}).$$
But \( K \geq 1 \). Hence, the recursive definition of \( m^{(K)} \) yields

\[
m^{(K)} = m \circ \left( \text{id}_A \otimes \underbrace{m^{(K-1)}}_{= m^{(K-2)} \circ (\text{id}_{A \otimes (K-1)} \otimes m \otimes \text{id}_{A^{\otimes (K-1)}})} \right) \\
= m \circ \left( \text{id}_A \otimes \left( m^{(K-2)} \circ (\text{id}_{A \otimes (K-1)} \otimes m \otimes \text{id}_{A^{\otimes (K-1)}}) \right) \right) \\
= m \circ \left( \text{id}_A \otimes \left( m^{(K-2)} \circ (\text{id}_{A^{\otimes (K-1)}} \otimes m \otimes \text{id}_{A^{\otimes (K-1)}}) \right) \right) \\
= m \circ \left( \text{id}_A \otimes m^{(K-2)} \circ (\text{id}_{A^{\otimes (K-1)}} \otimes m \otimes \text{id}_{A^{\otimes (K-1)}}) \right).
\]

(11.21)

But \( i \geq 1 \) and \( i \leq K - 1 \) together yield \( 1 \leq i \leq K - 1 \), so that \( K - 1 \geq 1 \). Thus, the recursive definition of \( m^{(K-1)} \) yields

\[
m^{(K-1)} = m \circ \left( \text{id}_A \otimes \underbrace{m^{(K-1)-1}}_{= m^{(K-2)}} \right) = m \circ \left( \text{id}_A \otimes m^{(K-2)} \right).
\]

(11.22)

Hence, (11.21) becomes

\[
m^{(K)} = m \circ \left( \text{id}_A \otimes m^{(K-2)} \circ (\text{id}_{A^{\otimes (K-1)}} \otimes m \otimes \text{id}_{A^{\otimes (K-1)}}) \right) \\
= m^{(K-1)} \circ \left( \text{id}_{A^{\otimes (K-1)}} \otimes m \otimes \text{id}_{A^{\otimes (K-1)}} \right).
\]

Thus, (11.20) is proven in Case 1.

Let us next consider Case 2. In this case, we have \( i \neq K - 1 \). Hence, \( i < K - 1 \) (since \( 0 \leq k - 1 \)), so that \( i \leq (K - 1) - 1 \) (since \( i \) and \( K - 1 \) are integers). Thus, we can apply Exercise 1.39(c) to \( K - 1 \) instead of \( k \) (because we have assumed that the claim of Exercise 1.39(c) holds for \( k = K - 1 \)). As a result, we obtain

\[
m^{(K-1)} = m^{((K-1)-1)} \circ \left( \text{id}_{A^{\otimes (K-1)}} \otimes m \otimes \text{id}_{A^{\otimes ((K-1)-1)}} \right) \\
= m^{(K-2)} \circ \left( \text{id}_{A^{\otimes (K-1)-1}} \right).
\]
But $K \geq 1$. Hence, Exercise 1.39(b) (applied to $k = K$) yields

$$m^{(K)} = m \circ \left( m^{(K-1)} \otimes id_A \right) = m \circ \left( m^{(K-2)} \circ (id_A \otimes 1) \otimes id_A \right) = m \circ \left( m^{(K-2)} \otimes id_A \right) \circ \left( id_A \otimes m \otimes id_A \otimes (K-1-1) \otimes id_A \right) = m \circ \left( m^{(K-2)} \otimes id_A \right) \circ \left( id_A \otimes m \otimes id_A \otimes (K-1-1) \otimes id_A \right)$$

(11.23)

But $0 \leq i \leq K - 1$ yields $1 \leq K - 1$, so that $K - 1 \geq 1$. Thus, Exercise 1.39(b) (applied to $k = K - 1$) yields

$$m^{(K-1)} = m \circ \left( m^{((K-1)-1)} \otimes id_A \right) = m \circ \left( m^{(K-2)} \otimes id_A \right).$$

Hence, (11.23) becomes

$$m^{(K)} = m \circ \left( m^{(K-2)} \otimes id_A \right) \circ \left( id_A \otimes m \otimes id_A \otimes (K-1-1) \otimes id_A \right) = m^{(K-1)} \circ \left( id_A \otimes m \otimes id_A \otimes (K-1-1) \right).$$

Thus, (11.20) is proven in Case 2.

Let us finally consider Case 3. In this case, we have neither $i \neq 0$ nor $i \neq K - 1$. Hence, we have both $i = 0$ and $i = K - 1$. Thus, $K - 1 = i = 0$, so that $K = 1$. Thus,

$$m^{(K)} = m^{(1)} = m \circ \left( id_A \otimes id_A \right) = m \circ id_A = m.$$ 

Compared with

$$m^{(K-1)} \circ (id_A \otimes m \otimes id_A \otimes (K-1-1)) = m^{(0)} \circ (id_A \otimes m \otimes id_A \otimes (K-1-1)) = id_A \otimes m = m,$$

this yields $m^{(K)} = m^{(K-1)} \circ (id_A \otimes m \otimes id_A \otimes (K-1-1))$. Thus, (11.20) is proven in Case 3.

We have now proven (11.20) in each of the three Cases 1, 2 and 3. Since these three Cases cover all possibilities, this shows that (11.20) always holds (for all $0 \leq i \leq K - 1$). In other words, the claim of Exercise 1.39(c) holds for $k = K$. This completes the induction step. The induction proof of the claim of Exercise 1.39(c) is therefore complete.
(d) Let \( k \geq 1 \). Applying Exercise 1.39(c) to \( i = 0 \), we obtain
\[
m^{(k)} = m^{(k-1)} \circ \left( \text{id}_A \otimes m \otimes \text{id}_A \otimes (k-1-0) \right) = m^{(k-1)} \circ (m \otimes \text{id}_A \otimes (k-1)) .
\]

Applying Exercise 1.39(c) to \( i = k-1 \), we obtain
\[
m^{(k)} = m^{(k-1)} \circ \left( \text{id}_A \otimes (k-1) \otimes m \otimes \text{id}_A \otimes (k-1-(k-1)) \right) = m^{(k-1)} \circ \left( \text{id}_A \otimes (k-1) \otimes m \otimes \text{id}_A \otimes (k-1) \right) = m^{(k-1)} \circ (\text{id}_A \otimes (k-1) \otimes m) .
\]

This solves Exercise 1.39(d).

11.9. Solution to Exercise 1.40. Solution to Exercise 1.40. A solution for Exercise 1.40 can be obtained by reversing all arrows (and renaming \( A, m \) and \( m^{(k)} \) by \( C, \Delta \) and \( \Delta^{(k)} \)) in the element-free solution of Exercise 1.39 that we gave above.

11.10. Solution to Exercise 1.42. Solution to Exercise 1.42. (a) We will solve Exercise 1.42(a) by induction over \( k \):

The induction base (\( k = 0 \)) requires us to prove that the map \( m^{(0)}_H : H^{\otimes (0+1)} \rightarrow H \) is a \( k \)-coalgebra homomorphism. But this is obvious, because \( m^{(0)}_H = \text{id}_H \) (by the definition of \( m^{(0)}_H \)).

Now, let us proceed to the induction step. Let \( K \) be a positive integer. We want to prove that the claim of Exercise 1.42(a) holds for \( k = K \). Assuming (as the induction hypothesis) that the claim of Exercise 1.42(a) holds for \( k = K-1 \).

By the axioms of a bialgebra, we know that \( m_H \) is a \( k \)-coalgebra homomorphism (since \( H \) is a \( k \)-bialgebra).

We have \( m^{(K)}_H = m_H \circ \left( \text{id}_H \otimes m^{(K-1)}_H \right) \) (by the recursive definition of \( m^{(K)}_H \)). We know that \( m^{(K-1)}_H : H^{\otimes ((K-1)+1)} \rightarrow H \) is a \( k \)-coalgebra homomorphism (since the claim of Exercise 1.42(a) holds for \( k = K-1 \)).

Of course, \( \text{id}_H : H \rightarrow H \) is also a \( k \)-coalgebra homomorphism. Exercise 1.13(b) (applied to \( C = H, C' = H, D = H^{\otimes ((K-1)+1)}, D' = H, f = \text{id}_H \) and \( g = m^{(K-1)}_H \)) thus yields that the map \( \text{id}_H \otimes m^{(K-1)}_H : H \otimes H^{\otimes ((K-1)+1)} \rightarrow H \otimes H \) is a \( k \)-coalgebra homomorphism. Since \( H \otimes H^{\otimes ((K-1)+1)} = H^{\otimes ((K-1)+1)+1} = H^{\otimes (K+1)} \), this rewrites as follows: The map \( \text{id}_H \otimes m^{(K-1)}_H : H^{\otimes (K+1)} \rightarrow H \otimes H \) is a \( k \)-coalgebra homomorphism. Hence, \( m_H \circ \left( \text{id}_H \otimes m^{(K-1)}_H \right) \) is the composition of two \( k \)-coalgebra homomorphisms (namely, of \( m_H \) and \( \text{id}_H \otimes m^{(K-1)}_H \)), hence a \( k \)-coalgebra homomorphism itself (since the composition of any two \( k \)-coalgebra homomorphisms is a \( k \)-coalgebra homomorphism). In other words, \( m^{(K)}_H \) is a \( k \)-coalgebra homomorphism (since \( m^{(K)}_H = m_H \circ \left( \text{id}_H \otimes m^{(K-1)}_H \right) \)). In other words, the claim of Exercise 1.42(a) holds for \( k = K \). This completes the induction step. Thus, Exercise 1.42(a) is solved by induction.

(b) The solution of Exercise 1.42(b) can be obtained from the solution of Exercise 1.42(a) by “reversing all arrows”. (The details of this are left to the reader.)

(d) Let \( \ell \in \mathbb{N} \). Exercise 1.42(a) (applied to \( \ell \) instead of \( k \)) yields that the map \( m^{(\ell)}_H : H^{\otimes (\ell+1)} \rightarrow H \) is a \( k \)-coalgebra homomorphism.

For every \( k \)-coalgebra \( C \), consider the map \( \Delta_C^{(k)} : C \rightarrow C^{\otimes (k+1)} \) (this is the map \( \Delta^{(k)} \) defined in Exercise 1.40). This map \( \Delta_C^{(k)} \) is clearly functorial in \( C \). By this we mean that if \( C \) and \( D \) are any two
k-coalgebras, and \( f : C \to D \) is any \( k \)-coalgebra homomorphism, then the diagram

\[
\begin{array}{c}
C \\
\downarrow \Delta^{(k)}_C \\
C \otimes (k+1) \\
\downarrow f \otimes (k+1) \\
D \otimes (k+1)
\end{array}
\]

commutes.\(^{278}\) We can apply this to \( C = H \otimes (t+1) \), \( D = H \) and \( f = m^{(t)}_H \) (since \( m^{(t)}_H \) is a \( k \)-coalgebra homomorphism). As a result, we conclude that the diagram

\[
\begin{array}{c}
H \otimes (t+1) \\
\downarrow \Delta^{(k)}_{H \otimes (t+1)} \\
(H \otimes (t+1))^\otimes (k+1) \\
\downarrow (m^{(t)}_H)^\otimes (k+1) \\
H \otimes (k+1)
\end{array}
\]

commutes. In other words, \( (m^{(t)}_H)^\otimes (k+1) \circ \Delta^{(k)}_{H \otimes (t+1)} = \Delta^{(k)}_H \circ m^{(t)}_H \). This solves Exercise 1.42(d).

(c) The solution of Exercise 1.42(c) can be obtained from the solution of Exercise 1.42(d) by “reversing all arrows”.

11.11. \textbf{Solution to Exercise 1.43.} Solution to Exercise 1.43. If \( k = 0 \), then the statement of Exercise 1.43 is clearly true.\(^{279}\) Hence, for the rest of this solution, we can WLOG assume that we don’t have \( k = 0 \). Assume this.

We shall show that

\[
(11.25) \quad f_1 \ast f_2 \ast \cdots \ast f_i = m^{(i-1)}_A \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_i) \circ \Delta^{(i-1)}_C
\]

for every \( i \in \{1, 2, \ldots, k\} \).

\textit{Proof of (11.25):} We will prove (11.25) by induction over \( i \):

\textit{Induction base:} If \( i = 1 \), then (11.25) holds.\(^{280}\) This completes the induction base.

\textit{Induction step:} Let \( I \in \{1, 2, \ldots, k\} \) be such that \( I < k \). Assume that (11.25) holds for \( i = I \). We must prove that (11.25) holds for \( i = I + 1 \).

\(^{278}\) This can be proven by induction over \( k \) in a completely straightforward manner.

\(^{279}\) Proof. Let \( k = 0 \). Then,

\[
f_1 \ast f_2 \ast \cdots \ast f_k = f_1 \ast f_2 \ast \cdots \ast f_0 = \text{(empty product in } (\text{Hom}(C, A), \ast)) \]

\[
= 1_{(\text{Hom}(C, A), \ast)} = u_A \ast C
\]

and

\[
m^{(k-1)}_A \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_k) \circ \Delta^{(k-1)}_C = m^{(0-1)}_A \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_0) \circ \Delta^{(k-1)}_C
\]

\[
= m^{(0-1)}_A = u_A = (\text{empty tensor product}) = \text{id}_k = \Delta^{(k-1)}_C
\]

\[
= u_A \circ \text{id}_k \circ \Delta^{(k-1)}_C = u_A \ast C.
\]

Hence, \( f_1 \ast f_2 \ast \cdots \ast f_k = u_A \ast C = m^{(k-1)}_A \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_k) \circ \Delta^{(k-1)}_C \), so that Exercise 1.43 is solved in the case when \( k = 0 \).

\(^{280}\) Proof. Assume that \( i = 1 \). Then, since \( i = 1 \), we have

\[
f_1 \ast f_2 \ast \cdots \ast f_i = f_1 \ast f_2 \ast \cdots \ast f_1 = f_1.
\]

Also, since \( i = 1 \), we have

\[
m^{(1-1)}_A \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_1) \circ \Delta^{(1-1)}_C = m^{(0-1)}_A \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_1) \circ \Delta^{(1-1)}_C
\]

\[
= m^{(0-1)}_A = u_A = \Delta^{(1-1)}_C = \text{id}_A \circ f_1 \circ \text{id}_C = f_1.
\]

Thus, \( f_1 \ast f_2 \ast \cdots \ast f_i = f_1 = m^{(i-1)}_A \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_i) \circ \Delta^{(i-1)}_C \). Hence, (11.25) holds, qed.
Denote the \( k \)-linear map \( f_1 \otimes f_2 \otimes \cdots \otimes f_I : C^{\otimes I} \to A^{\otimes I} \) by \( g \). Then, \( g = f_1 \otimes f_2 \otimes \cdots \otimes f_I \). But we assumed that (11.25) holds for \( i = I \). In other words,
\[
f_1 \ast f_2 \ast \cdots \ast f_I = m_A^{(I-1)} \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_I) \circ \Delta_C^{(I-1)} = m_A^{(I-1)} \circ g \circ \Delta_C^{(I-1)}.
\]
Now,
\[
f_1 \ast f_2 \ast \cdots \ast f_{I+1} = (f_1 \ast f_2 \ast \cdots \ast f_I) \ast f_{I+1}
\]
\[
= m_A \circ \left( (m_A^{(I-1)} \circ g \circ \Delta_C^{(I-1)}) \otimes (\text{id} \circ f_{I+1} \circ \text{id}_C) \right) \circ \Delta_C
\]
(by the definition of convolution)
\[
= m_A \circ \left( (m_A^{(I-1)} \circ g \circ \Delta_C^{(I-1)}) \otimes (\text{id} \circ f_{I+1} \circ \text{id}_C) \right) \circ \Delta_C
\]
\[
= m_A \circ \left( m_A^{(I-1)} \circ \text{id}_A \circ f \right) \circ g \circ f_{I+1} \circ \Delta_C
\]
\[
= m_A \circ \left( m_A^{(I-1)} \circ \text{id}_A \right) \circ g \circ f_{I+1} \circ \Delta_C
\]
(by Exercise 1.39(b), applied to \( k = I \))
\[
= m_A^{(I)} \circ g \circ f_{I+1} \circ \Delta_C
\]
(by Exercise 1.40(b), applied to \( k = I \))
\[
= m_A^{(I)} \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_{I+1}) \circ \Delta_C^{(I+1)-1}.
\]
In other words, (11.25) holds for \( i = I + 1 \). Thus, the induction step is complete. Therefore, (11.25) is proven by induction.

Recall that we don’t have \( k = 0 \). Hence, \( k > 0 \), so that \( k \in \{1, 2, \ldots, k\} \). Therefore, (11.25) (applied to \( i = k \)) yields
\[
f_1 \ast f_2 \ast \cdots \ast f_k = m_A^{(k-1)} \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_k) \circ \Delta_C^{(k-1)}.
\]
This solves Exercise 1.43.

11.12. Solution to Exercise 1.47. Solution to Exercise 1.47. The solution of this exercise, of course, proceeds by inverting all the arrows in the proof of Proposition 1.32, but this is complicated by the fact that said proof first has to be rewritten in terms of arrows (i.e., freed of all uses of elements)\(^{281}\). We give such a solution to Exercise 1.47 further below – see the solution of Exercise 1.48(c).

Alternatively, solutions of Exercise 1.47 can be found in many places: e.g., [1, Thm. 2.1.4(iv)], [86, III.(3.5)], [116, Prop. I.7.1 2]]

11.13. Solution to Exercise 1.48. Solution to Exercise 1.48. (a) We shall give two solutions of Exercise 1.48(a). The first of these solutions will be a (completely straightforward) generalization of the proof of Proposition 1.32 (which proposition, as we will see in solving Exercise 1.48(c), is a particular case of this exercise). The second solution will be element-free (i.e., it will only manipulate linear maps, but never refer to elements of \( C \) or \( A \)), but essentially is just a result of rewriting the first solution in element-free terms. One of the advantages of an element-free solution is the ease of applying it to more general contexts (such as

\(^{281}\)Rewriting a proof in terms of arrows is usually an exercise in category theory (or, rather, category practice); see the solution to Exercise 1.8 (or the solution to Exercise 1.59 further below) for how this is done (in a simple case).
We use Sweedler’s sumfree notations in the form ∆(r) = r ⊗ r. Thus,

(These definitions make sense, since each of r is a k-algebra anti-homomorphism (because r = r*). In other words, we need to prove that r = r* is a k-algebra anti-homomorphism. To prove this, we shall not fix a and b. Instead, we consider C ⊗ C as a k-coalgebra, and A as a k-algebra. Then, Hom(C ⊗ C, A) is an associative algebra with a convolution product ⊕ (to be distinguished from the convolution * on Hom(C, A)), having two-sided identity element u_A ⊗ C. We define three k-linear maps f, g and h from C ⊗ C to A as follows:

f(a ⊗ b) = r(a) r(b) for all a ∈ C and b ∈ C;
g(a ⊗ b) = τ(b) τ(a) for all a ∈ C and b ∈ C;
h(a ⊗ b) = τ(ab) for all a ∈ C and b ∈ C.

(11.26)

Once this equality is proven, we will then obtain

h = h ⊕ (u_A ⊗ C) (since u_A ⊗ C is the identity element of Hom(C ⊗ C, A), ⊕))
= h ⊕ (f ⊕ g) = (h ⊕ f) ⊕ g (by the associativity of the convolution ⊕)
= (u_A ⊗ C) ⊕ g = g (since u_A ⊗ C is the identity element of Hom(C ⊗ C, A), ⊕).

In order to prove (11.26), we evaluate the three maps h ⊕ f, u_A ⊗ C and f ⊕ g on pure tensors a ⊗ b ∈ C ⊗ C. We use Sweedler’s sumfree notations in the form ∆(a) = ∑(a) a_1 ⊗ a_2 and ∆(b) = ∑(b) b_1 ⊗ b_2; thus, ∆(ab) = ∑(a,b) a_1 b_1 ⊗ a_2 b_2 (since ∆ is a k-algebra homomorphism). We then obtain

(u_A ⊗ C) (a ⊗ b) = u_A (ε_C (a) ε_C (b)) = u_A (ε_C (ab)).

Furthermore,

(h ⊕ f) (a ⊗ b) = ∑(a,b) [h(a_1 ⊗ b_1) f(a_2 ⊗ b_2)]
= ∑(a,b) [τ(a_1 b_1) r(a_2) r(b_2)]
= ∑(a,b) [τ(a_1 b_1) r(a_2 b_2) = (τ ∗ r) (ab)]
= (u_A ⊗ C) (ab) = u_A (ε_C (ab))

(11.27)
and
\[(f \otimes g)(a \otimes b) = \sum_{(a),(b)} f(a_1 \otimes b_1) g(a_2 \otimes b_2) = \sum_{(a),(b)} r(a_1) r(b_1) \overline{r}(b_2) \overline{r}(a_2)\]
\[= \sum_{(a)} r(a_1) \left( \sum_{(b)} r(b_1) \overline{r}(b_2) \right) \overline{r}(a_2) = \sum_{(a)} r(a_1) (u_A \circ \epsilon_C)(b) \overline{r}(a_2)\]
\[= \sum_{(a)} r(a_1) \epsilon_C(b) \overline{r}(a_2) = \sum_{(a)} r(a_1) \epsilon_C(b) \epsilon_C(a) \overline{r}(a_2)\]
\[= (u_A \circ \epsilon_C)(a) \epsilon_C(b) = (u_A \circ \epsilon_C)(a)(u_A \circ \epsilon_C)(b) = (u_A \circ \epsilon_C)(ab)\]
\[= u_A(\epsilon_C(ab)).\]

These three results being all equal, we thus have shown that the maps \(h \otimes f, u_A \epsilon_{C \otimes C}\) and \(f \otimes g\) are equal on each pure tensor. Correspondingly, these maps must be identical (since they are \(k\)-linear). In other words, (11.26) holds. As explained above, this yields \(h = g\). Thus, every \(a \in C\) and \(b \in C\) satisfy
\[\overline{r}(ab) = h(a \otimes b) = g(a \otimes b) = \overline{r}(b) \overline{r}(a)\] (by the definition of \(g(a \otimes b)\)).

As we already know, this completes the solution of Exercise 1.48(a).

Second solution to Exercise 1.48(a). Let us first recall some linear-algebraic facts. One such fact states that if \(U, V, U'\) and \(V'\) are four \(k\)-modules and \(x : U \rightarrow U'\) and \(y : V \rightarrow V'\) are two \(k\)-linear maps, then
\[(y \otimes x) \circ T_{U',V'} = T_{U,V'} \circ (x \otimes y).\]

For every \(k\)-module \(U\), let \(\text{kan}_{1,U} : U \rightarrow U \otimes k\) and \(\text{kan}_{2,U} : U \rightarrow k \otimes U\) be the canonical \(k\)-module isomorphisms. These two isomorphisms are related to each other via the equalities
\[\text{kan}_{2,U}^{-1} \circ T_{U,k} = \text{kan}_{1,U}^{-1},\]
\[\text{kan}_{1,U}^{-1} \circ T_{k,U} = \text{kan}_{2,U}^{-1},\]
\[T_{k,U} \circ \text{kan}_{2,U} = \text{kan}_{1,U},\]
\[T_{U,k} \circ \text{kan}_{1,U} = \text{kan}_{2,U}\]
for every \(k\)-module \(U\). Moreover, every \(k\)-modules \(U\) and \(V\) satisfy
\[\text{id}_U \otimes \text{kan}_{1,V} = \text{kan}_{1,U \otimes V},\]
\[\text{kan}_{2,V} \otimes \text{id}_U = \text{kan}_{2,V \otimes U},\]
\[\text{id}_U \otimes \text{kan}_{1,V} = \text{kan}_{1,U \otimes V},\]
\[\text{kan}_{2,V} \otimes \text{id}_U = \text{kan}_{2,V \otimes U},\]
\[\text{id}_U \otimes \text{kan}_{2,V} = \text{kan}_{1,U \otimes V},\]
\[\text{id}_U \otimes \text{kan}_{2,V} = \text{kan}_{1,U \otimes V}.\]

Furthermore, if \(U\) and \(V\) are two \(k\)-modules and \(\alpha : U \rightarrow V\) is a \(k\)-linear map, then
\[\text{kan}_{1,V} \circ \alpha = (\alpha \otimes \text{id}_k) \circ \text{kan}_{1,U},\]
\[\text{kan}_{2,V} \circ \alpha = (\text{id}_k \circ \alpha) \circ \text{kan}_{2,U},\]
\[\text{kan}_{1,V}^{-1} \circ (\alpha \otimes \text{id}_k) = (\alpha \circ \text{kan}_{1,U}^{-1}),\]
\[\text{kan}_{2,V}^{-1} \circ (\text{id}_k \circ \alpha) = (\alpha \circ \text{kan}_{2,U}^{-1}).\]
Now, let us step to the actual solution of Exercise 1.48(a).
Let $\varphi = r^{*-1}$. Then, $\varphi = r^{*(-1)}$ is the $*$-inverse of $r$. Thus, $\varphi * r = 1_{(\text{Hom}(C, A), *)} = u_A \circ \epsilon_C$ and similarly $r * \varphi = u_A \circ \epsilon_C$.

Recall that $C$ is a $k$-coalgebra. By the axioms of a $k$-coalgebra, this shows that
\[
(\Delta_C \otimes \text{id}_C) \circ \Delta_C = (\text{id}_C \otimes \Delta_C) \circ \Delta_C,
\]
\[
(\text{id}_C \otimes \epsilon_C) \circ \Delta_C = \text{kan}_{1,C};
\]
\[
(\epsilon_C \otimes \text{id}_C) \circ \Delta_C = \text{kan}_{2,C}.
\]

Also, recall that $C$ is a $k$-algebra. By the axioms of a $k$-algebra, this shows that
\[
m_C \circ (m_C \otimes \text{id}_C) = m_C \circ (\text{id}_C \otimes m_C);
\]
\[
m_C \circ (\text{id}_C \otimes u_C) = \text{kan}^{-1}_{1,C};
\]
\[
m_C \circ (u_C \otimes \text{id}_C) = \text{kan}^{-1}_{2,C}.
\]

Also, recall that $C$ is a $k$-bialgebra. Due to the axioms of a $k$-bialgebra, this shows that $\Delta_C$ and $\epsilon_C$ are $k$-algebra homomorphisms, and that $m_C$ and $u_C$ are $k$-coalgebra homomorphisms.

Furthermore, $A$ is a $k$-algebra. By the axioms of a $k$-algebra, this shows that
\[
m_A \circ (m_A \otimes \text{id}_A) = m_A \circ (\text{id}_A \otimes m_A);
\]
\[
m_A \circ (\text{id}_A \otimes u_A) = \text{kan}^{-1}_{1,A};
\]
\[
m_A \circ (u_A \otimes \text{id}_A) = \text{kan}^{-1}_{2,A}.
\]

Since $\epsilon_C$ is a $k$-algebra homomorphism, we have $\epsilon_C \circ u_C = u_k = \text{id}_k$.

Recall that $\text{kan}_{1,k}$ is the canonical $k$-module isomorphism $k \to k \otimes k$. Hence, $u_C \otimes u_C = (u_C \otimes u_C) \circ \text{kan}_{1,k}$ (by the definition of the $k$-algebra $C \otimes C$). Also, $\text{kan}_{1,k} = \text{kan}_{2,k}$ (since each of $\text{kan}_{1,k}$ and $\text{kan}_{2,k}$ is the canonical $k$-module isomorphism $k \to k \otimes k$).

We know that $\varphi : C \to A$ is a $k$-algebra homomorphism. In other words, $r$ is a $k$-linear map satisfying $r \circ m_C = m_A \circ (r \otimes r)$ and $r \circ u_C = u_A$.

Now, we need to prove that $r^{*-1}$ is a $k$-algebra anti-homomorphism. In other words, we need to prove that $\varphi$ is a $k$-algebra anti-homomorphism (since $\varphi = r^{*(-1)}$). In other words, we need to prove that $\varphi$ satisfies the two equations $\varphi \circ m_C = m_A \circ (\varphi \otimes \varphi) \circ T_{C,C}$ and $\varphi \circ u_C = u_A$ (because these two equations are what makes $\varphi$ a $k$-algebra anti-homomorphism, according to the definition of a “$k$-algebra anti-homomorphism”).

Let us first prove the equality $\varphi \circ u_C = u_A$.

We have $r * \varphi = u_A \circ \epsilon_C$, so that $(r * \varphi) \circ u_C = u_A \circ \epsilon_C \circ u_C = u_A$. Hence,
\[
u_A = (r * \varphi) \circ u_C = m_A \circ (r \otimes \varphi) \circ \Delta_C \circ u_C
\]
(by the definition of convolution)
\[
= m_A \circ (r \otimes \varphi) \circ (u_C \otimes u_C) \circ \text{kan}_{1,k}
\]
(by $\Delta_C$ is a $k$-algebra homomorphism)
\[
= m_A \circ ((u_A \circ \varphi) \otimes (u_A \circ \varphi)) \circ \text{kan}_{2,k}
\]
(by $\text{id}_A \circ (\varphi \otimes \varphi)$)
\[
= m_A \circ ((u_A \circ \varphi) \otimes (u_A \circ \varphi)) \circ \text{kan}_{2,k}
\]
(by $\text{id}_A \circ (\varphi \otimes \varphi)$)
\[
= \varphi \circ u_C.
\]

Hence, $\varphi \circ u_C = u_A$. 
Now, it remains to prove the equality \( \tau \circ m_C = m_A \circ (\tau \otimes \tau) \circ T_{C,C} \). This is harder. Let us first make some preparations.

Recall that \( C \otimes C \) is a \( k \)-coalgebra (since \( C \) is a \( k \)-coalgebra), and \( A \) is an \( k \)-algebra. Thus, \( \text{Hom} (C \otimes C, A) \) is an associative algebra with respect to convolution. We shall denote the convolution on \( \text{Hom} (C \otimes C, A) \) by \( \otimes \) rather than by \( * \) (in order to distinguish it from the convolution \( * \) on \( \text{Hom} (C, A) \)). The algebra \( (\text{Hom} (C \otimes C, A), \otimes) \) has two-sided identity element \( u_A \circ \epsilon_{C \otimes C} \).

We define a \( k \)-linear map \( f : C \otimes C \rightarrow A \) by \( f = m_A \circ (r \otimes r) \).

We define a \( k \)-linear map \( g : C \otimes C \rightarrow A \) by \( g = m_A \circ T_{A,A} \circ (\tau \otimes \tau) \).

We define a \( k \)-linear map \( h : C \otimes C \rightarrow A \) by \( h = \tau \circ m_C \).

Clearly, \( f, g \) and \( h \) are three elements of \( \text{Hom} (C \otimes C, A) \). Our next goal is to prove that

\[
(11.47) \quad h \otimes f = u_A \circ \epsilon_{C \otimes C} = f \otimes g.
\]

Once this equality is proven, we will then obtain

\[
h = h \otimes (u_A \circ \epsilon_{C \otimes C}) = (h \otimes f) \circ g \quad \text{(by (11.44))}
\]

\[
= h \otimes (f \otimes g) = (h \otimes f) \circ g \quad \text{(by the associativity of the convolution \( \otimes \))}
\]

\[
= (u_A \circ \epsilon_{C \otimes C}) \circ g = g
\]

\[
(11.45) \quad (u_A \circ \epsilon_{C \otimes C}) \circ g = g
\]

\[
(\text{since } u_A \circ \epsilon_{C \otimes C} \text{ is the identity element of } (\text{Hom} (C \otimes C, A), \otimes)).
\]

So let us concentrate on proving (11.44). Let us first notice that

\[
(11.46) \quad u_A \circ \epsilon_C \circ m_C = u_A \circ \epsilon_{C \otimes C}.
\]

Next, we make the following observations:

- First, we notice that

\[
(11.47) \quad h \otimes f = m_A \circ (h \otimes f) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C).
\]

- Next, we notice that

\[
(11.48) \quad m_A \circ (h \otimes f) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C)
\]

\[
= m_A \circ ((\tau \circ m_C) \otimes (m_A \circ (r \otimes r))) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C).
\]

\[282\text{Proof of (11.46): We know that } m_C \text{ is a } k \text{-coalgebra homomorphism. Hence, } \epsilon_C \circ m_C = \epsilon_{C \otimes C}. \text{ Thus, } u_A \circ \epsilon_C \circ m_C = u_A \circ \epsilon_{C \otimes C}. \text{ This proves (11.46).}
\]

\[283\text{Proof of (11.47): By the definition of the convolution } h \otimes f, \text{ we have}
\]

\[
h \otimes f = m_A \circ (h \otimes f) \circ \Delta_{C \otimes C} \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C).
\]

\[
= m_A \circ ((\tau \circ m_C) \otimes (m_A \circ (r \otimes r))) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C).
\]

This proves (11.47).

\[284\text{Proof of (11.48): We have}
\]

\[
m_A \circ (h \otimes f) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C)
\]

\[
= m_A \circ ((\tau \circ m_C) \otimes (m_A \circ (r \otimes r))) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C).
\]

This proves (11.48).
• Next, we have

\[ m_A \circ ((\mathfrak{f} \circ m_C) \otimes (m_A \circ (r \otimes r))) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \]

(11.49)

\[ = m_A \circ ((\mathfrak{f} \circ m_C) \otimes (r \circ m_C)) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \]

(11.50)

• Furthermore, we have

\[ m_A \circ ((\mathfrak{f} \circ m_C) \otimes (r \circ m_C)) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \]

(11.51)

• Moreover, we have

\[ m_A \circ (\mathfrak{f} \otimes r) \circ \Delta_C \circ m_C = (\mathfrak{f} \ast r) \circ m_C. \]

(11.52)

Finally, we have

\[ (\mathfrak{f} \ast r) \circ m_C = u_A \circ \epsilon_C \circ m_C. \]

Now, we have

\[ h \otimes f = m_A \circ (h \otimes f) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \quad \text{(by (11.47))} \]

\[ = m_A \circ ((\mathfrak{f} \circ m_C) \otimes (m_A \circ (r \otimes r))) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \quad \text{(by (11.48))} \]

\[ = m_A \circ ((\mathfrak{f} \circ m_C) \otimes (r \circ m_C)) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \quad \text{(by (11.49))} \]

\[ = m_A \circ (\mathfrak{f} \otimes r) \circ \Delta_C \circ m_C \quad \text{(by (11.50))} \]

\[ = (\mathfrak{f} \ast r) \circ m_C \quad \text{(by (11.51))} \]

\[ = u_A \circ \epsilon_C \circ m_C \quad \text{(by (11.52))} \]

\[ = u_A \circ \epsilon_{C \otimes C} \quad \text{(by (11.46)).} \]

The first equality in (11.44) is thus proven.

We shall next show the second equality in (11.44).

285 Proof of (11.49): We have

\[ m_A \circ \left( (\mathfrak{f} \circ m_C) \otimes (r \circ m_C) \right) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \]

\[ = m_A \circ ((\mathfrak{f} \circ m_C) \otimes (r \circ m_C)) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) . \]

This proves (11.49).

286 Proof of (11.50): We have

\[ m_A \circ ((\mathfrak{f} \circ m_C) \otimes (r \circ m_C)) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \]

\[ = m_A \circ (\mathfrak{f} \otimes r) \circ (m_C \circ m_C) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \]

\[ = m_A \circ (\mathfrak{f} \otimes r) \circ m_{C \otimes C} \circ (\Delta_C \otimes \Delta_C) \]

\[ = m_A \circ (\mathfrak{f} \otimes r) \circ \Delta_C \circ m_C . \]

This proves (11.50).

287 Proof of (11.51): The definition of the convolution \( \mathfrak{f} \ast r \) yields \( \mathfrak{f} \ast r = m_A \circ (\mathfrak{f} \otimes r) \circ \Delta_C \). Hence, \( \mathfrak{f} \ast r \circ m_C = m_A \circ (\mathfrak{f} \otimes r) \circ \Delta_C \circ m_C . \)

288 Proof of (11.52): We have \( \mathfrak{f} \ast r \circ m_C = u_A \circ \epsilon_C \circ m_C . \) Thus, (11.52) is proven.
For every \( k \in \mathbb{N} \), let us define the map \( m^{(k)} : A^\otimes(k+1) \to A \) as in Exercise 1.39. We recall that these maps \( m^{(k)} \) are defined by induction over \( k \), with the induction base \( m^{(0)} = \text{id}_A \), and with the induction step

\[
m^{(k)} = m_A \circ \left( \text{id}_A \otimes m^{(k-1)} \right)
\]

for every \( k \geq 1 \).

Out of these maps, we shall only need \( m^{(0)} \), \( m^{(1)} \), \( m^{(2)} \) and \( m^{(3)} \). These satisfy the following formulae:

\[
m^{(0)} = \text{id}_A;
\]

\[
m^{(1)} = m_A;
\]

\[
m^{(2)} = m_A \circ (m_A \otimes \text{id}_A);
\]

\[
m^{(2)} = m_A \circ (\text{id}_A \otimes m_A);
\]

\[
m^{(3)} = m_A \circ (m_A \otimes m_A);
\]

\[
m^{(3)} = m^{(2)} \circ (\text{id}_A \otimes m_A \otimes \text{id}_A).
\]

Now, we make the following observations:

---

289 This is precisely the definition of these maps given in Exercise 1.39, with the only difference that \( m_A \) was denoted by \( m \) in that exercise.

290 Here are proofs for these formulae:

**Proof of (11.54):** The formula (11.54) follows immediately from the definition of \( m^{(0)} \).

**Proof of (11.55):** Applying (11.53) to \( k = 1 \), we obtain \( m^{(1)} = m_A \circ \left( \text{id}_A \otimes \frac{m^{(1-1)}}{m^{(0)}=\text{id}_A} \right) = m_A \circ (\text{id}_A \otimes \text{id}_A) = m_A \). This proves (11.55).

**Proof of (11.57):** Applying (11.53) to \( k = 2 \), we obtain \( m^{(2)} = m_A \circ \left( \text{id}_A \otimes \frac{m^{(2-1)}}{m^{(1)}=m_A} \right) = m_A \circ (\text{id}_A \otimes m_A) \). This proves (11.57).

**Proof of (11.56):** From (11.57), we obtain \( m^{(2)} = m_A \circ (\text{id}_A \otimes m_A) = m_A \circ (m_A \otimes \text{id}_A) \) (since \( m_A \circ (m_A \otimes \text{id}_A) = m_A \circ (\text{id}_A \otimes m_A) \)). This proves (11.56).

**Proof of (11.58):** Applying (11.53) to \( k = 3 \), we obtain

\[
m^{(3)} = m_A \circ \left( \text{id}_A \otimes \frac{m^{(3-1)}}{m^{(2)}=m_A \circ (m_A \otimes m_A)} \right) = m_A \circ ((\text{id}_A \circ \text{id}_A) \otimes (m_A \circ (m_A \otimes m_A)))
\]

\[
= m_A \circ (\text{id}_A \otimes (m_A \circ (m_A \otimes m_A)))
\]

\[
= m_A \circ (m_A \circ (m_A \otimes m_A))
\]

This proves (11.58).

**Proof of (11.59):** Applying (11.53) to \( k = 3 \), we obtain

\[
m^{(3)} = m_A \circ \left( \text{id}_A \otimes \frac{m^{(3-1)}}{m^{(2)}=m_A \circ (m_A \otimes m_A)} \right) = m_A \circ ((\text{id}_A \circ \text{id}_A) \otimes (m_A \circ (m_A \otimes m_A)))
\]

\[
= m_A \circ (\text{id}_A \otimes m_A) \circ (\text{id}_A \otimes (m_A \circ m_A \circ m_A))
\]

This proves (11.59).
• We have

\[(11.60) \quad f \otimes g = m_A \circ (f \otimes g) \circ (\text{id}_C \otimes T_C) \circ \text{id}_C \circ \Delta_C \otimes \Delta_C.\]

\[\text{Proof of (11.60):} \quad \text{By the definition of the convolution map } f \otimes g, \text{ we have}\]

\[f \otimes g = m_A \circ \left( f \otimes g \right) \circ \Delta_C \otimes \Delta_C = m_A \circ \left( f \otimes g \right) \circ (\text{id}_C \otimes T_C) \circ \text{id}_C \circ (\Delta_C \otimes \Delta_C).\]

This proves (11.60).

\[\text{Proof of (11.61):} \quad \text{The equality (11.29) (applied to } U = C, \quad V = C, \quad U' = A, \quad V' = A, \quad x = \tau \text{ and } y = \tau) \text{ yields}\]

\[f \otimes T_C = T_A, \quad g \circ (f \otimes g) \circ (\Delta_C \otimes \Delta_C).\]

\[\text{Hence, } g = m_A \circ (f \otimes g) \circ (\text{id}_C \otimes T_C) \circ \text{id}_C \circ (\Delta_C \otimes \Delta_C).\]

\[\text{Now,}\]

\[m_A \circ \left( f \otimes g \right) \circ \Delta_C \otimes \Delta_C = m_A \circ \left( f \otimes g \right) \circ (\text{id}_C \otimes T_C) \circ \text{id}_C \circ (\Delta_C \otimes \Delta_C).\]

\[\text{This proves (11.61).}\]

\[\text{Proof of (11.62):} \quad \text{We first notice that}\]

\[(id_C \otimes \text{id}_C \otimes T_C, C) \circ (id_C \otimes T_C, C) \circ \text{id}_C = id_C \otimes T_C, C, C.\]

In fact, the two maps \((id_C \otimes \text{id}_C \otimes T_C, C) \circ (id_C \otimes T_C, C) \circ \text{id}_C \circ \text{id}_C \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \circ \text{id}_C) \circ (\Delta_C \otimes \Delta_C).\]

\[\text{Therefore, they must be identical, so that (11.63) holds.}\]
We have

\[ m^{(2)} \circ (r \otimes (u_A \circ \epsilon_C) \otimes \tau) \circ (\text{id}_C \otimes T_{C,C}) \circ \left( \Delta_C \otimes \text{id}_C \right) \]

\[ = m_A \circ (\text{id}_A \otimes \text{kan}_{1,A}^{-1}) \circ (r \otimes \tau \otimes \epsilon_C) \circ \left( \Delta_C \otimes \text{id}_C \right). \]

(11.65)

Now,

\[ m^{(3)} \circ (r \otimes r \otimes \tau) \circ (\text{id}_C \otimes \text{id}_C \otimes T_{C,C}) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ \left( \Delta_C \otimes \Delta_C \right) \]

\[ = m^{(2)} \circ (\text{id}_A \otimes m_A \otimes \text{id}_A) \circ (r \otimes r \otimes \tau) \circ \left( \text{id}_C \otimes T_{C,C} \otimes \text{id}_C \right) \circ \left( \Delta_C \otimes \Delta_C \otimes \text{id}_C \right) \]

\[ = m^{(2)} \circ (\text{id}_A \otimes m_A \otimes \text{id}_A) \circ (r \otimes r \otimes \tau) \circ \left( \text{id}_C \otimes T_{C,C} \otimes \text{id}_C \right) \circ \left( \text{id}_C \otimes \Delta_C \otimes \text{id}_C \right) \circ \left( \Delta_C \otimes \text{id}_C \right). \]

Since

\[ (\text{id}_A \otimes m_A \otimes \text{id}_A) \circ (r \otimes r \otimes \tau) \]

\[ = (\text{id}_A \otimes m_A \otimes \text{id}_A) \circ (r \otimes (r \otimes \tau) \otimes \tau) = \left( \text{id}_A \circ r \otimes \left( m_A \circ (r \otimes \tau) \right) \otimes \left( \text{id}_A \circ r \right) \right) \]

\[ = r \circ (m_A \circ (r \otimes \tau)) \otimes \tau \]

and

\[ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ \left( \text{id}_C \otimes \text{id}_C \otimes \Delta_C \right) \]

\[ = (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ \left( \text{id}_C \otimes \text{id}_C \otimes \Delta_C \right) \]

\[ = (\text{id}_C \otimes \text{id}_C) \otimes \left( T_{C,C} \otimes \text{id}_C \right) \]

\[ = \left( \Delta_C \otimes \text{id}_C \right) \otimes \left( T_{C,C} \otimes \text{id}_C \right) \]

\[ = \left( \text{id}_C \otimes \Delta_C \otimes \text{id}_C \right) \otimes \left( T_{C,C} \otimes \text{id}_C \right) \]

\[ = \left( \text{id}_C \otimes \Delta_C \otimes \text{id}_C \right) \otimes \left( T_{C,C} \otimes \text{id}_C \right) \]

\[ = \left( \text{id}_C \otimes \Delta_C \otimes \text{id}_C \right) \otimes \left( T_{C,C} \otimes \text{id}_C \right) \]

\[ = \left( \text{id}_C \otimes \Delta_C \otimes \text{id}_C \right) \otimes \left( T_{C,C} \otimes \text{id}_C \right) \]

this becomes

\[ m^{(3)} \circ (r \otimes r \otimes \tau) \circ \left( \text{id}_C \otimes \text{id}_C \otimes T_{C,C} \right) \circ \left( \text{id}_C \otimes T_{C,C} \otimes \text{id}_C \right) \circ \left( \Delta_C \otimes \Delta_C \right) \]

\[ = m^{(2)} \circ \left( \text{id}_A \otimes m_A \otimes \text{id}_A \right) \circ \left( r \otimes r \otimes \tau \right) \circ \left( \text{id}_C \otimes T_{C,C} \otimes \text{id}_C \right) \circ \left( \text{id}_C \otimes \Delta_C \otimes \text{id}_C \right) \]

\[ = m^{(2)} \circ \left( r \otimes (m_A \circ (r \otimes \tau)) \otimes \tau \right) \circ \left( \text{id}_C \otimes T_{C,C} \otimes \text{id}_C \right) \circ \left( \text{id}_C \otimes \Delta_C \otimes \text{id}_C \right) \]

\[ = m^{(2)} \circ \left( r \otimes (m_A \circ (r \otimes \tau) \otimes \Delta_C) \otimes \tau \right) \circ \left( \text{id}_C \otimes T_{C,C} \otimes \text{id}_C \right) \circ \left( \text{id}_C \otimes \Delta_C \otimes \text{id}_C \right) \]

\[ = m^{(2)} \circ \left( r \otimes (m_A \circ (r \otimes \tau) \otimes \Delta_C) \otimes \tau \right) \circ \left( \text{id}_C \otimes T_{C,C} \otimes \text{id}_C \right) \circ \left( \text{id}_C \otimes \Delta_C \otimes \text{id}_C \right) \]

\[ = m^{(2)} \circ \left( r \otimes (m_A \circ (r \otimes \tau) \otimes \Delta_C) \otimes \tau \right) \circ \left( \text{id}_C \otimes T_{C,C} \otimes \text{id}_C \right) \circ \left( \text{id}_C \otimes \Delta_C \otimes \text{id}_C \right) \]

This proves (11.62).

Proof of (11.64): We have \( r \otimes \tau = u_A \circ \epsilon_C \) (according to the definition of convolution). Compared with \( r \otimes \tau = u_A \circ \epsilon_C \), this yields \( m_A \circ (r \otimes \tau) \circ \Delta_C = u_A \circ \epsilon_C \). Thus,

\[ m^{(2)} \circ \left( r \otimes (m_A \circ (r \otimes \tau) \otimes \Delta_C) \otimes \tau \right) \circ \left( \text{id}_C \otimes T_{C,C} \otimes \text{id}_C \right) \]

\[ = m^{(2)} \circ \left( r \otimes (u_A \circ \epsilon_C) \otimes \tau \right) \circ \left( \text{id}_C \otimes T_{C,C} \otimes \text{id}_C \right) \circ \left( \text{id}_C \otimes \Delta_C \otimes \text{id}_C \right). \]

This proves (11.64).
We have
\[ m_A \circ (\text{id}_A \circ \text{kan}^{-1}_{1,A}) \circ (r \otimes \varpi \otimes \epsilon_C) \circ (\Delta_C \otimes \text{id}_C) = \text{kan}^{-1}_{1,A} \circ ((m_A \circ (r \otimes \varpi) \circ \Delta_C) \otimes \epsilon_C). \]

Proof of (11.65): We have
\[
\begin{align*}
m^{(2)} \circ (r \otimes (u_A \circ \epsilon_C)) \circ (\Delta_C \otimes \text{id}_C) &= (m_A \circ (\text{id}_A \circ \text{kan}^{-1}_{1,A}) \circ (r \otimes \epsilon_C \otimes \varpi) \circ (\Delta_C \otimes \text{id}_C) \\
&= m_A \circ (\text{id}_A \circ m_A) \circ (r \otimes \epsilon_C \otimes \varpi) \circ (\Delta_C \otimes \text{id}_C) \\
&= m_A \circ (\text{id}_A \circ \text{kan}^{-1}_{1,A}) \circ (r \otimes \epsilon_C \otimes \varpi) \circ (\Delta_C \otimes \text{id}_C).
\end{align*}
\]

So that
\[
\begin{align*}
m^{(2)} \circ (r \otimes (u_A \circ \epsilon_C)) \circ (\Delta_C \otimes \text{id}_C) &= m_A \circ (\text{id}_A \circ \text{kan}^{-1}_{1,A}) \circ (r \otimes \epsilon_C \otimes \varpi) \circ (\Delta_C \otimes \text{id}_C) \\
&= m_A \circ (\text{id}_A \circ \text{kan}^{-1}_{1,A}) \circ (r \otimes \epsilon_C \otimes \varpi) \circ (\Delta_C \otimes \text{id}_C).
\end{align*}
\]

This proves (11.65).

Proof of (11.66): The equality (11.42) (applied to \( U = A \otimes A \), \( V = A \) and \( \alpha = m_A \)) yields \( \text{kan}^{-1}_{1,A} \circ (m_A \circ \text{id}_A) = m_A \circ \text{kan}^{-1}_{1,A} \otimes A \). But \( \text{id}_A \circ \text{kan}^{-1}_{1,A} = \text{kan}^{-1}_{1,A} \otimes A \) (according to (11.36), applied to \( U = A \) and \( V = A \)), and thus \( m_A \circ \text{id}_A \circ \text{kan}^{-1}_{1,A} \circ (r \otimes \varpi \otimes \epsilon_C) \circ (\Delta_C \otimes \text{id}_C) = m_A \circ \text{id}_A \circ \text{kan}^{-1}_{1,A} \circ (r \otimes \varpi \otimes \epsilon_C) \circ (\Delta_C \otimes \text{id}_C) \).

This proves (11.65).
• We have

\[
\text{(11.66)} \quad \text{kan}^{-1}_{1A} \circ ((m_A \circ (r \otimes \tau) \circ \Delta_C) \otimes \epsilon_C) = \text{kan}^{-1}_{1A} \circ ((u_A \circ \epsilon_C) \otimes \epsilon_C).
\]

297

• We have

\[
\text{(11.67)} \quad \text{kan}^{-1}_{1A} \circ ((u_A \circ \epsilon_C) \otimes \epsilon_C) = m_A \circ ((u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C)).
\]

298

• We have

\[
\text{(11.68)} \quad m_A \circ ((u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C)) = u_A \circ \epsilon_C \circ m_C.
\]

299

\[
m_A \circ \text{kan}^{-1}_{1A \otimes A} = \text{kan}^{-1}_{1A} \circ (m_A \otimes \text{id}_k). \text{ Hence,}
\]

\[
m_A \circ ((m_A \otimes \text{id}_k) \circ (r \otimes \tau \otimes \epsilon_C) \circ (u_A \circ \epsilon_C)) = \text{kan}^{-1}_{1A} \circ ((m_A \otimes \text{id}_k) \circ ((r \otimes \tau) \circ \epsilon_C) \circ (u_A \circ \epsilon_C)).
\]

This proves (11.66).

297 Proof of (11.67): The definition of convolution yields \(r \ast \tau = m_A \circ (r \otimes \tau) \circ \Delta_C\). Compared with \(r \ast \tau = u_A \circ \epsilon_C\), this yields \(m_A \circ (r \otimes \tau) \circ \Delta_C = u_A \circ \epsilon_C\). Hence, \(\text{kan}^{-1}_{1A} \circ ((m_A \circ (r \otimes \tau) \circ \Delta_C) \otimes \epsilon_C) = \text{kan}^{-1}_{1A} \circ ((u_A \circ \epsilon_C) \otimes \epsilon_C)\). This proves (11.67).

298 Proof of (11.68): We have

\[
(u_A \circ \epsilon_C) \otimes \epsilon_C = (u_A \circ \epsilon_C) \otimes (\text{id}_k \circ \epsilon_C) = (u_A \otimes \text{id}_k) \circ (\epsilon_C \otimes \epsilon_C),
\]

so that

\[
\text{kan}^{-1}_{1A} \circ ((u_A \circ \epsilon_C) \otimes \epsilon_C) = \text{kan}^{-1}_{1A} \circ (u_A \otimes \text{id}_k) \circ (\epsilon_C \otimes \epsilon_C).
\]

Compared with

\[
m_A \circ ((u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C)) = (u_A \circ (u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C))
\]

\[
= \text{id}_A \otimes (u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C) = \text{kan}^{-1}_{1A} \circ (u_A \otimes \text{id}_k) \circ (\epsilon_C \otimes \epsilon_C),
\]

this yields \(\text{kan}^{-1}_{1A} \circ ((u_A \circ \epsilon_C) \otimes \epsilon_C) = m_A \circ ((u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C))\). Thus, (11.68) is solved.

299 Proof of (11.69): We know that \(u_A\) is a \(k\)-algebra homomorphism (indeed, this is an easy fact that holds whenever \(A\) is a \(k\)-algebra). Since the two maps \(u_A\) and \(\epsilon_C\) are \(k\)-algebra homomorphisms, their composition \(u_A \circ \epsilon_C\) is a \(k\)-algebra homomorphism. Consequently, \((u_A \circ \epsilon_C) \circ m_C = m_A \circ ((u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C))\) (in fact, this is one of the axioms a \(k\)-algebra homomorphism has to satisfy), so that \(m_A \circ ((u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C)) = (u_A \circ \epsilon_C) \circ m_C = u_A \circ \epsilon_C \circ m_C\). This proves (11.69).
Now,
\[
f \otimes g = m_A \circ (f \otimes g) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \quad \text{(by (11.60))}
\]
\[
= m^{(3)} \circ (r \otimes r \otimes r) \circ (\text{id}_C \otimes \text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \quad \text{(by (11.61))}
\]
\[
= m^{(2)} \circ (r \otimes (m_A \circ (r \otimes r) \otimes \Delta_C) \otimes \text{id}_C) \circ (\Delta_C \otimes \text{id}_C) \circ (\Delta_C \otimes \text{id}_C) \quad \text{(by (11.62))}
\]
\[
= m_A \circ (\text{id}_A \otimes \text{kan}_{1,A}^{-1}) \circ (r \otimes r \otimes \epsilon_C) \circ (\Delta_C \otimes \text{id}_C) \quad \text{(by (11.65))}
\]
\[
= \text{kan}_{1,A}^{-1} \circ ((m_A \circ (r \otimes r) \otimes \Delta_C) \otimes \epsilon_C) \quad \text{(by (11.66))}
\]
\[
= m_A \circ ((u_A \circ \epsilon_C) \otimes \epsilon_C) \quad \text{(by (11.67))}
\]
\[
= m_A \circ ((u_A \circ \epsilon_C) \otimes (u_A \circ \epsilon_C)) \quad \text{(by (11.68))}
\]
\[
= u_A \circ \epsilon_C \circ m_C \quad \text{(by (11.69))}
\]
\[
= u_A \circ \epsilon_C \otimes C \quad \text{(by (11.46))}.
\]

The second equality in (11.44) is thus proven.

Thus, both equalities in (11.44) are proven. Hence, (11.44) is proven. As we have already seen, this yields (11.45). In other words, \( h = g \). But \( h = F \circ m_C \), so that \( F \circ m_C = h = g = m_A \circ T_{A,A} \circ (F \otimes F) \).

Finally, (11.29) (applied to \( U = C \), \( V = C \), \( U' = A \), \( V' = A \), \( x = F \) and \( y = F \) yields \( F \circ m_C = m_A \circ T_{A,A} \circ (F \otimes F) = T_{A,A} \circ (F \otimes F) \).

Thus, \( T_{A,A} \circ (F \otimes F) = \sum_{(a), (b)} h(a_1 \otimes b_1) f(a_2 \otimes b_2) \).

As we know, this completes the solution of Exercise 1.48(a).

[Remark: The second solution of Exercise 1.48(a) has been obtained more or less straightforwardly from the first solution by rewriting it in an element-free fashion. First of all, the maps \( F, G \) and \( h \) introduced in the second solution are precisely the maps \( F, G \) and \( h \) introduced in the first solution, just rewritten in an element-free way. Also, for example, the equalities (11.47), (11.48), (11.49), (11.50), (11.51) and (11.52) have been found by rewriting the six equality signs in the computation (11.27) in an element-free way: For instance, the second equality sign in (11.27) stands for the equality
\[
\sum_{(a), (b)} h(a_1 \otimes b_1) f(a_2 \otimes b_2) = \sum_{(a), (b)} \tau(a_1 b_1) r(a_2) r(b_2)
\]
for all \( a \in C \) and \( b \in C \),

i.e., for the following equality of maps:
\[
\left( \text{the } k\text{-linear map } C \otimes C \to A \text{ sending every } a \otimes b \text{ to } \sum_{(a), (b)} h(a_1 \otimes b_1) f(a_2 \otimes b_2) \right)
\]
\[
= \left( \text{the } k\text{-linear map } C \otimes C \to A \text{ sending every } a \otimes b \text{ to } \sum_{(a), (b)} \tau(a_1 b_1) r(a_2) r(b_2) \right).
\]

But upon rewriting these two maps without referring to elements, this takes the form
\[
m_A \circ (h \otimes f) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C)
\]
\[
= m_A \circ ((F \circ m_C) \otimes (m_A \circ (r \otimes r))) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C)
\]
(because the \( k\)-linear map \( C \otimes C \to A \) sending every \( a \otimes b \) to \( \sum_{(a), (b)} h(a_1 \otimes b_1) f(a_2 \otimes b_2) \) is \( m_A \circ (h \otimes f) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \), and the \( k\)-linear map \( C \otimes C \to A \) sending every \( a \otimes b \) to \( \sum_{(a), (b)} \tau(a_1 b_1) r(a_2) r(b_2) \) is \( m_A \circ ((F \circ m_C) \otimes (m_A \circ (r \otimes r))) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \)); and this is precisely the equality (11.48). The process of rewriting a map in an element-free way is not completely deterministic (e.g., we could also have rewritten the \( k\)-linear map \( C \otimes C \to A \) sending every \( a \otimes b \) to \( \sum_{(a), (b)} \tau(a_1 b_1) r(a_2) r(b_2) \) as \( m^{(3)} \circ ((r \circ m_C) \otimes r) \circ (\text{id}_C \otimes T_{C,C} \otimes \text{id}_C) \circ (\Delta_C \otimes \Delta_C) \), since \( m^{(3)} \) is the \( k\)-linear map \( A \otimes A \otimes A \to A \) sending every \( a \otimes b \otimes c \) to \( abc \)), but all possible results (for a given map) can be
reduced to each other using purely linear-algebraic formulae\(^{300}\) and various forms of the general associativity law (such as \((11.58)\) and \((11.59)\)\(^{301}\). The equalities \((11.47)\), \((11.48)\), \((11.49)\), \((11.50)\), \((11.51)\) and \((11.52)\) are somewhat more complicated to check than the corresponding equality signs in \((11.27)\), because of the fact that one and the same map can be written in different forms whose equivalence needs to be proven. However, this additional complexity is straightforward to surmount: each equality, when rewritten in element-free terms, can be proven by the same arguments as the corresponding equality that uses elements, combined with purely linear-algebraic formulae like \((11.29)\) and \((11.30)\) and various forms of the general associativity law (such as \((11.58)\) and \((11.59)\)). (There is also a way how to do element-free computations without such added difficulty, using string diagrams\(^{302}\).

See the solutions of Exercise 1.8 and of Exercise 1.59 for other examples of how a proof that uses elements can be rewritten in an element-free fashion.]

(b) We have solved Exercise 1.48(a) in an element-free fashion (in the Second solution to Exercise 1.48(a))\(^{303}\). Thus, by “reversing all arrows” in this solution of Exercise 1.48(a), we can obtain a solution to the dual of Exercise 1.48(a). Consequently, the dual of Exercise 1.48(a) holds.

But it is easy to see that the notion of a \(k\)-coalgebra anti-homomorphism is dual to the notion of a \(k\)-algebra anti-homomorphism (i.e., is defined from the latter notion by “reversing all arrows”), and the notion of convolution is dual to itself. Hence, the dual of Exercise 1.48(a) is the following exercise:

Exercise \(A\): If \(C\) is a \(k\)-bialgebra, if \(A\) is a \(k\)-coalgebra, and if \(r: A \to C\) is a \(\ast\)-invertible \(k\)-coalgebra homomorphism, then prove that the \(\ast\)-inverse \(\ast^{-1}\) of \(r\) is a \(k\)-coalgebra anti-homomorphism.

Exercise 1.48(b) immediately follows from this Exercise \(A\) (applied to \(A\) and \(C\) instead of \(C\) and \(A\)). Exercise 1.48(b) is thus solved.

(c) Alternative proof of Proposition 1.32 using Exercise 1.48(a). Let \(A\) be a Hopf algebra. The antipode \(S\) of this Hopf algebra \(A\) is defined as the \(\ast\)-inverse of the identity map \(id_A\); thus, \(S = id_A^{\ast(-1)}\). Therefore, Exercise 1.48(a) (applied to \(C = A\) and \(r = id_A\)) yields that the \(\ast\)-inverse \(id_A^{\ast(-1)}\) of \(id_A\) is a \(k\)-algebra anti-homomorphism. In other words, \(S\) is a \(k\)-algebra anti-homomorphism (since \(S\) is the \(\ast\)-inverse \(id_A^{\ast(-1)}\) of \(id_A\)). In other words, \(S\) is a \(k\)-algebra anti-endomorphism of \(A\). This proves Proposition 1.32.

Alternative solution of Exercise 1.47 using Exercise 1.48(b). This is analogous to the proof of Proposition 1.32 using Exercise 1.48(a) just shown, but instead of Exercise 1.48(a) we now need to use Exercise 1.48(b).

(d) Alternative proof of Corollary 1.34 using Proposition 1.46. Let \(A\) be a commutative Hopf algebra. Then, the \(k\)-linear map \(id_A : A \to A\) is \(\ast\)-invertible (since \(A\) is a Hopf algebra), and its \(\ast\)-inverse \(id_A^{\ast(-1)}\) is the antipode \(S\) of \(A\). That is, \(id_A^{\ast(-1)} = S\). Applying Exercise 1.48(a) to \(C = A\) and \(r = id_A\), we now conclude that \(id_A^{\ast(-1)}\) is a \(k\)-algebra anti-homomorphism \(A \to A\) (since \(id_A\) is a \(k\)-algebra homomorphism \(A \to A\)). Since a \(k\)-algebra anti-homomorphism \(A \to A\) is the same thing as a \(k\)-algebra homomorphism \(A \to A\) (because \(A\) is commutative), this yields that \(id_A^{\ast(-1)}\) is a \(k\)-algebra homomorphism \(A \to A\). In other words, \(S\) is a \(k\)-algebra homomorphism \(A \to A\) (since \(id_A^{\ast(-1)} = S\)). Now, Proposition 1.46(a) (applied to \(H = A\) and \(\alpha = S\)) yields \(S \circ S = S^\ast(-1)\). But \(S^\ast(-1) = id_A\) (since \(S = id_A^{\ast(-1)}\)). Hence, \(S^2 = S \circ S = S^\ast(-1) = id_A\). This proves Corollary 1.34. Exercise 1.48(d) is solved.

\(^{300}\)By “purely linear-algebraic formulae”, I mean formulae such as \((11.29)\) and \((11.30)\) and the identity \((\beta \circ \alpha) \circ (\beta' \circ \alpha') = (\beta \circ \beta') \circ (\alpha \circ \alpha')\) for tensor products of compositions of maps. This kind of formulae hold in any tensor category (at least if we follow the abuse of notation that allows us to treat the tensor product as associative).

\(^{301}\)The reason why we need to use the general associativity law is that, in the first solution, we used unparenthesized products of more than one element of \(A\) (for example, the expression \(\sum_{(a,b)} r(a_1 b_1) r(a_2) r(b_2)\) contains the unparenthesized product \(r(a_1 b_1) r(a_2) r(b_2)\) of three factors). If we would parenthesize all such products in such a way that no more than two factors are ever multiplied at the same time (e.g., we could replace \(r(a_1 b_1) r(a_2) r(b_2)\) by \(r(a_1 b_1) (r(a_2) r(b_2))\)), and if we would explicitly use the (non-general) associativity law \((xy)z = x(yz)\) to switch between these parenthesizations, then we would not have to use general associativity any more when rewriting the proof in an element-free fashion (but, of course, the proof would be longer).

\(^{302}\)See http://ncatlab.org/nlab/show/string+diagram and the references therein.

\(^{303}\)To be fully honest, this second solution was not entirely element-free, since we proved the auxiliary equality \((11.63)\) using elements. But for the purposes of what we are going to do (reversing arrows), this is not problematic, since this auxiliary equality \((11.63)\) can be easily shown to hold with all arrows reversed (the proof is more or less the same as for the original \((11.63)\)).
(c) Let $A$ be a cocommutative Hopf algebra. Then, the $k$-linear map $\id_A : A \to A$ is $\ast$-invertible (since $A$ is a Hopf algebra), and its $\ast$-inverse $\id_A^{(-1)}$ is the antipode $S$ of $A$. That is, $\id_A^{(-1)} = S$. Applying Exercise 1.48(b) to $C = A$ and $r = \id_A$, we now conclude that $\id_A^{(-1)}$ is a $k$-coalgebra anti-homomorphism $A \to A$ (since $\id_A$ is a $k$-coalgebra homomorphism $A \to A$). Since a $k$-coalgebra anti-homomorphism $A \to A$ is the same thing as a $k$-coalgebra homomorphism $A \to A$ (because $A$ is cocommutative), this yields that $\id_A^{(-1)}$ is a $k$-coalgebra homomorphism $A \to A$. In other words, $S$ is a $k$-coalgebra homomorphism $A \to A$ (since $\id_A^{(-1)} = S$). Now, Proposition 1.46(b) (applied to $H = A$, $C = A$ and $\gamma = S$) yields $S \circ S = S^{\ast(-1)}$. But $S^{\ast(-1)} = \id_A$ (since $S = \id_A^{(-1)}$). Hence, $S^2 = S \circ S = S^{\ast(-1)} = \id_A$. This solves Exercise 1.48(e).

11.14. Solution to Exercise 1.49. Solution to Exercise 1.49. (a) Consider a map $P$ satisfying the given assumption. Consider also the antipode $S$ of $A$. We know from Exercise 1.47 that $S$ is a coalgebra anti-endomorphism; thus, $(S \otimes S) \Delta = T \circ \Delta \circ S$ and $\epsilon \circ S = \epsilon$, with $T$ being the twist map $A \otimes A \to A \otimes A$, $c \otimes d \mapsto d \otimes c$. We also know from Proposition 1.32 that $S$ is an algebra anti-endomorphism; this can be rewritten as $m \circ (S \otimes S) = S \circ m \circ T$ and $S \circ u = u$.

Now, we know that $m \circ (P \otimes \id) \circ T \circ \Delta = u \circ \epsilon$ (indeed, this is just an element-free rewriting of the assumption that every $a \in A$ satisfies $\sum_{(a)} P(a_2) \cdot a_1 = u(\epsilon(a))$). Now,

$$
(P \circ S) \ast = m \circ ((P \circ S) \otimes S) \Delta = m \circ (P \otimes \id) \circ (S \otimes S) \Delta = (P \otimes \id) \circ (S \otimes S) \Delta = u \circ \epsilon S = u \circ \epsilon,
$$

so that $P \circ S$ is a left $\ast$-inverse to $S$. Since $S$ is $\ast$-invertible with $\ast$-inverse $\id$, this yields that $P \circ S = \id$. Furthermore,

$$
S \ast (S \circ P) = m \circ (S \otimes (S \circ P)) \Delta = m \circ (S \otimes S) \circ (\id \otimes P) \circ \Delta
$$

$$
= (S \otimes S) \circ (\id \otimes P) \circ \Delta = (S \otimes S) \circ \Delta = S \circ m \circ (P \otimes \id) \circ T \circ \Delta = S \circ u \circ \epsilon = u \circ \epsilon,
$$

whence $S \circ P$ is a right $\ast$-inverse to $S$. Since $S$ is $\ast$-invertible with $\ast$-inverse $\id$, this yields that $S \circ P = \id$. Combined with $P \circ S = \id$, this yields that $S$ is invertible and its inverse is $P$.

(b) Similar to the solution for (a), the details being left to the reader.

(c) Let $A$ be a graded connected Hopf algebra. Just as a left $\ast$-inverse $S$ to $\id_A$ has been constructed in the proof of Proposition 1.36, we could construct a $k$-linear map $P : A \to A$ such that every $a \in A$ satisfies

$$
\sum_{(a)} P(a_2) \cdot a_1 = u(\epsilon(a)).
$$

By part (a), this yields that the antipode of $A$ is invertible.

11.15. Solution to Exercise 1.51. Solution to Exercise 1.51. We will be maximally explicit in this solution; in particular, we will not regard inclusions as identities even in cases where we would usually do that. We will only use the notations $\Delta$ and $\epsilon$ to denote the maps $\Delta_C$ and $\epsilon_C$ (not the maps $\Delta_D$ and $\epsilon_D$, which we will introduce later).

We know that $D$ is a direct summand of $C$ as a $k$-module. In other words, there exists a $k$-submodule $E$ of $C$ such that $C = D \oplus E$. Fix such an $E$.

Let $i : D \to C$ be the canonical inclusion map. Let $p : D \oplus E \to D$ be the canonical projection from the direct sum $D \oplus E$ onto its summand $D$. Notice that $p$ is a $k$-linear map from $D \oplus E = C$ to $D$, and satisfies $p \circ i = \id_D$. Hence, $i \circ p \circ i = i$. 

= \id_D
In the statement of the exercise, we have assumed that \( \Delta(D) \subset C \otimes D \). Since we don’t want to abuse notation, we have to rewrite this as \( \Delta(D) \subset (\text{id}_C \otimes i)(C \otimes D) \) (because the \( k \)-submodule of \( C \otimes C \) spanned by tensors of the form \( c \otimes d \) with \( c \in C \) and \( d \in D \) is precisely \( (\text{id}_C \otimes i)(C \otimes D) \)). Similarly, \( \Delta(D) \subset (\text{id}_D \otimes C)(D \otimes C) \).

Now,

\[
(11.70) \quad (i \otimes i) ((p \otimes p)(x)) = x \quad \text{for every } x \in \Delta(D).
\]

Proof of (11.70): Let \( x \in \Delta(D) \). Then, \( x \in \Delta(D) \subset (\text{id}_C \otimes i)(C \otimes D) \). Hence, there exists some \( y \in C \otimes D \) such that \( x = (\text{id}_C \otimes i)(y) \). Consider this \( y \). We have

\[
(i \otimes i) ((p \otimes p)(x)) = (i \otimes i) ((p \otimes p)((\text{id}_C \otimes i)(y))) = ((i \otimes i) \circ (p \otimes p) \circ (\text{id}_C \otimes i))(y)\]

\[
= (i \circ (p \circ \text{id}_C) \otimes (p \circ \text{id}_C))(y) = (((i \circ p) \circ \text{id}_C) \otimes (i \circ \text{id}_C))(y)\]

\[
= (((i \circ p) \circ \text{id}_C) \otimes (i \circ \text{id}_C))(y) = ((i \circ p) \otimes \text{id}_C)((i \circ \text{id}_C))(y).
\]

On the other hand, \( x \in \Delta(D) \subset (i \otimes \text{id}_C)(D \otimes C) \). Thus, there exists some \( z \in D \otimes C \) such that \( x = (i \otimes \text{id}_C)(z) \). Consider this \( z \). We have

\[
(i \otimes i) ((p \otimes p)(x)) = ((i \circ p) \otimes \text{id}_C)(z) = (i \circ p)(\text{id}_C \otimes i)(x) = (i \circ p)(\text{id}_C \otimes i)(y) = (i \otimes p) \circ \text{id}_C((i \circ \text{id}_C)(y)).
\]

This proves (11.70).

Next, let us define a \( k \)-linear map \( \Delta_D : D \to D \otimes D \) by

\[
\Delta_D = (p \otimes p) \circ \Delta \circ i.
\]

Let us also define a \( k \)-linear map \( \epsilon_D : D \to k \) by

\[
\epsilon_D = \epsilon \circ i.
\]

We will show that \((D, \Delta_D, \epsilon_D)\) is a \( k \)-coalgebra.\(^{304}\) But first, let us see that

\[
(11.71) \quad (i \otimes i) \circ \Delta_D = \Delta \circ i.
\]

Proof of (11.71): Let \( d \in D \). Then, \( i(d) = d \) (since \( i \) is just an inclusion map), so that \( \Delta(D) \subset \Delta(i(d)) \) \( (\text{by } (11.70), \text{applied to } x = \Delta(i(d))) \). Now,

\[
(i \otimes i) \circ \Delta_D(d) = ((i \otimes i) \circ (p \otimes p) \circ \Delta \circ i)(d) = (i \otimes i)((p \otimes p)(\Delta(i(d))) = \Delta(i(d)) = (\Delta \circ i)(d).
\]

\(^{304}\)Recall that we are not going to abbreviate \( \Delta_D \) and \( \epsilon_D \) by \( \Delta \) and \( \epsilon \); thus, \( \Delta \) and \( \epsilon \) still mean the maps \( \Delta_C \) and \( \epsilon_C \).
Now, forget that we fixed $d$. We thus have shown that every $d \in D$ satisfies $(i \otimes i) \circ \Delta_D (d) = (\Delta \circ i) (d)$. In other words, $(i \otimes i) \circ \Delta_D = \Delta \circ i$. This proves (11.71).

Now, let us check that $(D, \Delta_D, \epsilon_D)$ is a $k$-coalgebra. In order to do so, we must check that the diagrams

$$\text{(11.72)}$$

are commutative (where the canonical isomorphisms $k \otimes D \to D$ and $D \otimes k \to D$ are used in (11.73)). Let us start with the first diagram.

Since the diagram (1.3) for $C$ is commutative (as $C$ is a $k$-coalgebra), we have $(\Delta \otimes \text{id}_C) \circ \Delta = (\text{id}_C \otimes \Delta) \circ \Delta$. But

$$(i \otimes i) \circ (\Delta_D \otimes \text{id}_D) \circ \Delta_D = \left( (i \otimes i) \circ (\Delta_D \otimes \text{id}_D) \circ \Delta_D = (\Delta \circ i) \otimes (i \circ \text{id}_D) \circ \Delta_D \right) \circ \Delta_D = (\Delta \otimes \text{id}_C) \circ (i \circ i) \circ \Delta_D = (\Delta \otimes \text{id}_C) \circ \Delta \circ i.$$

Using this and the analogously provable identity $(i \otimes i \otimes i) \circ (\Delta_D \otimes \Delta_D) \circ \Delta_D = (\text{id}_C \otimes \Delta) \circ \Delta \circ i$, we obtain

$$\left( (i \otimes i \otimes i) \circ (\Delta_D \otimes \text{id}_D) \circ \Delta_D = (\Delta \otimes \text{id}_C) \circ (i \circ i) \circ \Delta_D = (\Delta \otimes \text{id}_C) \circ \Delta \circ i = (i \otimes i \otimes i) \circ (\text{id}_D \otimes \Delta_D) \circ \Delta_D. \right)$$

We can cancel the left $i \otimes i \otimes i$ factor from this equation\textsuperscript{305}, and thus obtain $(\Delta_D \otimes \text{id}_D) \circ \Delta_D = (\text{id}_D \otimes \Delta_D) \circ \Delta_D$. In other words, the diagram (11.72) is commutative.

Let us now check that the diagram (11.73) is commutative. We will only prove this for its left square, leaving the (completely analogous) right square to the reader. For every $k$-module $V$, we let $\text{can}_V$ denote the canonical $k$-module isomorphism $V \otimes k \to V$. The upper left horizontal arrow on diagram (11.73) is precisely $\text{can}_D$. Notice that $\text{can}_C \circ (\text{id}_C \otimes \epsilon) \circ \Delta = \text{id}_C$ due to the commutativity of the diagram (1.4) (which, of course, commutes since $C$ is a coalgebra).

Now, any two $k$-modules $V$ and $W$ and any $k$-linear map $f : V \to W$ satisfy

$$f \circ \text{can}_V = \text{can}_W \circ (f \otimes \text{id}_k)$$

\textsuperscript{305} because $i \otimes i \otimes i$ is left-invertible:

$$(p \otimes p \otimes p) \circ (i \otimes i \otimes i) = (p \circ i) \otimes (p \otimes i) \otimes (p \circ i) = \text{id}_D \otimes \text{id}_D \otimes \text{id}_D = \text{id}_{D \otimes D \otimes D}.$$
(this is just trivial linear algebra). This (applied to \( V = D, W = C \) and \( f = i \)) yields \( i \circ \text{can}_D = \text{can}_C \circ (i \otimes \text{id}_k) \). Hence,

\[
i \circ \text{can}_D \circ (\text{id}_D \otimes \epsilon_D) \circ \Delta_D = \text{can}_C \circ (i \otimes \text{id}_k) \circ (\text{id}_D \otimes \epsilon_D) \circ \Delta_D = \text{can}_C \circ (i \circ \text{id}_D) \otimes (\text{id}_k \circ \epsilon_D) \circ \Delta_D
\]

\[
= \text{can}_C \circ ((\text{id}_C \circ i) \otimes (\epsilon \circ i)) \circ \Delta_D = \text{can}_C \circ (\text{id}_C \otimes \epsilon) \circ ((i \otimes i) \circ \Delta_D)
\]

\[
= \text{can}_C \circ (\text{id}_C \otimes \epsilon) \circ \Delta \circ i = \text{id}_C \circ i = i.
\]

We can cancel the \( i \) factors from this equation (because \( i \) is left-invertible: \( p \circ i = \text{id}_D \)), and thus are left with \( \text{can}_D \circ (\text{id}_D \otimes \epsilon_D) \circ \Delta_D = \text{id}_D \). This means precisely that the left square of (11.73) is commutative. As we said, this completes the verification of the fact that \((D, \Delta_D, \epsilon_D)\) must be a \( k \)-coalgebra. We will denote this \( k \)-coalgebra simply by \( D \).

Now, the diagrams

\[
\begin{align*}
\begin{array}{ccc}
D & \xrightarrow{i} & C \\
\Delta_D & & \Delta \\
D \otimes D & \xrightarrow{i \otimes i} & C \otimes C
\end{array}
\end{align*}
\]

are commutative\(^{306}\). Hence, \( i \) is a \( k \)-coalgebra homomorphism (by the definition of a \( k \)-coalgebra homomorphism). In other words, the canonical inclusion map \( D \to C \) is a \( k \)-coalgebra homomorphism (since \( i \) is the canonical inclusion map \( D \to C \)).

So we know that \( D \) is a \( k \)-coalgebra such that \( D \subset C \) and such that the canonical inclusion map \( D \to C \) is a \( k \)-coalgebra homomorphism. In other words, \( D \) is a \( k \)-subcoalgebra of \( C \) (by the definition of a \( k \)-subcoalgebra). This solves the exercise.

11.16. Solution to Exercise 1.52. Solution to Exercise 1.52. We will identify the tensor products \( K \otimes K, C \otimes K \) and \( K \otimes C \) with the corresponding \( k \)-submodules of the tensor product \( C \otimes C \). (We can afford to do this since \( k \) is a field.)

We will only use the notations \( \Delta \) and \( \epsilon \) to denote the maps \( \Delta_C \) and \( \epsilon_C \).

Notice that \( C \) is a free \( k \)-module. Hence, tensoring with \( C \) is an exact functor, thus a left-exact functor.

By the recursive definition of \( \Delta^{(1)} \), we have

\[
\Delta^{(1)} = \left( \text{id}_C \otimes \Delta^{(1-1)} \right) \circ \Delta = \left( \text{id}_C \otimes \text{id}_C \right) \circ \Delta = \Delta.
\]

By the recursive definition of \( \Delta^{(2)} \), we have

\[
\Delta^{(2)} = \left( \text{id}_C \otimes \Delta^{(1)} \right) \circ \Delta = \left( \text{id}_C \otimes \Delta \right) \circ \Delta = \left( \Delta \otimes \text{id}_C \right) \circ \Delta
\]

(by the axioms of a coalgebra).

(a) Applying Exercise 1.40(b) to \( k = 3 \), we obtain

\[
\Delta^{(3)} = \left( \Delta^{(2)} \otimes \text{id}_C \right) \circ \Delta,
\]

so that

\[
(11.74) \quad \left( \Delta^{(2)} \otimes \text{id}_C \right) \circ \Delta = \Delta^{(3)} = \left( \text{id}_C \otimes \Delta^{(2)} \right) \circ \Delta
\]

(by the recursive definition of \( \Delta^{(3)} \)).
Let \( \bar{f} = (\text{id}_C \otimes f \otimes \text{id}_C) \circ \Delta^{(2)} : C \to C \otimes U \otimes C \). Then, \( \ker \bar{f} = \ker ((\text{id}_C \otimes f \otimes \text{id}_C) \circ \Delta^{(2)}) = K \). But

\[
\bar{f} \otimes \text{id}_C = (\text{id}_C \otimes f \otimes \text{id}_C) \circ \Delta^{(2)}
\]

\[
= \left( (\text{id}_C \otimes f \otimes \text{id}_C) \circ \Delta^{(2)} \right) \otimes \text{id}_C = \left( (\text{id}_C \otimes f \otimes \text{id}_C) \otimes \text{id}_C \right) \circ \left( \Delta^{(2)} \otimes \text{id}_C \right)
\]

\[
= (\text{id}_C \otimes f \otimes \text{id}_{C \otimes C}) \circ \left( \Delta^{(2)} \otimes \text{id}_C \right),
\]

so that

\[
\left( \bar{f} \otimes \text{id}_C \right) \circ \Delta
\]

\[
= (\text{id}_C \otimes f \otimes \text{id}_{C \otimes C}) \circ \left( \Delta^{(2)} \otimes \text{id}_C \right)
\]

\[
= (\text{id}_C \otimes f \otimes \text{id}_{C \otimes C}) \circ \left( \Delta^{(2)} \otimes \text{id}_C \right) \circ \Delta
\]

\[
= \left( \text{id}_C \otimes \left( f \otimes \text{id}_{C \otimes C} \right) \circ \left( \Delta^{(2)} \otimes \text{id}_C \right) \right) \circ \Delta
\]

\[
= \left( \text{id}_C \otimes \left( f \otimes \text{id}_{C \otimes C} \circ \left( \text{id}_C \otimes \Delta \right) \right) \circ \Delta \right)
\]

\[
= \left( \text{id}_C \otimes \left( f \otimes \text{id}_C \circ \left( \text{id}_C \otimes \Delta \right) \right) \circ \Delta \right)
\]

\[
= \left( f \otimes \text{id}_C \circ \left( \text{id}_C \otimes \Delta \right) \right) \circ \Delta
\]

\[
= \left( f \otimes \text{id}_C \circ \left( \text{id}_C \otimes \Delta \right) \right) \circ \Delta^{(2)}
\]

\[
= \left( f \otimes \text{id}_C \right) \circ \Delta^{(2)}
\]

\[
= \left( f \otimes \text{id}_C \right) \circ \bar{f}.
\]

Hence,

\[
\ker \left( (\bar{f} \otimes \text{id}_C) \circ \Delta \right) = \ker \left( (\text{id}_C \otimes \text{id}_U \otimes \Delta) \circ \bar{f} \right) \supset \ker \bar{f} = K,
\]

so that

\[
K \subset \ker \left( (\bar{f} \otimes \text{id}_C) \circ \Delta \right) = \Delta^{-1} \left( \ker \left( \bar{f} \otimes \text{id}_C \right) \right),
\]

so that

\[
\Delta (K) \subset \ker \left( \bar{f} \otimes \text{id}_C \right) = \left( \ker \bar{f} \right) \otimes C
\]

(since tensoring with \( C \) is a left-exact functor)

\[
= K \otimes C.
\]

Similarly, \( \Delta (K) \subset C \otimes K \). But \( K \) is a direct summand of \( C \) as a \( k \)-module (since \( k \) is a field). Hence, Exercise 1.51 (applied to \( D = K \)) yields that there is a canonically defined \( k \)-coalgebra structure on \( K \) which makes \( K \) a subcoalgebra of \( C \). In other words, \( K \) is a \( k \)-subcoalgebra of \( C \). This solves Exercise 1.52(a).
(b) Let \( E \) be a \( k \)-subcoalgebra of \( C \) which is a subset of \( \ker f \). We must show that \( E \) is a subset of \( K \). We have \( E \subset \ker f \), so that \( f(E) = 0 \). Since \( E \) is a \( k \)-subcoalgebra of \( C \), we have \( \Delta(E) \subset E \otimes E \), and

\[
\Delta^{(2)}(E) = (\text{id}_C \otimes \Delta) \circ \Delta(E) = (\text{id}_C \otimes \Delta)
\]

\[
\subset (\text{id}_C \otimes \Delta)(E \otimes E) = \text{id}_C(E) \otimes \Delta(E) \subset E \otimes E \otimes E
\]

and thus

\[
\left( (\text{id}_C \otimes f \otimes \text{id}_C) \circ \Delta^{(2)} \right)(E) = (\text{id}_C \otimes f \otimes \text{id}_C)
\]

\[
\subset (\text{id}_C \otimes f \otimes \text{id}_C)(E \otimes E \otimes E)
\]

\[
= \text{id}_C(E) \otimes f(E) \otimes \text{id}_C(E) = 0.
\]

Hence, \( E \subset \ker ((\text{id}_C \otimes f \otimes \text{id}_C) \circ \Delta^{(2)}) = K \), which means that \( E \) is a subset of \( K \). This solves Exercise 1.52(b).

Remark: Exercise 1.52(a) would not hold if we allowed \( k \) to be an arbitrary commutative ring rather than a field. (For a stupid counterexample, try \( k = \mathbb{Z} \), \( C = k \) and \( U = \mathbb{Z}/2 \), with \( f \) being the canonical projection.) It might be an interesting question to figure out how much freedom we can allow without breaking correctness. Here is one case which definitely works: If \( k \) is a principal ideal domain and \( C \) and \( U \) are finite free \( k \)-modules, then Exercise 1.52 is correct. (In fact, our above solution works in this case, after one notices that \( (\text{id}_C \otimes f \otimes \text{id}_C) \circ \Delta^{(2)} \) is a homomorphism of finite free \( k \)-modules, and the kernel of every homomorphism of finite free \( k \)-modules over a principal ideal domain is a direct summand of its domain.)

11.17. Solution to Exercise 1.53. Solution to Exercise 1.53.

(a) Here is Takeuchi’s argument: We know that the map \( h|_{C_0} \in \text{Hom}(C_0, A) \) is \( \ast \)-invertible; let \( \bar{g} \) be its \( \ast \)-inverse. Extend \( \bar{g} \) to a \( k \)-linear map \( g : C \rightarrow A \) by defining it as 0 on every \( C_n \) for \( n > 0 \). It is then easy to see that \( (h \ast g)|_{C_0} = (g \ast h)|_{C_0} = (ue)|_{C_0} \). This allows us to assume WLOG that \( h|_{C_0} = (ue)|_{C_0} \) (because once we know that \( h \ast g \) and \( g \ast h \) are \( \ast \)-invertible, it follows that so is \( h \)). Assuming this, we conclude that \( h - ue \) annihilates \( C_0 \). Define \( f \) as \( h - ue \). Now, we can proceed as in the proof of Proposition 1.44 to show that \( \sum_{k \geq 0} (-1)^k f^{*k} \) is a well-defined linear map \( C \rightarrow A \) and a two-sided \( \ast \)-inverse for \( h \). Thus, \( h \) is \( \ast \)-invertible, and part (a) of the exercise is proven.

An alternative proof proceeds by mimicking the proof of Proposition 1.36, again by first assuming WLOG that \( h|_{C_0} = (ue)|_{C_0} \).

(b) Apply part (a) to \( C = A \) and the map \( \text{id}_A : A \rightarrow A \).

(c) Applying part (b), we see that \( A \) is a Hopf algebra (since \( A_0 = k \) is a Hopf algebra) in the setting of Proposition 1.36. Okay, one still has to check that the antipode is unique, but this is pretty much trivial.

11.18. Solution to Exercise 1.54. Solution to Exercise 1.54. (a) Let \( I_n = I \cap A_n \) for every \( n \in \mathbb{N} \). Then,

\[
I = \bigoplus_{n \geq 0} (I \cap A_n) = \bigoplus_{n \geq 0} I_n.
\]

Since \( I \) is a two-sided coideal, we have \( \epsilon(I) = 0 \).

We now will prove that

\[
(11.75) \quad \text{every } n \in \mathbb{N} \text{ satisfies } I_n = 0.
\]

Proof of (11.75): Let us prove (11.75) by strong induction over \( n \):
Let \( N \in \mathbb{N} \). Assume that (11.75) holds for every \( n \in \mathbb{N} \) satisfying \( n < N \). We must then prove that (11.75) holds for \( n = N \).

We know that (11.75) holds for every \( n \in \mathbb{N} \) satisfying \( n < N \). In other words,

(11.76) for every \( n \in \mathbb{N} \) satisfying \( n < N \), we have \( I_n = 0 \).

We have \( I_N = I \cap A_N \subset I \) and thus \( \epsilon \left( \frac{I_N}{I} \right) \subset \epsilon (I) = 0 \), hence \( \epsilon (I_N) = 0 \).

Let \( \pi_N : A \otimes A \to (A \otimes A)_N \) be the projection from the graded \( k \)-module \( A \otimes A \) to its \( N \)-th graded component \( (A \otimes A)_N \). Then,

(11.77) \( \pi_N (t) = t \) for every \( t \in (A \otimes A)_N \),

and

(11.78) \( \pi_N (t) = 0 \) for every \( \ell \in \mathbb{N} \setminus \{N\} \) and every \( t \in (A \otimes A)_\ell \).

As a consequence,

(11.79) every \( (n,m) \in \mathbb{N}^2 \) such that \( n + m \neq N \) satisfy \( \pi_N (A_n \otimes A_m) = 0 \).

Let \( i \in I_N \) be arbitrary. Then, \( i \in A_N \), so that \( \Delta (i) \in \Delta (A_N) \subset (A \otimes A)_N \) (since \( \Delta \) is a graded map). Thus, \( \pi_N (\Delta (i)) = \Delta (i) \) (by (11.77), applied to \( t = \Delta (i) \)). On the other hand, since \( i \in I_N \subset I \), we have

\[
\Delta (i) \in \Delta (I) \subset \bigoplus_{n \geq 0} I_n \otimes \bigoplus_{m \geq 0} A_m + \bigoplus_{m \geq 0} A_m \otimes \bigoplus_{n \geq 0} I_n = \bigoplus_{(m,n) \in \mathbb{N}^2} I_n \otimes A_m + \bigoplus_{(m,n) \in \mathbb{N}^2} A_m \otimes I_n.
\] (since \( I \) is a two-sided coideal)
Thus,

\[
\pi_N (\Delta (i))
\in \pi_N \left( \sum_{(n,m) \in \mathbb{N}^2} I_n \otimes A_m + \sum_{(m,n) \in \mathbb{N}^2} A_m \otimes I_n \right)
\]

\[
= \sum_{(m,n) \in \mathbb{N}^2} \pi_N (I_n \otimes A_m) + \sum_{(m,n) \in \mathbb{N}^2} \pi_N (A_m \otimes I_n)
\]

\[
= \sum_{(m,n) \in \mathbb{N}^2; \ n+m=N} \pi_N (I_n \otimes A_m) + \sum_{(m,n) \in \mathbb{N}^2; \ n+m \neq N} \pi_N (I_n \otimes A_m)
\]

\[
= \sum_{(m,n) \in \mathbb{N}^2; \ n+m=N} \pi_N (I_n \otimes A_m) + \sum_{(m,n) \in \mathbb{N}^2; \ m+n=N} \pi_N (A_m \otimes I_n)
\]

\[
= \sum_{(m,n) \in \mathbb{N}^2; \ n+m=N} \sum_{n \geq N} \pi_N \left( I_n \otimes A_m \right) + \sum_{(m,n) \in \mathbb{N}^2; \ n+m=N; \ n \geq N} \pi_N \left( I_n \otimes A_m \right)
\]

\[
= \sum_{(m,n) \in \mathbb{N}^2; \ n+m=N; \ n < N} \pi_N \left( I_n \otimes A_m \right) + \sum_{(m,n) \in \mathbb{N}^2; \ n+m=N; \ n \geq N} \pi_N \left( I_n \otimes A_m \right)
\]

\[
= \sum_{(m,n) \in \mathbb{N}^2; \ n+m=N} \pi_N \left( I_n \otimes A_m \right) + \sum_{(m,n) \in \mathbb{N}^2; \ n+m=N; \ n \geq N} \pi_N \left( I_n \otimes A_m \right)
\]

\[
= I_N \otimes A_0 + A_0 \otimes I_N
\]

\[
= I_N \otimes k_1 A + k_1 A \otimes I_N.
\]
Since \( \pi_N (\Delta (i)) = \Delta (i) \), this rewrites as \( \Delta (i) \in I_N \otimes k 1_A + k 1_A \otimes I_N \). In other words, there exists some \( y \in I_N \otimes k 1_A \) and some \( z \in k 1_A \otimes I_N \) such that \( \Delta (i) = y + z \). Consider these \( y \) and \( z \).

Since \( y \in I_N \otimes k 1_A = I_N \otimes 1_A \), there exists a \( j \in I_N \) such that \( y = j \otimes 1_A \). Consider this \( j \).

Since \( z \in k 1_A \otimes I_N = 1_A \otimes I_N \), there exists a \( k \in I_N \) such that \( z = 1_A \otimes k \). Consider this \( k \).

Now,

\[
\Delta (i) = \frac{y}{=j \otimes 1_A} + \frac{z}{=1_A \otimes k} = j \otimes 1_A + 1_A \otimes k.
\]

Applying the map \( \epsilon \otimes \text{id} : A \otimes A \to A \) to both sides of this equation, we obtain

\[
(\epsilon \otimes \text{id}) (\Delta (i)) = (\epsilon \otimes \text{id}) (j \otimes 1_A + 1_A \otimes k) = \epsilon (j) \quad \text{id} (1_A) + \epsilon (1_A) \text{id} (k) = 0 \text{id} (1_A) + k = k.
\]

Since \((\epsilon \otimes \text{id}) (\Delta (i)) = i \) (by the axioms of a coalgebra), this rewrites as \( i = k \).

Similarly, applying \( \text{id} \otimes \epsilon \) to both sides of the equation \( \Delta (i) = j \otimes 1_A + 1_A \otimes k \) and simplifying, we obtain \( i = j \).

Now,

\[
\Delta (i) = \frac{j}{=i} \otimes 1_A + \frac{1_A \otimes k}{=i} = i \otimes 1_A + 1_A \otimes i.
\]

Hence, \( i \) is primitive. In other words, \( i \in \mathfrak{p} \). Combined with \( i \in I_N = I \cap A_N \subseteq I \), this yields \( i \in I \cap \mathfrak{p} = 0 \), so that \( i = 0 \).

Now, forget that we fixed \( i \). We thus have shown that every \( i \in I_N \) satisfies \( i = 0 \). In other words, \( I_N = 0 \).

In other words, \((11.75)\) holds for \( n = N \). This completes the induction proof of \((11.75)\).

(b) By Exercise 1.20(a), we know that \( \ker f \) is a two-sided coideal of \( A \). It further satisfies \( \ker f = \bigoplus_{n \geq 0} ((\ker f) \cap A_n) \) (since \( f \) is graded). If \( f |_{\mathfrak{p}} \) is injective, then \( (\ker f) \cap \mathfrak{p} = 0 \), and thus part (a) of the current exercise (applied to \( I = \ker f \)) yields that \( \ker f = 0 \), so that \( f \) is injective. This solves part (b) of the exercise.

(c) The solution of part (c) proceeds precisely as the solution of part (b), except that instead of using Exercise 1.20(a) we now must use Exercise 1.20(b).

11.19. Solution to Exercise 1.58. Solution to Exercise 1.58.

(a) Let \( A \) be any associative \( k \)-algebra. Define a \( k \)-bilinear map \([\cdot, \cdot] : A \times A \to A\) by setting

\[
[a, b] = ab - ba \quad \text{for all } a, b \in A.
\]

We must prove that this \( k \)-bilinear map \([\cdot, \cdot]\) makes \( A \) into a Lie algebra. In order to do so, it is clearly enough to prove that

\[
(11.80) \quad [x, x] = 0 \quad \text{for all } x \in A,
\]

and that

\[
(11.81) \quad [x, [y, z]] + [z, [x, y]] + [y, [z, x]] = 0 \quad \text{for all } x, y, z \in A.
\]

Proof of \((11.80)\): Every \( x \in A \) satisfies

\[
[x, x] = xx - x x \quad \text{(by the definition of } [x, x])
\]

\[
= 0.
\]

This proves \((11.80)\).
Proof of (11.81): Every $x, y, z \in A$ satisfy

\[
\begin{align*}
[x, [y, z]] &= yz - yx \quad (by \text{ the definition of } [x, [y, z]]) \\
[z, [x, y]] &= xy - yz \quad (by \text{ the definition of } [z, [x, y]]) \\
[y, [z, x]] &= xz - zx \quad (by \text{ the definition of } [y, [z, x]])
\end{align*}
\]

\[
\begin{align*}
&= x \quad \text{(by the definition of } [y, z]) \\
&= -y \quad \text{(by the definition of } [z, y]) \\
&= x + z \quad \text{(by the definition of } [x, y]) \\
&= x + y \quad \text{(by the definition of } [x, y]) \\
&= z - y \quad \text{(by the definition of } [z, x]) \\
&= y - x \quad \text{(by the definition of } [y, x])
\end{align*}
\]

\[
= x(yz - yx) - (yz - xy) x + z(xy - yx) - (xy - yx) z + y(zx - zx) - (zx - zx) y
\]

\[
= xyz - xzy - yzx + yzx - zyx - xyz + yzx + yzx - yzx - zyx + xzy
\]

\[
= 0.
\]

This proves (11.81).

Now, both (11.80) and (11.81) are proven. Hence, the $k$-module $A$ endowed with the $k$-bilinear map $[\cdot, \cdot]$ satisfies the axioms of a Lie algebra, and therefore is a Lie algebra. This solves part (a) of the exercise.

(b) Let $A$ be a bialgebra. Let $p$ be the set of all primitive elements of $A$. Then, $p$ is a $k$-submodule of $A$ (because it is easy to see that $0 \in p$, that $\lambda a \in p$ for every $\lambda \in k$ and $a \in p$, and that $a + b \in p$ for all $a \in p$ and $b \in p$). A simple computation shows that every $x \in p$ and $y \in p$ satisfy $[x, y] \in p$. Hence, $[p, p] \subseteq p$ (since $p$ is a $k$-submodule of $A$). Therefore, $p$ is a Lie subalgebra of $A$. This solves part (b) of the exercise.

(c) For every subset $S$ of a $k$-module $U$, we let $\langle S \rangle$ denote the $k$-submodule of $U$ spanned by $S$.

We have defined $J$ as the two-sided ideal of $T(p)$ generated by all elements $xy - yx - [x, y]$ for $x, y \in p$.

In other words,

\[
\begin{align*}
(11.82) \quad J &= T(p) \cdot (xy - yx - [x, y] \mid x, y \in p) \cdot T(p) \\
&\supseteq k \\
(11.83) \quad \supseteq k \cdot (xy - yx - [x, y] \mid x, y \in p) \cdot k = \langle xy - yx - [x, y] \mid x, y \in p \rangle.
\end{align*}
\]

Thus,

\[
(11.84) \quad xy - yx - [x, y] \in J \quad \text{for all } x, y \in p.
\]

It is also easy to show that

\[
(11.85) \quad xy - yx - [x, y] \text{ is a primitive element of } T(p) \text{ for all } x, y \in p.
\]

308 From this it is easy to obtain

\[
(11.86) \quad \Delta ([xy - yx - [x, y] \mid x, y \in p]) \subseteq J \otimes T(p) + T(p) \otimes J.
\]

307 Proof. Let $x$ and $y$ be elements of $A$. Then, $\Delta [x, y] = \Delta 1 \otimes [x, y] + [x, y] \otimes 1$ (by (1.10)). In other words, the element $[x, y]$ of $A$ is primitive. In other words, $[x, y] \in p$ (since $p$ is the set of all primitive elements of $A$), qed.

308 Proof. Let $x$ and $y$ be elements of $A$. Then, $\Delta (xy - yx) = 1 \otimes (xy - yx) + (xy - yx) \otimes 1$ (this can be proven just as in the proof of (1.10)). But $[x, y]$ is an element of $p$, and thus (by the definition of the comultiplication of $T(p)$) satisfies $\Delta [x, y] = \Delta 1 \otimes [x, y] + [x, y] \otimes 1$ in $T(p) \otimes T(p)$. Now, since $\Delta$ is the $k$-linear map, we have

\[
\Delta (xy - yx - [x, y]) = \Delta (xy - yx) - \Delta [x, y]
\]

\[
= 1 \otimes (xy - yx) + (xy - yx) \otimes 1 \otimes [x, y] + [x, y] \otimes 1
\]

\[
= (1 \otimes (xy - yx) + (xy - yx) \otimes 1) - (1 \otimes [x, y] + [x, y] \otimes 1)
\]

\[
= (xy - yx) \otimes 1 - [x, y] \otimes 1 + 1 \otimes (xy - yx) - 1 \otimes [x, y]
\]

\[
= (xy - yx) \otimes 1 + 1 \otimes (xy - yx) - [x, y] \otimes 1
\]

\[
= 1 \otimes (xy - yx) - [x, y] \otimes 1.
\]

In other words, the element $xy - yx - [x, y]$ of $T(p)$ is primitive. This proves (11.85).
We already know that preserved under taking quotients. Hence, commutativity of the diagrams (1.8) and the commutativity of the diagrams (1.19); these axioms are clearly

\[ \Delta (p) \]  

Thus, \( (11.86) \) is proven.

Also, \( \epsilon (x - y x - [x, y] \mid x, y \in p) \otimes \epsilon (T(p)) \subset \epsilon (T(p)) \cdot \epsilon (x - y x - [x, y] \mid x, y \in p) \cdot \epsilon (T(p)) \)

(11.87)  

\[ J \otimes T(p) + T(p) \otimes J. \]

Also,  

(11.88)  

\[ \epsilon (x - y x - [x, y] \mid x, y \in p) = 0. \]

Apply to both sides of the equality (11.82), we obtain  

\[
\Delta (p) = \Delta (T(p) \cdot (x y - y x - [x, y] \mid x, y \in p) \cdot T(p))
\]

\[
\subset (T(p) \otimes T(p)) \cdot (J \otimes T(p) + T(p) \otimes J) \cdot T(p) \otimes T(p)
\]

\[ = (T(p) \otimes T(p)) \cdot (J \otimes T(p)) \cdot T(p) \otimes T(p) + (T(p) \otimes T(p)) \cdot (T(p) \otimes J) \cdot (T(p) \otimes T(p))
\]

\[ = (T(p) \cdot J \cdot T(p)) \otimes (T(p) \cdot T(p) \cdot T(p))
\]

(11.87)  

\[ J \otimes T(p) + T(p) \otimes J. \]

Thus, \( \epsilon (11.86) \) is proven.

Applying the map \( \epsilon \) to both sides of the equality (11.82), we obtain  

\[
\epsilon (J) = \epsilon (T(p) \cdot (x y - y x - [x, y] \mid x, y \in p) \cdot T(p)) \subset \epsilon (T(p)) \cdot \epsilon (x y - y x - [x, y] \mid x, y \in p) \cdot \epsilon (T(p))
\]

(11.88)  

\[ = 0. \]

Thus, \( \epsilon (11.86) = 0 \). Combined with (11.87), this yields that \( J \) is a two-sided coideal of \( T(p) \). Thus, the quotient \( T(p) / J \) becomes a k-coalgebra. Also, \( T(p) / J \) is a k-algebra, since \( J \) is a two-sided ideal of \( T(p) \).

Now we want to show that \( T(p) / J \) has a structure of a cocommutative k-bialgebra inherited from \( T(p) \). We already know that \( T(p) / J \) is a k-algebra and a k-coalgebra, with both structures being inherited from \( T(p) \). The remaining axioms of a cocommutative k-bialgebra that need to be checked for \( T(p) / J \) are the commutativity of the diagrams (1.8) and the commutativity of the diagrams (1.19); these axioms are clearly preserved under taking quotients. Hence, \( T(p) / J \) is a cocommutative k-bialgebra, with its structure being inherited from \( T(p) \). In other words, \( U(p) \) is a cocommutative k-bialgebra, with its structure being inherited from \( T(p) \) (since \( U(p) = T(p) / J \)). This solves part (c) of the exercise.

Proof. Let \( x, y \in p \) be arbitrary. Then, \( x y - y x - [x, y] \) is a primitive element of \( T(p) \) (by (11.85)). In other words,

\[
\Delta (x y - y x - [x, y]) = 1 \otimes (x y - y x - [x, y]) + (x y - y x - [x, y]) \otimes 1
\]

\[ = (x y - y x - [x, y]) \otimes 1 + 1 \otimes (x y - y x - [x, y])
\]

(11.84)  

\[ \in J \otimes T(p) + T(p) \otimes J. \]

Now forget that we fixed \( x, y \). We thus have proven that \( \Delta (x y - y x - [x, y]) \in J \otimes T(p) + T(p) \otimes J \) for all \( x, y \in p \). Since \( J \otimes T(p) + T(p) \otimes J \) is a k-submodule of \( T(p) \otimes T(p) \), this yields that \( \Delta ((x y - y x - [x, y]) \mid x, y \in p) \subset J \otimes T(p) + T(p) \otimes J. \) Thus, (11.86) is proven.

Proof. Let \( x, y \in p \). Then, \( x y - y x \in p^{\otimes 2} \) and thus \( \epsilon (x y - y x) = 0 \) (by the definition of the comultiplication \( \epsilon \) on \( T(p) \)). Also, \( [x, y] \in p^{\otimes 1} \) and thus \( \epsilon ([x, y]) = 0 \) (again by the definition of the comultiplication \( \epsilon \) on \( T(p) \)). Since \( \epsilon \) is k-linear, we have

\[
\epsilon (x y - y x - [x, y]) = \epsilon (x y - y x) - \epsilon ([x, y]) = 0 - 0 = 0.
\]

Now, forget that we fixed \( x, y \). We thus have shown that \( \epsilon (x y - y x - [x, y]) = 0 \) for all \( x, y \in p \). By linearity, this yields \( \epsilon (x y - y x - [x, y] \mid x, y \in p) = 0 \), so that (11.88) is proven.
(d) It is easy to see that
\[(11.89)\]
\[S(\langle xy - yx - [x, y] \mid x, y \in \mathfrak{p} \rangle) \subseteq J.\]

But recall that if $A$ and $B$ are two $k$-algebras and $\varphi : A \rightarrow B$ is a $k$-algebra homomorphism, then any $k$-submodules $V_1$, $V_2$, ..., $V_n$ of $A$ satisfy
\[\varphi (V_1 V_2 \cdots V_n) = \varphi (V_1) \cdot \varphi (V_2) \cdot \cdots \cdot \varphi (V_n).\]

Similarly, if $A$ and $B$ are two $k$-algebras and $\varphi : A \rightarrow B$ is a $k$-algebra antihomomorphism, then any $k$-submodules $V_1$, $V_2$, ..., $V_n$ of $A$ satisfy
\[\varphi (V_1 V_2 \cdots V_n) = \varphi (V_n) \cdot \varphi (V_{n-1}) \cdots \cdot \varphi (V_1).\]

Applying this to $A = T(\mathfrak{p})$, $B = T(\mathfrak{p})$, $\varphi = S$, $n = 3$, $V_1 = T(\mathfrak{p})$, $V_2 = \langle xy - yx - [x, y] \mid x, y \in \mathfrak{p} \rangle$ and $V_3 = T(\mathfrak{p})$, we obtain
\[S(T(\mathfrak{p}) \cdot (xy - yx - [x, y] \mid x, y \in \mathfrak{p}) \cdot T(\mathfrak{p})) = S(T(\mathfrak{p})) \cdot S((xy - yx - [x, y] \mid x, y \in \mathfrak{p})) \cdot S(T(\mathfrak{p}))\]
\[\subseteq T(\mathfrak{p}) \cdot J \cdot T(\mathfrak{p}) \subseteq J\]
(since $J$ is a two-sided ideal). Due to (11.82), this rewrites as $S(J) \subseteq J$. Hence, the $k$-linear map $S : T(\mathfrak{p}) \rightarrow T(\mathfrak{p})$ induces a $k$-linear map $S : T(\mathfrak{p}) / J \rightarrow T(\mathfrak{p}) / J$ on the quotient $k$-modules. This resulting map
$S : T(\mathfrak{p}) / J \rightarrow T(\mathfrak{p}) / J$ is an antipode for the $k$-bialgebra $T(\mathfrak{p}) / J$ (in fact, the diagram (1.12) with $A$ and $S$ replaced by $T(\mathfrak{p}) / J$ and $S$ commutes, because the diagram (1.12) with $A$ replaced by $T(\mathfrak{p})$ commutes). Hence, the $k$-bialgebra $T(\mathfrak{p}) / J$ has an antipode, and thus is a Hopf algebra. In other words, $U(\mathfrak{p})$ is a Hopf algebra (since $U(\mathfrak{p}) = T(\mathfrak{p}) / J$). Since we already know that $U(\mathfrak{p})$ is cocommutative, this yields that $U(\mathfrak{p})$ is a cocommutative Hopf algebra. This solves part (d) of the exercise.

11.20. **Solution to Exercise 1.59.** Solution to Exercise 1.59. Recall that our abstract definition of a $k$-algebra (Definition 1.1) and our definition of a $k$-coalgebra (Definition 1.6) differ from each other only in the directions of the arrows. More precisely, reversing all arrows in the former definition yields the latter definition. Similarly, our definition of a cocommutative $k$-coalgebra is obtained by reversing all arrows in our abstract definition of a commutative $k$-algebra, and our definition of a $k$-coalgebra homomorphism is obtained by reversing all arrows in our abstract definition of a $k$-algebra homomorphism. Hence, the statements of parts (a) and (b) of this exercise can be obtained from each other by reversing all arrows. Therefore, if we can solve part (b) of this exercise in an element-free way\(^{312}\), then we can clearly apply the same argument “with all arrows reversed” (and, of course, with $A$, $m_A$ and $u_A$ replaced by $C$, $\Delta_C$ and $\epsilon_C$) to solve part (a). Hence, in order to solve this exercise, it is enough to find an element-free solution to its part (b). Let us do this now.

Let us first show how to solve part (b) using computations with elements (i.e., not in an element-free way). This is very easy. We need to prove the following statements:

**Statement 1:** If a $k$-algebra $A$ is commutative, then its multiplication $m_A : A \otimes A \rightarrow A$ is a $k$-algebra homomorphism.

\(^{311}\) _Proof_. Let $x, y \in \mathfrak{p}$. Then, $xy - yx - [x, y]$ is a primitive element of $T(\mathfrak{p})$ (by (11.85)). Hence, Proposition 1.37 (applied to $xy - yx - [x, y]$ instead of $x$) yields
\[S(xy - yx - [x, y]) = -\underbrace{xy - yx - [x, y]}_{\in J} \in J = J\]
(by (11.84))

(since $J$ is a two-sided ideal).

Now, forget that we fixed $x, y$. We thus have shown that $S(xy - yx - [x, y]) \in J$ for all $x, y \in \mathfrak{p}$. Since $J$ is a $k$-module, this yields that $S\langle xy - yx - [x, y] \mid x, y \in \mathfrak{p} \rangle \subseteq J$. Thus, (11.89) is proven.

\(^{312}\) By an “element-free” argument, we mean an argument which only talks about linear maps, but never talks about elements of modules such as $A$ and $A \otimes A$. For instance, the Second solution of Exercise 1.29(a) that we gave above was element-free, whereas the First solution of Exercise 1.29(a) (which we also showed above) was not element-free (since it involved elements $c$ and $d$ of $C$ and $D$). A synonym for “element-free argument” is “argument by pure diagram chasing”, although it is not required that one actually draws any diagrams in the argument (commutative diagrams are just shortcuts for identities between maps).
Statement 2: If a \( \mathbf{k} \)-algebra \( A \) has the property that its multiplication \( m_A : A \otimes A \to A \) is a \( \mathbf{k} \)-algebra homomorphism, then \( A \) is commutative.

Proof of Statement 1 using computations with elements: Assume that a \( \mathbf{k} \)-algebra \( A \) is commutative. We need to prove that \( m_A : A \otimes A \to A \) is a \( \mathbf{k} \)-algebra homomorphism. To do so, it is enough to show that \( m_A \) preserves products and that \( m_A \) maps the unity of \( A \otimes A \) to the unity of \( A \).

Let us show that \( m_A \) preserves products. This means proving that \( m_A(uv) = m_A(u) \cdot m_A(v) \) for all \( u \in A \otimes A \) and \( v \in A \otimes A \). So let \( u \in A \otimes A \) and \( v \in A \otimes A \). Since the equality in question \( (m_A(uv) = m_A(u) \cdot m_A(v)) \) is linear in each of \( u \) and \( v \), we can WLOG assume that \( u \) and \( v \) are pure tensors. Having made this assumption, we can write \( u = a \otimes b \) for some \( a \in A \) and \( b \in A \), and we can write \( v = c \otimes d \) for some \( c \in A \) and \( d \in A \). Now,

\[
m_A \left( \frac{u}{a \otimes b} \right) \cdot m_A \left( \frac{v}{c \otimes d} \right) = m_A (a \otimes b) \cdot m_A (c \otimes d) = (ab) \cdot (cd).
\]

But\(^{313}\)

\[
(11.90) \quad (ac) (bd) = ((ac) b) d = \left( a \begin{pmatrix} b \end{pmatrix} \right) d = (a (bc)) d = (ab) c = (ab) (cd).
\]

Hence, altogether,

\[
m_A(uv) = (ac) (bd) = (ab) (cd) = m_A(u) \cdot m_A(v).
\]

Thus, we have proven that \( m_A \) preserves products. In order to prove that \( m_A \) maps the unity of \( A \otimes A \) to the unity of \( A \), we recall that the former unity is \( 1_A \otimes 1_A \) and the latter unity is \( 1_A \), which satisfy \( m_A(1_A \otimes 1_A) = 1_A \cdot 1_A = 1_A \). We have thus shown that \( m_A \) preserves products and that \( m_A \) maps the unity of \( A \otimes A \) to the unity of \( A \). In other words, \( m_A \) is a \( \mathbf{k} \)-algebra homomorphism, and Statement 1 is proven.

Proof of Statement 2 using computations with elements: Assume that a \( \mathbf{k} \)-algebra \( A \) has the property that its multiplication \( m_A : A \otimes A \to A \) is a \( \mathbf{k} \)-algebra homomorphism. Let \( a \) and \( b \) be elements of \( A \). Then,

\[
m_A \left( \frac{(1_A \otimes b) (a \otimes 1_A)}{= (1_A \otimes b) \cdot (a \otimes 1_A)} \right) = m_A (a \otimes b) = ab.
\]

Comparing this with

\[
m_A((1_A \otimes b) (a \otimes 1_A)) = m_A(1_A \otimes b) m_A(a \otimes 1_A) \quad (\text{since } m_A \text{ is a } \mathbf{k} \text{-algebra homomorphism})
\]

\[
= 1_A \cdot a = a = a \cdot 1_A = a
\]

this becomes \( ab = ba \). We thus have shown that \( ab = ba \) for all \( a \in A \) and \( b \in A \). In other words, \( A \) is commutative, and thus Statement 2 is proven.

We thus have solved part (b) of the exercise using computations with elements. But we want an element-free solution of part (b). It turns out that we can obtain such a solution from our above solution by a more or less straightforward rewriting procedure. Let us show how this works.

Again, we need to prove Statements 1 and 2 made above.

\(^{313}\)In the following computation, we are deliberately being painstakingly slow and writing down every single step, including every application of associativity. This is to simplify our job later on (when we will translate this computation to an element-free argument).
**Element-free proof of Statement 1:** Assume that a \(k\)-algebra \(A\) is commutative. We need to prove that \(m_A : A \otimes A \rightarrow A\) is a \(k\)-algebra homomorphism. To do so, it is enough to show that the diagrams

\[
\begin{array}{ccc}
A \otimes A & \xrightarrow{m_A} & A \\
\uparrow{m_{A \otimes A}} & & \downarrow{m_A} \\
A \otimes A \otimes A \otimes A & \xrightarrow{m_A \otimes m_A} & A \otimes A
\end{array}
\] and

\[
\begin{array}{ccc}
A \otimes A & \xrightarrow{m_A} & A \\
\uparrow{u_A \otimes u_A} & & \downarrow{u_A} \\
A \otimes A \otimes A \otimes A & \xrightarrow{m_A \otimes m_A} & A \otimes A
\end{array}
\]

are commutative. In other words, it is enough to show that

\[
(11.91) \quad m_A \circ m_{A \otimes A} = m_A \circ (m_A \otimes m_A)
\]

and

\[
(11.92) \quad u_A = m_A \circ u_{A \otimes A}.
\]

Let us prove (11.91) first. If we were allowed to compute with elements, then we could prove (11.91) by evaluating both sides of (11.91) at a pure tensor \(a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A\); this would leave us with the task of showing that \((ac)(bd) = (ab)(cd)\), which we already have done in the computation which proved (11.90). However, we are not allowed to do this, because we want this proof to be element-free. But what we can do is computing with maps instead of elements. We just need to replace the computation which proved (11.90) by a computation which uses maps instead of elements. If we replace every expression in (11.90) by the \(k\)-linear map which sends every \(a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A\) to said expression, then we obtain:

\[
\begin{align*}
\text{(the } k\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (ac)(bd)) & = m_A \circ (m_A \otimes m_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) ; \\
\text{(the } k\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } ((ac)b)d) & = m_A \circ (m_A \otimes \text{id}_A) \circ (m_A \otimes \text{id}_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) ; \\
\text{(the } k\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (a(cb))d) & = m_A \circ (m_A \otimes \text{id}_A) \circ (\text{id}_A \otimes m_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) ; \\
\text{(the } k\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (a(bc))d) & = m_A \circ (m_A \otimes \text{id}_A) \circ (\text{id}_A \otimes m_A \otimes \text{id}_A) ; \\
\text{(the } k\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } ((ab)c)d) & = m_A \circ (m_A \otimes \text{id}_A) \circ (m_A \otimes \text{id}_A \otimes \text{id}_A) ; \\
\text{(the } k\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (ab)(cd)) & = m_A \circ (m_A \otimes m_A) .
\end{align*}
\]

We are not yet done, because we still are using elements (in describing the maps). So we should rewrite the maps appearing in the computation (11.93) in such a way that no elements occur in them anymore. Denoting by \(T\) the twist map \(A \otimes A \rightarrow A \otimes A\) (sending every \(a \otimes b\) to \(b \otimes a\)) \footnote{We do not count the use of this map \(T\) as a use of elements (even though we just defined it using elements). Twist maps like \(T\) are one of the basic features of tensor products (along with associativity isomorphisms \((U \otimes V) \otimes W \rightarrow U \otimes (V \otimes W)\), which we are suppressing, and with trivial isomorphisms of the form \(k \otimes U \rightarrow U\), and their use is allowed in element-free arguments. They don’t interfere with “reversing the arrows” because arrows like \(T\) are very easy to reverse.}, we have

\[
\begin{align*}
\text{(the } k\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (ac)(bd)) & = m_A \circ (m_A \otimes m_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) ; \\
\text{(the } k\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } ((ac)b)d) & = m_A \circ (m_A \otimes \text{id}_A) \circ (m_A \otimes \text{id}_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) ; \\
\text{(the } k\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (a(cb))d) & = m_A \circ (m_A \otimes \text{id}_A) \circ (\text{id}_A \otimes m_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) ; \\
\text{(the } k\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (a(bc))d) & = m_A \circ (m_A \otimes \text{id}_A) \circ (\text{id}_A \otimes m_A \otimes \text{id}_A) ; \\
\text{(the } k\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } ((ab)c)d) & = m_A \circ (m_A \otimes \text{id}_A) \circ (m_A \otimes \text{id}_A \otimes \text{id}_A) ; \\
\text{(the } k\text{-linear map sending every } a \otimes b \otimes c \otimes d \in A \otimes A \otimes A \otimes A \text{ to } (ab)(cd)) & = m_A \circ (m_A \otimes m_A) .
\end{align*}
\]
Hence, the computation (11.93) rewrites as

\[
\begin{align*}
    m_A \circ (m_A \otimes m_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) \\
    = m_A \circ (m_A \otimes \text{id}_A) \circ (m_A \otimes \text{id}_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) \\
    = m_A \circ (m_A \otimes \text{id}_A) \circ (\text{id}_A \otimes m_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) \\
    = m_A \circ (m_A \otimes \text{id}_A) \circ (\text{id}_A \otimes m_A \otimes \text{id}_A) \\
    = m_A \circ (m_A \otimes \text{id}_A) \circ (m_A \otimes \text{id}_A \otimes \text{id}_A).
\end{align*}
\]

(11.94)

However, now that we are no longer talking about elements, our computation has become significantly harder to follow. For example, the first equality in the computation (11.94) is far less obvious than the corresponding equality \((ac)(bd) = ((ac)b)d\) in the computation (11.90). So, in order to justify the computation (11.94), we need to recall where exactly we used associativity or commutativity in (11.93), and translate these uses into element-free language. Let us do this step by step:

**The first equality in (11.94):** The first equality in (11.94) is simply an element-free way to state \((ac)(bd) = ((ac)b)d\) for all \(a, b, c, d \in A\). On the level of elements, this follows from associativity to the elements \(ac\) and \(b\). For example, the first equality in the computation (11.94) is far less obvious than the corresponding

\[
\begin{align*}
    &m_A \circ (m_A \otimes m_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) \\
    &\quad = m_A \circ (m_A \otimes \text{id}_A) \circ (m_A \otimes \text{id}_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) \\
    &\quad = m_A \circ (m_A \otimes \text{id}_A) \circ (\text{id}_A \otimes m_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) \\
    &\quad = m_A \circ (m_A \otimes \text{id}_A) \circ (m_A \otimes \text{id}_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A).
\end{align*}
\]

Thus, we have found an element-free proof of the first equality in (11.94).

**The second equality in (11.94):** The second equality in (11.94) is simply an element-free way to state \(((ac)b)d = (a(cb))d\) for all \(a, b, c, d \in A\). On the level of elements, this follows from associativity to the elements \(a, c\) and \(b\) of \(A\), and then multiplying with \(d\) on the right. In other words, this follows from applying the associativity law \(m_A \circ (\text{id}_A \otimes m_A) = m_A \circ (m_A \otimes \text{id}_A)\) to the tensor \(ac \otimes b \otimes d = (m_A \otimes \text{id}_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A)\) \((a \otimes b \otimes c \otimes d)\). Therefore, the first equality in (11.94) should follow from \(m_A \circ (\text{id}_A \otimes m_A) = m_A \circ (m_A \otimes m_A)\) by composition with the map \((m_A \otimes \text{id}_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A)\) on the right. And indeed, this is how it is proven:

\[
\begin{align*}
    &m_A \circ (m_A \otimes m_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) \\
    &\quad = m_A \circ (\text{id}_A \otimes m_A) \circ (m_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) \\
    &\quad = m_A \circ (\text{id}_A \otimes m_A) \circ (m_A \otimes \text{id}_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A) \\
    &\quad = m_A \circ (m_A \otimes \text{id}_A) \circ (m_A \otimes \text{id}_A \otimes \text{id}_A) \circ (\text{id}_A \otimes T \otimes \text{id}_A).
\end{align*}
\]

Thus, we have obtained an element-free proof of the second equality in (11.94).

**The third equality in (11.94):** The third equality in (11.94) is simply an element-free way to state \((a(cb))d = (a(bc))d\) for all \(a, b, c, d \in A\). On the level of elements, this follows from applying commutativity to the elements \(c\) and \(b\) of \(A\), then multiplying with \(a\) on the left, and then multiplying with \(d\) on the
right. In other words, this follows from applying the commutativity law \( m_A = m_A \circ T \) to the second and third tensorands of the tensor \( a \otimes c \otimes b \otimes d = (id_A \otimes T \otimes id_A) (a \otimes b \otimes c \otimes d) \), and then applying \( m_A \circ (m_A \otimes id_A) \).

In other words, this follows from applying the equality \( id_A \otimes m_A \otimes id_A = id_A \otimes (m_A \circ T) \otimes id_A \) to the tensor \( (id_A \otimes T \otimes id_A) (a \otimes b \otimes c \otimes d) \), and then applying \( m_A \circ (m_A \otimes id_A) \).

Therefore, the third equality in (11.94) should follow from \( m_A = m_A \circ T \) by tensoring both sides with id_A on the left and on the right and then composing them with the map id_A \( \otimes T \otimes id_A \) on the right and with the map \( m_A \circ (m_A \otimes id_A) \) on the left. And indeed, this is how it is proven:

\[
\begin{align*}
&= m_A \circ (m_A \otimes id_A) \circ (id_A \otimes T \otimes id_A) \\
&= m_A \circ (m_A \otimes id_A) \circ (id_A \otimes id_A) \\
&= m_A \circ (m_A \otimes id_A) \circ (id_A \otimes id_A) \\
&= m_A \circ (m_A \otimes id_A) \circ (id_A \otimes m_A \otimes id_A). \\
\end{align*}
\]

Thus, we have obtained an element-free proof of the third equality in (11.94).

The fourth equality in (11.94): The fourth equality in (11.94) is simply an element-free way to state \((a(bc))d = ((ab)c)d\) for all \(a, b, c, d \in A\). On the level of elements, this follows from applying associativity to the elements \(a, b\) and \(c, d\) of \(A\), and then multiplying with \(d\) on the right. In other words, this follows from applying the associativity law \( m_A \circ (id_A \otimes m_A) = m_A \circ (m_A \otimes id_A) \) to the first three tensorands of the tensor \(a \otimes b \otimes c \otimes d\), and then applying the map \(m_A\). In other words, this follows from applying the equality \((m_A \circ (id_A \otimes m_A)) \otimes id_A = (m_A \circ (m_A \otimes id_A)) \otimes id_A\) to the tensor \(a \otimes b \otimes c \otimes d\), and then applying the map \(m_A\).

Therefore, the fourth equality in (11.94) should follow from \(m_A \circ (id_A \otimes m_A) = m_A \circ (m_A \otimes id_A)\) by tensoring both sides with \(id_A\) on the right and then composing them with the map \(m_A\) on the left. And indeed, this is how it is proven:

\[
\begin{align*}
&= m_A \circ (m_A \otimes id_A) \circ (id_A \otimes m_A \otimes id_A) \\
&= m_A \circ (m_A \otimes id_A) \circ (id_A \otimes m_A \otimes id_A) \\
&= m_A \circ (m_A \otimes id_A) \circ (id_A \otimes m_A \otimes id_A). \\
\end{align*}
\]

Thus, we have obtained an element-free proof of the fourth equality in (11.94).

The fifth equality in (11.94): The fifth equality in (11.94) is simply an element-free way to state \(((ab)c)d = (ab)(cd)\) for all \(a, b, c, d \in A\). On the level of elements, this follows from applying associativity to the elements \(ab, c\) and \(d\) of \(A\). In other words, this follows from applying the associativity law \(m_A \circ (m_A \otimes id_A) = m_A \circ (id_A \otimes m_A)\) to the tensor \(ab \otimes c \otimes d = (m_A \otimes id_A \otimes id_A)(a \otimes b \otimes c \otimes d)\). Therefore, the fifth equality in (11.94) should follow from \(m_A \circ (m_A \otimes id_A) = m_A \circ (id_A \otimes m_A)\) by composing both sides of this equality with \(m_A \otimes id_A \otimes id_A\) on the right. And indeed, this is how it is proven:

\[
\begin{align*}
&= m_A \circ (m_A \otimes id_A) \circ (id_A \otimes m_A \otimes id_A) \\
&= m_A \circ (m_A \otimes id_A) \circ (id_A \otimes m_A \otimes id_A) \\
&= m_A \circ (m_A \otimes m_A). \\
\end{align*}
\]

Thus, we have obtained an element-free proof of the fifth equality in (11.94).
Now, all five equalities in the computation (11.94) are proven without reference to elements. Hence, we have shown $m_A \circ (m_A \otimes m_A) \circ (id_A \otimes T \otimes id_A) = m_A \circ (m_A \otimes m_A)$ in an element-free way. In other words, $m_A \circ m_{A \otimes A} = m_A \circ (m_A \otimes m_A)$ is proven in an element-free way (because the definition of $m_{A \otimes A}$ yields $m_{A \otimes A} = (m_A \otimes m_A) \circ (id_A \otimes T \otimes id_A)$). In other words, (11.91) is proven.

It remains to prove (11.92). We leave this very simple proof to the reader (noticing that it does not require the commutativity of $A$).

Now we know that both (11.91) and (11.92) hold. Statement 1 is thus proven in an element-free way.

We leave it to the reader to prove Statement 2 in an element-free way. Altogether, Statements 1 and 2 have now been proven. Therefore, part (b) of the exercise has been solved in an element-free way. Therefore, a solution of part (a) can be obtained mechanically by “reversing all arrows”.

Of course, part (a) can alternatively be solved using Sweedler’s notation.

11.21. Solution to Exercise 1.61. Solution to Exercise 1.61. For every $1 \leq i < j \leq k$, let $t_{i,j}$ be the transposition in $S_k$ which transposes $i$ with $j$. It is well-known that the symmetric group $S_k$ is generated by the transpositions $t_{i,i+1}$ with $i$ ranging over $\{1, 2, \ldots, k-1\}$.

But let us notice that $(\rho(\pi)) \circ (\rho(\psi)) = \rho(\pi \psi)$ for any two elements $\pi$ and $\psi$ of $S_k$. Hence, the set of all $\pi \in S_k$ satisfying $m^{(k-1)} \circ (\rho(\pi)) = m^{(k-1)}$ is closed under multiplication. Since this set also contains the trivial permutation $id = 1_{S_k} \in S_k$ and is closed under taking inverses (this is easy to check), this shows that this set is a subgroup of $S_k$. Therefore, if this set contains a set of generators of $S_k$, then this set must be the whole $S_k$. Hence, if we can show that this set contains the transposition $t_{i,i+1}$ for every $i \in \{1, 2, \ldots, k-1\}$, then it will follow that this set must be the whole $S_k$ (because the group $S_k$ is generated by the transpositions $t_{i,i+1}$ with $i$ ranging over $\{1, 2, \ldots, k-1\}$), and this will entail that every $\pi \in S_k$ satisfies $m^{(k-1)} \circ (\rho(\pi)) = m^{(k-1)}$, which will solve the problem. Hence, in order to complete this solution, it is enough to check that the set of all $\pi \in S_k$ satisfying $m^{(k-1)} \circ (\rho(\pi)) = m^{(k-1)}$ contains the transposition $t_{i,i+1}$ for every $i \in \{1, 2, \ldots, k-1\}$. In other words, it is enough to check that

$$m^{(k-1)} \circ (\rho(t_{i,i+1})) = m^{(k-1)} \quad \text{ for all } i \in \{1, 2, \ldots, k-1\}. \tag{11.95}$$

Proof of (11.95): Let $i \in \{1, 2, \ldots, k-1\}$. Let $T$ denote the twist map $A \otimes A \rightarrow A \otimes A$ sending every pure tensor $a \otimes b$ to $b \otimes a$. Any $k$ vectors $v_1, v_2, \ldots, v_k$ in $A$ satisfy

\[
(\rho(t_{i,i+1}))(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = t_{i,i+1}(v_1 \otimes v_2 \otimes \cdots \otimes v_k) = \left(\frac{v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1} \otimes v_{i+1} \otimes v_i \otimes v_{i+2} \otimes \cdots \otimes v_k}{= id_A \otimes id_A \otimes (v_1 \otimes v_2 \otimes \cdots \otimes v_{i-1})} \right) \otimes T(v_i \otimes v_{i+1}) \otimes id_{A \otimes A \otimes (k-1-i)}(v_{i+2} \otimes v_{i+3} \otimes \cdots \otimes v_k) = (id_A \otimes T \otimes id_{A \otimes A \otimes (k-1-i)})(v_1 \otimes v_2 \otimes \cdots \otimes v_k).
\]

In other words, the two maps $\rho(t_{i,i+1}) \circ id_{A \otimes (i-1)} \otimes T \otimes id_{A \otimes (k-1-i)}$ and $id_{A \otimes (i-1)} \otimes T \otimes id_{A \otimes (k-1-i)}$ are equal to each other on every pure tensor. Being $k$-linear maps, these two maps must therefore be identical, i.e., we have $\rho(t_{i,i+1}) = id_{A \otimes (i-1)} \otimes T \otimes id_{A \otimes (k-1-i)}$.

But recall that $A$ is commutative, whence the diagram (1.18) commutes. Thus, $m \circ T = m$.

We have $i \in \{1, 2, \ldots, k-1\}$, so that $1 \leq i \leq k-1$ and therefore $0 \leq i-1 \leq (k-1) - 1$. Hence, in particular, $k-1 \geq 1 \geq 0$. We can thus apply Exercise 1.39(c) to $k-1$ and $i-1$ instead of $k$ and $i$. As a result, we obtain

$$m^{(k-1)} = m^{((k-1)-1)} \circ (id_{A \otimes (i-1)} \otimes m \otimes id_{A \otimes ((k-1) - 1 - (i-1))}) = m^{((k-1)-1)} \circ (id_{A \otimes (i-1)} \otimes m \otimes id_{A \otimes (k-1-i)}).$$
Hence,

\[
m^{(k-1)} \circ (\rho (t_{i,i+1})) = m^{(k-1)} \circ (\id_A \otimes (\rho T) \otimes \id_A) = (\id_A \otimes (m T) \otimes \id_A) = m^{(k-1)}.
\]

Thus, (11.95) is proven. This completes the solution of Exercise 1.61.

11.22. **Solution to Exercise 1.62.** *Solution to Exercise 1.62.* Here is the statement:

**Exercise.** Let \( C \) be a cocommutative \( k \)-coalgebra, and let \( k \in \mathbb{N} \). The symmetric group \( \mathfrak{S}_k \) acts on the \( k \)-fold tensor power \( C^\otimes k \) by permuting the tensor factors: \( \sigma (v_1 \otimes v_2 \otimes \cdots \otimes v_k) = v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \cdots \otimes v_{\sigma^{-1}(k)} \) for all \( v_1, v_2, \ldots, v_k \in C \) and \( \sigma \in \mathfrak{S}_k \). For every \( \pi \in \mathfrak{S}_k \), denote by \( \rho (\pi) \) the action of \( \pi \) on \( C^\otimes k \) (this is an endomorphism of \( C^\otimes k \)). Show that every \( \pi \in \mathfrak{S}_k \) satisfies \( (\rho (\pi)) \circ \Delta^{(k-1)} = \Delta^{(k-1)} \). (Recall that \( \Delta^{(k-1)} : C \to C^\otimes k \) is defined as in Exercise 1.40 for \( k \geq 1 \), and by \( \Delta^{(-1)} = \epsilon : C \to k \) for \( k = 0 \).

The solution of this exercise can be obtained from the above solution of Exercise 1.61 by reversing all arrows (and replacing \( A, m \) and \( m^{(i)} \) by \( C, \Delta \) and \( \Delta^{(i)} \)).

11.23. **Solution to Exercise 1.63.** *Solution to Exercise 1.63.* (a) Let \( H \) be a \( k \)-bialgebra and \( A \) be a commutative \( k \)-algebra. Let \( f \) and \( g \) be two \( k \)-algebra homomorphisms \( H \to A \).

Since \( A \) is commutative, the map \( m_A : A \otimes A \to A \) is a \( k \)-algebra homomorphism (according to Exercise 1.59(b)). Also, \( f \otimes g \) is a \( k \)-algebra homomorphism (by Exercise 1.13(a), since \( f \) and \( g \) are \( k \)-algebra homomorphisms). Finally, the axioms of a \( k \)-bialgebra show that \( \Delta_H : H \to H \otimes H \) is a \( k \)-algebra homomorphism (since \( H \) is a \( k \)-bialgebra). Now, the definition of convolution yields that \( f * g = m_A \circ (f \otimes g) \circ \Delta_H \). Thus, \( f * g \) is a composition of three \( k \)-algebra homomorphisms (namely, of \( m_A, f \otimes g \) and \( \Delta_H \)), and hence a \( k \)-algebra homomorphism itself. This solves Exercise 1.63(a).

(b) Exercise 1.63(b) can be solved by straightforward induction over \( k \) using (in the induction step) the result of Exercise 1.63(a) and (in the induction base) the fact that \( u_A \circ \epsilon_H \) is a \( k \)-algebra homomorphism (since \( u_A \) and \( \epsilon_H \) are \( k \)-algebra homomorphisms). The details are left to the reader.

(c) Let \( H \) be a Hopf algebra, and let \( A \) be a commutative \( k \)-algebra. Let \( f : H \to A \) be a \( k \)-algebra homomorphism.

Proposition 1.32 shows that the antipode \( S \) of \( H \) is an algebra anti-endomorphism. Combined with the fact that \( f \) is a \( k \)-algebra homomorphism, this yields that the composition \( f \circ S \) is a \( k \)-algebra anti-homomorphism. But since algebra anti-homomorphisms \( H \to A \) are the same as algebra homomorphisms \( H \to A \) (since \( A \) is commutative), this yields that \( f \circ S \) is a \( k \)-algebra homomorphism. It remains to prove
that \( f \circ S \) is \( \ast \)-inverse to \( f \). But this is easy: The definition of convolution yields

\[
(f \circ S) \ast f = m_A \circ \left( (f \circ S) \otimes f \right) \circ \Delta_H = m_A \circ ((f \circ S) \otimes (f \circ \text{id})) \circ \Delta_H
\]

\[
= m_A \circ (f \otimes f) \circ (S \otimes \text{id}) \circ \Delta_H = f \circ m_H \circ (S \otimes \text{id}) \circ \Delta_H
\]

(since \( f \) is a \( k \)-algebra homomorphism)

\[
= \phi \circ \Delta_H = u_A \circ \epsilon_H,
\]

(since \( f \) is a \( k \)-algebra homomorphism)

and similarly \( f \ast (f \circ S) = u_A \circ \epsilon_H \), so that \( f \circ S \) is \( \ast \)-inverse to \( f \). This solves Exercise 1.63(c).

(d) Let \( A \) be a commutative \( k \)-algebra. Then, the map \( m_A : A \otimes A \to A \) is a \( k \)-algebra homomorphism (according to Exercise 1.59(b)).

We need to prove that \( m_A^{(k)} \) is a \( k \)-algebra homomorphism (since \( m^{(k)} = m_A^{(k)} \)). To do so, it suffices to adapt the solution of Exercise 1.42(a) with only very minor changes (mainly we have to change \( H \) into \( A \), replace “coalgebra” by “algebra” and use Exercise 1.13(a) instead of using Exercise 1.13(b)). The details of this adaptation are left to the reader.

(c) Let \( C' \) and \( C \) be two \( k \)-coalgebras. Let \( \gamma : C \to C' \) be a \( k \)-coalgebra homomorphism. Let \( A \) and \( A' \) be two \( k \)-algebras. Let \( \alpha : A \to A' \) be a \( k \)-algebra homomorphism. Let \( f_1, f_2, \ldots, f_k \) be several \( k \)-linear maps \( C' \to A \).

Define a map \( \varphi : \text{Hom}(C', A) \to \text{Hom}(C, A') \) by setting \( \varphi(f) = \alpha \circ f \circ \gamma \) for every \( f \in \text{Hom}(C', A) \). As was proven in the proof of Proposition 1.46, this map \( \varphi \) is a \( k \)-algebra homomorphism (\( \text{Hom}(C', A'), \ast ) \to (\text{Hom}(C, A), \ast) \). Thus,

\[
\varphi(f_1 \ast f_2 \ast \cdots \ast f_k) = \left( \begin{array}{c}
\varphi(f_1) \\
\varphi(f_2) \\
\vdots \\
\varphi(f_k)
\end{array} \right) = \left( \begin{array}{c}
\alpha \circ f_1 \circ \gamma \\
\alpha \circ f_2 \circ \gamma \\
\vdots \\
\alpha \circ f_k \circ \gamma
\end{array} \right)
\]

(by the definition of \( \varphi \)).

Hence,

\[
(\alpha \circ f_1 \circ \gamma) \ast (\alpha \circ f_2 \circ \gamma) \ast \cdots \ast (\alpha \circ f_k \circ \gamma) = \varphi(f_1 \ast f_2 \ast \cdots \ast f_k) = \alpha \circ (f_1 \ast f_2 \ast \cdots \ast f_k)
\]

(by the definition of \( \varphi \)). This solves Exercise 1.63(e).

(f) Let \( H \) be a commutative \( k \)-bialgebra. Let \( k \) and \( \ell \) be two nonnegative integers. Then, Exercise 1.63(b) (applied to \( A = H \) and \( f_i = \text{id}_H \)) yields that \( \text{id}_H \ast \text{id}_H \ast \cdots \ast \text{id}_H \) is a \( k \)-algebra homomorphism \( H \to H \).

Since \( \text{id}_H \ast \text{id}_H \ast \cdots \ast \text{id}_H = \text{id}_H^k \), this shows that \( \text{id}_H^k \) is a \( k \)-algebra homomorphism \( H \to H \). We can thus apply Exercise 1.63(e) to \( H, H, H, H, \ell, \text{id}_H, \text{id}_H^k \) and \( \text{id}_H \) instead of \( C, C', A, A', k, f_i, \alpha \) and \( \gamma \). As a result, we obtain

\[
\text{id}_H^k \circ \left( \text{id}_H \ast \text{id}_H \ast \cdots \ast \text{id}_H \right) \circ \text{id}_H = \left( \text{id}_H^k \circ \text{id}_H \circ \text{id}_H \right) \ast \left( \text{id}_H^k \circ \text{id}_H \circ \text{id}_H \right) \ast \cdots \ast \left( \text{id}_H^k \circ \text{id}_H \circ \text{id}_H \right).
\]

Since \( \text{id}_H^k \circ \text{id}_H \circ \text{id}_H = \text{id}_H^k \) and \( \text{id}_H \ast \text{id}_H \ast \cdots \ast \text{id}_H = \text{id}_H^k \), this rewrites as

\[
\text{id}_H^k \circ \text{id}_H^\ell = \left( \text{id}_H^k \ast \text{id}_H^k \ast \cdots \ast \text{id}_H^k \right)^\ell = \text{id}_H^{k \ell}.
\]

This solves Exercise 1.63(f).

(g) Let \( H \) be a commutative \( k \)-bialgebra.
First, it is easy to see that
\begin{equation}
\text{id}^k_H \text{ is a } \mathbf{k}\text{-algebra homomorphism } H \to H \text{ for every } k \in \mathbb{Z}.
\end{equation}

Furthermore, we have
\begin{equation}
\text{id}^{(-k)}_H = \text{id}^k_H \circ S \text{ for every } k \in \mathbb{Z}.
\end{equation}

Now, fix two integers \( k \) and \( \ell \). Due to (11.96), we know that \( \text{id}^k_H \) is a \( \mathbf{k}\)-algebra homomorphism \( H \to H \). Hence, if \( \ell \) is nonnegative, then we can prove \( \text{id}^k_H \circ \text{id}^\ell_H = \text{id}^{k\ell}_H \) just as we did in the solution to Exercise 1.63(f) (and thus finish the solution to Exercise 1.63(g)). Hence, for the rest of this solution, we WLOG assume that \( \ell \) is not nonnegative. Thus, \( \ell < 0 \), so that \( -\ell \) is nonnegative. We can thus apply Exercise 1.63(e) to \( H, H, H, H, -\ell, \text{id}_H, \text{id}^k_H \) and \( \text{id}^{\ell}_H \) instead of \( C, C', A, A', f_i, \alpha \) and \( \gamma \). As a result, we obtain
\[
\text{id}^k_H \circ \left( \text{id}^k_H \circ \text{id}^k_H \cdots \circ \text{id}^k_H \right) \circ \text{id}_H = \left( \text{id}^k_H \circ \text{id}^k_H \circ \text{id}^k_H \right) \cdots \left( \text{id}^k_H \circ \text{id}^k_H \circ \text{id}^k_H \right).
\]

Since \( \text{id}^k_H \circ \text{id}^k_H = \text{id}^{2k}_H \) and \( \text{id}^k_H \circ \text{id}^k_H \cdots \circ \text{id}^k_H = \text{id}^{(-\ell)}_H \), this rewrites as
\[
\text{id}^k_H \circ \text{id}^{(-\ell)}_H = \text{id}^{2k}_H \circ \text{id}^{(-\ell)}_H = \left( \text{id}^k_H \right)^{\ell} = \text{id}^{k(-\ell)}_H = \text{id}^{(k\ell)}_H.
\]

Now, \( \text{id}^\ell_H = \text{id}^{(-\ell)}_H = \text{id}^{(-\ell)}_H \circ S \) (by (11.97), applied to \( -\ell \) instead of \( k \)), and thus
\[
\text{id}^k_H \circ \text{id}^\ell_H = \text{id}^{2k}_H \circ \text{id}^{(-\ell)}_H = \text{id}^{k(-\ell)}_H = \text{id}^{(-k\ell)}_H.
\]

Compared with \( \text{id}^{(k\ell)}_H = \text{id}^{(k(-\ell))}_H = \text{id}^{(k\ell)}_H \) (by (11.97), applied to \( -k\ell \) instead of \( k \)), this yields
\[
\text{id}^k_H \circ \text{id}^\ell_H = \text{id}^{(k\ell)}_H. \text{ This solves Exercise 1.63(g).}
\]

(h) The dual of Exercise 1.63(a) is the following exercise:

If \( H \) is a \( \mathbf{k}\)-bialgebra and \( C \) is a cocommutative \( \mathbf{k}\)-coalgebra, and if \( f \) and \( g \) are two \( \mathbf{k}\)-coalgebra homomorphisms \( C \to H \), then prove that \( fg \) also is a \( \mathbf{k}\)-coalgebra homomorphism \( C \to H \).

The solution of this exercise is obtained from our above solution of Exercise 1.63(a) by “reversing arrows” (and replacing “algebra” by “coalgebra”, and applying Exercise 1.59(a) instead of Exercise 1.59(b), and using Exercise 1.13(b) instead of Exercise 1.13(a)).

The dual of Exercise 1.63(b) is the following exercise:

Proof of (11.96): Let \( k \in \mathbb{Z} \). If \( k \) is nonnegative, then (11.96) can be proven just as in the solution to Exercise 1.63(f). Hence, for the rest of this proof of (11.96), we assume WLOG that \( k \) is not nonnegative. Thus, \( k < 0 \), so that \( -k \) is nonnegative. Hence, Exercise 1.63(b) (applied to \( H \) and \( -k \) and \( \text{id}_H \) instead of \( A, f_i \) and \( \alpha \)) yields that \( \text{id}^k_H \circ \text{id}^k_H \times \cdots \times \text{id}^k_H \) is a \( \mathbf{k}\)-algebra homomorphism \( H \to H \). Since \( \text{id}^k_H \times \cdots \times \text{id}^k_H = \text{id}^{(-k)}_H \), this shows that \( \text{id}^{(-k)}_H \) is a \( \mathbf{k}\)-algebra homomorphism \( H \to H \). Thus, Exercise 1.63(c) (applied to \( A = H \) and \( f = \text{id}^{(-k)}_H \)) yields that \( \text{id}^{(-k)}_H \circ S : H \to H \) is again a \( \mathbf{k}\)-algebra homomorphism, and is a \( \ast \)-inverse to \( \text{id}^{(-k)}_H \).

Since \( \text{id}^{(-k)}_H \circ S \) is a \( \ast \)-inverse to \( \text{id}^{(-k)}_H \), we have \( \text{id}^{(-k)}_H \circ S = \left( \text{id}^{(-k)}_H \right)^{\ast} = \text{id}^{(-k)(-1)}_H = \text{id}^{(-k)}_H \). Hence, \( \text{id}^{(-k)}_H \) is a \( \mathbf{k}\)-algebra homomorphism (since \( \text{id}^{(-k)}_H \circ S \) is a \( \mathbf{k}\)-algebra homomorphism), and thus (11.96) is proven.

Proof of (11.97): Let \( k \in \mathbb{Z} \). We know that \( \text{id}^k_H \) is a \( \mathbf{k}\)-algebra homomorphism \( H \to H \) (according to (11.96)). Thus, Exercise 1.63(c) (applied to \( A = H \) and \( f = \text{id}^k_H \)) yields that \( \text{id}^k_H \circ S : H \to H \) is again a \( \mathbf{k}\)-algebra homomorphism, and is a \( \ast \)-inverse to \( \text{id}^k_H \).

Since \( \text{id}^k_H \circ S \) is a \( \ast \)-inverse to \( \text{id}^k_H \), we have \( \text{id}^k_H \circ S = \left( \text{id}^k_H \right)^{\ast} = \text{id}^{(k(-1))} = \text{id}^{(k(-1))} \). This proves (11.97).
If $H$ is a $k$-bialgebra and $C$ is a cocommutative $k$-coalgebra, and if $f_1, f_2, \ldots, f_k$ are several $k$-coalgebra homomorphisms $C \to H$, then prove that $f_1 \ast f_2 \ast \cdots \ast f_k$ also is a $k$-coalgebra homomorphism $C \to H$.

This can be solved by induction over $k$ in the same way as Exercise 1.63(b) (but now using the dual of Exercise 1.63(a) instead of Exercise 1.63(a) itself).

The dual of Exercise 1.63(c) is the following exercise:

If $H$ is a Hopf algebra and $C$ is a cocommutative $k$-coalgebra, and if $f : C \to H$ is a $k$-coalgebra homomorphism, then prove that $S \circ f : C \to H$ (where $S$ is the antipode of $H$) is again a $k$-coalgebra homomorphism, and is a $\ast$-inverse to $f$.

A solution of this can be obtained by reversing all arrows in the above solution of Exercise 1.63(c) (and using Exercise 1.47 in lieu of Proposition 1.32).

The dual of Exercise 1.63(d) is the following exercise:

If $C$ is a cocommutative $k$-coalgebra, then show that $\Delta^{(k)}$ is a $k$-coalgebra homomorphism for every $k \in \mathbb{N}$. (The map $\Delta^{(k)} : C \to C \otimes (k+1)$ is defined as in Exercise 1.40.)

This can be solved just as we solved Exercise 1.63(d), but again with all arrows reversed (and referring to Exercise 1.59(a) and Exercise 1.42(b) instead of Exercise 1.59(b) and Exercise 1.42(a), respectively).

The dual of Exercise 1.63(e) is Exercise 1.63(e) itself (up to renaming objects and maps).

The dual of Exercise 1.63(f) is the following exercise:

If $H$ is a cocommutative $k$-bialgebra, and $k$ and $\ell$ are two nonnegative integers, then prove that $\text{id}_H^f \circ \text{id}_H^k = \text{id}_H^{f \ast k}$.

This can be solved just as we solved Exercise 1.63(f), but again with all arrows reversed.

The dual of Exercise 1.63(g) is the following exercise:

If $H$ is a cocommutative $k$-Hopf algebra, and $k$ and $\ell$ are two integers, then prove that $\text{id}_H^f \circ \text{id}_H^k = \text{id}_H^{f \ast k}$.

This can be solved just as we solved Exercise 1.63(g), but again with all arrows reversed.

11.24. Solution to Exercise 1.65. Solution to Exercise 1.65. (a) Here is a messy computational solution:

Fix $a \in A$, and assume WLOG that $a$ is homogeneous of degree $n \in \mathbb{N}$. Let us prove that $(S \ast E)(a)$ is primitive.

In fact, using Sweedler’s notation, we can write $\Delta(a) = \sum_{(a)} a_1 \otimes a_2$ with all $a_1$ and $a_2$ homogeneous. Thus,

$$(S \ast E)(a) = \sum_{(a)} S(a_1) \cdot \underbrace{E(a_2)}_{=(\deg a_2) \cdot a_2} = \sum_{(a)} (\deg a_2) S(a_1) \cdot a_2.$$
Hence,

$$\Delta \left( (S \ast E) (a) \right) = \sum_{(a)} (\deg a_2) \Delta \left( S (a_1) \cdot a_2 \right) = \sum_{(a)} (\deg a_2) \sum_{(a_1) (a_2)} (S (a_1)_1) \cdot (a_2)_1 \otimes (S (a_1)_2) \cdot (a_2)_2$$

$$= \sum_{(a)} (\deg a_2) \sum_{(a_1) (a_2)} (S ((a_1)_2)) \cdot (a_2)_1 \otimes (S ((a_1)_1)) \cdot (a_2)_2$$

(here we used \((S (a_1)_1) \otimes (S (a_1)_2)) = S((a_1)_2) \otimes S((a_1)_1)\),
which is a consequence of Exercise 1.47)

$$= \sum_{(a)} (\deg a_2) \sum_{(a_1) (a_2)} \left( (\deg a_3 + \deg a_4) S (a_2) \right) \cdot S (a_1) \cdot a_3 \otimes S (a_1) \cdot a_4$$

$$= \sum_{(a)} (\deg a_2) (S (a_1) \cdot a_2 \otimes S (a_3)) \cdot a_4 + \sum_{(a)} (\deg a_4) (S (a_2) \cdot a_3 \otimes S (a_1) \cdot a_4)$$

(by the cocommutativity of \(A\))

$$= (S (a_1) \cdot a_2 \otimes 1 + 1 \otimes (S (a_1) \cdot a_2)$$

This (slightly unclean but easily formalizable) computation shows that \((S \ast E) (a)\) is primitive, and similarly the same can be shown for \((E \ast S) (a)\). This proves (a).

There is, however, a nicer proof: A coderivation of a \(k\)-coalgebra \((C, \Delta, \epsilon)\) is defined as a \(k\)-linear map \(F : C \rightarrow C\) such that \(\Delta \circ F = (F \otimes \text{id} + \text{id} \otimes F) \circ \Delta\). (The reader can check that this axiom is the result of writing the axiom for a derivation in element-free terms and reversing all arrows. Nothing less should be expected.) It is easy to see (by checking on each homogeneous component) that \(E\) is a coderivation. Hence, it will be enough to check that \((S \ast f) (a)\) and \((f \ast S) (a)\) are primitive whenever \(f : A \rightarrow A\) is a coderivation and \(a \in A\). So fix a coderivation \(f : A \rightarrow A\). Notice that the antipode \(S\) of \(A\) is a coalgebra anti-endomorphism (by Exercise 1.47), thus a coalgebra endomorphism (because coalgebra anti-endomorphisms of a cocommutative coalgebra are precisely the same as coalgebra endomorphisms). Moreover, \(\Delta : A \rightarrow A \otimes A\) is a coalgebra homomorphism\(^{317}\).

Applying (1.17) to \(A \otimes A\), \(A, A, \Delta, \text{id}_A, S\) and \(f\) instead of \(A', C, C', \alpha, \gamma, f\) and \(g\), we obtain

$$\Delta \circ (S \ast f) = \frac{\Delta \circ S}{(\Delta \circ f)} \ast \frac{\Delta \circ f}{(\Delta \circ S) \circ (S \ast f)} \ast \frac{\Delta \circ f}{(\Delta \circ S) \otimes (S \ast f)} \ast \frac{\Delta}{(\Delta \circ (S \ast f)) \circ \Delta}$$

$$= ((S \otimes S) \circ (f \otimes \text{id}) \circ \Delta) = (S \otimes S) \ast (f \otimes \text{id} \otimes f) \circ \Delta$$

by (1.17), applied to \(A \otimes A, A \otimes A, A \otimes A \otimes A, A \otimes A,\)

\(\text{id}, \Delta, S \otimes S\) and \(f \otimes \text{id} \otimes f\) instead of \(A', C, C', \alpha, \gamma, f\) and \(g\)

(11.98)

But Exercise 1.29(a) yields \((S \otimes S) \ast (f \otimes \text{id}) = (S \ast f) \otimes (S \ast f) \ast \text{id} \otimes (S \ast f) \ast \text{id}\) and similarly \((S \otimes S) \ast (\text{id} \otimes f) = \text{id} \otimes (S \ast f) \ast \text{id}\)

(11.98) becomes

$$\Delta \circ (S \ast f) = ((S \otimes S) \ast (f \otimes \text{id}) \circ \Delta + ((S \otimes S) \ast (\text{id} \otimes f)) \circ \Delta$$

$$= (S \ast f) \otimes \text{id} \ast \text{id} \otimes (S \ast f) \ast \text{id} \otimes (S \ast f) \circ \Delta$$

$$= (S \ast f) \otimes \text{id} \ast (S \ast f) \otimes \text{id} \ast (S \ast f) \circ \Delta.$$

\(^{317}\)In fact, Exercise 1.59(a) (applied to \(C = A\)) shows that \(A\) is cocommutative if and only if \(\Delta : A \rightarrow A \otimes A\) is a coalgebra homomorphism. But we know that \(A\) is cocommutative, so that \(\Delta : A \rightarrow A \otimes A\) is a coalgebra homomorphism.
Hence, every \( a \in A \) satisfies (using Sweedler’s notation)
\[
(\Delta \circ (S \ast f))(a) = (((S \ast f) \otimes \text{id}) \circ \Delta + (\text{id} \otimes (S \ast f)) \circ \Delta)(a)
= ((S \ast f) \otimes \text{id})(\Delta(a)) + (\text{id} \otimes (S \ast f))(\Delta(a))
= \sum_{(a)} (S \ast f)(a_1) \otimes (\text{id})(a_2) + \sum_{(a)} (\text{id})(a_1) \otimes (S \ast f)(a_2)
= \sum_{(a)} (S \ast f)(a_1) \epsilon \otimes (a_2) + \sum_{(a)} \epsilon(a_1) (S \ast f)(a_2)
= \sum_{(a)} = (S \ast f)(a_1) = (S \ast f)(a)
= (S \ast f)(a) \otimes 1 + 1 \otimes (S \ast f)(a).
\]

In other words, for every \( a \in A \), the element \((S \ast f)(a)\) is primitive. Similarly the same can be shown for \((f \ast S)(a)\), and so we are done.

(b) is a very simple computation. (Alternatively, the \((S \ast E)(p) = E(p)\) part follows from applying part (c) to \( a = 1 \), and similarly one can show \((E \ast S)(p) = E(p)\).)

(c) This is computational again: It is straightforward to check that \( E \) is a derivation of the algebra \( A \). Now,
\[
\Delta(\text{ap}) = \Delta(a) \Delta(p) = \left( \sum_{(a)} a_1 \otimes a_2 \right) (p \otimes 1 + 1 \otimes p)
= \sum_{(a)} a_1 p \otimes a_2 + \sum_{(a)} a_1 \otimes a_2 p,
\]
so that
\[
(S \ast E)(\text{ap}) = \sum_{(a)} S(a_1 p) E(a_2) + \sum_{(a)} S(a_1) E(a_2 p)
= \sum_{(a)} S(p) S(a_1) E(a_2) + \sum_{(a)} S(a_1) (E(a_2) p + a_2 E(p))
= -p \sum_{(a)} S(a_1) E(a_2) + \sum_{(a)} S(a_1) E(a_2 p) + \sum_{(a)} S(a_1) a_2 E(p)
= -p (S \ast E)(a_1) + (S \ast E)(a_2) + \sum_{(a)} \epsilon(a_1) E(p)
= [(S \ast E)(a_1), p] + \epsilon(a_1) E(p),
\]
thus proving part (c).

(d) Assume that \( A \) is connected and that \( Q \) is a subring of \( k \). Let \( B \) be the \( k \)-subalgebra of \( A \) generated by \( p \). In order to prove part (d), we need to show that \( A \subseteq B \). Clearly, \( p \subseteq B \).

Consider the grading \( A = \bigoplus_{n \geq 0} A_n \) of \( A \). Now, we need to prove that \( A \subseteq B \). In order to prove this, it is clearly enough to show that \( A_n \subseteq B \) for every \( n \in \mathbb{N} \). We will prove this by induction over \( n \). The induction base is obvious (in fact, \( A_0 = k \cdot 1_A \subseteq B \)), so we can proceed to the induction step. Let \( n \in \mathbb{N} \) be positive, and assume that we have shown that \( A_m \subseteq B \) for every nonnegative integer \( m < n \). We need to prove that \( A_n \subseteq B \).
Let \( a \in A_n \). We know that \((S \ast E)(a)\) is primitive (by part (a)), so that \((S \ast E)(a) \in \mathfrak{p} \subset B\). On the other hand, \( a \in A_n \) shows that

\[
\Delta(a) \in \Delta(A_n) \subset (A \otimes A)_n \quad \text{(since } \Delta \text{ is a graded map)}
\]

\[
= \sum_{i=0}^{n} A_i \otimes A_{n-i} = A_0 \otimes A_n + \sum_{i=1}^{n} A_i \otimes A_{n-i}.
\]

We can thus write \( \Delta(a) \) in the form \( \Delta(a) = u + v \) for some \( u \in A_0 \otimes A_n \) and \( v \in \sum_{i=1}^{n} A_i \otimes A_{n-i} \). Consider these \( u \) and \( v \).

We are first going to prove that \( u = 1 \otimes a \). Indeed, applying \( \epsilon \otimes \text{id} \) to the equality \( \Delta(a) = u + v \), we obtain \( (\epsilon \otimes \text{id}) (\Delta(a)) = (\epsilon \otimes \text{id})(u + v) = (\epsilon \otimes \text{id})(u) + (\epsilon \otimes \text{id})(v) \). But since \( (\epsilon \otimes \text{id})(\Delta(a)) = a \) (by the commutativity of the diagram (1.4)) and \( (\epsilon \otimes \text{id})(v) = 0 \) (because

\[
(\epsilon \otimes \text{id})(v) \in (\epsilon \otimes \text{id}) \left( \sum_{i=1}^{n} A_i \otimes A_{n-i} \right)
\]

\[
= \sum_{i=1}^{n} \epsilon(A_i) \otimes A_{n-i} = \sum_{i=1}^{n} 0 \otimes A_{n-i} = 0
\]

), this rewrites as \( a = (\epsilon \otimes \text{id})(u) + 0 \). In other words, \( a = (\epsilon \otimes \text{id})(u) \). But the element \( u \) has the form \( u = 1_A \otimes u' \) for some \( u' \in A_n \) (because \( u \in A_0 \otimes A_n = k \cdot 1_A \otimes A_n = 1_A \otimes A_n \)). This \( u' \in A_n \) can be recovered by \( u' = (\epsilon \otimes \text{id})(u) \) (since it satisfies \( (\epsilon \otimes \text{id}) \left( \begin{array}{c} u \\ =1_A \otimes u' \end{array} \right) = (\epsilon \otimes \text{id}) (1_A \otimes u') = (\epsilon (1_A) u') = u' \)), and thus simply equals \( a \) (since \( a = (\epsilon \otimes \text{id})(u) \)). Thus, \( u = 1_A \otimes a \).

Now,

\[
(S \ast E) (a) = (m \circ (S \otimes E) \circ \Delta)(a) = (m \circ (S \otimes E)) \left( \Delta(a) \right)_{=u+v}
\]

\[
= (m \circ (S \otimes E))(u + v) = (m \circ (S \otimes E))(u) + (m \circ (S \otimes E))(v).
\]

Since

\[
(m \circ (S \otimes E)) \left( \begin{array}{c} u \\ =1_A \otimes a \end{array} \right) = (m \circ (S \otimes E))(1_A \otimes a) = m \left( S(1_A) \otimes E(a) \right)_{=S(1_A) \otimes E(a)} = S(1_A) \otimes E(a)_{=1_A} = na
\]

\[
= m \left( S(1_A) \otimes E(a) \right)_{=1_A} = na_{=na} \quad \text{(since } a \in A_n \text{)}
\]


and

\[
(m \circ (S \otimes E)) \left( \sum_{i=1}^{n} A_i \otimes A_{n-i} \right)
\]

\[\in (m \circ (S \otimes E)) \left( \sum_{i=1}^{n} A_i \otimes A_{n-i} \right) = \sum_{i=1}^{n} (m \circ (S \otimes E)) (A_i \otimes A_{n-i})
\]

\[= \sum_{i=1}^{n} m ((S \otimes E) (A_i \otimes A_{n-i})) = \sum_{i=1}^{n} m (S (A_i) \otimes E (A_{n-i}))
\]

\[= \sum_{i=1}^{n-1} S (A_i) \cdot E (A_{n-i}) = \sum_{i=1}^{n-1} S (A_i) \cdot E (A_{n-i}) + S (A_n) \cdot E (A_0)
\]

(by the definition of \(E\))

\[\in \sum_{i=1}^{n-1} S (A_i) \cdot E (A_{n-i}) \subseteq A_i \otimes A_{n-i} \subseteq A_n \otimes E (A_{n-i})
\]

\[\subseteq \sum_{i=1}^{n-1} B \cdot B \subseteq B
\]

(since \(B \) is a subring of \(k\), we can divide this by \(n\) and obtain \(a \in B\).

We have thus shown that \(a \in B\) for every \(a \in A_n\). Hence, \(A_n \subseteq B\). This completes the induction step.

Thus, we know that \(A_n \subseteq B\) for every \(n \in \mathbb{N}\). Consequently, \(A \subseteq B\), and the proof of (d) is complete.

This solution of part (d) is not the most generalizable one – for instance, (d) also holds if \(A\) is connected filtered instead of connected graded, and then a different argument is necessary. This is a part of the Cartier-Milnor-Moore theorem, and appears e.g. in \([45, \S 3.2]\).

(e) If \(a \in T(V)\) is homogeneous of positive degree and \(p \in V\), then part (c) yields

\[
(S \ast E) (ap) = [(S \ast E) (a), p] + (S \ast E) (v) \in na + B,
\]

so that

\[
na \in (S \ast E) (a) + B \subseteq B + B \subseteq B.
\]

Since \(n\) is positive and \(\mathbb{Q}\) is a subring of \(k\), we can divide this by \(n\) and obtain \(a \in B\).

11.25. Solution to Exercise 1.67. Solution to Exercise 1.67. (a) Let \(u\) be the map \(\epsilon_C \circ s : k \rightarrow C^*\). Let \(m\) be the map \(\Delta_C^* \circ \rho_{C,C} : C^* \otimes C^* \rightarrow C^*\). Our goal is to prove that \(C^*\), endowed with \(m\) as the associative operation and \(u\) as the unity map, is a \(k\)-algebra. In order to achieve this, we need to prove that the diagrams (1.3) and (1.4) with \(A\), \(m\) and \(u\) replaced by \(C^*\), \(m\) and \(u\) are commutative.
Let us first prove that the diagram (1.3) with $A$, $m$ and $u$ replaced by $C^*$, $m$ and $u$ is commutative. In other words, let us prove that the diagram

\begin{equation}
\begin{array}{c}
C^* \\
\downarrow m \quad \downarrow m \quad \downarrow m \\
C^* \\
\end{array}
\end{equation}

is commutative.

Indeed, consider the diagram

\begin{equation}
\begin{array}{c}
C^* \\
\downarrow m \quad \downarrow m \quad \downarrow m \\
C^* \\
\end{array}
\end{equation}

We are going to show that this diagram is commutative. In order to do so, we will show that its little squares and triangles are commutative.

The triangle

\begin{equation}
\begin{array}{c}
C^* \\
\downarrow m \quad \downarrow m \quad \downarrow m \\
C^* \\
\end{array}
\end{equation}

is commutative, since $m = \Delta_C^* \circ \rho_{C,C}$. For the same reason, the triangle

\begin{equation}
\begin{array}{c}
C^* \\
\downarrow m \quad \downarrow m \quad \downarrow m \\
C^* \\
\end{array}
\end{equation}
is commutative. The commutativity of the triangles

\[
\begin{align*}
&\xymatrix{C^* \otimes C^* \ar[rr]^{\rho_{C,C}} \ar[dr]_{m} && (C \otimes C)^* \ar[dr]^{\Delta_C} & \\
& & C^* \ar[rr]^{\Delta_C} & \ar@{..>}[ur]_m & \ar@{..>}[ur]_m & \ar@{..>}[ur]_m \\
& \xymatrix{C^* \otimes C^* \ar[rr]^{\rho_{C,C}} \ar[dr]_{m} && (C \otimes C)^* \ar[dr]^{\Delta_C} & \\
& & C^* \ar[rr]^{\Delta_C} & \ar@{..>}[ur]_m & \ar@{..>}[ur]_m & \ar@{..>}[ur]_m 
}\end{align*}
\]

and

\[
\begin{align*}
&\xymatrix{C^* \otimes C^* \ar[rr]^{\rho_{C,C}} \ar[dr]_{m} && (C \otimes C)^* \ar[dr]^{\Delta_C} & \\
& & C^* \ar[rr]^{\Delta_C} & \ar@{..>}[ur]_m & \ar@{..>}[ur]_m & \ar@{..>}[ur]_m \\
& \xymatrix{C^* \otimes C^* \ar[rr]^{\rho_{C,C}} \ar[dr]_{m} && (C \otimes C)^* \ar[dr]^{\Delta_C} & \\
& & C^* \ar[rr]^{\Delta_C} & \ar@{..>}[ur]_m & \ar@{..>}[ur]_m & \ar@{..>}[ur]_m 
}\end{align*}
\]

also clearly follows from \(m = \Delta_C^* \circ \rho_{C,C}\). The commutativity of the square

\[
\begin{align*}
&\xymatrix{C^* \otimes C^* \otimes C^* \ar[rr]^{\rho_{C,C} \otimes \text{id}} \ar[dr]_{\text{id} \otimes \rho_{C,C}} && (C \otimes C^* \otimes C)^* \ar[dr]^{\rho_{C,C} \otimes \Delta_C} & \\
& & (C \otimes C^*)^* \ar[rr]^{\rho_{C,C} \otimes \Delta_C} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C}} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C}} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C}} \\
& \xymatrix{C^* \otimes C^* \otimes C^* \ar[rr]^{\rho_{C,C} \otimes \text{id}} \ar[dr]_{\text{id} \otimes \rho_{C,C}} && (C \otimes C^*)^* \ar[dr]^{\rho_{C,C} \otimes \Delta_C} & \\
& & (C \otimes C^*)^* \ar[rr]^{\rho_{C,C} \otimes \Delta_C} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C}} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C}} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C}} 
}\end{align*}
\]

is a basic linear-algebraic fact. The commutativity of the square

\[
\begin{align*}
&\xymatrix{(C \otimes C)^* \otimes C^* \ar[rr]^{\rho_{C,C} \otimes \text{id}} \ar[dr]_{\text{id} \otimes \rho_{C,C}} && C^* \otimes (C \otimes C)^* \ar[dr]^{\rho_{C,C} \otimes \Delta_C} & \\
& & (C \otimes C)^* \ar[rr]^{\rho_{C,C} \otimes \Delta_C} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C}} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C}} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C}} \\
& \xymatrix{(C \otimes C)^* \otimes C^* \ar[rr]^{\rho_{C,C} \otimes \text{id}} \ar[dr]_{\text{id} \otimes \rho_{C,C}} && C^* \otimes (C \otimes C)^* \ar[dr]^{\rho_{C,C} \otimes \Delta_C} & \\
& & (C \otimes C)^* \ar[rr]^{\rho_{C,C} \otimes \Delta_C} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C}} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C}} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C}} 
}\end{align*}
\]

is a particular case of the following linear-algebraic fact: If \(X, Y\) and \(B\) are three \(k\)-modules and \(f : Y \to X\) is a \(k\)-linear map, then the diagram

\[
\begin{align*}
&\xymatrix{X^* \otimes B^* \ar[rr]^{\rho_{X,B}} \ar[dr]_{\rho_{Y,B}} && (X \otimes B)^* \ar[dr]^{(f \otimes \text{id})^*} & \\
& & (Y \otimes B)^* \ar[rr]^{(f \otimes \text{id})^*} & \ar@{..>}[ur]_{(f \otimes \text{id})^*} & \ar@{..>}[ur]_{(f \otimes \text{id})^*} & \ar@{..>}[ur]_{(f \otimes \text{id})^*} \\
& \xymatrix{X^* \otimes B^* \ar[rr]^{\rho_{X,B}} \ar[dr]_{\rho_{Y,B}} && (X \otimes B)^* \ar[dr]^{(f \otimes \text{id})^*} & \\
& & (Y \otimes B)^* \ar[rr]^{(f \otimes \text{id})^*} & \ar@{..>}[ur]_{(f \otimes \text{id})^*} & \ar@{..>}[ur]_{(f \otimes \text{id})^*} & \ar@{..>}[ur]_{(f \otimes \text{id})^*} 
}\end{align*}
\]

is commutative.\(^{318}\) Similarly, the commutativity of the square

\[
\begin{align*}
&\xymatrix{(C \otimes C) \otimes (C \otimes C)^* \ar[rr]^{\rho_{C,C} \otimes (\text{id} \otimes \Delta_C)} \ar[dr]_{\text{id} \otimes \rho_{C,C} \otimes \Delta_C} && (C \otimes C)^* \ar[dr]^{\rho_{C,C} \otimes \Delta_C} & \\
& & (C \otimes C)^* \ar[rr]^{\rho_{C,C} \otimes \Delta_C} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C} \otimes \Delta_C} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C} \otimes \Delta_C} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C} \otimes \Delta_C} \\
& \xymatrix{(C \otimes C) \otimes (C \otimes C)^* \ar[rr]^{\rho_{C,C} \otimes (\text{id} \otimes \Delta_C)} \ar[dr]_{\text{id} \otimes \rho_{C,C} \otimes \Delta_C} && (C \otimes C)^* \ar[dr]^{\rho_{C,C} \otimes \Delta_C} & \\
& & (C \otimes C)^* \ar[rr]^{\rho_{C,C} \otimes \Delta_C} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C} \otimes \Delta_C} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C} \otimes \Delta_C} & \ar@{..>}[ur]_{\text{id} \otimes \rho_{C,C} \otimes \Delta_C} 
}\end{align*}
\]

\(^{318}\)This is part of the reason why \(\rho_{U,V}\) is functorial in \(U\).
can be shown. Finally, the square

\[
\begin{array}{ccc}
(C \otimes C \otimes C)^* & \xymatrix{ & (C \otimes C)^* \ar[dl] \ar[dr] \ar[d]}
\end{array}
\]

is commutative, because it is obtained by dualizing the commutative diagram (1.3).

We thus have shown that every little square and every little triangle of the diagram (11.100) is commutative. Hence, the whole diagram (11.100) is commutative. In particular, the square formed by the four long curved arrows in (11.100) is commutative. But this square is precisely the diagram (11.99). Hence, we have shown that the diagram (11.99) is commutative. In other words, the diagram (1.3) with \( A, m \) and \( u \) replaced by \( C^*, m \) and \( u \) is commutative.

It now remains to prove that the diagram (1.4) with \( A, m \) and \( u \) replaced by \( C^*, m \) and \( u \) is commutative. In other words, it remains to prove that the diagram

\[
\begin{array}{ccc}
\xymatrix{ C^* \otimes k & C^* \otimes C^* & k \otimes C^* \ar[ll] \ar[lll]}
\end{array}
\]

is commutative. We will only prove the commutativity of the left square of this diagram (since the right square is analogous). That is, we will only prove the commutativity of the square

(11.101)

\[
\begin{array}{ccc}
\xymatrix{ C^* \otimes k & C^* \ar[ll] \ar[lll]}
\end{array}
\]

Indeed, the proof is similar to our proof of (11.99), but instead of the big diagram (11.100) we now have the diagram

(11.102)

in which the arrow \( C^* \to (C \otimes k)^* \) is the adjoint map of the canonical isomorphism \( C \otimes k \to C \) (and in which we identify \( k^* \) with \( k \)). The commutativity of the little triangles and squares is again easily proven (the square

\[
\begin{array}{ccc}
\xymatrix{ (C \otimes k)^* & C^* \ar[dl] \ar[dr] \ar[d]}
\end{array}
\]

(11.101)
is commutative by virtue of being the dual of the left square in (1.4). The commutativity of the “2-gon”

\[
\begin{array}{c}
C^* \\
\downarrow \text{id}^* \\
C^*
\end{array}
\]

simply says that id* = id, which is obvious. Thus, everything in (11.102) commutes. By following the “outer quadrilateral” of (11.102), we obtain precisely the commutativity of (11.101). This completes our solution of Exercise 1.67 (a).

Remark: Our solution was not the simplest one (by far). We could have saved much work by doing certain abuses of notation (such as identifying s with the identity map). We could also solve part (a) very easily if we had solved part (b) first. In the above solution, we have avoided all such shortcuts.

(b) Let \( C \) be a \( k \)-coalgebra. Let us notice that

\[
\rho_{C,C} (f \otimes g) = m_k \circ (f \otimes g) \quad \text{for all } f \in C^* \text{ and } g \in C^*.
\]

Now, we want to prove that the \( k \)-algebra structure defined on \( C^* \) in part (a) is precisely the one defined on \( \text{Hom}(C,k) = C^* \) in Definition 1.27 applied to \( A = k \). In order to do this, it is clearly enough to show that the product of any two elements of \( C^* \) with respect to the former \( k \)-algebra structure equals their product with respect to the latter \( k \)-algebra structure. In other words, it is enough to prove that for any \( f \in C^* \) and \( g \in C^* \), we have

\[
(\text{the product } fg \text{ with respect to the } k \text{-algebra } C^* \text{ defined in part (a)})
\]

\[
= (\text{the product } fg \text{ with respect to the } k \text{-algebra } \text{Hom}(C,k) \text{ defined in Definition 1.27 applied to } A = k).
\]

But this follows from the following computation:

\[
\begin{align*}
&= m_{C^*} (f \otimes g) = (\Delta_C \circ \rho_{C,C}) (f \otimes g) = \Delta_C \left( \rho_{C,C} (f \otimes g) \right) \\
&= m_C \circ (f \otimes g) = m_k \circ (f \otimes g) \circ \Delta_C \quad \text{(by the definition of } \Delta_C) \\
&= f \ast g \quad \text{(since } f \ast g = m_k \circ (f \otimes g) \circ \Delta_C \text{ (by the definition of } f \ast g)) \\
&= (\text{the product } fg \text{ with respect to the } k \text{-algebra } \text{Hom}(C,k) \text{ defined in Definition 1.27 applied to } A = k).
\end{align*}
\]

Thus, part (b) of the exercise is solved.

(c) Let \( C \) be a graded \( k \)-coalgebra. Let \( C = \bigoplus_{n \geq 0} C_n \) be its decomposition into homogeneous components. Then, for every \( n \in \mathbb{N} \), we identify \( (C_n)^* \) with a \( k \)-submodule of \( C^* \), namely with the \( k \)-submodule \( \{ f \in C^* \mid f(C_p) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq n \} \). Thus,

\[
(C_n)^* = \{ f \in C^* \mid f(C_p) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq n \}
\]

for every \( n \in \mathbb{N} \).

By the definition of \( C^n \), we have \( C^n = \bigoplus_{n \geq 0} (C_n)^* = \sum_{n \geq 0} (C_n)^* \).

Now let \( a \) and \( b \) be two nonnegative integers. We shall show that \( (C_n)^* (C_b)^* \subset (C_{a+b})^* \).

Indeed, let \( x \in (C_n)^* \) and \( y \in (C_b)^* \) be arbitrary. We have

\[
\begin{align*}
x \in (C_n)^* = \{ f \in C^* \mid f(C_p) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq a \}
\end{align*}
\]

(by (11.104), applied to \( n = a \)). In other words, \( x \) is an element of \( C^* \) such that

\[
x(C_p) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq a.
\]

\footnote{Proof of (11.103): Let \( f \in C^* \) and \( g \in C^* \). Then, the definition of \( \rho_{C,C} \) shows that \( \rho_{C,C} (f \otimes g) \) is the composition \( C \otimes C \xrightarrow{f \otimes g} k \otimes k \xrightarrow{m_k} k \) of the map \( f \otimes g \) with the canonical isomorphism \( k \otimes k \xrightarrow{m_k} k \). In other words, \( \rho_{C,C} (f \otimes g) = m_k(f \otimes g) \). This proves (11.103).}
Similarly, $y$ is an element of $C^*$ such that
\begin{equation}
y(C_p) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq b.
\end{equation}

Also, \eqref{eq:11.104} (applied to $n = a + b$) yields
\begin{equation}
(C_{a+b})^* = \{ f \in C^* \mid f(C_p) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq a + b \}.
\end{equation}

Now, part (b) of this exercise yields that the $k$-algebra structure defined on $C^*$ in part (a) is precisely the one defined on $\text{Hom}(C, k) = C^*$ in Definition 1.27 applied to $A = k$. Thus, the product of the $k$-algebra $C^*$ is precisely the convolution product $\ast$ on $\text{Hom}(C, k) = C^*$. Hence,
\begin{equation}
xy = x \ast y = m_k \circ (x \otimes y) \circ \Delta_C.
\end{equation}

Now, fix $p \in \mathbb{N}$ such that $p \neq a + b$. Then, $\Delta_C(C_p) \subset \sum_{q=0}^{p} C_q \otimes C_{p-q}$ (since the coalgebra $C$ is graded). Now, every $q \in \{0, 1, \ldots, p\}$ satisfies
\begin{equation}
(x \otimes y)(C_{q} \otimes C_{p-q}) = 0.
\end{equation}

\textbf{(Proof of \eqref{eq:11.108}):} Let $q \in \{0, 1, \ldots, p\}$. Then we must have either $q \neq a$ or $p - q \neq b$ (or both), because otherwise we would have $p = \underbrace{q + (p - q)}_{=a} = \underbrace{a + b}$ which would contradict $p \neq a + b$. If $q \neq a$, then $x(C_q) = 0$ (by \eqref{eq:11.105}) and therefore $(x \otimes y)(C_q \otimes C_{p-q}) = x(C_q) \otimes y(C_{p-q}) = 0$. If $p - q \neq b$, then \eqref{eq:11.106} yields $y(C_{p-q}) = 0$, and thus $(x \otimes y)(C_q \otimes C_{p-q}) = x(C_q) \otimes y(C_{p-q}) = 0$. Hence, in either case, we have $(x \otimes y)(C_q \otimes C_{p-q}) = 0$, and so \eqref{eq:11.108} is proven.

Now,
\begin{equation}
(C_p) = \left( m_k \circ (x \otimes y) \circ \Delta_C \right)(C_p) = m_k \left( (x \otimes y) \left( \sum_{q=0}^{p} C_q \otimes C_{p-q} \right) \right) \subset \sum_{q=0}^{p} m_k \left( (x \otimes y)(C_q \otimes C_{p-q}) \right) = \sum_{q=0}^{p} m_k (0) = 0.
\end{equation}

In other words, $(xy)(C_p) = 0$.

Now, forget that we fixed $x$. We thus have shown that $(xy)(C_p) = 0$ for all $p \in \mathbb{N}$ satisfying $p \neq a + b$. In other words,
\begin{equation}
xy \in \{ f \in C^* \mid f(C_p) = 0 \text{ for all } p \in \mathbb{N} \text{ satisfying } p \neq a + b \} = (C_{a+b})^*
\end{equation}
(by \eqref{eq:11.107}).

Now, forget that we fixed $x$ and $y$. We thus have shown that $xy \in (C_{a+b})^*$ for all $x \in (C_a)^*$ and $y \in (C_b)^*$. This yields $(C_a)^*(C_b)^* \subset (C_{a+b})^*$ (since $(C_{a+b})^*$ is a $k$-module).

But this has been proven for all $a \in \mathbb{N}$ and $b \in \mathbb{N}$. From this, it is easy to conclude that $C^o C^o \subset C^o$ (since $C^o = \sum_{n \geq 0} (C_n)^*$). Hence, the $k$-submodule $C^o$ of $C^*$ is closed under multiplication. Since we also have $1_{C^*} \in C^o$ (in fact, it is very easy to see that $1_{C^*} = \epsilon_C \in (C_0)^* \subset \sum_{n \geq 0} (C_n)^* = C^o$), this shows that $C^o$ is a $k$-subalgebra of $C^*$. This solves part (c) of the exercise.

(d) Let $C$ and $D$ be two $k$-coalgebras. Let $f : C \to D$ be a homomorphism of $k$-coalgebras. Then, the two diagrams
\begin{equation}
\begin{array}{ccc}
C & \xrightarrow{f} & D \\
\Delta_C & \downarrow & \Delta_D \ \\
C \otimes C & \xrightarrow{f \otimes f} & D \otimes D
\end{array}
\end{equation}
and
\begin{equation}
\begin{array}{ccc}
C & \xrightarrow{f} & D \\
\epsilon_C & \downarrow & \epsilon_D \ \\
\k & \xrightarrow{1} & \k
\end{array}
\end{equation}
are commutative (since $f$ is a homomorphism of $\mathbf{k}$-coalgebras).

We need to prove that $f^* : D^* \to C^*$ is a homomorphism of $\mathbf{k}$-algebras. In other words, we need to prove that the two diagrams

\[
\begin{array}{ccc}
D^* & \xrightarrow{f^*} & C^* \\
m_{D^*} & & m_{C^*} \\
D^* \otimes D^* & \xrightarrow{f^* \otimes f^*} & C^* \otimes C^*
\end{array}
\quad \text{and} \quad
\begin{array}{ccc}
D^* & \xrightarrow{f^*} & C^* \\
u_{D^*} & & u_{C^*} \\
D^* \otimes D^* & \xrightarrow{f^* \otimes f^*} & C^* \otimes C^*
\end{array}
\]

are commutative.

By the definition of the $\mathbf{k}$-algebra $D^*$, we have $m_{D^*} = \Delta_D \circ \rho_{D,D}$. Thus, the left triangle (with two vertical edges and one curved edge) in the diagram

\[
\begin{array}{ccc}
D^* & \xrightarrow{f^*} & C^* \\
m_{D^*} & & m_{C^*} \\
(D \otimes D)^* & \xrightarrow{(f \otimes f)^*} & (C \otimes C)^*
\end{array}
\]

are commutative. Similarly, the right triangle in (11.111) also commutes. The upper rectangle commutes because it is obtained from the first diagram in (11.109) by dualizing. The lower triangle in (11.111) commutes by basic linear algebra. Hence, the whole diagram (11.111) commutes. But the outer rim of the diagram (11.111) is exactly the first diagram in (11.110). Thus, the first diagram in (11.110) commutes. The (even simpler) task of proving the commutativity of the second diagram in (11.110) is left to the reader. Thus, $f^*$ is a homomorphism of $\mathbf{k}$-algebras. Part (d) of the exercise is solved.

(c) This is an exercise in linear algebra (it has nothing to do with Hopf algebras). Here is a rough sketch of how it is solved: Let $U = \bigoplus_{n \geq 0} U_n$ be the decomposition of $U$ into its homogeneous components, and let $V = \bigoplus_{n \geq 0} V_n$ be the decomposition of $V$ into its homogeneous components. Notice that every $V_n$ is finite free (since $V$ is of finite type). Graded $\mathbf{k}$-linear maps $U \to V$ can be regarded as elements of $\prod_{n \geq 0} \text{Hom} (U_n, V_n)$ (because a graded $\mathbf{k}$-linear map $U \to V$ restricts to a $\mathbf{k}$-linear map $U_n \to V_n$ for every $n \in \mathbb{N}$, and is uniquely determined by the totality of these restrictions). Similarly, graded $\mathbf{k}$-linear maps $V^o \to U^o$ can be regarded as elements of $\prod_{n \geq 0} \text{Hom} ((V_n)^*, (U_n)^*)$. For every $n \in \mathbb{N}$, there is a canonical $\mathbf{k}$-module isomorphism $\text{Hom} (U_n, V_n) \to \text{Hom} ((V_n)^*, (U_n)^*)$ which sends every $\varphi \in \text{Hom} (U_n, V_n)$ to $\varphi^* \in \text{Hom} ((V_n)^*, (U_n)^*)$ (since $V_n$ is finite free). Taking the product of these isomorphisms, we obtain a $\mathbf{k}$-module isomorphism from $\prod_{n \geq 0} \text{Hom} (U_n, V_n)$ to $\prod_{n \geq 0} \text{Hom} ((V_n)^*, (U_n)^*)$, that is, a $\mathbf{k}$-module isomorphism from the $\mathbf{k}$-module of all graded $\mathbf{k}$-linear maps $U \to V$ to the $\mathbf{k}$-module of all graded $\mathbf{k}$-linear maps $V^o \to U^o$. It is easy to see that this isomorphism sends every $f : U \to V$ to $f^* : V^o \to U^o$. Hence, there is a 1-to-1 correspondence between graded $\mathbf{k}$-linear maps $U \to V$ and graded $\mathbf{k}$-linear maps $V^o \to U^o$ given by $f \mapsto f^*$ (namely, this isomorphism). This solves Exercise 1.67 (e).

(f) Let $f : C \to D$ be a graded $\mathbf{k}$-linear map. We need to prove that $f : C \to D$ is a $\mathbf{k}$-coalgebra map if and only if $f^* : D^o \to C^o$ is a $\mathbf{k}$-algebra map. In other words, we need to prove the following two assertions:

Assertion 1: If $f : C \to D$ is a $\mathbf{k}$-coalgebra map, then $f^* : D^o \to C^o$ is a $\mathbf{k}$-algebra map.

Assertion 2: If $f^* : D^o \to C^o$ is a $\mathbf{k}$-algebra map, then $f : C \to D$ is a $\mathbf{k}$-coalgebra map.

We start by proving Assertion 1. One way to prove it proceeds by repeating the solution of Exercise 1.67 (d), except that $D^*$, $(D \otimes D)^*$, $C^*$ and $(C \otimes C)^*$ are replaced by $D^o$, $(D \otimes D)^o$, $C^o$ and $(C \otimes C)^o$ (where, of course, the map $\rho_{D,D} : D^o \to (D \otimes D)^o$ has to be interpreted as the restriction of the map $\rho_{D,D} : D^* \to (D \otimes D)^*$ to the $\mathbf{k}$-submodule $D^o$ of $D^*$, and similarly for the other maps). This replacement can be done completely robotically, and thus is left to the reader. Another way to prove Assertion 1 is by realizing that it follows immediately from Exercise 1.67 (d) (since $f^* : D^o \to C^o$ is a restriction of the map $f^* : D^* \to C^*$). Either way, we do not end up using the assumption that $D$ is of finite type. However, we will need this assumption in our proof of Assertion 2.
Now, let us prove Assertion 2. Assume that \( f^*: D^o \rightarrow C^o \) is a \( k \)-algebra map. We want to show that \( f: C \rightarrow D \) is a \( k \)-coalgebra map. In other words, we want to show that the two diagrams (11.109) commute. Let us start with the left one of these diagrams.

The graded \( k \)-module \( D \) is of finite type, and therefore the map \( \rho_{D,D}^*: D^o \otimes D^* \rightarrow (D \otimes D)^o \) (a restriction of the map \( \rho_{D,D}^* : D^* \otimes D^* \rightarrow (D \otimes D)^* \)) is an isomorphism. Its inverse \( \rho_{D,D}^{-1} : (D \otimes D)^o \rightarrow D^o \otimes D^o \) is therefore well-defined. We can thus form the (asymmetric!) diagram

\[
\begin{array}{ccc}
D^o & \xrightarrow{f^*} & C^o \\
\Delta^*_D & \swarrow m_{D^*} & \searrow m_{C^*}
\end{array}
\]

(The arrows labelled \( m_{C^*} \) and \( m_{D^*} \) could just as well have been labelled \( m_{C^o} \) and \( m_{D^o} \), since the multiplication maps \( m_{C^o} \) and \( m_{D^o} \) are restrictions of \( m_{C^*} \) and \( m_{D^*} \).) The two triangles in (11.112) commute due to \( m_{D^*} = \Delta^*_D \circ \rho_{D,D}^* \) and \( m_{C^*} = \Delta^*_C \circ \rho_{C,C}^* \). The upper quadrilateral in (11.112) commutes because \( f^* \) is a \( k \)-algebra homomorphism, and the lower quadrilateral in (11.112) commutes because of the commutativity of the diagram

\[
\begin{array}{ccc}
D^o \otimes D^o & \xrightarrow{f^* \otimes f^*} & C^o \otimes C^o \\
\downarrow \rho_{D,D} & & \downarrow \rho_{C,C} \\
(D \otimes D)^o & \xrightarrow{(f \otimes f)^*} & (C \otimes C)^o
\end{array}
\]

(this follows from standard linear algebra). Hence, the whole diagram (11.112) commutes. Removing the two interior nodes of this diagram, we obtain the commutative diagram

\[
\begin{array}{ccc}
D^o & \xrightarrow{f^*} & C^o \\
\Delta^*_D & \swarrow & \Delta^*_C \\
(D \otimes D)^o & \xrightarrow{(f \otimes f)^*} & (C \otimes C)^o
\end{array}
\]

This does not immediately yield the commutativity of the first diagram in (11.109) (because we cannot revert taking dual \( k \)-modules), so we are not yet done. But we are close.

We have

\[
(\Delta_D \circ f)^* = f^* \circ \Delta^*_D = \Delta^*_C \circ (f \otimes f)^* \quad \text{(since the diagram (11.113) commutes)}
\]

as maps from \( (D \otimes D)^o \) to \( C^o \). But a general linear-algebraic fact states that if \( U \) and \( V \) are two graded \( k \)-modules such that \( V \) is of finite type, and if \( \alpha \) and \( \beta \) are two graded \( k \)-linear maps \( U \rightarrow V \) such that \( \alpha^* = \beta^* \) as maps from \( V^o \) to \( U^o \), then \( \alpha = \beta \). Applying this to \( U = C, V = D \otimes D, \alpha = \Delta_D \circ f \) and \( \beta = (f \otimes f) \circ \Delta_C \), we obtain \( \Delta_D \circ f = (f \otimes f) \circ \Delta_C \) (since \( D \otimes D \) is of finite type) and since \( (\Delta_D \circ f)^* = ((f \otimes f) \circ \Delta_C)^* \). In other words, the first diagram in (11.109) is commutative. The reader can verify that so is the second diagram in (11.109) (once again, this is the easier part). So we have shown that both diagrams in (11.109) are commutative, and thus \( f \) is a \( k \)-coalgebra map. This proves Assertion 2.

Now that Assertions 1 and 2 are both proven, part (f) of the exercise is solved.

\[\text{20} \text{Beware: we don’t have an inverse of the non-restricted map } \rho_{D,D}^*: D^* \otimes D^* \rightarrow (D \otimes D)^*.\]

\[\text{21} \text{This follows immediately from Exercise 1.67 (e).}\]

\[\text{22} \text{This is because } D \text{ is of finite type.}\]
11.26. Solution to Exercise 1.70. Solution to Exercise 1.70. We have Sym$(V) \cong k[x]$. Thus, $(x^k)_{k \in \mathbb{N}}$ is a basis of the $k$-module Sym$(V)$. Therefore, $(x^k \otimes x^\ell)_{(k,\ell) \in \mathbb{N}^2}$ is a basis of the $k$-module Sym$(V) \otimes$ Sym$(V)$.

On the other hand, $(f^{(i)})_{i \in \mathbb{N}}$ is a basis of the $k$-module $(\text{Sym}(V))^o$ (namely, the dual basis to the graded basis $(x^k)_{k \in \mathbb{N}}$ of Sym$(V)$). Hence, $(f^{(i)} \otimes f^{(\ell)})_{(i,\ell) \in \mathbb{N}^2}$ is a basis of the $k$-module $(\text{Sym}(V))^o \otimes (\text{Sym}(V))^o$.

We denote by $m$ the multiplication map of Sym$(V)$. We denote by $u$ the unity map of Sym$(V)$ (that is, the map $k \to \text{Sym}(V)$ which sends $1_k$ to $1_{\text{Sym}(V)}$). We denote by $\Delta$ the comultiplication of Sym$(V)$. We denote by $\epsilon$ the counit of Sym$(V)$. We denote by $S$ the antipode of Sym$(V)$.

We denote by $m(\text{Sym}(V))^o$, $u(\text{Sym}(V))^o$, $\Delta(\text{Sym}(V))^o$, $\epsilon(\text{Sym}(V))^o$ and $S(\text{Sym}(V))^o$ the maps analogous to $m$, $u$, $\Delta$, $\epsilon$ and $S$ but defined for the Hopf algebra $(\text{Sym}(V))^o$ instead of Sym$(V)$.

(a) Clearly, $(x^i \cdot x^j = x^{i+j}$ for all $i \in \mathbb{N}$ and $j \in \mathbb{N}$.

We have $\Delta_{T(V)} (x) = 1 \otimes x + x \otimes 1$ in $T(V) \otimes T(V)$ (by the definition of $\Delta_{T(V)}$). Projecting this equality down onto Sym$(V) \otimes$ Sym$(V)$, we obtain $\Delta_{\text{Sym}(V)} (x) = x \otimes 1 + 1 \otimes x$ (since Sym$(V)$ is a quotient Hopf algebra of $T(V)$). But we defined $\Delta$ to mean the comultiplication of Sym$(V)$. We thus have $\Delta = \Delta_{\text{Sym}(V)}$, so that $\Delta (x) = \Delta_{\text{Sym}(V)} (x) = x \otimes 1 + 1 \otimes x$. Thus, $x$ is a primitive element of Sym$(V)$. Proposition 1.37 thus yields $S (x) = -x$.

Proposition 1.32 yields that the antipode $S$ of Sym$(V)$ is an algebra anti-endomorphism of Sym$(V)$. Since an algebra anti-endomorphism of Sym$(V)$ means the same as an algebra endomorphism of Sym$(V)$ (because Sym$(V)$ is commutative), this yields that $S$ is an algebra endomorphism of Sym$(V)$. Consequently, for any $n \in \mathbb{N}$, we have

$$S (x^n) = \left( \frac{S (x)}{x \to -x} \right)^n = \left( -x \right)^n = (-1)^n x^n. \quad (11.14)$$

It remains to prove that $\Delta (x^n) = \sum_{i+j=n} \binom{n}{i} x^i \otimes x^j$ for every $n \in \mathbb{N}$ (where the summation sign “$\sum_{i+j=n}$” is shorthand for “$\sum_{(i,j) \in \mathbb{N}^2}$”). Let $n \in \mathbb{N}$. The $k$-algebra Sym$(V) \otimes$ Sym$(V)$ is commutative (since Sym$(V)$ is commutative), and thus the binomial formula can be applied in it. We know that Sym$(V)$ is a bialgebra, and thus $\Delta$ is a $k$-algebra homomorphism (by the axioms of a bialgebra). Hence,

$$\Delta (x^n) = \left( \Delta (x) \right)^n \left. \frac{\Delta (x)}{x \to 1 + x} \right|_{x \to 1 + x} = \sum_{\ell=0}^{n} \binom{n}{\ell} \frac{(x \otimes 1)^\ell (1 \otimes x)^{n-\ell}}{x^\ell \otimes 1} \frac{1 \otimes x^{n-\ell}}{1 \otimes x^{n-\ell}} = \sum_{\ell=0}^{n} \binom{n}{\ell} x^\ell \otimes x^{n-\ell} = \sum_{i+j=n} \binom{n}{i} x^i \otimes x^j \quad (11.15)$$

(here, we have substituted $(i, j)$ for $(\ell, n - \ell)$ in the sum). This completes the solution of Exercise 1.70(a).

(b) The definition of $f^{(i)}$ shows that

$$f^{(i)} (x^j) = \delta_{i,j} \quad \text{for all } i \in \mathbb{N} \text{ and } j \in \mathbb{N}. \quad (11.16)$$

But the definition of the $k$-algebra $(\text{Sym}(V))^o$ shows that the multiplication map $m(\text{Sym}(V))^o$ of the $k$-algebra $(\text{Sym}(V))^o$ is adjoint to the comultiplication map $\Delta$ of the $k$-coalgebra Sym$(V)$. In other words, we have

$$m(\text{Sym}(V))^o (a) \cdot b)_{\text{Sym}(V)} = (a \otimes \Delta (b))_{\text{Sym}(V) \otimes \text{Sym}(V)} \quad (11.17)$$

for any $a \in (\text{Sym}(V))^o \otimes (\text{Sym}(V))^o$ and any $b \in \text{Sym}(V)$. Thus, any $p \in (\text{Sym}(V))^o$, $q \in (\text{Sym}(V))^o$ and $b \in \text{Sym}(V)$ satisfy

$$(pq, b)_{\text{Sym}(V)} = (p \otimes q, \Delta (b))_{\text{Sym}(V) \otimes \text{Sym}(V)} \quad (11.18)$$
On the other hand, the definition of the $k$-coalgebra $(\text{Sym}(V))^o$ shows that the comultiplication map $\Delta(\text{Sym}(V))^o$ of the $k$-coalgebra $(\text{Sym}(V))^o$ is adjoint to the multiplication map $m$ of the $k$-algebra $\text{Sym}(V)$. In other words, we have

\begin{equation}
(\Delta(\text{Sym}(V))^o(a),b)_{\text{Sym}(V)\otimes \text{Sym}(V)} = (a,m(b))_{\text{Sym}(V)} \tag{11.119}
\end{equation}

for any $a \in (\text{Sym}(V))^o$ and any $b \in \text{Sym}(V) \otimes \text{Sym}(V)$. Thus, any $a \in (\text{Sym}(V))^o$, $p \in \text{Sym}(V)$ and $q \in \text{Sym}(V)$ satisfy

\begin{equation}
(\Delta(\text{Sym}(V))^o(a),p \otimes q)_{\text{Sym}(V)\otimes \text{Sym}(V)} = (a,pq)_{\text{Sym}(V)} \tag{11.120}
\end{equation}

The definition of the $k$-Hopf algebra $(\text{Sym}(V))^o$ shows that the antipode $S(\text{Sym}(V))^o$ of the $k$-Hopf algebra $(\text{Sym}(V))^o$ is adjoint to the antipode $S$ of the $k$-Hopf algebra $\text{Sym}(V)$. In other words, we have

\begin{equation}
(S(\text{Sym}(V))^o(a),b)_{\text{Sym}(V)} = (a,S(b))_{\text{Sym}(V)} \tag{11.121}
\end{equation}

for any $a \in (\text{Sym}(V))^o$ and $b \in \text{Sym}(V)$.

Let us also notice a simple fact about sums: If $u$, $v$ and $n$ are three nonnegative integers, if $\mathfrak{A}$ is an additive abelian group, and if $(t_{i,j})_{(i,j) \in \mathbb{N}^2}$ is a family of elements of $\mathfrak{A}$, then

\begin{equation}
\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j} = t_{u,v} \delta_{u+v,n} \tag{11.122}
\end{equation}

(where, again, “$\sum_{i+j=n}$” is shorthand for “$\sum_{(i,j) \in \mathbb{N}^2}$”).

Now, we have

\begin{equation}
f(u)f(v) = \begin{pmatrix} u+v \\ u \end{pmatrix} f(u+v) \quad \text{for every } u \in \mathbb{N} \text{ and } v \in \mathbb{N}. \tag{11.123}
\end{equation}

---

**Proof of (11.118):** Let $p \in (\text{Sym}(V))^o$, $q \in (\text{Sym}(V))^o$ and $b \in \text{Sym}(V)$. Then, $m(\text{Sym}(V))^o(p \otimes q) = pq$ (since $m(\text{Sym}(V))^o$ is the multiplication map of the algebra $(\text{Sym}(V))^o$). Thus,

\[
\begin{pmatrix} pq \\ m(\text{Sym}(V))^o(p \otimes q) \end{pmatrix} = \begin{pmatrix} m(\text{Sym}(V))^o(p \otimes q),b \end{pmatrix}_{\text{Sym}(V)} = (p \otimes q,\Delta(b))_{\text{Sym}(V)\otimes \text{Sym}(V)}
\]

(by (11.17), applied to $a = p \otimes q$). This proves (11.118).

**Proof of (11.120):** Let $a \in (\text{Sym}(V))^o$, $p \in \text{Sym}(V)$ and $q \in \text{Sym}(V)$. Then, $m(p \otimes q) = pq$ (since $m$ is the multiplication map of the algebra $\text{Sym}(V)$). But (11.119) (applied to $b = p \otimes q$) yields

\[
(\Delta(\text{Sym}(V))^o(a),p \otimes q)_{\text{Sym}(V)\otimes \text{Sym}(V)} = \begin{pmatrix} a,m(p \otimes q) \end{pmatrix}_{\text{Sym}(V)} = (a,pq)_{\text{Sym}(V)}.
\]

This proves (11.120).

**Proof of (11.122):** Let $u$, $v$ and $n$ be three nonnegative integers. Let $\mathfrak{A}$ be an additive abelian group. Let $(t_{i,j})_{(i,j) \in \mathbb{N}^2}$ be a family of elements of $\mathfrak{A}$.

Let us first assume that $u + v \neq n$. Each addend of the sum $\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j}$ contains the factor $\delta_{u,i} \delta_{v,j}$, which is 0 unless we have both $u = i$ and $v = j$. Thus, each addend of this sum vanishes unless it satisfies both $u = i$ and $v = j$ at the same time. Since no addend of this sum can satisfy both $u = i$ and $v = j$ at the same time (because such an addend would then also satisfy $u + v = i + j = n$, which would contradict $u + v \neq n$), this shows that each addend of this sum vanishes.

Consequently, the sum itself must vanish. That is, we have $\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j} = 0$. Compared with $t_{u,v} \delta_{u+v,n} = 0$ (since $u+v \neq n$), this yields $\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j} = t_{u,v} \delta_{u+v,n}$. Hence, (11.122) is proven under the assumption that $u + v \neq n$.

Therefore, for the rest of our proof of (11.122), we can WLOG assume that we don’t have $u + v \neq n$. Thus, $u + v = n$. Again, each addend of the sum $\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j}$ contains the factor $\delta_{u,i} \delta_{v,j}$, which is 0 unless we have both $u = i$ and $v = j$. Thus, each addend of this sum vanishes unless it satisfies both $u = i$ and $v = j$ at the same time. But there exists exactly one addend of the sum $\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j}$ which satisfies both $u = i$ and $v = j$: namely, the addend for $(i,j) = (u,v)$ (and this is indeed an addend because we have $u + v = n$). This addend equals $t_{u,v} \delta_{u+v,n} = t_{u,v}$. Because of this, and because all other
Comparing this with (\(t_{addends of our sum vanish, we thus conclude that the whole sum must equal 266\)) DARIJ GRINBERG AND VICTOR REINER

Let \(k \in \mathbb{N}\). Then,

\[
\left( f(u) f(v) \right) \left( x^k \right) = \left( f(u) f(v), x^k \right)_{\text{Sym}(V)} = \left( f(u) \otimes f(v), \Delta \left( x^k \right) \right)_{\text{Sym}(V) \otimes \text{Sym}(V)}
\]

(by (11.118), applied to \(p = f(u), q = f(v)\) and \(b = x^k\))

\[
= \left( f(u) \otimes f(v), \sum_{i+j=k} \binom{k}{i} x^i \otimes x^j \right)_{\text{Sym}(V) \otimes \text{Sym}(V)}
\]

\[
= \sum_{i+j=k} \binom{k}{i} \left( f(u) \otimes f(v), x^i \otimes x^j \right)_{\text{Sym}(V) \otimes \text{Sym}(V)}
\]

(by the definition of the bilinear form \(\langle \cdot, \cdot \rangle_{\text{Sym}(V) \otimes \text{Sym}(V)}\))

\[
= \sum_{i+j=k} \binom{k}{i} \delta_{u,i} \delta_{v,j} = \binom{k}{u} \delta_{u+v,k}
\]

(by (11.122), applied to \(n = k, A = k\) and \(i,j = \binom{k}{i}\))

\[
= \left( \frac{u + v}{u} \right) \delta_{u+v,k}
\]

because the equality \(\binom{k}{u} \delta_{u+v,k} = \binom{u + v}{u} \delta_{u+v,k}\) holds in the case when \(u + v = k\) (obviously) and in the case when \(u + v \neq k\) (since both sides of this equality are 0 in this case (due to \(\delta_{u+v,k} = 0\))

Comparing this with

\[
\left( \left( \frac{u + v}{u} \right) f(u+v) \right) \left( x^k \right) = \left( \frac{u + v}{u} \right) \left( f(u+v) \left( x^k \right) \right) = \left( \frac{u + v}{u} \right) \delta_{u+v,k},
\]

(by (11.116), applied to \(u+v\) and \(k\) instead of \(i\) and \(j\))

we obtain \(f(u) f(v) \left( x^k \right) = \left( \left( \frac{u + v}{u} \right) f(u+v) \right) \left( x^k \right).

Let us now forget that we fixed \(k\). We thus have shown that \(f(u) f(v) \left( x^k \right) = \left( \left( \frac{u + v}{u} \right) f(u+v) \right) \left( x^k \right)\) for every \(k \in \mathbb{N}\). In other words, the two \(k\)-linear maps \(f(u) f(v)\) and \(\left( \frac{u + v}{u} \right) f(u+v)\) are equal to each other on the basis \((x^k)_{k \in \mathbb{N}}\) of the \(k\)-module \(\text{Sym}(V)\). Therefore, these two \(k\)-linear maps must be identical (because if two \(k\)-linear maps from one and the same domain are equal to each other on a basis of this domain, then addends of our sum vanish, we thus conclude that the whole sum must equal \(t_{u,v}\). That is, we have \(\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j} = t_{u,v}\).

Compared with \(t_{u,v} = t_{u,v}\), this yields \(\sum_{i+j=n} t_{i,j} \delta_{u,i} \delta_{v,j} = t_{u,v} \delta_{u+v,n}\). Hence, (11.122) is proven.
these $k$-linear map maps must be identical). In other words, $f(u)f(v) = \left(\begin{array}{c} u + v \\ u \end{array}\right)f(u+v)$. Thus, (11.123) is proven.

Thus, we have shown that

$$f^{(i)}f^{(j)} = \left(\begin{array}{c} i+j \\ i \end{array}\right)f^{(i+j)}$$

for every $i \in \mathbb{N}$ and $j \in \mathbb{N}$.

(In fact, (11.124) follows from (11.123) by renaming the variables $u$ and $v$ as $i$ and $j$.)

We shall now show that

$$\Delta((\text{Sym}(V))^o)(f^{(n)}) = \sum_{i+j=n} f^{(i)} \otimes f^{(j)}$$

for every $n \in \mathbb{N}$

(where, again, “$\sum_{i+j=n}$” is shorthand for “$\sum_{(i,j) \in \mathbb{N}^2}$.”)

Proof of (11.125): Let $n \in \mathbb{N}$. We identify $(\text{Sym}(V))^o \otimes (\text{Sym}(V))^o$ with $(\text{Sym}(V) \otimes \text{Sym}(V))^o$ (since $\text{Sym}(V)$ is a graded $k$-module of finite type); thus, an element of $(\text{Sym}(V))^o \otimes (\text{Sym}(V))^o$ can be regarded as a $k$-linear map $\text{Sym}(V) \otimes \text{Sym}(V) \rightarrow k$. In particular, $\Delta((\text{Sym}(V))^o)(f^{(n)})$ thus becomes a $k$-linear map $\text{Sym}(V) \otimes \text{Sym}(V) \rightarrow k$.

Fix $(k, \ell) \in \mathbb{N}^2$. Thus, $k \in \mathbb{N}$ and $\ell \in \mathbb{N}$. Then,

$$\left(\Delta((\text{Sym}(V))^o)(f^{(n)})\right)(x^k \otimes x^\ell) = \left(\Delta((\text{Sym}(V))^o)(f^{(n)})\right)(x^{k+\ell})_{\text{Sym}(V)}$$

(by (11.120), applied to $a = f^{(n)}$, $p = x^k$ and $q = x^\ell$)

$$= \sum_{i+j=n} f^{(i)}(x^k) \otimes f^{(j)}(x^\ell)$$

(by (11.126), applied to $i = n$ and $j = k + \ell$).

Compared with

$$\left(\sum_{i+j=n} f^{(i)} \otimes f^{(j)}\right)(x^k \otimes x^\ell) = \sum_{i+j=n} f^{(i)}x^k \otimes f^{(j)}x^\ell$$

$$= \sum_{i+j=n} \delta_{i,k} \delta_{j,\ell} = \sum_{i+j=n} \delta_{k,i} \delta_{\ell,j} = \sum_{i+j=n} 1 \delta_{k,i} \delta_{\ell,j} = \sum_{i+j=n} \delta_{k+i,\ell+j} = \delta_{n,k+\ell}$$

(by (11.122), applied to $k$, $\ell$, $k$ and $1$ instead of $u$, $v$, $\mathfrak{A}$ and $t_{i,j}$)

this yields

$$\left(\Delta((\text{Sym}(V))^o)(f^{(n)})\right)(x^k \otimes x^\ell) = \left(\sum_{i+j=n} f^{(i)} \otimes f^{(j)}\right)(x^k \otimes x^\ell).$$

Let us now forget that we fixed $(k, \ell)$. We thus have shown that (11.126) holds for every $(k, \ell) \in \mathbb{N}^2$. In other words, the two $k$-linear maps $\Delta((\text{Sym}(V))^o)(f^{(n)})$ and $\sum_{i+j=n} f^{(i)} \otimes f^{(j)}$ (from $\text{Sym}(V) \otimes \text{Sym}(V)$ to $k$) are equal to each other on the basis $(x^k \otimes x^\ell)_{(k,\ell) \in \mathbb{N}^2}$ of the $k$-module $\text{Sym}(V) \otimes \text{Sym}(V)$. Therefore, these two $k$-linear maps must be identical (because if two $k$-linear maps from one and the same domain are equal to each other on a basis of this domain, then these $k$-linear map maps must be identical). In other words, $\Delta((\text{Sym}(V))^o)(f^{(n)}) = \sum_{i+j=n} f^{(i)} \otimes f^{(j)}$. Thus, (11.125) is proven.

Finally, let us show that

$$S((\text{Sym}(V))^o)(f^{(n)}) = (-1)^n f^{(n)}$$

for every $n \in \mathbb{N}$. 
Proof of (11.127): Let \( n \in \mathbb{N} \). Let \( k \in \mathbb{N} \). Then,

\[
\left( S_{(\text{Sym}(V))^n} \left( f^{(n)} \right) \right) (x^k) = \left( S_{(\text{Sym}(V))^n} \left( f^{(n)} \right), x^k \right)_{\text{Sym}(V)} = \begin{pmatrix} f^{(n)} & S (x^k) \end{pmatrix} = (-1)^k x^k
\]

(by (11.114), applied to \( k \) instead of \( n \))

\[
\text{Sym}(V)
\]

\[
= (f^{(n)}, (-1)^k x^k)_{\text{Sym}(V)} = f^{(n)} \left( (-1)^k x^k \right)
\]

\[
= (-1)^k f^{(n)} (x^k)
\]

(by (11.116), applied to \( i = n \) and \( j = k \))

\[
= (-1)^k \delta_{n,k}.
\]

Since \((-1)^k \delta_{n,k} = (-1)^n \delta_{n,k}\), this becomes

\[
\left( S_{(\text{Sym}(V))^n} \left( f^{(n)} \right) \right) (x^k) = (-1)^n \delta_{n,k}.
\]

Compared with

\[
\left( (-1)^n f^{(n)} \right) (x^k) = (-1)^n f^{(n)} (x^k) = (-1)^n \delta_{n,k},
\]

(this yields \( S_{(\text{Sym}(V))^n} \left( f^{(n)} \right) (x^k) = (-1)^n f^{(n)} (x^k) \).

Let us now forget that we fixed \( k \). We thus have shown that \( S_{(\text{Sym}(V))^n} \left( f^{(n)} \right) (x^k) = (-1)^n f^{(n)} (x^k) \) for every \( k \in \mathbb{N} \). In other words, the two \( k \)-linear maps \( S_{(\text{Sym}(V))^n} \left( f^{(n)} \right) \) and \((-1)^n f^{(n)}\) are equal to each other on the basis \( (x^k)_{k \in \mathbb{N}} \) of the \( k \)-module \( \text{Sym}(V) \). Therefore, these two \( k \)-linear maps must be identical (because if two \( k \)-linear maps from one and the same domain are equal to each other on a basis of this domain, then these \( k \)-linear map maps must be identical). In other words, \( S_{(\text{Sym}(V))^n} \left( f^{(n)} \right) = (-1)^n f^{(n)} \).

Thus, (11.127) is proven.

Now, all of the identities (11.124), (11.125) and (11.127) are proven; thus, Exercise 1.70(b) is solved.

Before we start the solution of Exercise 1.70(c), let us make a few more simple observations.

We have

\[
\epsilon (x^k) = \delta_{k,0} \quad \text{for every } k \in \mathbb{N}.
\]

(11.128)

\[327\]

The definition of the \( k \)-algebra \( (\text{Sym}(V))^n \) shows that the unity map \( u_{(\text{Sym}(V))} \) of the \( k \)-algebra \( (\text{Sym}(V))^n \) is adjoint to the counit map \( \epsilon \) of the \( k \)-coalgebra \( \text{Sym}(V) \). In other words, we have

\[
\left( u_{(\text{Sym}(V))} ^{n} \left( a \right), b \right)_{\text{Sym}(V)} = (a, \epsilon (b))_{k}
\]

for any \( a \in k \) and any \( b \in \text{Sym}(V) \). Thus,

\[
u_{(\text{Sym}(V))}^{n} \left( a \right) = af^{(0)} \quad \text{for every } a \in k.
\]

(11.130)

\[327\]
The definition of the \(k\)-coalgebra \((\text{Sym}(V))^o\) shows that the counit map \(\epsilon_{\text{Sym}(V)}\) of the \(k\)-algebra \((\text{Sym}(V))^o\) is adjoint to the unity map \(u\) of the \(k\)-algebra \(\text{Sym}(V)\). In other words, we have

\[
(\epsilon_{\text{Sym}(V)})(a), b) = (a, u(b))_{\text{Sym}(V)}
\]

for any \(a \in \text{Sym}(V)\) and any \(b \in k\). Thus,

\[
\epsilon_{\text{Sym}(V)}(f^{(k)}) = \delta_{k,0} \quad \text{for every } k \in \mathbb{N}.
\]

(c) For the time being, let us not assume that \(\mathbb{Q}\) is a subring of \(k\). Instead, we shall introduce a map and prove some of its properties which hold for every \(k\).

We define a \(k\)-linear map \(\Phi : \text{Sym}(V) \rightarrow (\text{Sym}(V))^o\) by

\[
(\Phi(x^k) = k!f^{(k)} \quad \text{for every } k \in \mathbb{N}).
\]

---

328 Proof of (11.130): Let \(a \in k\). Then, every \(b \in \text{Sym}(V)\) satisfies

\[
(u_{\text{Sym}(V)})(a) \cdot (b) = (u_{\text{Sym}(V)})(a), b)_{\text{Sym}(V)} = (a, \epsilon(b))_k \quad \text{(by (11.129))}
\]

\[
= a \cdot \epsilon(b) \quad \text{(by the definition of the form } (\cdot, \cdot)_k). \tag{11.131}
\]

Now, every \(k \in \mathbb{N}\) satisfies

\[
(u_{\text{Sym}(V)})(a) \left(x^k\right) = a \cdot \epsilon\left(x^k\right) \quad \text{(by (11.131), applied to } b = x^k)
\]

\[
= a \cdot \delta_{k,0} \quad \text{(by (11.128))}
\]

\[
= a \cdot f^{(0)}(x^k) \quad \text{(since (11.116) (applied to } i=0 \text{ and } j=k)}
\]

In other words, the two \(k\)-linear maps \(u_{\text{Sym}(V)}(a)\) and \(af^{(0)}\) are equal to each other on the basis \((x^k)_{k \in \mathbb{N}}\) of the \(k\)-module \(\text{Sym}(V)\). Therefore, these two \(k\)-linear maps must be identical (because if two \(k\)-linear maps from one and the same domain are equal to each other on a basis of this domain, then these \(k\)-linear maps must be identical). In other words, \(u_{\text{Sym}(V)}(a) = af^{(0)}\). Thus, (11.130) is proven.

329 Proof of (11.133): Let \(k \in \mathbb{N}\). The map \(u\) is the unity map of the \(k\)-algebra \(\text{Sym}(V)\); thus, \(u(1) = 1_{\text{Sym}(V)} = x^0\).

The definition of the bilinear form \((\cdot, \cdot)_k\) yields \(\epsilon_{\text{Sym}(V)}(f^{(k)}), 1) = \epsilon_{\text{Sym}(V)}(f^{(k)}) \cdot 1 = \epsilon_{\text{Sym}(V)}(f^{(k)}) = \epsilon_{\text{Sym}(V)}(f^{(k)}), \) so that

\[
\epsilon_{\text{Sym}(V)}(f^{(k)}) = \epsilon_{\text{Sym}(V)}(f^{(k)}), 1)_{k} = \left(f^{(k)}, u(1)\right)_{\text{Sym}(V)} = \left(f^{(k)} - x^0\right)_{\text{Sym}(V)} \quad \text{(by (11.132), applied to } a = f^{(k)} \text{ and } b = 1)
\]

\[
= \left(f^{(k)}, x^0\right)_{\text{Sym}(V)} = \delta_{k,0} \quad \text{(by (11.116), applied to } i = k \text{ and } j = 0).}
\]

This proves (11.133).
(This is well-defined, since $(x^k)_k \in \mathbb{N}$ is a basis of the $k$-module $\text{Sym} (V)$.) We can now see that this $k$-linear map $\Phi$ satisfies $\Phi \circ m = m_{(\text{Sym} (V))} \circ (\Phi \otimes \Phi)$ \hfill \text{330} and $\Phi \circ u = u_{(\text{Sym} (V))}$ \hfill \text{331}. Hence, $\Phi$ is a $k$-algebra

\textbf{Proof.} Fix $(k, \ell) \in \mathbb{N}^2$. Recall that $m$ is the multiplication map of the $k$-algebra $\text{Sym} (V)$. Thus, $m(x^k \otimes x^\ell) = x^k x^\ell = x^{k+\ell}$. Also, $m_{(\text{Sym} (V))}$ is the multiplication map of the $k$-algebra $(\text{Sym} (V))$. Thus, $m_{(\text{Sym} (V))} (f(k) \otimes f(\ell)) = f(k)f(\ell) = \binom{k+\ell}{k} f(k+\ell)$ (by (11.123), applied to $u = k$ and $v = \ell$). Now,

$$\Phi \circ m (x^k \otimes x^\ell) = \Phi \left( m_{(\text{Sym} (V))} (x^k \otimes x^\ell) \right) = \Phi (x^{k+\ell}) = (k+\ell)! f(k+\ell) \quad \text{by (11.134), applied to } k + \ell \text{ instead of } k.$$

Compared with

$$m_{(\text{Sym} (V))} \circ (\Phi \otimes \Phi) (x^k \otimes x^\ell) = m_{(\text{Sym} (V))} \left( \frac{\Phi (x^k) \otimes \Phi (x^\ell)}{\Phi (x^k) \otimes \Phi (x^\ell)} \right) = m_{(\text{Sym} (V))} \left( \frac{\Phi (x^k) \otimes \Phi (x^\ell)}{k! \ell f(k) \otimes f(\ell)} \right) \quad \text{by (11.134)},$$

$$= \frac{k! \ell! m_{(\text{Sym} (V))} (f(k) \otimes f(\ell))}{(k+\ell)! f(k+\ell)} \quad \text{since the map } m_{(\text{Sym} (V))} \text{ is } k\text{-linear}$$

$$= \frac{k! \ell! \binom{k+\ell}{k} f(k+\ell)}{(k+\ell)! f(k+\ell)} \quad \text{by (11.133)}$$

this yields $(\Phi \circ m) (x^k \otimes x^\ell) = m_{(\text{Sym} (V))} \circ (\Phi \otimes \Phi) (x^k \otimes x^\ell)$.

Let us now forget that we fixed $(k, \ell)$. We thus have shown that $(\Phi \circ m) (x^k \otimes x^\ell) = m_{(\text{Sym} (V))} \circ (\Phi \otimes \Phi) (x^k \otimes x^\ell)$ for every $(k, \ell) \in \mathbb{N}^2$. In other words, the two $k$-linear maps $\Phi \circ m$ and $m_{(\text{Sym} (V))} \circ (\Phi \otimes \Phi)$ (from $\text{Sym} (V)$ to $\text{Sym} (V)$) are equal to each other on the basis $(x^k \otimes x^\ell)_R \in \mathbb{N}^2$ of the $k$-module $\text{Sym} (V)$. Therefore, these two $k$-linear maps must be identical (because if two $k$-linear maps from one and the same domain are equal to each other on a basis of this domain, then these $k$-linear map maps must be identical). In other words, $\Phi \circ m = m_{(\text{Sym} (V))} \circ (\Phi \otimes \Phi)$.Qed.

\textbf{Proof.} The map $u$ is the unity map of the $k$-algebra $\text{Sym} (V)$; thus, $u (1) = 1_{\text{Sym} (V)} = x^0$. Now, every $a \in k$ satisfies

$$\Phi (u_1) \left( \frac{a}{a} \right) = \Phi (u) (a \cdot 1) = a \cdot (\Phi \circ u) (1) \quad \text{since the map } \Phi \circ u \text{ is } k\text{-linear}$$

$$= a \cdot \Phi (u_1) \left( \frac{a}{a} \right) = a \cdot \frac{\Phi (x^0)}{x^0} = a \cdot 0! f(0) = a f(0) = u_{(\text{Sym} (V))} (a) \quad \text{by (11.130)}.$$

Thus, $\Phi \circ u = u_{(\text{Sym} (V))}$, qed.
homomorphism. Also, the \( k \)-linear map \( \Phi \) satisfies \((\Phi \otimes \Phi) \circ \Delta = \Delta_{(\text{Sym}(V))^o} \circ \Phi \) \(^{332}\) and \( \epsilon = \epsilon_{(\text{Sym}(V))^o} \circ \Phi \) \(^{333}\). Hence, \( \Phi \) is a \( k \)-coalgebra homomorphism. Thus, \( \Phi \) is a \( k \)-bialgebra homomorphism (since \( \Phi \) is both a \( k \)-algebra homomorphism and a \( k \)-coalgebra homomorphism), and therefore a Hopf algebra homomorphism (according to Proposition 1.46(c), applied to \( H_1 = \text{Sym}(V) \), \( H_2 = (\text{Sym}(V))^o \), \( S_1 = S \) and \( S_2 = S_{(\text{Sym}(V))^o} \)).

\(^{332}\)Proof. Let \( k \in \mathbb{N} \). We have

\[
\begin{align*}
((\Phi \otimes \Phi) \circ \Delta)(x^k) &= (\Phi \otimes \Phi) \left( \sum_{i+j=k} \binom{k}{i} x^i \otimes x^j \right) \\
&= \sum_{i+j=k} \binom{k}{i} \Phi(x^i) \otimes \Phi(x^j) \\
&= \sum_{i+j=k} \binom{k}{i} i! f^{(i)} \otimes j! f^{(j)} \\
&= \sum_{i+j=k} \binom{k}{i} \frac{k!}{i!(k-i)!} f^{(i)} \otimes f^{(j)} \\
&= \sum_{i+j=k} \binom{k}{i} \delta_{i,j} = k! \sum_{i+j=k} f^{(i)} \otimes f^{(j)}. \\
\end{align*}
\]

Compared with

\[
\begin{align*}
\Delta_{(\text{Sym}(V))^o} \circ \Phi)(x^k) &= \Delta_{(\text{Sym}(V))^o} \left( \sum_{i+j=k} \binom{k}{i} \Phi(x^i) \otimes x^j \right) \\
&= \sum_{i+j=k} \binom{k}{i} x^i \otimes x^j \\
&= \sum_{i+j=k} \binom{k}{i} x^i \otimes x^j \\
&= \sum_{i+j=k} \binom{k}{i} \frac{k!}{i!(k-i)!} f^{(i)} \otimes f^{(j)}. \\
\end{align*}
\]

this yields \((\Phi \otimes \Phi) \circ \Delta)(x^k) = \Delta_{(\text{Sym}(V))^o} \circ \Phi)(x^k)\).

Let us now forget that we fixed \( k \). We thus have shown that \((\Phi \otimes \Phi) \circ \Delta)(x^k) = \Delta_{(\text{Sym}(V))^o} \circ \Phi)(x^k)\) for every \( k \in \mathbb{N} \). In other words, the two \( k \)-linear maps \((\Phi \otimes \Phi) \circ \Delta\) and \( \Delta_{(\text{Sym}(V))^o} \circ \Phi\) are equal to each other on the basis \( x^k \in \mathbb{N} \) of the \( k \)-module Sym \((V)\). Therefore, these two \( k \)-linear maps must be identical (because if two \( k \)-linear maps from one and the same domain are equal to each other on a basis of this domain, then these \( k \)-linear map must be identical). In other words, \((\Phi \otimes \Phi) \circ \Delta = \Delta_{(\text{Sym}(V))^o} \circ \Phi\). Qed.

\(^{333}\)Proof. Let \( k \in \mathbb{N} \). We have

\[
\begin{align*}
(\epsilon_{(\text{Sym}(V))^o} \circ \Phi)(x^k) &= \epsilon_{(\text{Sym}(V))^o} \left( \sum_{i+j=k} \binom{k}{i} \Phi(x^i) \otimes x^j \right) \\
&= \sum_{i+j=k} \binom{k}{i} \epsilon_{(\text{Sym}(V))^o}(x^i) \otimes x^j \\
&= \sum_{i+j=k} \delta_{i,j} \epsilon_{(\text{Sym}(V))^o}(x^i) \otimes x^j \\
&= \sum_{i+j=k} \delta_{i,j} \epsilon_{(\text{Sym}(V))^o}(x^i) \otimes x^j \\
&= \sum_{i+j=k} \delta_{i,j} \epsilon_{(\text{Sym}(V))^o}(x^i) \otimes x^j. \\
\end{align*}
\]

But it is easy (by treating the cases \( k = 0 \) and \( k \neq 0 \) separately) to see that \( k! \delta_{k,0} = \delta_{k,0} \). Hence,

\[
(\epsilon_{(\text{Sym}(V))^o} \circ \Phi)(x^k) = k! \delta_{k,0} = \delta_{k,0} = \epsilon(x^k) \quad \text{(by (11.128))}.
\]

In other words, \( \epsilon(x^k) = (\epsilon_{(\text{Sym}(V))^o} \circ \Phi)(x^k) \).
The map $\Phi$ is also graded. Thus, $\Phi$ is a homomorphism of graded $k$-Hopf algebras (since $\Phi$ is graded and a Hopf algebra homomorphism).

Now, let us assume that $Q$ is a subring of $k$. We define a $k$-linear map $\Psi : (\text{Sym} (V))^\circ \to \text{Sym} (V)$ by

$$
(11.135) 
\left( \Psi \left( f^{(k)} \right) \right) = \frac{x^k}{k!} \quad \text{for every } k \in \mathbb{N}.
$$

(This is well-defined, since $(f^{(k)})_{k \in \mathbb{N}}$ is a basis of the $k$-module $(\text{Sym} (V))^\circ$.) It is clear that the maps $\Phi$ and $\Psi$ are mutually inverse. Thus, the map $\Phi$ is invertible, and its inverse is $\Psi$.

The map $\Phi$ is an invertible homomorphism of graded $k$-Hopf algebras and therefore an isomorphism of graded $k$-Hopf algebras (since every invertible homomorphism of graded $k$-Hopf algebras is an isomorphism of graded $k$-Hopf algebras). Hence, the inverse of $\Phi$ is also an isomorphism of graded $k$-Hopf algebras. In other words, $\Psi$ is an isomorphism of graded $k$-Hopf algebras (since the inverse of $\Phi$ is $\Psi$).

But every $n \in \mathbb{N}$ satisfies $\Psi (f^{(n)}) = \frac{x^n}{n!}$ (according to (11.135), applied to $k = n$). Hence, $\Psi$ is the $k$-linear map $\text{Sym} (V) \to \text{Sym} (V)$ sending $f^{(n)} \mapsto \frac{x^n}{n!}$. Thus, the $k$-linear map $\text{Sym} (V) \to \text{Sym} (V)$ sending $f^{(n)} \mapsto \frac{x^n}{n!}$ is an isomorphism of graded $k$-Hopf algebras (since $\Psi$ is an isomorphism of graded $k$-Hopf algebras). This solves Exercise 1.70(c).

(d) For the time being, let us not assume that $Q$ is a subring of $k$. Instead, let us first prove a formula which does not require any restrictions on $k$.

Namely, let us show that

$$
(11.136) 
\left( f^{(1)} \right)^m = m! f^{(m)} \quad \text{for every } m \in \mathbb{N}.
$$

Proof of (11.136): Let us consider the $k$-linear map $\Phi : \text{Sym} (V) \to (\text{Sym} (V))^\circ$ defined in our solution to Exercise 1.70(c) above. Then, $\Phi$ is a $k$-algebra homomorphism. (In fact, this was proven in our solution to Exercise 1.70(c) above.) Applying (11.134) to $k = 1$, we obtain $\Phi (x^1) = \frac{1!}{1!} f^{(1)} = f^{(1)}$, so that

$$
\Phi (\text{Sym} (V)) = \Phi (k \cdot x^n) \subset k \cdot f^{(n)}. 
$$

(by (11.134), applied to $k = n$)

$$
= k \cdot n! f^{(n)} \subset k \cdot f^{(n)} = (\text{Sym} (V))^\circ_n \quad \text{(since } ((\text{Sym} (V))^\circ_n) = k \cdot f^{(n)}). 
$$

Now, let us forget that we fixed $k$. We thus have shown that $\Phi (\text{Sym} (V)) \subset ((\text{Sym} (V))^\circ_n)$ for every $n \in \mathbb{N}$. In other words, the map $\Phi$ is graded, qed.

Indeed, the map $\Phi \circ \Psi$ sends every $f^{(k)}$ to $f^{(k)}$ and thus is the identity map, whereas the map $\Psi \circ \Phi$ sends every $x^k$ to $x^k$ and therefore is the identity map as well.
\[ f^{(1)} = \Phi (x^1) = \Phi (x). \] Thus, for every \( m \in \mathbb{N} \), we have
\[ \left( \frac{f^{(1)}}{\Phi(x)} \right)^m = \Phi (x^m) = m!f^{(m)} \quad \text{(since } \Phi \text{ is a } k\)-algebra homomorphism) \]

This proves (11.136).

Now, let us assume that \( k \) is a field of characteristic \( p > 0 \). Then, \( p \cdot 1_k = 0 \). But
\[ (p - 1)!p \cdot 1_k = 0. \]
Applying (11.136) to \( m = p \), we obtain
\[ \left( f^{(1)} \right)^p = p!f^{(p)} = p! \cdot 1_k \cdot f^{(p)} = 0. \]

Now, it remains to show that there can be no Hopf isomorphism \((\text{Sym}(V))^\circ \to \text{Sym}(V)\). In fact, we can prove something stronger: Namely, there exists no algebra isomorphism \((\text{Sym}(V))^\circ \to \text{Sym}(V)\). This is because the algebra \((\text{Sym}(V))^\circ\) has a nonzero nilpotent element (namely, \( f^{(1)} \) is nonzero and satisfies \( (f^{(1)})^p = 0 \)), whereas the algebra \(\text{Sym}(V)\) contains no nonzero nilpotent elements (in fact, \(\text{Sym}(V) \cong k[x]\) is isomorphic to a polynomial ring over the field \(k\), and therefore is an integral domain, which yields that it contains no nonzero nilpotent elements). This concludes the solution of Exercise 1.70(d).

11.27. **Solution to Exercise 1.71.** *Solution to Exercise 1.71.* (a) Let \( \Delta' : k[x] \to k[x,y] \) be the map sending every polynomial \( f(x_1, x_2, \ldots, x_n) \in k[x] \) to \( f(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \in k[x,y] \). Our goal is to prove that \( \Delta' = \Delta_{\text{Sym}(V)} \), where \( \Delta_{\text{Sym}(V)} \) is the usual coproduct on \( \text{Sym}(V) \) (part of the coalgebra structure obtained by regarding \( \text{Sym}(V) \) as a quotient of \( T(V) \) as in Exercise 1.21), and where we are identifying \( \text{Sym}(V) \) with \( k[x] \) and \( \text{Sym}(V) \otimes \text{Sym}(V) \) with \( k[x,y] \) along the isomorphisms given at the beginning of the exercise.

Notice first that \( \Delta_{\text{Sym}(V)} \) is a \( k \)-algebra homomorphism \( \text{Sym}(V) \to \text{Sym}(V) \otimes \text{Sym}(V) \) (because the axioms of a \( k \)-bialgebra require that the coproduct of any \( k \)-bialgebra \( A \) is a \( k \)-algebra homomorphism \( A \to A \otimes A \)). On the other hand, the map \( \Delta' \) is a \( k \)-algebra homomorphism\(^{336}\). Hence, the equality that we are trying to prove, namely \( \Delta' = \Delta_{\text{Sym}(V)} \), is an equality between two \( k \)-algebra homomorphisms. It is well-known that in order to prove such an equality, it is enough to verify it on a generating set of the domain of these homomorphisms\(^{337}\); i.e., it is enough to pick out a generating set of its domain, and check that for every element \( s \) of the generating set, the images of \( s \) under the two sides of the equality are equal to each other. In our case, the \( k \)-algebra homomorphisms \( \Delta' \) and \( \Delta_{\text{Sym}(V)} \) have domain \( k[x] \), and as a generating set of this \( k \)-algebra \( k[x] \) we can pick the set \( \{x_1, x_2, \ldots, x_n\} \). We then have to check that for every element \( s \) of this generating set, the images of \( s \) under \( \Delta' \) and \( \Delta_{\text{Sym}(V)} \) are equal to each other.

So let \( s = \{x_1, x_2, \ldots, x_n\} \) be arbitrary. Then, \( s = x_i \) for some \( i \in \{1, 2, \ldots, n\} \). Consider this \( i \). We have \( x_i \in V \), and thus \( x_i \) is a primitive element of \( T(V) \) (by the definition of the coalgebra structure on \( T(V) \)). This shows that \( \Delta_T(V)(x_i) = 1 \otimes x_i + x_i \otimes 1 \). Since \( \text{Sym}(V) \) is a quotient bialgebra of \( T(V) \), this shows that \( \Delta_{\text{Sym}(V)}(x_i) = 1 \otimes x_i + x_i \otimes 1 \) as well. Under our identification of \( \text{Sym}(V) \otimes \text{Sym}(V) \) with \( k[x,y] \), the element \( 1 \otimes x_i \) of \( \text{Sym}(V) \otimes \text{Sym}(V) \) equals the element \( y_i \) of \( k[x,y] \), and the element \( x_i \otimes 1 \) of \( \text{Sym}(V) \otimes \text{Sym}(V) \) equals the element \( x_i \) of \( k[x,y] \). Hence, \( \Delta_{\text{Sym}(V)}(x_i) = 1 \otimes x_i + x_i \otimes 1 \) rewrites as \( \Delta_{\text{Sym}(V)}(x_i) = y_i + x_i \).

But by the definition of \( \Delta' \), we have \( \Delta'(x_i) = x_i = x_i + y_1 \). Therefore, \( \Delta'(x_i) = x_i + y_i + x_i = \Delta_{\text{Sym}(V)}(x_i) \). Hence, the images of \( s \) under \( \Delta' \) and \( \Delta_{\text{Sym}(V)} \) are equal to each other. This is exactly what we needed to prove, and so the solution to Exercise 1.71(a) is complete.

\(^{336}\)In fact, by its very definition, it is an evaluation homomorphism, meaning a map from a polynomial ring to a commutative \( k \)-algebra \( A \) which sends every polynomial \( f \) to the evaluation \( f(a_1, a_2, \ldots, a_l) \) for some given tuple \( (a_1, a_2, \ldots, a_l) \) of elements \( A \). Such maps are always \( k \)-algebra homomorphisms.

\(^{337}\)In our case, the domain is \( k[x] \).
We will solve the remaining parts of the exercise starting with (c), then continuing with (d) and finally solving (b). This order has the advantage of requiring the least amount of work.

(c) We introduce a few notations. As long as we are using the topologist’s conventions, an element \( a \) of a graded \( \mathbf{k} \)-module is said to be odd if \( a \) is a sum of homogeneous elements of odd degree, and an element \( b \) of a graded \( \mathbf{k} \)-module is said to be even if \( b \) is a sum of homogeneous elements of even degree. By assumption, every element of \( V \) is odd. Notice that 0 is both even and odd, and a graded \( \mathbf{k} \)-module can (in general) contain vectors which are neither even nor odd.

It is easy to see that if \( P \) and \( Q \) are two graded \( \mathbf{k} \)-modules, and \( p \in P \) and \( q \in Q \) are two elements, then

\[
(11.137) \quad T(p \otimes q) = q \otimes p \quad \text{if } p \text{ is even and } q \text{ is even;}
\]

\[
(11.138) \quad T(p \otimes q) = q \otimes p \quad \text{if } p \text{ is even and } q \text{ is odd;}
\]

\[
(11.139) \quad T(p \otimes q) = q \otimes p \quad \text{if } p \text{ is odd and } q \text{ is even;}
\]

\[
(11.140) \quad T(p \otimes q) = -q \otimes p \quad \text{if } p \text{ is odd and } q \text{ is odd.}
\]

Now, let \( x \in V \). Then, \( x \) is odd (since every element of \( V \) is odd). The definition of \( \Delta \) enforces that \( \Delta \) is an algebra homomorphism \( T(V) \to T(V) \otimes T(V) \), but the meaning of this depends on the \( \mathbf{k} \)-algebra \( T(V) \otimes T(V) \), and therefore on the twist map \( T \), because the multiplication map of the \( \mathbf{k} \)-algebra \( T(V) \otimes T(V) \) is

\[
m_{T(V) \otimes T(V)} = (m_{T(V)} \otimes m_{T(V)}) \circ (id \otimes T \otimes id).
\]

The fact that we used the topologist’s convention (1.7) in this definition means that this twist map \( T \) is as in (1.7). Thus, (11.137), (11.138), (11.139) and (11.140) apply. Since \( x \) is odd, we can apply (11.140) to \( P = T(V), Q = T(V), p = x \) and \( q = x \), and obtain \( T(x \otimes x) = -x \otimes x \). Thus, in the \( \mathbf{k} \)-algebra

\[
\begin{align*}
T \left( \sum_{i \in I} p_i \otimes \sum_{j \in J} q_j \right) &= T \left( \sum_{i \in I} p_i \otimes \sum_{j \in J} q_j \right) = T \left( \sum_{i \in I} \sum_{j \in J} p_i \otimes q_j \right) \\
&= \sum_{i \in I} \sum_{j \in J} (-1)^{\deg(p_i) \cdot \deg(q_j)} q_j \otimes p_i \\
&= \sum_{i \in I} \sum_{j \in J} (-1)^{\deg(p_i) \cdot \deg(q_j)} q_j \otimes p_i \\
&= -q \otimes p.
\end{align*}
\]

This proves (11.140).

\[^{338}\text{Proof.} \text{ We will only prove (11.140); the other three identities are analogous.}
\]

Let \( P \) and \( Q \) be two graded \( \mathbf{k} \)-modules. Let \( p \in P \) and \( q \in Q \). Assume that \( p \) is odd and \( q \) is odd. Since \( p \) is odd, we know that \( p \) is a sum of homogeneous elements of \( P \) of odd degree; in other words, we can write \( p = \sum_{i \in I} p_i \), where \( I \) is a finite set and each \( i \in I \) has \( p_i \in P \) homogeneous of odd degree. Similarly, we can write \( q = \sum_{j \in J} q_j \), where \( J \) is a finite set and each \( j \in J \) has \( q_j \in Q \) homogeneous of odd degree. Using these \( p_i \) and \( q_j \), we now find
Similarly, we can compute

\[(1 \otimes x) \cdot (1 \otimes x) = 1 \otimes x^2,\]

\[(x \otimes 1) \cdot (1 \otimes x) = x \otimes x,\]

\[(x \otimes 1) \cdot (x \otimes 1) = x^2 \otimes 1.\]

(Since \(1 \in T(V)\) is even, we have to use (11.139), (11.137) and respectively (11.138) instead of (11.140) here; thus, we incur no negative signs.)

Since \(\Delta\) is a \(k\)-algebra homomorphism, we have

\[
\Delta \left( x^2 \right) = \left( \frac{\Delta (x)}{=1 \otimes x + x \otimes 1} \right)^2 = (1 \otimes x + x \otimes 1)^2
\]

\[
= (1 \otimes x + x \otimes 1) \cdot (1 \otimes x + x \otimes 1)
\]

\[
= (1 \otimes x) \cdot (1 \otimes x) + (1 \otimes x) \cdot (x \otimes 1) + (x \otimes 1) \cdot (1 \otimes x) + (x \otimes 1) \cdot (x \otimes 1)
\]

\[
= 1 \otimes x^2 + (-x \otimes x) + x \otimes x + x^2 \otimes 1 = 1 \otimes x^2 + x^2 \otimes 1.
\]

This solves Exercise 1.71(c).

(d) Let us first check that the ideal\(^{339}\) \(J\) of \(T(V)\) is a graded \(k\)-submodule of \(T(V)\). This is not obvious! We know that

\[(11.141)\]

\[
\text{if an ideal } I \text{ of a graded } k\text{-algebra } A \text{ is generated by homogeneous elements, then } I \text{ is a graded } k\text{-submodule of } A
\]

\[340\]

So we should concentrate on showing that the ideal \(J\) is generated by homogeneous elements.

In fact, we know that \(J\) is the ideal of \(T(V)\) generated by \(\{x^2\}_{x \in V}\). Thus,

\[(11.142)\]

\[
x^2 \in J \text{ for every } x \in V.
\]

\[339\]By “ideal”, we always mean “two-sided ideal”, unless we explicitly say “left ideal” or “right ideal”.

\[340\]This is a well-known fact in the case when \(A\) is commutative (and the topologist’s convention is not used), but it is proven exactly the same way in the general case.
Also, any two elements \( x \) and \( y \) of \( V \) satisfy

\[
xy + yx = \left( \sum_{\alpha \in J} \alpha \right)^2 - \sum_{\beta \in J} \beta^2 - \sum_{\gamma \in J} \gamma^2 \quad \text{(since } (x + y)^2 = x^2 + xy + yx + y^2)\]

(by (11.142), applied to \( x + y \) instead of \( x \))

\[
\in J - J - J \subset J \quad \text{(since } J \text{ is a } k\text{-module)}.
\]

Now, define a subset \( G \) of \( T(V) \) by

\[
G = \{ x^2 \mid x \in V; \ x \text{ is homogeneous} \} \cup \{ xy + yx \mid x \in V; \ y \in V; \ x \text{ and } y \text{ are homogeneous} \}.
\]

(11.144)

It is clear that all elements of \( G \) are homogeneous elements of \( T(V) \). By the definition of \( G \), we have

\[
x^2 \in G \quad \text{for every homogeneous } x \in V,
\]

(11.145)

and for the same reason we have

\[
xy + yx \in G \quad \text{for any two homogeneous elements } x \text{ and } y \text{ of } V.
\]

(11.146)

We shall now show that the ideal generated by \( G \) is \( J \).

Indeed, let \( g \in G \) be arbitrary. We will now show that \( g \in J \). We have

\[
g \in G = \{ x^2 \mid x \in V; \ x \text{ is homogeneous} \} \cup \{ xy + yx \mid x \in V; \ y \in V; \ x \text{ and } y \text{ are homogeneous} \}.
\]

Hence, either \( g \) has the form \( g = x^2 \) for some homogeneous \( x \in V \), or \( g \) has the form \( g = xy + yx \) for two homogeneous elements \( x \) and \( y \) of \( V \). In the first of these two cases, it is clear that \( g = x^2 \in J \) (by (11.142)). In the second of these cases, we have \( g = xy + yx \in J \) (by (11.143)). Hence, we have shown that \( g \in J \) in either case. This proves that \( g \in J \).

Now, forget that we fixed \( g \). We have thus proven that \( g \in J \) for every \( g \in G \). In other words, \( J \) contains \( G \) as a subset. Since \( J \) is an ideal of \( T(V) \), we thus have

\[
J \supset (\text{the smallest ideal of } T(V) \text{ containing } G \text{ as a subset}) = \text{(the ideal of } T(V) \text{ generated by } G).
\]

(11.147)

Let us now show the reverse inclusion. Let \( x \in V \). Then, \( x \) is a sum of homogeneous elements of \( V \) (because every element of a graded \( k\)-module is a sum of homogeneous elements). In other words, we can write \( x \) in the form \( x = \sum_{i \in I} x_i \), where \( J \) is a finite set and each \( i \in I \) has \( x_i \in V \) homogeneous. Consider
this \( I \) and these \( x_i \). Squaring both sides of the equality \( x = \sum_{i \in I} x_i \), we obtain

\[
x^2 = \left( \sum_{i \in I} x_i \right)^2 = \sum_{(i,j) \in I^2} x_i x_j = \sum_{(i,j) \in I^2; i<j} x_i x_j + \sum_{(i,j) \in I^2; i=j} x_i x_j + \sum_{(i,j) \in I^2; i>j} x_i x_j = \sum_{(i,j) \in I^2; i>j} x_i x_j
\]

(here, we substituted \( (j,i) \) for \( (i,j) \) in the sum)

\[
= \sum_{(i,j) \in I^2; i>j} x_j x_i + \sum_{(i,j) \in I^2; i=j} x_j x_i = \sum_{(i,j) \in I^2; i>j} x_j x_i + \sum_{(i,j) \in I^2; i=j} x_j x_i + \sum_{i \in I} \sum_{i>j} G
\]

(by (11.145), applied to \( x_i \) instead of \( x \))

\[
\in \sum_{(i,j) \in I^2; i>j} G + \sum_{(i,j) \in I^2; i=j} G \subset \text{(the ideal of } T(V) \text{ generated by } G) .
\]

Now, forget that we fixed \( x \). We thus have shown that \( x^2 \in \text{(the ideal of } T(V) \text{ generated by } G) \) for every \( x \in V \). Hence, \( \text{(the ideal of } T(V) \text{ generated by } G) \) contains the elements \( x^2 \) for all \( x \in V \). Since \( \text{(the ideal of } T(V) \text{ generated by } G) \) is an ideal, this yields that

\[
\text{(the ideal of } T(V) \text{ generated by } G) \supset \text{(the smallest ideal of } T(V) \text{ containing the elements } x^2 \text{ for all } x \in V) = \text{(the ideal generated by } \{ x^2 \}_{x \in V} ) = J .
\]

Combined with (11.147), this yields

\[
J = \text{(the ideal of } T(V) \text{ generated by } G) .
\]

The ideal \( J \) is thus generated by \( G \). Thus, the ideal \( J \) of \( T(V) \) is generated by homogeneous elements of \( T(V) \) (since all elements of \( G \) are homogeneous elements of \( T(V) \)). Therefore, (11.141) (applied to \( J \) and \( T(V) \) instead of \( I \) and \( A \)) yields that \( J \) is a graded \( k \)-submodule of \( T(V) \).

Now, let us show that \( \Delta(J) \subset J \otimes T(V) + T(V) \otimes J \).

If \( A \) and \( B \) are two graded \( k \)-algebras (in the topologist’s sense) and \( P \) and \( Q \) are two ideals of \( A \) and \( B \) which are graded \( k \)-submodules of \( A \) and \( B \), then \( P \otimes Q \) is an ideal of \( A \otimes B \) \footnote{Proof. Since \( P \) is an ideal of \( A \), we see that \( P \) is a \( k \)-submodule of \( A \) satisfying \( m_A(P \otimes A) \subset P \) and \( m_A(A \otimes P) \subset P \). Since \( Q \) is an ideal of \( B \), we see that \( Q \) is a \( k \)-submodule of \( B \) satisfying \( m_B(Q \otimes B) \subset Q \) and \( m_B(B \otimes Q) \subset Q \).} . Applying this to \( A = T(V), B = T(V), P = J \) and \( Q = T(V) \), we conclude that \( J \otimes T(V) \) is an ideal of \( T(V) \otimes T(V) \). Similarly, \( T(V) \otimes J \) is an ideal of \( T(V) \otimes T(V) \) as well. The sum \( J \otimes T(V) + T(V) \otimes J \) of these two ideals therefore is an ideal of \( T(V) \otimes T(V) \), too. As a consequence, \( \Delta^{-1}(J \otimes T(V) + T(V) \otimes J) \) is an
ideal of $T(V)$ (because $\Delta$ is a $k$-algebra homomorphism\textsuperscript{342}, and the preimage of an ideal under a $k$-algebra homomorphism is always an ideal).

Now, every $x \in V$ satisfies

$$
\Delta (x^2) = \frac{1}{\epsilon T(V)} \otimes \frac{x^2}{\epsilon T(V)} + \frac{x^2}{\epsilon T(V)} \otimes \frac{1}{\epsilon T(V)} \quad (\text{by Exercise 1.71(c)})
$$

so that $x^2 \in \Delta^{-1} (J \otimes T(V) + T(V) \otimes J)$. So the ideal $\Delta^{-1} (J \otimes T(V) + T(V) \otimes J)$ contains the elements $x^2$ for all $x \in V$. Hence,

$$
\Delta^{-1} (J \otimes T(V) + T(V) \otimes J) \supset \{ \text{the smallest ideal which contains the elements } x^2 \text{ for all } x \in V \}
$$

$$
= \{ \text{the ideal generated by } \{ x^2 \}_{x \in V} \} = J.
$$

Thus, $\Delta(J) \subset J \otimes T(V) + T(V) \otimes J$.

It remains to prove that $\epsilon(J) = 0$. This is similar to the above argument but much simpler. Since $\epsilon$ is a $k$-algebra homomorphism, its kernel $\ker \epsilon$ is an ideal of $T(V)$. Since $\epsilon$ is a $k$-algebra homomorphism, every $x \in V$ satisfies $\epsilon (x^2) = \left( \epsilon(x) \right)^2 = 0$ and thus $x^2 \in \ker \epsilon$. Thus, the ideal $\ker \epsilon$ contains the elements $x^2$ for all $x \in V$. Hence,

$$
\ker \epsilon \supset \{ \text{the smallest ideal which contains the elements } x^2 \text{ for all } x \in V \}
$$

$$
= \{ \text{the ideal generated by } \{ x^2 \}_{x \in V} \} = J.
$$

This yields $\epsilon(J) = 0$. Combined with $\Delta(J) \subset J \otimes T(V) + T(V) \otimes J$, this shows that $J$ is a two-sided coideal of $T(V)$. Since $J$ is also a two-sided ideal, this shows that the quotient $T(V)/J$ inherits a $k$-bialgebra structure from $T(V)$. This quotient $T(V)/J$ is $\wedge V$, and so we obtain a $k$-bialgebra structure on $\wedge V$. The $k$-bialgebra $\wedge V$ obtained this way is graded (because it is the quotient of the graded $k$-bialgebra $T(V)$ by the graded ideal $J$) and connected (since its 0-th graded component is $\wedge^0 V = k$ \textsuperscript{343}), therefore a Hopf algebra (by Proposition 1.36\textsuperscript{344}). Thus, $\wedge V$ is a connected graded Hopf algebra, and therefore the

\begin{itemize}
\item But since $Q$ is a graded $k$-submodule of $B$, there is a topologist’s twist map $T : Q \otimes A \to A \otimes Q$, which is the restriction of the twist map $T : B \otimes A \to A \otimes B$. Thus, $T(Q \otimes A) \subset A \otimes Q$. Now,

$$
\begin{aligned}
&\left( m_{A \otimes B} \circ (id \otimes T \otimes id) \right)
\left( \left( P \otimes Q \otimes (A \otimes B) \right) \right)

&= \left( (m_A \otimes m_B) \circ (id \otimes T \otimes id) \right)
\left( \left( P \otimes Q \otimes A \otimes B \right) \right)

&= \left( m_A \otimes m_B \right)
\left( \left( [id \otimes T \otimes id] \left( P \otimes Q \otimes A \otimes B \right) \right) \right)

&= \left( m_A \otimes m_B \right)
\left( \left( \left( P \otimes T(Q \otimes A) \otimes B \right) \right) \right)

&= \left( m_A \otimes m_B \right)
\left( \left( P \otimes A \otimes Q \otimes B \right) \right)

&= \left( m_A \otimes m_B \right)
\left( P \otimes A \otimes Q \otimes B \right)

&= \left( m_A \otimes m_B \right)
\left( P \otimes A \otimes Q \otimes B \right)

&= m_A \left( P \otimes A \otimes Q \otimes B \right)

&= m_A \left( P \otimes A \otimes Q \otimes B \right)

&= m_B \left( Q \otimes B \right)

&= P \otimes Q.
\end{aligned}
\end{itemize}

Similarly, $m_{A \otimes B} \left( (A \otimes B) \otimes (P \otimes Q) \right) \subset P \otimes Q$ (but here, we need to use $T(B \otimes P) \subset P \otimes B$ instead of $T(Q \otimes A) \subset A \otimes Q$). These two inclusions prove that $P \otimes Q$ is an ideal of $A \otimes B$, qed.

We could have done this proof using by working with elements, but that way we would have to take care of the fact that the twist $T$ is the topologist’s one and comes with signs. The way we have done it, we were almost entirely untroubled by this fact.\textsuperscript{342}\textsuperscript{by the definition of $\Delta$}

\textsuperscript{343}Here we are using the fact that $V_0 = 0$ (which is a consequence of the fact that $V$ is concentrated in odd degrees).

\textsuperscript{344}Here it helps to notice that our use of the topologist’s sign convention does not invalidate the proof of Proposition 1.36; in fact, no changes are necessary to that proof! (This might not be too surprising, given that said proof made no use of the bialgebra axioms (1.8).)
comultiplication $\Delta_{\wedge V}$ of $\wedge V$ is part of a connected graded Hopf algebra structure on $\wedge V$. In other words, the comultiplication $\Delta_{\wedge V}$ makes the $k$-algebra $\wedge V$ into a connected graded Hopf algebra.

Let us recall that the comultiplication $\Delta_{T(V)}$ of the $k$-bialgebra $T(V)$ satisfies

$$\Delta_{T(V)}(x) = 1 \otimes x + x \otimes 1 \quad \text{ for every } x \in V$$

(by the definition of $\Delta_{T(V)}$). Since the $k$-bialgebra $\wedge V$ was defined as a quotient of $T(V)$, we can project this equality down on $(\wedge V) \otimes (\wedge V)$, and thus conclude that the comultiplication $\Delta_{\wedge V}$ of the $k$-bialgebra $\wedge V$ satisfies

$$(11.148) \quad \Delta_{\wedge V}(x) = 1 \otimes x + x \otimes 1 \quad \text{ for every } x \in V.$$  

Thus, every $i \in \{1, 2, \ldots, n\}$ satisfies

$$(11.149) \quad \Delta_{\wedge V}(x_i) = 1 \otimes x_i + x_i \otimes 1$$

(by (11.148), applied to $x = x_i$). But our identification of $(\wedge V) \otimes (\wedge V)$ with $\wedge (V \oplus V)$ equates $1 \otimes x_i$ with $y_i$, and equates $x_i \otimes 1$ with $x_i$ for every $i \in \{1, 2, \ldots, n\}$. Hence, for every $i \in \{1, 2, \ldots, n\}$, the equality (11.149) rewrites as

$$(11.150) \quad \Delta_{\wedge V}(x_i) = 1 \otimes x_i + x_i \otimes 1 = y_i + x_i = x_i + y_i$$

in $(\wedge V) \otimes (\wedge V) = \wedge (V \oplus V)$.

Recall that we made $\wedge V$ into a $k$-bialgebra by viewing it as the quotient $T(V)/J$. Now, it remains to prove that the coproduct on $\wedge V$ which is part of this $k$-bialgebra structure on $T(V)$ is the same as the one defined in Exercise 1.71(b).\footnote{We should be careful because we have not yet proven that the latter coproduct satisfies the axioms for a coproduct (we have not solved Exercise 1.71(b) yet).} In order to do so, we shall prove the following claim: If $\Delta_{\wedge V}$ denotes the coproduct on $\wedge V$ obtained by regarding $\wedge V$ as the quotient $k$-bialgebra $T(V)/J$, then

$$(11.151) \quad \Delta_{\wedge V}\left(\sum_{i_1 < \cdots < i_d} c_{i_1, \ldots, i_d} x_{i_1} \wedge \cdots \wedge x_{i_d}\right) = \sum_{i_1 < \cdots < i_d} c_{i_1, \ldots, i_d} (x_{i_1} + y_{i_1}) \wedge \cdots \wedge (x_{i_d} + y_{i_d})$$

for every family $(c_{i_1, \ldots, i_d})_{i_1 < \cdots < i_d}$ of elements of $k$ indexed by the strictly increasing sequences $(i_1 < \cdots < i_d)$ of elements of $\{1, 2, \ldots, n\}$.

Proof of (11.151): Let $(c_{i_1, \ldots, i_d})_{i_1 < \cdots < i_d}$ be a family of elements of $k$ indexed by the strictly increasing sequences $(i_1 < \cdots < i_d)$ of elements of $\{1, 2, \ldots, n\}$. The multiplication in $\wedge V$ is given by the wedge product, so that we have $x_{i_1} \cdots x_{i_d} = x_{i_1} \wedge \cdots \wedge x_{i_d}$ for every strictly increasing sequence $(i_1 < \cdots < i_d)$ of elements of $\{1, 2, \ldots, n\}$.

But the comultiplication $\Delta_{\wedge V}$ is a $k$-algebra homomorphism (by the axioms of a $k$-bialgebra, since $\wedge V$ is a $k$-bialgebra), and thus we have

$$\Delta_{\wedge V}\left(\sum_{i_1 < \cdots < i_d} c_{i_1, \ldots, i_d} x_{i_1} \cdots x_{i_d}\right) = \sum_{i_1 < \cdots < i_d} c_{i_1, \ldots, i_d} \frac{\Delta_{\wedge V}(x_{i_1}) \cdots \Delta_{\wedge V}(x_{i_d})}{= x_{i_1} + y_{i_1} \quad \text{by (11.150), applied to } i=i_1} \cdots \frac{\Delta_{\wedge V}(x_{i_d})}{= x_{i_d} + y_{i_d} \quad \text{by (11.150), applied to } i=i_d}$$

$$= \sum_{i_1 < \cdots < i_d} c_{i_1, \ldots, i_d} (x_{i_1} + y_{i_1}) \cdots (x_{i_d} + y_{i_d}) = (x_{i_1} + y_{i_1}) \wedge \cdots \wedge (x_{i_d} + y_{i_d})$$

(since multiplication in $\wedge (V \oplus V)$ is given by the wedge product)

$$= \sum_{i_1 < \cdots < i_d} c_{i_1, \ldots, i_d} (x_{i_1} + y_{i_1}) \wedge \cdots \wedge (x_{i_d} + y_{i_d}).$$

Since $x_{i_1} \cdots x_{i_d} = x_{i_1} \wedge \cdots \wedge x_{i_d}$ for every strictly increasing sequence $(i_1 < \cdots < i_d)$ of elements of $\{1, 2, \ldots, n\}$, this rewrites as follows:

$$\Delta_{\wedge V}\left(\sum_{i_1 < \cdots < i_d} c_{i_1, \ldots, i_d} x_{i_1} \wedge \cdots \wedge x_{i_d}\right) = \sum_{i_1 < \cdots < i_d} c_{i_1, \ldots, i_d} (x_{i_1} + y_{i_1}) \wedge \cdots \wedge (x_{i_d} + y_{i_d}).$$
Hence, there exists a map satisfying this property. Such a map is furthermore unique (because the defining property determines its value on every element of the form \( \sum_{i_1, \ldots, i_d} c_{i_1, \ldots, i_d} x_{i_1} \cdots x_{i_d} \)), and therefore the map \( \Delta_A \) (because our map \( \Delta_A \) satisfies the defining property of the map (1.25)), and therefore makes \( \wedge V \) into a connected graded Hopf algebra (since we know that the comultiplication \( \Delta_A \) makes the \( k \)-algebra \( \wedge V \) into a connected graded Hopf algebra). This solves Exercise 1.71(b).

But the coproduct on \( \wedge V \) inherited from \( T(V) \) is \( \Delta_A \), and as we know, this \( \Delta_A \) is exactly the map (1.25), i.e., the coproduct defined in Exercise 1.71(b). Hence, the coproduct on \( \wedge V \) inherited from \( T(V) \) is the coproduct defined in Exercise 1.71(b). Thus, Exercise 1.71(d) is also solved. The solution to Exercise 1.71 is thus complete.

11.28. **Solution to Exercise 1.72**. Solution to Exercise 1.72. Define a \( k \)-linear map \( \rho_{U,V} : U^* \otimes V^* \to (U \otimes V)^* \) for any two \( k \)-modules \( U \) and \( V \) as in Exercise 1.67. Recall that this \( \rho_{U,V} \) is a \( k \)-module isomorphism if both \( U \) and \( V \) are finite free. Hence, \( \rho_{A,A} \) is a \( k \)-module isomorphism.

Also, basic linear algebra shows that if \( U, V, U' \) and \( V' \) are four \( k \)-modules and if \( \alpha : U \to U' \) and \( \beta : V \to V' \) are two \( k \)-linear maps, then

\[
\rho_{U,V} \circ (\alpha \otimes \beta^*) = (\alpha \otimes \beta)^* \circ \rho_{U',V'}.
\]

This (applied to \( U = C, V = C, U' = A, V' = A, \alpha = f \) and \( \beta = g \)) yields

\[
\rho_{C,C} \circ (f^* \otimes g^*) = (f \otimes g)^* \circ \rho_{A,A}.
\]

The definition of convolution yields both

\[
f \ast g = m_A \circ (f \otimes g) \circ \Delta_C
\]

and

\[
f^* \ast g^* = m_C \circ (f^* \otimes g^*) \circ \Delta_A
\]

\[
= m_C \circ (f^* \otimes g^*) \circ \rho_{A,C} \circ (f \otimes g)^* \circ \rho_{C,A}^{-1} \circ m_A
\]

\[
= \Delta_C^* \circ \rho_{C,C} \circ (f^* \otimes g^*) \circ \rho_{A,A}^{-1} \circ m_A
\]

\[
= \Delta_C^* \circ \rho_{C,C} \circ (f \otimes g)^* \circ \rho_{A,A} \circ m_A
\]

\[
= \Delta_C^* \circ (f \otimes g)^* \circ m_A
\]

\[
= \left( m_A \circ (f \otimes g) \circ \Delta_C \right)^*
\]

\[
= (f \ast g)^*.
\]

This solves Exercise 1.72.

11.29. **Solution to Exercise 1.74**. Solution to Exercise 1.74. Our goal is to prove Proposition 1.73.

Let \( m_{\omega} \) denote the \( k \)-linear map \( T(V) \otimes T(V) \to T(V) \) which sends every \( a \otimes b \) to \( a \omega b \). Let \( u \) denote the unit map of \( T(V) \) (that is, the \( k \)-linear map \( k \to T(V) \) sending 1 to \( 1_{T(V)} \)). Let \( S \) denote the antipode of the Hopf algebra \( T(V) \). Then, the result that we have to prove boils down to the statement that the \( k \)-module \( T(V) \), endowed with the multiplication \( m_{\omega} \), the unit \( u \), the comultiplication \( \Delta_{\omega} \) and the counit \( \epsilon \), becomes a commutative Hopf algebra with the antipode \( S \). Since we already know that \( m_{\omega}, u, \Delta_{\omega}, \epsilon \) and \( S \) are \( k \)-linear, this latter statement will immediately follow once we can show that the following diagrams commute:

- the diagrams (1.1) and (1.2), with \( A \) and \( m \) replaced by \( T(V) \) and \( m_{\omega} \);
- the diagrams (1.3) and (1.4), with \( C \) and \( \Delta \) replaced by \( T(V) \) and \( \Delta_{\omega} \);
• the diagrams (1.8), with \( A, m \) and \( \Delta \) replaced by \( T(V), m_\mu \) and \( \Delta_\mu \);
• the diagram (11.12), with \( A, m \) and \( \Delta \) replaced by \( T(V), m_\mu \) and \( \Delta_\mu \);
• the diagram (11.18), with \( A \) and \( m \) replaced by \( T(V) \) and \( m_\mu \).

Out of all these statements, we will only prove the commutativity of the first diagram in (1.8), since all other diagrams are similar but only easier.

So we must prove the commutativity of the first diagram in (1.8), with \( A, m \) and \( \Delta \) replaced by \( T(V), m_\mu \) and \( \Delta_\mu \). In other words, we must show that the diagram

\[
\begin{array}{c}
T(V) \\
\downarrow \Phi \\
T(V) \\
\end{array}
\]

commutes, where \( T : T(V) \otimes T(V) \to T(V) \otimes T(V) \) is the twist map sending every \( a \otimes b \) to \( b \otimes a \) (and being \( \mathbf{k} \)-linear at that). Let us prove this now.

We need to prove that the diagram (11.153) commutes. In other words, we need to prove the identity

\[
(m_\mu \otimes m_\mu) \circ (\text{id} \otimes T \otimes \text{id}) \circ \left( \Delta_\mu \otimes \Delta_\mu \right) = \Delta_\mu \circ m_\mu.
\]

By linearity, it is clearly enough to verify this identity only on the pure tensors in \( T(V) \otimes T(V) \); that is, it is enough to check that every \( a \in T(V) \) and \( b \in T(V) \) satisfy

\[
(m_\mu \otimes m_\mu) \circ (\text{id} \otimes T \otimes \text{id}) \circ \left( \Delta_\mu \otimes \Delta_\mu \right) (a \otimes b) = \left( \Delta_\mu \circ m_\mu \right) (a \otimes b).
\]

So let \( a \in T(V) \) and \( b \in T(V) \) be arbitrary. All we need now is to prove (11.155). By linearity again, we can WLOG assume that \( a \) and \( b \) have the form \( a = v_1 v_2 \cdots v_p \) and \( b = v_{p+1} v_{p+2} \cdots v_{p+q} \) for some \( p, q \in \mathbb{N} \), \( v_1, v_2, \ldots, v_{p+q} \in V \) (since \( T(V) \) is spanned as a \( \mathbf{k} \)-module by pure tensors). Assume this, and define \( W \) to be the free \( \mathbf{k} \)-module with basis \( \{x_1, x_2, \ldots, x_{p+q}\} \). Let \( A \) be the tensor algebra \( T(W) \) of this \( \mathbf{k} \)-module \( W \). Then, \( W \) is a finite free \( \mathbf{k} \)-module, and so we know from Example 1.69 (applied to \( W \) instead of \( V \)) that the graded dual \( A^\circ \) of its tensor algebra \( A = T(W) \) is a Hopf algebra whose basis \( \{y(i_1,i_2,\ldots,i_\ell)\} \) is indexed by words in the alphabet \( I := \{1,2,\ldots,p+q\} \).

Notice that \( \{y(i_1,i_2,\ldots,i_\ell) \otimes y(j_1,j_2,\ldots,j_m)\}_{\ell, m \in \mathbb{N}, (i_1,i_2,\ldots,i_\ell) \in I^\ell, (j_1,j_2,\ldots,j_m) \in I^m} \) is a \( \mathbf{k} \)-module basis of \( A^\circ \otimes A^\circ \) (since \( \{y(i_1,i_2,\ldots,i_\ell)\} \) is a basis of \( A^\circ \)). Relabelling this basis, we see that

\[
\{y(i_1,i_2,\ldots,i_\ell) \otimes y(i_{\ell+1},i_{\ell+2},\ldots,i_{\ell+m})\}_{\ell, m \in \mathbb{N}, (i_1,i_2,\ldots,i_{\ell+m}) \in I_{\ell+m}} \quad \text{is a \( \mathbf{k} \)-module basis of \( A^\circ \otimes A^\circ \).}
\]

We can define a \( \mathbf{k} \)-linear map \( \phi : A^\circ \to T(V) \) by setting

\[
\phi\left(y(i_1,i_2,\ldots,i_\ell)\right) = v_{i_1} v_{i_2} \cdots v_{i_\ell} \quad \text{for every} \quad \ell \in \mathbb{N} \text{ and} \quad (i_1,i_2,\ldots,i_\ell) \in I^\ell
\]

(because \( \{y(i_1,i_2,\ldots,i_\ell)\} \) is a basis of \( A^\circ \)). Consider this map \( \phi \).

Notice that the definition of \( \phi \) yields \( \phi\left(y(1,2,\ldots,p)\right) = v_1 v_2 \cdots v_p = a \), and similarly \( \phi\left(y(p+1,p+2,\ldots,p+q)\right) = b \).

Thus,

\[
\phi\left(y(1,2,\ldots,p)\right) \otimes \phi\left(y(p+1,p+2,\ldots,p+q)\right) = \phi\left(y(1,2,\ldots,p+1,p+2,\ldots,p+q)\right).
\]

\[
\in A^\circ \otimes A^\circ.
\]
In other words, \( a \otimes b \) lies in the image of the map \( \phi \otimes \phi \).

Notice that we already know that \( A^o \) is a \( k \)-bialgebra, and thus it satisfies all the axioms of a bialgebra; in particular, the diagrams (1.8), with \( A, m \) and \( \Delta \) replaced by \( A^o, m_{A^o} \) and \( \Delta_{A^o} \), commute. In particular, the first of these diagrams commutes. In other words, the diagram

\[
\begin{array}{c}
\Delta_{A^o} \otimes \Delta_{A^o} \\
\downarrow \downarrow \\
A^o \otimes A^o \\
\downarrow \downarrow \\
m_{A^o} \\
\end{array}
\]

commutes.

Now, we claim that

\[
\begin{align*}
\phi \circ m_{A^o} &= m_{\omega} \circ (\phi \otimes \phi), \\
\phi \circ u_{A^o} &= u, \\
(\phi \otimes \phi) \circ \Delta_{A^o} &= \Delta_{\omega} \circ \phi, \\
\epsilon_{A^o} &= \epsilon \circ \phi, \\
\phi \circ S_{A^o} &= S \circ \phi.
\end{align*}
\]

We are going to only prove the two equalities (11.157) and (11.159), leaving the (simpler!) proofs of the other three equalities (11.158), (11.160) and (11.161) to the reader. (Only the equalities (11.157) and (11.159) will be used in the proof of the commutativity of the first diagram in (1.8); the other are used for the other diagrams.)

Proof of (11.157): We need to prove the equality (11.157). Since both sides of this equality (11.157) are \( k \)-linear maps, it is enough to prove this equality on a \( k \)-basis of \( A^o \otimes A^o \). Let us pick the basis \( \{ y_{(i_1, i_2, \ldots, i_\ell)} \otimes y_{(i_{\ell+1}, i_{\ell+2}, \ldots, i_{\ell+m})} \} \) for \( A^o \otimes A^o \); it thus is enough to prove the equality (11.157) on this basis, i.e., to prove that

\[
(\phi \circ m_{A^o})_\omega (y_{(i_1, i_2, \ldots, i_\ell)} \otimes y_{(i_{\ell+1}, i_{\ell+2}, \ldots, i_{\ell+m})}) = (m_{\omega} \circ (\phi \otimes \phi))_\omega (y_{(i_1, i_2, \ldots, i_\ell)} \otimes y_{(i_{\ell+1}, i_{\ell+2}, \ldots, i_{\ell+m})})
\]
for every $\ell \in \mathbb{N}$, $m \in \mathbb{N}$ and $(i_1, i_2, \ldots, i_{\ell+m}) \in I^{\ell+m}$. So let us fix $\ell \in \mathbb{N}$, $m \in \mathbb{N}$ and $(i_1, i_2, \ldots, i_{\ell+m}) \in I^{\ell+m}$.

Comparing

$$
(\phi \circ m_{A^o}) \left( y(i_1, i_2, \ldots, i_\ell) \otimes y(i_{\ell+1}, i_{\ell+2}, \ldots, i_{\ell+m}) \right)
= \phi \left( \sum_{\sigma \in \text{Sh}_{\ell,m}} y(i_{\sigma(1)}, \ldots, i_{\sigma(\ell+1)}) \cdot y(i_{\sigma(\ell+2)}, \ldots, i_{\sigma(\ell+m)}) \right) = \phi \left( \sum_{\sigma \in \text{Sh}_{\ell,m}} y(i_{\sigma(1)}, \ldots, i_{\sigma(\ell+1)}) \right) 
= \sum_{\sigma \in \text{Sh}_{\ell,m}} \phi \left( y(i_{\sigma(1)}, \ldots, i_{\sigma(\ell+1)}) \right)
= \sum_{\sigma \in \text{Sh}_{\ell,m}} v_{i_{\sigma(1)}}, v_{i_{\sigma(2)}}, \ldots, v_{i_{\sigma(\ell+m)}}
$$

with

$$
(m_{\text{shuffle}} \circ (\phi \otimes \phi)) \left( y(i_1, i_2, \ldots, i_\ell) \otimes y(i_{\ell+1}, i_{\ell+2}, \ldots, i_{\ell+m}) \right)
= m_{\text{shuffle}} \left( \phi \left( y(i_1, i_2, \ldots, i_\ell) \otimes y(i_{\ell+1}, i_{\ell+2}, \ldots, i_{\ell+m}) \right) \otimes \phi \left( y(i_1, i_2, \ldots, i_\ell) \otimes y(i_{\ell+1}, i_{\ell+2}, \ldots, i_{\ell+m}) \right) \right)
= m_{\text{shuffle}} \left( \phi \left( y(i_1, i_2, \ldots, i_\ell) \otimes y(i_{\ell+1}, i_{\ell+2}, \ldots, i_{\ell+m}) \right) \otimes \phi \left( y(i_1, i_2, \ldots, i_\ell) \otimes y(i_{\ell+1}, i_{\ell+2}, \ldots, i_{\ell+m}) \right) \right)
= \sum_{\sigma \in \text{Sh}_{\ell,m}} v_{i_{\sigma(1)}}, v_{i_{\sigma(2)}}, \ldots, v_{i_{\sigma(\ell+m)}}
$$

we obtain $(\phi \circ m_{A^o}) \left( y(i_1, i_2, \ldots, i_\ell) \otimes y(i_{\ell+1}, i_{\ell+2}, \ldots, i_{\ell+m}) \right) = (m_{\text{shuffle}} \circ (\phi \otimes \phi)) \left( y(i_1, i_2, \ldots, i_\ell) \otimes y(i_{\ell+1}, i_{\ell+2}, \ldots, i_{\ell+m}) \right)$. Thus, (11.162) is proven, and this establishes the equality (11.157).

Proof of (11.159): Now we must prove the equality (11.159). By linearity, it is enough to verify this equality on a $k$-basis of $A^o$. We will use the basis $\{ y(i_1, i_2, \ldots, i_\ell) \}$; then, we need to check the equality

$$
((\phi \otimes \phi) \circ \Delta_{A^o}) \left( y(i_1, i_2, \ldots, i_\ell) \right) = (\Delta_{A^o} \circ \phi) \left( y(i_1, i_2, \ldots, i_\ell) \right)
$$

holds for every $\ell \in \mathbb{N}$ and $(i_1, i_2, \ldots, i_\ell) \in I^\ell$. This equality follows by comparing

$$
((\phi \otimes \phi) \circ \Delta_{A^o}) \left( y(i_1, i_2, \ldots, i_\ell) \right) = (\phi \otimes \phi) \left( \Delta_{A^o} \left( y(i_1, i_2, \ldots, i_\ell) \right) \right)
= (\phi \otimes \phi) \left( \sum_{j=0}^{\ell} y(i_1, \ldots, i_j) \otimes y(i_{j+1}, \ldots, i_\ell) \right)
= (\phi \otimes \phi) \left( \sum_{j=0}^{\ell} y(i_1, \ldots, i_j) \otimes y(i_{j+1}, \ldots, i_\ell) \right)
= \sum_{j=0}^{\ell} (\phi \otimes \phi) \left( y(i_1, \ldots, i_j) \otimes y(i_{j+1}, \ldots, i_\ell) \right)
= \sum_{j=0}^{\ell} v_{i_1}, v_{i_2}, \ldots, v_{i_j} \otimes v_{i_{j+1}}, v_{i_{j+2}}, \ldots, v_{i_\ell}
$$
with

\[(\Delta_{\omega} \circ \varphi) (y_{i_1, i_2, \ldots, i_{\ell}}) = \Delta_{\omega} \left( \varphi \left( y_{i_1, i_2, \ldots, i_{\ell}} \right) \right) = \Delta_{\omega} \left( v_{i_1} v_{i_2} \cdots v_{i_{\ell}} \right) = \sum_{k=0}^{\ell} (v_{i_1} v_{i_2} \cdots v_{i_k}) \otimes (v_{i_{k+1}} v_{i_{k+2}} \cdots v_{i_{\ell}}) \quad \text{(by the definition of } \Delta_{\omega})\]

Thus, (11.159) is proven.

Now, we can derive (11.155) in a very straightforward way from the commutativity of (11.156) using the equalities (11.159) and (11.157): Consider the diagram

where \( T \) is shorthand for \( T(V) \). The large pentagon in this diagram is commutative (because it is the diagram (11.156), which is known to commute), and so are all five quadrilaterals\(^{347}\). This does not automatically yield the commutativity of the small pentagon, but it yields that all paths from the \( A^o \otimes A^o \) at the top of the diagram to the \( T \otimes T \) one row above the very bottom give the same map; in particular, we have

\[(m_{\omega} \otimes m_{\omega}) \circ (\text{id} \otimes T \otimes \text{id}) \circ (\Delta_{\omega} \otimes \Delta_{\omega}) \circ (\varphi \otimes \varphi) = (\Delta_{\omega} \circ m_{\omega}) \circ (\varphi \otimes \varphi).\]

Thus, the two maps \((m_{\omega} \otimes m_{\omega}) \circ (\text{id} \otimes T \otimes \text{id}) \circ (\Delta_{\omega} \otimes \Delta_{\omega})\) and \(\Delta_{\omega} \circ m_{\omega}\) are equal to each other on the image of the map \(\varphi \otimes \varphi\). Since \(a \otimes b\) lies in the image of the map \(\varphi \otimes \varphi\), this yields that these two maps

\(^{347}\text{In fact:}\)

- the northeastern quadrilateral commutes because of (11.157);
- the southeastern quadrilateral commutes because of (11.159);
- the northwestern quadrilateral commutes because \n\[
(\varphi \otimes \varphi \otimes \varphi \otimes \varphi) \circ (\Delta_{A^o} \otimes \Delta_{A^o}) = (\varphi \otimes \varphi) \circ (\Delta_{A^o} \otimes \Delta_{A^o}) \quad \text{(by (11.159))}
\]
\[
= (\Delta_{\omega} \circ \varphi) \otimes (\Delta_{\omega} \circ \varphi) = (\Delta_{\omega} \otimes \Delta_{\omega}) \circ (\varphi \otimes \varphi);\]
- the southwestern quadrilateral commutes for a similar reason (but using (11.157) instead of (11.159));
- the western quadrilateral commutes as a consequence of simple linear algebra.
are equal to each other on \( a \otimes b \). In other words, (11.155) holds. As we have said above, this completes our proof of Proposition 1.73.

11.30. **Solution to Exercise 2.2.** Solution to Exercise 2.2. We have \( f \in R(x) \). Thus, \( f \) is a formal power series of bounded degree. In other words, \( f = \sum_\alpha c_\alpha x^\alpha \) (with the sum ranging over all weak compositions \( \alpha \)) for some elements \( c_\alpha \) in \( k \) such that there exists a \( d \in \mathbb{N} \) such that every \( \alpha \) satisfying \( \deg(x^\alpha) > d \) must satisfy \( c_\alpha = 0 \). Consider these \( c_\alpha \) and this \( d \).

We have written \( f \) as the sum \( \sum_\alpha c_\alpha x^\alpha \). Now, substituting \( a_1, a_2, \ldots, a_k, 0, 0, \ldots \) for \( x_1, x_2, x_3, \ldots \) in \( f \) maps all but finitely many of the terms \( c_\alpha x^\alpha \) appearing in this sum to 0. In fact:

- all terms \( c_\alpha x^\alpha \) such that the monomial \( x^\alpha \) contains at least one of the variables \( x_{k+1}, x_{k+2}, x_{k+3}, \ldots \) (that is, such that for some integer \( i > k \), the \( i \)-th entry of \( \alpha \) is nonzero) become 0 under our substitution (because the substitution takes each of the variables \( x_{k+1}, x_{k+2}, x_{k+3}, \ldots \) to 0);
- all terms \( c_\alpha x^\alpha \) with \( \deg(x^\alpha) > d \) become 0 under our substitution (because we know that \( c_\alpha = 0 \) for every \( \alpha \) satisfying \( \deg(x^\alpha) > d \));
- the remaining terms (that is, the terms \( c_\alpha x^\alpha \) satisfying neither of the preceding two conditions) might not get sent to 0, but there are only finitely many such terms\(^{348}\).

Hence, our substitution maps all but finitely many of the terms \( c_\alpha x^\alpha \) appearing in the sum \( \sum_\alpha c_\alpha x^\alpha \) to 0. Since \( \sum_\alpha c_\alpha x^\alpha = f \), this rewrites as follows: Our substitution maps all but finitely many of the terms of \( f \) to 0. This solves Exercise 2.2.

11.31. **Solution to Exercise 2.16.** Solution to Exercise 2.16. Let us first make an auxiliary observation. Namely, let us show that every \( \nu \in \text{Par}_n \) and every \( k \in \mathbb{N} \) satisfy

\[
(11.163) \quad (\nu^t)_1 + (\nu^t)_2 + \ldots + (\nu^t)_k = \sum_{j=1}^{\infty} \min \{ \nu_j, k \}.
\]

(Note that the right hand side of (11.163) is well-defined because every sufficiently high \( j \) satisfies \( \min \{ \nu_j, k \} = 0 \).)

**Proof of (11.163):** This is best seen by double-counting boxes in a Ferrers diagram. The left hand side of (11.163) counts the boxes in the first \( k \) rows of the Ferrers diagram of \( \nu^t \). Since the Ferrers diagram of \( \nu^t \) is obtained from that of \( \nu \) by exchanging rows for columns (i.e., reflecting the diagram across its main diagonal), it is clear that this is the same as counting the boxes in the first \( k \) columns of the Ferrers diagram of \( \nu \). But these latter boxes can also be counted row-by-row: For every \( j \in \{1, 2, 3, \ldots\} \), the number of boxes in the first \( k \) columns of the Ferrers diagram of \( \nu \) that lie in row \( j \) is \( \min \{ \nu_j, k \} \). Thus, the total amount of boxes is \( \sum_{j=1}^{\infty} \min \{ \nu_j, k \} \), which is precisely the right hand side of (11.163). Thus, (11.163) is proven.

Let us now come to the solution of the exercise. We need to prove that

\[
(11.164) \quad (\text{if } \lambda \triangleright \mu, \text{ then } \mu^t \triangleright \lambda^t)
\]

and

\[
(11.165) \quad (\text{if } \mu^t \triangleright \lambda^t, \text{ then } \lambda \triangleright \mu).
\]

**Proof of (11.165):** Assume that \( \mu^t \triangleright \lambda^t \). By the definition of the dominance order, this means that

\[
(11.166) \quad (\mu^t)_1 + (\mu^t)_2 + \ldots + (\mu^t)_k \geq (\lambda^t)_1 + (\lambda^t)_2 + \ldots + (\lambda^t)_k \quad \text{for all } k \in \{1, 2, \ldots, n\}.
\]

This inequality holds for \( k = 0 \) as well (because both sides are 0 when \( k = 0 \)), and thus it holds for all \( k \in \{0, 1, \ldots, n\} \). We can further rewrite this inequality by applying (11.163): The left hand side becomes \( \sum_{j=1}^{\infty} \min \{ \mu_j, k \} \), and the right hand side becomes \( \sum_{j=1}^{\infty} \min \{ \lambda_j, k \} \). As a result, we obtain

\[
(11.167) \quad \sum_{j=1}^{\infty} \min \{ \mu_j, k \} \geq \sum_{j=1}^{\infty} \min \{ \lambda_j, k \} \quad \text{for all } k \in \{0, 1, \ldots, n\}.
\]

\(^{348}\)In fact, these are the terms for which \( x^\alpha \) is a monomial which has degree \( \leq d \) and contains no other variables than \( x_1, x_2, \ldots, x_k \). Clearly, there are only finitely many such monomials, and thus only finitely many such terms.
Now, let \( k \in \{1, 2, \ldots, n\} \) be arbitrary. We are going to show that \( \lambda_1 + \lambda_2 + \ldots + \lambda_k \geq \mu_1 + \mu_2 + \ldots + \mu_k \).

First of all, \( \lambda \in \text{Par}_n \), so that \( n = |\lambda| \geq \lambda_k \). Thus, \( \lambda_k \in \{0, 1, \ldots, n\} \). Therefore, (11.167) (applied to \( \lambda_k \) instead of \( k \)) yields that

\[
(11.168) \quad \sum_{j=1}^{\infty} \min \{\mu_j, \lambda_k\} \geq \sum_{j=1}^{\infty} \min \{\lambda_j, \lambda_k\}.
\]

But the right hand side of this inequality can be rewritten as follows:

\[
\sum_{j=1}^{\infty} \min \{\lambda_j, \lambda_k\} = \sum_{j=1}^{k} \min \{\lambda_j, \lambda_k\} + \sum_{j=k+1}^{\infty} \min \{\lambda_j, \lambda_k\} = \lambda_k + \sum_{j=k+1}^{\infty} \lambda_j = k\lambda_k + \sum_{j=k+1}^{\infty} \lambda_j = k\lambda_k + (|\lambda| - (\lambda_1 + \lambda_2 + \ldots + \lambda_k)) = k\lambda_k + n - (\lambda_1 + \lambda_2 + \ldots + \lambda_k).
\]

(11.169)

Meanwhile, the left hand side can be bounded from above:

\[
\sum_{j=1}^{\infty} \min \{\mu_j, \lambda_k\} = \sum_{j=1}^{k} \min \{\mu_j, \lambda_k\} + \sum_{j=k+1}^{\infty} \min \{\mu_j, \lambda_k\} \leq \sum_{j=1}^{k} \mu_j + \sum_{j=k+1}^{\infty} \mu_j = k\lambda_k + \sum_{j=k+1}^{\infty} \mu_j = k\lambda_k + (|\mu| - (\mu_1 + \mu_2 + \ldots + \mu_k)) = k\lambda_k + n - (\mu_1 + \mu_2 + \ldots + \mu_k).
\]

(11.170)

Now, (11.168) yields

\[
0 \leq \sum_{j=1}^{\infty} \min \{\mu_j, \lambda_k\} - \sum_{j=1}^{\infty} \min \{\lambda_j, \lambda_k\} \leq k\lambda_k + n - (\mu_1 + \mu_2 + \ldots + \mu_k) - (\lambda_1 + \lambda_2 + \ldots + \lambda_k) = (\lambda_1 + \lambda_2 + \ldots + \lambda_k) - (\mu_1 + \mu_2 + \ldots + \mu_k)
\]

In other words, \( \lambda_1 + \lambda_2 + \ldots + \lambda_k \geq \mu_1 + \mu_2 + \ldots + \mu_k \).

Now, forget that we fixed \( k \). We thus have shown that

\[
\lambda_1 + \lambda_2 + \ldots + \lambda_k \geq \mu_1 + \mu_2 + \ldots + \mu_k \quad \text{for all } k \in \{1, 2, \ldots, n\}.
\]

In other words, \( \lambda \triangleright \mu \). This proves (11.165).

Proof of (11.164): Assume that \( \lambda \triangleright \mu \). We have \( (\lambda_1^{t})^{t} = \lambda \triangleright \mu = (\mu_1^{t})^{t} \). Thus, we can apply (11.165) to \( \mu^{t} \) and \( \lambda^{t} \) instead of \( \lambda \) and \( \mu \). As a result, we obtain \( \mu^{t} \triangleright \lambda^{t} \). This proves (11.164).

Combining (11.164) and (11.165), we obtain the equivalence of the two assertions \( \lambda \triangleright \mu \) and \( \mu^{t} \triangleright \lambda^{t} \). This solves Exercise 2.16.
11.32. Solution to Exercise 2.22. Solution to Exercise 2.22. (a) This is straightforward. If \( \lambda \) and \( \mu \) are two partitions such that \( \mu \subseteq \lambda \), and if \( \mathcal{L} \) is any total order on the positive integers, then we say that an assignment \( T \) of entries in \( \{1, 2, 3, \ldots \} \) to the cells of the Ferrers diagram of \( \lambda/\mu \) is an \( \mathcal{L} \text{-column-strict tableau} \) if it is \( \mathcal{L} \)-increasing left-to-right in rows, and strictly \( \mathcal{L} \)-increasing top-to-bottom in columns. (This definition of \( \mathcal{L} \text{-column-strict tableaux} \) clearly extends the definition given in Remark 2.12 for tableaux of shape \( \lambda \).) Now, the analogue of Proposition 2.13 is the following statement:

**Proposition 11.1.** Let \( \lambda \) and \( \mu \) be two partitions such that \( \mu \subseteq \lambda \). Then, for any total order \( \mathcal{L} \) on the positive integers,

\[
 s_{\lambda/\mu} = \sum_T x^{\text{cont}(T)}
\]

as \( T \) runs through all \( \mathcal{L} \text{-column-strict tableaux} \) of shape \( \lambda/\mu \).

The proof of Proposition 11.1 is completely analogous to the proof of Proposition 2.13.

(b) If \( \alpha \) and \( \beta \) are two partitions such that \( \beta \subseteq \alpha \), then let \( Y (\alpha/\beta) \) denote the skew Ferrers diagram \( \alpha/\beta \).

This is a finite set of cells in \( \{1, 2, 3, \ldots \} \) \(^2\).

We know that the skew Ferrers diagram \( \lambda'/\mu' \) can be obtained from the skew Ferrers diagram \( \lambda/\mu \) by a \( 180^\circ \) rotation. In other words, there exists a \( 180^\circ \) rotation \( r \) such that \( r(\lambda/(\lambda/\mu)) = \lambda/(\lambda'/\mu') \). Consider this \( r \).

Let \( \mathcal{L} \) be the total order on the set of all positive integers which is defined by \( \cdots <_\mathcal{L} 3 <_\mathcal{L} 2 <_\mathcal{L} 1 \) (in other words, let \( \mathcal{L} \) be the reverse of the usual total order on the set of all positive integers). Recall the definition of \( \mathcal{L} \text{-column-strict tableaux} \) that we gave in the solution of part (a) of this exercise. According to this definition, an \( \mathcal{L} \text{-column-strict tableau} \) of shape \( \lambda'/\mu' \) is an assignment of entries in \( \{1, 2, 3, \ldots \} \) to the cells of the Ferrers diagram of \( \lambda'/\mu' \) which is \( \mathcal{L} \)-increasing left-to-right in rows, and strictly \( \mathcal{L} \)-increasing top-to-bottom in columns. Since “weakly \( \mathcal{L} \)-increasing” is the same as “weakly decreasing” (because \( \mathcal{L} \) is the reverse of the usual total order on the set of all positive integers), and since “strictly \( \mathcal{L} \)-increasing” is the same as “strictly decreasing” (for the same reason), this rewrites as follows: An \( \mathcal{L} \text{-column-strict tableau} \) of shape \( \lambda'/\mu' \) is an assignment of entries in \( \{1, 2, 3, \ldots \} \) to the cells of the Ferrers diagram of \( \lambda'/\mu' \) which is weakly \( \mathcal{L} \)-increasing left-to-right in rows, and strictly \( \mathcal{L} \)-increasing top-to-bottom in columns. Hence, if \( T \) is an \( \mathcal{L} \text{-column-strict tableau} \) of shape \( \lambda'/\mu' \), then \( T \circ r \) (this composition is well-defined\(^{349}\)) is an assignment of entries in \( \{1, 2, 3, \ldots \} \) to the cells of the Ferrers diagram of \( \lambda/\mu \) which is weakly decreasing right-to-left in rows, and strictly decreasing top-to-bottom in columns. In other words, if \( T \) is an \( \mathcal{L} \text{-column-strict tableau} \) of shape \( \lambda'/\mu' \), then \( T \circ r \) is a column-strict tableau (in the usual sense) of shape \( \lambda/\mu \). Hence, we have constructed a map

\[
\{ \mathcal{L} \text{-column-strict tableaux} \text{ of shape } \lambda'/\mu' \} \rightarrow \{ \text{column-strict tableaux of shape } \lambda/\mu \},
\]

which sends every \( T \) to \( T \circ r \). This map is easily seen to be a bijection. Therefore, we can substitute \( T \circ r \) for \( T \) in the sum

\[
\sum_{T \text{ is a column-strict tableau of shape } \lambda'/\mu'} x^{\text{cont}(T)}.
\]

We thus obtain

\[
(11.17) \quad \sum_{T \text{ is a column-strict tableau of shape } \lambda'/\mu'} x^{\text{cont}(T)} = \sum_{T \text{ is an } \mathcal{L} \text{-column-strict tableau of shape } \lambda'/\mu'} x^{\text{cont}(T)} = \sum_{T \text{ is an } \mathcal{L} \text{-column-strict tableau of shape } \lambda'/\mu'} x^{\text{cont}(T)}.
\]

But Proposition 11.1 (applied to \( \lambda' \) and \( \mu' \) instead of \( \lambda \) and \( \mu \)) yields

\[
 s_{\lambda'/\mu'} = \sum_T x^{\text{cont}(T)}
\]

\(^{349}\)because \( T \) is an assignment of entries in \( \{1, 2, 3, \ldots \} \) to the cells of the Ferrers diagram of \( \lambda'/\mu' \), that is, a map \( Y(\lambda'/\mu') \rightarrow \{1, 2, 3, \ldots \} \), whereas \( r \) is a map sending \( Y(\lambda/\mu) \) to \( Y(\lambda'/\mu') \)

\(^{350}\)because \( r \) is a \( 180^\circ \) rotation, and thus interchanges “left-to-right” with “right-to-left” and interchanges “top-to-bottom” with “bottom-to-top”
as $T$ runs through all $L$-column-strict tableaux of shape $\lambda'/\mu'$. In other words,

$$s_{\lambda'/\mu'} = \sum_{T \text{ is an } L\text{-column-strict tableau of shape } \lambda'/\mu'} x^{\text{cont}(T)} = \sum_{T \text{ is a column-strict tableau of shape } \lambda'/\mu'} x^{\text{cont}(T)}$$

(by (11.171)).

But the definition of $s_{\lambda'/\mu'}$ yields $s_{\lambda'/\mu'} = \sum_T x^{\text{cont}(T)}$, where the sum ranges over all column-strict tableaux $T$ of shape $\lambda'/\mu'$. In other words,

$$s_{\lambda'/\mu'} = \sum_{T \text{ is a column-strict tableau of shape } \lambda'/\mu'} x^{\text{cont}(T)}.$$

Compared with (11.172), this yields $s_{\lambda'/\mu'} = s_{\lambda'/\mu}$. This solves part (b) of the exercise.

11.33. Solution to Exercise 2.23. Solution to Exercise 2.23. If $\varphi$ and $\psi$ are two partitions such that $\psi \subseteq \varphi$, then let $Y(\varphi/\psi)$ denote the skew Ferrers diagram $\varphi/\psi$ (this is a subset of $\{1, 2, 3, \ldots\}^2$).

Whenever $Z$ is a subset of $\mathbb{Z}^2$, we define a column-strict $Z$-tableau to be an assignment of entries in $\{1, 2, 3, \ldots\}$ to the elements of $Z$ which is weakly increasing left-to-right in rows and strictly increasing top-to-bottom in columns. It is clear that if $\varphi$ and $\psi$ are two partitions such that $\psi \subseteq \varphi$, then a column-strict tableau of shape $\varphi/\psi$ is the same as a column-strict $Y(\varphi/\psi)$-tableau. We define the notation $x^{\text{cont}(T)}$ (and therefore, $x^{\text{cont}(T)}$) for a column-strict $Z$-tableau $T$ in the same way as it is defined for a column-strict tableau of shape $\lambda'/\mu'$ (for some partitions $\lambda$ and $\mu$).

The following is now more or less obvious:

**Lemma 11.2.** Let $\varphi$ and $\psi$ be two partitions such that $\psi \subseteq \varphi$. Let $Z$ be a subset of $\mathbb{Z}^2$. Assume that the skew Ferrers diagram $\varphi/\psi$ can be obtained from $Z$ by parallel translation. Then,

$$s_{\varphi/\psi} = \sum_{T \text{ is a column-strict } Z\text{-tableau}} x^{\text{cont}(T)}.$$

**Proof of Lemma 11.2.** Let $R$ be the parallel translation which sends the set $Z$ to $Y(\varphi/\psi)$.

The definition of $s_{\varphi/\psi}$ yields $s_{\varphi/\psi} = \sum_T x^{\text{cont}(T)}$ for a column-strict tableau $T$ of shape $\varphi/\psi$. It remains to prove that the right hand side of this equality equals $\sum_T x^{\text{cont}(T)}$. To achieve this, it is clearly enough to find a bijection

$$\Gamma : (\text{the set of all column-strict tableaux of shape } \varphi/\psi) \to (\text{the set of all column-strict } Z\text{-tableaux})$$

which satisfies

$$\left(x^{\text{cont}(\Gamma(T))}\right) = x^{\text{cont}(T)}$$

for every column-strict tableau $T$ of shape $\varphi/\psi$.

But this is very easy: The bijection $\Gamma$ sends every column-strict tableau $T$ of shape $\varphi/\psi$ to the column-strict $Z$-tableau $T \circ R$. (The notation $T \circ R$ makes sense because $T$, being a column-strict tableau of shape $\varphi/\psi$, is an assignment of entries in $\{1, 2, 3, \ldots\}$ to the cells of the skew Ferrers diagram $\varphi/\psi$, that is, a map $Y(\varphi/\psi) \to \{1, 2, 3, \ldots\}$. Visually speaking, $T \circ R$ is the result of moving the tableau $T$ so that it takes up the cells of $Z$ rather than the cells of $\varphi/\psi$.) Lemma 11.2 is proven.

Now, let us return to the solution of the exercise. Clearly, the subsets $F_{\text{rows} \leq k}$ and $F_{\text{rows} > k}$ of $F$ are disjoint, and their union is $F$.

Lemma 11.2 (applied to $\alpha$, $\beta$ and $F_{\text{rows} \leq k}$ instead of $\varphi$, $\psi$ and $Z$) yields

$$s_{\alpha/\beta} = \sum_{T \text{ is a column-strict } F_{\text{rows} \leq k}\text{-tableau}} x^{\text{cont}(T)} = \sum_{P \text{ is a column-strict } F_{\text{rows} \leq k}\text{-tableau}} x^{\text{cont}(P)}$$

(here, we renamed the summation index $T$ as $P$). Also, Lemma 11.2 (applied to $\gamma$, $\delta$ and $F_{\text{rows} > k}$ instead of $\varphi$, $\psi$ and $Z$) yields

$$s_{\gamma/\delta} = \sum_{T \text{ is a column-strict } F_{\text{rows} > k}\text{-tableau}} x^{\text{cont}(T)} = \sum_{Q \text{ is a column-strict } F_{\text{rows} > k}\text{-tableau}} x^{\text{cont}(Q)}$$
(here, we renamed the summation index $T$ as $Q$). Multiplying the identities (11.173) and (11.174), we obtain

\[
s_{\alpha/\beta} s_{\gamma/\delta} = \left( \sum_{P \text{ is a column-strict } F_{\text{rows}} \leq k \text{-tableau}} x^{cont(P)} \right) \left( \sum_{Q \text{ is a column-strict } F_{\text{rows}} > k \text{-tableau}} x^{cont(Q)} \right)
\]

\[
= \sum_{P \text{ is a column-strict } Q \text{ is a column-strict } F_{\text{rows}} \leq k \text{-tableau}} \sum_{F_{\text{rows}} > k \text{-tableau}} x^{cont(P)} x^{cont(Q)}
\]

(11.175)

\[
= \sum_{(P,Q) \in \{ \text{the set of all column-strict } F_{\text{rows}} \leq k \text{-tableaux}\} \times \{ \text{the set of all column-strict } F_{\text{rows}} > k \text{-tableaux}\}} x^{cont(P)} x^{cont(Q)}.
\]

On the other hand, the definition of $s_{\lambda/\mu}$ yields

\[
s_{\lambda/\mu} = \sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu} x^{cont(T)}.
\]

(11.176)

Our goal is to prove that the left-hand side of (11.176) equals the left-hand side of (11.175). For this, it is clearly enough to show that the right-hand side of (11.176) equals the right-hand side of (11.175). But to achieve this, it clearly suffices to exhibit a bijection

\[
\Phi: (\text{the set of all column-strict tableaux of shape } \lambda/\mu) \\
\rightarrow (\text{the set of all column-strict } F_{\text{rows}} \leq k \text{-tableaux}) \times (\text{the set of all column-strict } F_{\text{rows}} > k \text{-tableaux})
\]

which has the property that

(11.177)

\[
\text{if } (P,Q) = \Phi(T) \text{ for some column-strict tableau } T \text{ of shape } \lambda/\mu, \text{ then } x^{cont(P)} x^{cont(Q)} = x^{cont(T)}.
\]

We claim that such a bijection $\Phi$ can be defined by

(11.178)

\[
\Phi(T) = (T |_{F_{\text{rows}} \leq k}, T |_{F_{\text{rows}} > k}) \text{ for every column-strict tableau } T \text{ of shape } \lambda/\mu.
\]

Indeed, it is clear that we can define a map

\[
\Phi: (\text{the set of all column-strict tableaux of shape } \lambda/\mu) \\
\rightarrow (\text{the set of all column-strict } F_{\text{rows}} \leq k \text{-tableaux}) \times (\text{the set of all column-strict } F_{\text{rows}} > k \text{-tableaux})
\]

by (11.178), and that this map $\Phi$ satisfies (11.177). All that remains to be proven is that this map $\Phi$ is a bijection. It is clear that $\Phi$ is injective, so we only need to prove that $\Phi$ is surjective.

Let $(P,Q) \in (\text{the set of all column-strict } F_{\text{rows}} \leq k \text{-tableaux}) \times (\text{the set of all column-strict } F_{\text{rows}} > k \text{-tableaux})$ be arbitrary. We are going to prove that $(P,Q)$ lies in the image of $\Phi$.

Define a map $T: F \rightarrow \{1,2,3,\ldots\}$ by setting

\[
(T(p)) = \begin{cases} P(p), & \text{if } p \in F_{\text{rows}} \leq k; \\ Q(p), & \text{if } p \in F_{\text{rows}} > k \end{cases} \quad \text{for all } p \in F.
\]

This map $T$ is clearly well-defined (since the subsets $F_{\text{rows}} \leq k$ and $F_{\text{rows}} > k$ of $F$ are disjoint, and their union is $F$), and thus is an assignment of entries in $\{1,2,3,\ldots\}$ to the cells of the skew Ferrers diagram $\lambda/\mu$ (since $F$ is the set of those cells). It furthermore satisfies $T |_{F_{\text{rows}} \leq k} = P$ and $T |_{F_{\text{rows}} > k} = Q$. We shall now show that this assignment $T$ is a column-strict tableau of shape $\lambda/\mu$.

This rests on the following observation:

Assertion A: Let $c$ and $d$ be two cells lying in $F$. Assume that the cells $c$ and $d$ either lie in one and the same row, or lie in one and the same column. Then, either both $c$ and $d$ belong to $F_{\text{rows}} \leq k$, or both $c$ and $d$ belong to $F_{\text{rows}} > k$.

Assertion A is an easy consequence of our assumption that $\mu_k \geq \lambda_{k+1}$. Now, we want to prove that $T$ is a column-strict tableau of shape $\lambda/\mu$. To do so, we need to check that $T$ is weakly increasing.

\[^{351} \text{Proof of Assertion A: Assume the contrary. Then, one of the cells } c \text{ and } d \text{ belongs to } F_{\text{rows}} \leq k, \text{ whereas the other belongs to } F_{\text{rows}} > k. \text{ We WLOG assume that } c \text{ belongs to } F_{\text{rows}} \leq k, \text{ whereas } d \text{ belongs to } F_{\text{rows}} > k \text{ (since otherwise, we can simply switch } c \text{ with } d).\]
left-to-right in rows, and strictly increasing top-to-bottom in columns. We will only prove the latter part of this statement, as the former part is proven analogously. So we are going to show that $T$ is strictly increasing top-to-bottom in columns. In other words, we are going to show that if $c$ and $d$ are two cells of $\lambda/\mu$ lying in one and the same column, with $d$ lying strictly further south than $c$, then $T(c) < T(d)$. Indeed, consider two such cells $c$ and $d$. Assertion A shows that either both $c$ and $d$ belong to $F_{\text{rows} \leq k}$, or both $c$ and $d$ belong to $F_{\text{rows} > k}$. Let us WLOG assume that we are in the first of these two cases (the other case is exactly analogous). Then, both $c$ and $d$ belong to $F_{\text{rows} \leq k}$, so that we have $T(c) = (T|_{F_{\text{rows} \leq k}})(c)$ and $T(d) = (T|_{F_{\text{rows} \leq k}})(d)$. Since $T|_{F_{\text{rows} \leq k}} = P$, these two equalities rewrite as $T(c) = P(c)$ and $T(d) = P(d)$. But since $P$ is strictly increasing top-to-bottom in columns (because $P$ is a column-strict tableau), we have $P(c) < P(d)$, and thus $T(c) = P(c) < P(d) = T(d)$. Thus, we have proven that $T(c) < T(d)$. This completes the proof that $T$ is a column-strict tableau of shape $\lambda/\mu$. Hence, $\Phi(T)$ is well-defined, and the definition of $\Phi(T)$ shows that $\Phi(T) = \left( \frac{T|_{F_{\text{rows} \leq k}}}{P}, \frac{T|_{F_{\text{rows} > k}}}{Q} \right) = (P,Q)$. Thus, $(P,Q)$ lies in the image of $\Phi$.

Now, let us forget that we fixed $(P,Q)$. We thus have shown that every $(P,Q) \in \left( \text{the set of all column-strict } F_{\text{rows} \leq k}\text{-tableaux} \right) \times \left( \text{the set of all column-strict } F_{\text{rows} > k}\text{-tableaux} \right)$ lies in the image of $\Phi$. In other words, the map $\Phi$ is surjective, which (as we know that $\Phi$ is injective) yields that $\Phi$ is a bijection. As explained above, this completes the solution of Exercise 2.23.

11.34. Solution to Exercise 2.25. Solution to Exercise 2.25. As usual, let $T$ denote the twist map $\Lambda \otimes \Lambda \rightarrow \Lambda \otimes \Lambda$ (that is, the $k$-linear map sending every $c \otimes d \in \Lambda \otimes \Lambda$ to $d \otimes c$). By the definition of “cocommutative”, we know that the Hopf algebra $\Lambda$ is cocommutative if and only if the diagram

$\begin{array}{ccc} 
\Lambda \otimes \Lambda & \xrightarrow{T} & \Lambda \otimes \Lambda \\
\Delta & \downarrow & \Delta \\
\Lambda & \xrightarrow{\Delta} & \Lambda \otimes \Lambda 
\end{array}$

commutes. Hence, in order to solve Exercise 2.25(a), it is enough to check that the diagram (11.179) commutes.

The set $\{h_n\}_{n=1,2,...}$ generates the $k$-algebra $\Lambda$ (due to Proposition 2.28). In other words, the set $\{h_1, h_2, h_3, ... \}$ is a generating set of the $k$-algebra $\Lambda$.

By the axioms of a bialgebra, the comultiplication $\Delta$ of $\Lambda$ is a $k$-algebra homomorphism (since $\Lambda$ is a bialgebra). Hence, $T \circ \Delta$ also is a $k$-algebra homomorphism (since $T$ and $\Delta$ are $k$-algebra homomorphisms).

Write the cell $c$ in the form $(a,b)$, so that $c$ lies in row $a$ and column $b$. Write the cell $d$ in the form $(a',b')$, so that $d$ lies in row $a'$ and column $b'$.

We have $(a,b) = c \in F_{\text{rows} \leq k}$, so that $a \leq k$ (by the definition of $F_{\text{rows} \leq k}$). We have $(a',b') = d \in F_{\text{rows} > k}$, so that $a' > k$ (by the definition of $F_{\text{rows} > k}$).

We have $a \leq k < a'$ (since $a' > k$), hence $c \neq a'$. Thus, the cells $c$ and $d$ lie in different rows (since the cell $c$ lies in row $a$, whereas the cell $d$ lies in row $a'$). As a consequence, the cells $c$ and $d$ must lie in one and the same column (since we assumed that the cells $c$ and $d$ either lie in one and the same row, or lie in one and the same column). In other words, $b = b'$ (since the cell $c$ lies in column $b$, while the cell $d$ lies in column $b'$).

Now, $(a,b) = c \in F = Y(\lambda/\mu)$. By the definition of $Y(\lambda/\mu)$, this shows that $\mu_a < b \leq \lambda_a$. Similarly, $\mu_{a'} < b' \leq \lambda_{a'}$.

We have $a' > k$, hence $a' \geq k + 1$ (since $a'$ and $k$ are integers), thus $k + 1 \leq a'$.

Since $\mu$ is a partition, we have $\mu_1 \geq \mu_2 \geq \mu_3 \geq ...$. Hence, $\mu_a \geq \mu_1$ (since $a \leq k$). Thus, $\mu_a \geq \mu_k \geq \lambda_k + 1$. Since $\lambda$ is a partition, we have $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq ...$. Thus, $\lambda_{k+1} \geq \lambda_{a'}$ (since $k + 1 \leq a'$). Thus, $\mu_a \geq \lambda_{k+1} \geq \lambda_{a'}$, so that $\lambda_{a'} \leq \mu_a$ and thus $b' \leq \lambda_{a'} \leq \mu_a < b = b'$, which is absurd. This contradiction proves that our assumption was wrong, and thus Assertion A is proven.
Every positive integer \( n \) satisfies

\[
(T \circ \Delta) (h_n) = T \left( \sum_{i+j=n} h_i \otimes h_j \right) = T \left( \sum_{i+j=n} h_j \otimes h_i \right) = \Delta h_n
\]

(by the definition of the twist map \( T \))

\[
= \sum_{i+j=n} h_i \otimes h_j = \Delta h_n
\]

(by Proposition 2.24(iii))

In other words, for every positive integer \( n \), the two maps \( T \circ \Delta \) and \( \Delta \) are equal to each other on the element \( h_n \). In other words, the two maps \( T \circ \Delta \) and \( \Delta \) are equal to each other on the set \( \{ h_1, h_2, h_3, \ldots \} \). Hence, the two maps \( T \circ \Delta \) and \( \Delta \) are equal to each other on a generating set of the \( k \)-algebra \( \Lambda \) (since the set \( \{ h_1, h_2, h_3, \ldots \} \) is a generating set of the \( k \)-algebra \( \Lambda \)). Since these two maps \( T \circ \Delta \) and \( \Delta \) are \( k \)-algebra homomorphisms, this shows that the two maps \( T \circ \Delta \) and \( \Delta \) must be identical (because if two \( k \)-algebra homomorphisms with the same domain and the same target are equal to each other on a generating set of their domain, then these two homomorphisms must be identical). In other words, \( T \circ \Delta = \Delta \). Hence, the diagram (11.179) commutes. As we know, this shows that the Hopf algebra \( \Lambda \) is cocommutative. This solves Exercise 2.25(a).

(b) Let \( \lambda \) and \( \nu \) be two partitions. We have shown above that \( T \circ \Delta = \Delta \). Hence, \( \Delta = T \circ \Delta \). Applying both sides of this equality to \( s_{\lambda/\nu} \), we obtain

\[
\Delta s_{\lambda/\nu} = (T \circ \Delta) (s_{\lambda/\nu}) = T \left( \sum_{\mu \subseteq \lambda, \nu \subseteq \mu} s_{\mu/\nu} \otimes s_{\lambda/\mu} \right) = T \left( \sum_{\mu \subseteq \lambda, \nu \subseteq \mu} s_{\mu/\nu} \otimes s_{\lambda/\mu} \right)
\]

(by the definition of the twist map \( T \)).

This solves Exercise 2.25(b).

11.35. Solution to Exercise 2.26. Solution to Exercise 2.26. (a) Whenever \( \alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots) \) is a weak composition satisfying \( \alpha_i = 0 \) for every \( i > n \), the monomial \( x^\alpha \) is a monomial in \( k[x_1, x_2, \ldots, x_n] \). This will be often used in the following.

Recall that \( s_{\lambda/\mu} \) is defined as \( s_{\lambda/\mu} = \sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu} x^{\text{cont}(T)} \). Now, \( s_{\lambda/\mu} (x_1, x_2, \ldots, x_n) \) is the result of substituting \( x_1, x_2, \ldots, x_n, 0, 0, 0, \ldots \) for \( x_1, x_2, x_3, \ldots \) in \( s_{\lambda/\mu} \). This substitution has the following effect on any given monomial \( x^\alpha \):

- if none of the indeterminates \( x_{n+1}, x_{n+2}, x_{n+3}, \ldots \) occur in this monomial \( x^\alpha \), then the monomial \( x^\alpha \) stays fixed;
• otherwise, the monomial $x^n$ goes to 0.

Hence, the effect of this substitution on the power series $s_{\lambda/\mu} = \sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu} x^{\text{cont}(T)}$ is that:

• every addend in the sum $\sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu} x^{\text{cont}(T)}$ for which none of the indeterminates $x_{n+1}, x_{n+2}, x_{n+3}, \ldots$ occur in the monomial $x^{\text{cont}(T)}$ stays fixed;

• all other addends go to 0.

The result of the substitution is therefore $\sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu; \text{none of the indeterminates } x_{n+1}, x_{n+2}, x_{n+3}, \ldots \text{ occur in the monomial } x^{\text{cont}(T)}} s_{\lambda/\mu} (x_1, x_2, \ldots, x_n) = \sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu; \text{none of the indeterminates } x_{n+1}, x_{n+2}, x_{n+3}, \ldots \text{ occur in the monomial } x^{\text{cont}(T)}} x^{\text{cont}(T)}$.

But this rewrites as

$$s_{\lambda/\mu} (x_1, x_2, \ldots, x_n) = \sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu; \text{all entries of } T \text{ belong to } \{1, 2, \ldots, n\}} x^{\text{cont}(T)}$$

(b) Let $\lambda$ be a partition having more than $n$ parts. We have to prove that $s_{\lambda} (x_1, x_2, \ldots, x_n) = 0$. Since $s_{\lambda} (x_1, x_2, \ldots, x_n) = s_{\lambda/\emptyset} (x_1, x_2, \ldots, x_n) = \sum_{T \text{ is a column-strict tableau of shape } \lambda/\emptyset; \text{all entries of } T \text{ belong to } \{1, 2, \ldots, n\}} x^{\text{cont}(T)}$, applying Exercise 2.26(a), this goal will clearly be achieved if we can show that the sum $\sum_{T \text{ is a column-strict tableau of shape } \lambda/\emptyset; \text{all entries of } T \text{ belong to } \{1, 2, \ldots, n\}} x^{\text{cont}(T)}$ is empty, i.e., that there exists no column-strict tableau $T$ of shape $\lambda/\emptyset$ such that all entries of $T$ belong to $\{1, 2, \ldots, n\}$.

Assume the contrary. Thus, there exists a column-strict tableau $T$ of shape $\lambda/\emptyset$ such that all entries of $T$ belong to $\{1, 2, \ldots, n\}$. This tableau has more than $n$ rows (since the partition $\lambda$ has more than $n$ parts), and thus the first column of this tableau must have more than $n$ entries. These entries must be strictly increasing top-to-bottom (since the entries of a column-strict tableau strictly increase top-to-bottom along columns) and hence be pairwise distinct, but at the same time (like all entries of $T$) they must belong to $\{1, 2, \ldots, n\}$. So we have found more than $n$ entries which are pairwise distinct and belong to $\{1, 2, \ldots, n\}$. This contradicts the fact that the set $\{1, 2, \ldots, n\}$ does not have more than $n$ pairwise distinct elements. This contradiction concludes our proof, and Exercise 2.26(b) is solved.

11.36. Solution to Exercise 2.38. Solution to Exercise 2.38. Recall that $(s_\kappa)_{\kappa \in \text{Par}}$ is a basis of the $k$-module $\Lambda$. Hence, $(s_\kappa \otimes s_\lambda)_{(\kappa, \lambda) \in \text{Par} \times \text{Par}}$ is a basis of the $k$-module $\Lambda \otimes \Lambda$. We will refer to this basis as the “Schur basis” of $\Lambda \otimes \Lambda$.

Now, the first diagram in (1.8) commutes when $A$ is set to $\Lambda$ (since $\Lambda$ is a bialgebra). In other words,

$$(11.180) \quad \Delta \circ m = (m \otimes m) \circ (\text{id} \otimes T \otimes \text{id}) \circ (\Delta \otimes \Delta) : \Lambda \otimes \Lambda \rightarrow \Lambda \otimes \Lambda.$$

Fix four partitions $\varphi$, $\psi$, $\kappa$ and $\lambda$. Applying both sides of the equality (11.180) to $s_\varphi \otimes s_\psi$, we obtain

$$(\Delta \circ m) (s_\varphi \otimes s_\psi) = ((m \otimes m) \circ (\text{id} \otimes T \otimes \text{id}) \circ (\Delta \otimes \Delta)) (s_\varphi \otimes s_\psi).$$
Comparing the coefficients before $s_k \otimes s_\lambda$ in this equality, we obtain
\[ \sum_{\rho \in \Par} c_{\rho,\lambda,\varphi,\psi}^\rho = \sum_{(\alpha,\beta,\gamma,\delta) \in \Par^4} c_{\alpha,\gamma,\beta,\delta}^\alpha c_{\alpha,\beta}^\beta c_{\gamma,\delta}^\gamma c_{\varphi,\psi}^\psi \]
(after a straightforward computation using (2.26) and Corollary 2.35). This solves the exercise.

11.37. Solution to Exercise 2.39. Solution to Exercise 2.39. Before we step to the solution of this problem, let us make some general observations.

- Every two partitions $\lambda$ and $\mu$ satisfy $s_{\lambda/\mu} = \sum_\nu c_{\mu,\nu}^\lambda s_\nu$, where the sum ranges over all partitions $\nu$ (according to Remark 2.37). In other words, every two partitions $\lambda$ and $\mu$ satisfy
\[ s_{\lambda/\mu} = \sum_{\nu \in \Par} c_{\mu,\nu}^\lambda s_\nu. \]  
(11.181)

- On the other hand, the first equality in (2.26) yields
\[ s_\mu s_\nu = \sum_{\lambda \in \Par} c_{\mu,\nu}^\lambda s_\lambda \] for any two partitions $\mu$ and $\nu$.
(11.182)

(a) Let $\mu$ be a partition. We have
\[ \sum_{\lambda \in \Par} s_\lambda (x) \cdot s_{\lambda/\mu} (y) = \sum_{\lambda \in \Par} s_\lambda (x) \cdot \sum_{\nu \in \Par} c_{\mu,\nu}^\lambda s_\nu (y). \]  
(11.183)

On the other hand, (2.24) yields
\[ \prod_{i,j=1}^\infty (1 - x_i y_j)^{-1} = \sum_{\lambda \in \Par} s_\lambda (x) \cdot s_\lambda (y) = \sum_{\nu \in \Par} s_\nu (x) \cdot s_\nu (y) \]
(here, we renamed the summation index $\lambda$ as $\nu$). Multiplying this equality with $s_\mu (x)$, we obtain
\[ s_\mu (x) \cdot \prod_{i,j=1}^\infty (1 - x_i y_j)^{-1} = s_\mu (x) \cdot \sum_{\nu \in \Par} s_\nu (x) \cdot s_\nu (y) = \sum_{\nu \in \Par} s_\nu (x) \cdot s_\nu (y) = \sum_{\nu \in \Par} \left( \sum_{\lambda \in \Par} c_{\mu,\nu}^\lambda s_\lambda (x) \right) s_\nu (y) \]
\[ = \sum_{\nu \in \Par} \left( \sum_{\lambda \in \Par} c_{\mu,\nu}^\lambda s_\lambda (x) \right) s_\nu (y) \]
\[ = \sum_{\lambda \in \Par} s_\lambda (x) \cdot \sum_{\nu \in \Par} c_{\mu,\nu}^\lambda s_\nu (y). \]

Compared with (11.183), this yields
\[ \sum_{\lambda \in \Par} s_\lambda (x) \cdot s_{\lambda/\mu} (y) = s_\mu (x) \cdot \prod_{i,j=1}^\infty (1 - x_i y_j)^{-1}. \]

This solves Exercise 2.39(a).

(b) Let $k[[x, y, z, w]]$ denote the ring
\[ k[[x_1, x_2, x_3, ..., y_1, y_2, y_3, ..., z_1, z_2, z_3, ..., w_1, w_2, w_3, ...]]. \]
This ring clearly contains $k[[x, y]]$ as a subring. We will use the obvious abbreviations for variable sets: $x = (x_1, x_2, x_3, ...)$, $(x, z) = (x_1, x_2, x_3, ..., z_1, z_2, z_3, ...)$, etc.
Every partition \( \lambda \) satisfies

\[
\begin{align*}
s_{\lambda}(y, x) &= s_{\lambda}(x, y) = \sum_{\mu \leq \lambda} s_{\mu}(x) s_{\lambda/\mu}(y) \quad \text{(by (2.13))} \\
&= \sum_{\mu \in \text{Par}} s_{\mu}(x) s_{\lambda/\mu}(y) - \sum_{\mu \in \text{Par}; \mu \not\subseteq \lambda} s_{\mu}(x) s_{\lambda/\mu}(y) = \sum_{\mu \in \text{Par}} s_{\mu}(x) s_{\lambda/\mu}(y) - \sum_{\mu \in \text{Par}; \mu \not\subseteq \lambda} s_{\mu}(x) = 0
\end{align*}
\]

Applying this equality to the variables \( z \) and \( x \) instead of \( x \) and \( y \), we obtain

\[
(11.184) \quad s_{\lambda}(x, z) = \sum_{\mu \in \text{Par}} s_{\mu}(x) s_{\lambda/\mu}(x).
\]

Applying this equality to the variables \( y \) and \( w \) instead of \( x \) and \( z \), we obtain

\[
(11.185) \quad s_{\lambda}(y, w) = \sum_{\mu \in \text{Par}} s_{\mu}(w) s_{\lambda/\mu}(y) = \sum_{\nu \in \text{Par}} s_{\nu}(w) s_{\lambda/\nu}(y)
\]

(here, we renamed the summation index \( \mu \) as \( \nu \)) for any partition \( \lambda \).

Now,

\[
\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \prod_{a \in (x, x); b \in (y, y)} (1 - a b)^{-1} = \sum_{\lambda \in \text{Par}} \underbrace{s_{\lambda}(x, z)}_{(by \ (11.184))} \underbrace{s_{\lambda}(y, w)}_{(by \ (11.185))} = \sum_{\mu \in \text{Par}} s_{\mu}(x) s_{\lambda/\mu}(y)
\]

(by (2.24), applied to the variable sets \( (x, z) \) and \( (y, w) \) instead of \( x \) and \( y \))

\[
\begin{align*}
&= \sum_{\lambda \in \text{Par}} \left( \sum_{\mu \in \text{Par}} s_{\mu}(z) s_{\lambda/\mu}(x) \right) \left( \sum_{\nu \in \text{Par}} s_{\nu}(w) s_{\lambda/\nu}(y) \right) \\
&= \sum_{\mu \in \text{Par}} \sum_{\nu \in \text{Par}} s_{\mu}(z) s_{\nu}(w) \left( \sum_{\lambda \in \text{Par}} s_{\lambda/\mu}(x) s_{\lambda/\nu}(y) \right)
\end{align*}
\]

(11.186)

On the other hand, every partition \( \kappa \) satisfies

\[
(11.187) \quad \sum_{\lambda \in \text{Par}} s_{\lambda}(x) s_{\lambda/\kappa}(y) = s_{\kappa}(x) \cdot \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}
\]

(by Exercise 2.39(a), applied to \( \kappa \) instead of \( \mu \)). Applying this to the variable sets \( w \) and \( x \) instead of \( x \) and \( y \), we obtain

\[
\begin{align*}
\sum_{\lambda \in \text{Par}} s_{\lambda}(w) s_{\lambda/\kappa}(x) &= s_{\kappa}(w) \cdot \prod_{i,j=1}^{\infty} (1 - w_i x_j)^{-1} \\
&= \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}
\end{align*}
\]

(11.188)

(here, we renamed the index \( (i,j) \) as \( (j,i) \))

\[
\begin{align*}
&= s_{\kappa}(w) \cdot \prod_{j,i=1}^{\infty} (1 - x_i w_j)^{-1} = s_{\kappa}(w) \cdot \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}
\end{align*}
\]
for every partition $\kappa$. But applying both sides of the identity \((11.187)\) to the variable set $z$ instead of $x$, and renaming the summation index $\lambda$ as $\mu$ on the left hand side of this equality, we obtain

\[(11.189) \quad \sum_{\mu \in \text{Par}} s_\mu(z) s_{\mu/\kappa}(y) = s_\kappa(z) \cdot \prod_{i,j=1}^\infty (1 - z_i y_j)^{-1}.
\]

Now, applying \((2.24)\) to the variable sets $z$ and $w$ instead of $x$ and $y$, we obtain

\[\prod_{i,j=1}^\infty (1 - z_i y_j)^{-1} = \sum_{\lambda \in \text{Par}} s_\lambda(z) s_\lambda(w) = \sum_{\kappa \in \text{Par}} s_\kappa(z) s_\kappa(w).
\]

Hence,

\[\prod_{i,j=1}^\infty (1 - x_i y_j)^{-1} \prod_{i,j=1}^\infty (1 - x_i w_j)^{-1} \prod_{i,j=1}^\infty (1 - z_i y_j)^{-1} \prod_{i,j=1}^\infty (1 - z_i w_j)^{-1}
= \sum_{\kappa \in \text{Par}} s_\kappa(z) s_\kappa(w)
= \sum_{\lambda \in \text{Par}} s_\lambda(z) s_\lambda(y) \prod_{i,j=1}^\infty (1 - x_i y_j)^{-1}
= \sum_{\mu \in \text{Par}} \sum_{\lambda \in \text{Par}} s_\mu(z) s_{\lambda/\kappa}(w) \left( s_{\lambda/\mu}(x) s_{\mu/\kappa}(y) \right) \prod_{i,j=1}^\infty (1 - x_i y_j)^{-1}
= \sum_{\mu \in \text{Par}} \sum_{\nu \in \text{Par}} s_\mu(z) s_{\nu}(w) \left( s_{\nu/\mu}(x) s_{\mu/\nu}(y) \right) \prod_{i,j=1}^\infty (1 - x_i y_j)^{-1}
\]

(here, we renamed the summation indices $\lambda$ and $\kappa$ as $\nu$ and $\rho$, respectively). Comparing this with \((11.186)\), we obtain

\[
\sum_{\mu \in \text{Par}} \sum_{\nu \in \text{Par}} s_\mu(z) s_{\nu}(w) \left( s_{\lambda/\mu}(x) s_{\lambda/\nu}(y) \right) \prod_{i,j=1}^\infty (1 - x_i y_j)^{-1}.
\]

We can regard this as an identity in the ring $\textbf{k} [[x,y]][[z,w]]$ of formal power series in the variables $(z,w) = (z_1, z_2, z_3, ..., w_1, w_2, w_3, ...)$ over the ring $\textbf{k} [[x,y]]$. Extracting the coefficients in front of $s_\alpha(z) s_\beta(w)$ in this identity (this is well-defined because the family $(s_\mu(z) s_\nu(w))_{(\mu,\nu) \in \text{Par} \times \text{Par}} \in ((\textbf{k} [[x,y]])[[z,w]])_{\text{Par} \times \text{Par}}$ is $\textbf{k} [[x,y]]$-linearly independent), we obtain

\[\sum_{\lambda \in \text{Par}} s_{\lambda/\alpha}(x) s_{\lambda/\beta}(y) = \left( \sum_{\rho \in \text{Par}} s_{\beta/\rho}(x) s_{\alpha/\rho}(y) \right) \prod_{i,j=1}^\infty (1 - x_i y_j)^{-1}.
\]

This solves Exercise 2.39(b).
11.38. **Solution to Exercise 2.41.**  
Solution to Exercise 2.41. We know that \((s_\lambda)_{\lambda \in \text{Par}}\) is a basis of the \(k\)-module \(\Lambda\), but we can also say something more specific: For every \(n \in \mathbb{N}\), the family \((s_\lambda)_{\lambda \in \text{Par}_n}\) is a basis of the \(k\)-module \(\Lambda_n\).  
The basis \((s_\lambda)_{\lambda \in \text{Par}}\) of \(\Lambda\) is orthonormal with respect to the Hall inner product. In other words,

\[(s_\lambda, s_\mu) = \delta_{\lambda, \mu} \quad \text{for any partitions } \lambda \text{ and } \mu.\]  

(a) Let \(n\) and \(m\) be two distinct nonnegative integers. Let \(f \in \Lambda_n\) and \(g \in \Lambda_m\).  
We need to prove that \((f,g) = 0\). Since this equality is \(k\)-linear in \(f\), we can WLOG assume that \(f\) is an element of the basis \((s_\lambda)_{\lambda \in \text{Par}_n}\) of \(\Lambda_n\). In other words, we can WLOG assume that \(f = s_\lambda\) for some \(\lambda \in \text{Par}_n\).  
Assume this, and similarly assume that \(g = s_\mu\) for some \(\mu \in \text{Par}_m\). These two partitions \(\lambda\) and \(\mu\) must be distinct (because their sizes differ: \(|\lambda| = n \neq m = |\mu|\)), and so they satisfy \(\delta_{\lambda, \mu} = 0\). Now,

\[
(f, g) = (s_\lambda, s_\mu) = \delta_{\lambda, \mu} \quad \text{(by (11.190))}
\]

\[
= 0.
\]

This solves Exercise 2.41(a).  
(b) Let \(n \in \mathbb{N}\) and \(f \in \Lambda_n\).  
We need to prove that \((h_n, f) = f(1)\). Since this equality is \(k\)-linear in \(f\), we can WLOG assume that \(f\) is an element of the basis \((s_\lambda)_{\lambda \in \text{Par}_n}\) of \(\Lambda_n\). In other words, we can WLOG assume that \(f = s_\lambda\) for some \(\lambda \in \text{Par}_n\). Assume this.  

We must be in one of the two cases:  
Case 1: We have \(\ell(\lambda) \leq 1\).  
Case 2: We have \(\ell(\lambda) > 1\).  
Let us consider Case 1 first. In this case, \(\ell(\lambda) \leq 1\). Hence, \(\lambda = (n)\) (since \(\lambda \in \text{Par}_n\)), so that \(s_\lambda = s_{(n)}\).  
Hence, \(f = s_\lambda = s_{(n)} = h_n\) and thus \(f(1) = h_n(1) = 1\). Compared with

\[
\begin{pmatrix}
-\frac{h_n}{s_{(n)}} & \frac{f}{s_{(n)}}
\end{pmatrix} = \begin{pmatrix}
(s_{(n)}, s_{(n)})
\end{pmatrix} = \delta_{(n), (n)} \quad \text{(by (11.190))}
\]

\[
= 1,
\]

this yields \((h_n, f) = f(1)\). Hence, \((h_n, f) = f(1)\) is proven in Case 1.  
Let us now consider Case 2. In this case, \(\ell(\lambda) > 1\), so the partition \(\lambda\) has more than 1 part. Exercise 2.26(b) (applied to 1 instead of \(n\)) thus yields \(s_\lambda(x_1) = 0\). But \(s_\lambda(1)\) can be seen as the result of substituting 1 for \(x_1\) in \(s_\lambda(x_1)\), and therefore must be 0 as well (since \(s_\lambda(x_1) = 0\)). Thus, we have \(s_\lambda(1) = 0\). Since

---

352Proof. We have \(s_\lambda \in \Lambda_{|\lambda|} = \Lambda_n\) for every \(\lambda \in \text{Par}_n\). Thus, \((s_\lambda)_{\lambda \in \text{Par}_n}\) is a family of elements of \(\Lambda_n\). This family is \(k\)-linearly independent (because it is part of the basis \((s_\lambda)_{\lambda \in \text{Par}}\) of \(\Lambda\)). It remains to prove that this family spans the \(k\)-module \(\Lambda_n\).  
Let \(f \in \Lambda_n\). We can write \(f\) in the form \(f = \sum_{\lambda \in \text{Par}} a_\lambda s_\lambda\) for some family \((a_\lambda)_{\lambda \in \text{Par}} \in k^{\text{Par}}\) of scalars. Consider this \((a_\lambda)_{\lambda \in \text{Par}}\) and notice that

\[
f = \sum_{\lambda \in \text{Par}} a_\lambda s_\lambda = \sum_{\lambda \in \text{Par}_n} a_\lambda s_\lambda + \sum_{\lambda \in \text{Par}; \lambda \notin \text{Par}_n} a_\lambda s_\lambda,
\]

so that \(f - \sum_{\lambda \in \text{Par}_n} a_\lambda s_\lambda = \sum_{\lambda \notin \text{Par}_n} a_\lambda s_\lambda\). The left hand side of this latter equality is homogeneous of degree \(n\) (since \(f\) and all the \(s_\lambda\) with \(\lambda \in \text{Par}_n\) are homogeneous of degree \(n\)), while the right hand side is a sum of homogeneous elements of degrees different from \(n\). So the only way these two sides can be equal is if they both are 0. In particular, this shows that the left hand side is 0. In other words, \(f - \sum_{\lambda \in \text{Par}_n} a_\lambda s_\lambda = 0\), so that \(f = \sum_{\lambda \in \text{Par}_n} a_\lambda s_\lambda\). Hence, \(f\) is a \(k\)-linear combination of the \(s_\lambda\) for \(\lambda \in \text{Par}_n\).  
Since we have proven this for every \(f \in \Lambda_n\), we thus conclude that the family \((s_\lambda)_{\lambda \in \text{Par}_n}\) spans the \(k\)-module \(\Lambda_n\), qed.
f = s_\lambda, we now have f(1) = s_\lambda(1) = 0. Compared with
\[
\begin{pmatrix}
  h_n \\
  \vdots
\end{pmatrix}
\begin{pmatrix}
  f(n)
\end{pmatrix}
= \begin{pmatrix}
  s(n) \\
  \vdots
\end{pmatrix}
= \delta(n, \lambda) 
\]  
(by (11.190))
\[
= 0 \quad \text{(since } n \neq \lambda \text{ (because } \ell((n)) = 1 < \ell(\lambda)),
\]
this yields \( (h_n, f) = f(1) \). Hence, \( (h_n, f) = f(1) \) is proven in Case 2.

Now, \( (h_n, f) = f(1) \) is proven in both Cases, which solves Exercise 2.41(b).

11.39. Solution to Exercise 2.46. Solution to Exercise 2.46. (a) This is a well-known fact; for a proof, see Theorem 5.3 in Keith Conrad, Universal Identities I, http://www.math.uconn.edu/~kconrad/blurbs/.

(b) Consider the endomorphism of the \( k \)-module \( A \) defined by sending \( \gamma_i \) to \( \beta_i \) for every \( i \in I \). This endomorphism is well-defined (since \( (\gamma_i)_{i \in I} \) is a basis of \( A \)) and surjective (since the family \( (\beta_i)_{i \in I} \) spans \( A \)), therefore is a \( k \)-module isomorphism (according to Exercise 2.46(a)). As a consequence, it must send the basis \( (\gamma_i)_{i \in I} \) of \( A \) to a basis of \( A \) (because a \( k \)-module isomorphism sends any basis to a basis). Since it sends the basis \( (\gamma_i)_{i \in I} \) to \( (\beta_i)_{i \in I} \), this yields that \( (\beta_i)_{i \in I} \) must be a basis of \( A \). This solves Exercise 2.46(b).

11.40. Solution to Exercise 2.47. Solution to Exercise 2.47. (a) This can be solved by following the proof of Corollary 2.45, but in doing so one has to be careful about how one obtains the invertibility of \( A \): it is no longer a consequence of \( A \) being a transition matrix between two bases (because we do not know in advance that \( (u_\lambda)_{\lambda \in \text{Par}} \) is a basis). Instead, one has to argue as in the footnote: The matrices \( A \) and \( B \) are block-diagonal, with each block corresponding to the partitions of size \( n \) for a given \( n \in \mathbb{N} \)\(^{353} \). In particular, they are block-diagonal matrices with each block being a square matrix of finite size. It is known that if such a matrix is right-invertible, then it is left-invertible\(^{354} \); therefore, since \( AB = I \), we conclude that \( A \) is invertible.

(b) We know that for every partition \( \lambda \), the symmetric functions \( h_\lambda \) and \( m_\lambda \) are two homogeneous elements of \( \Lambda \), both of degree \( |\lambda| \). We also know (from Proposition 2.43) that
\[
\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \text{Par}} h_\lambda(x) m_\lambda(y).
\]

Compared with (2.24), this yields \( \sum_{\lambda \in \text{Par}} s_\lambda(x) s_\lambda(y) = \sum_{\lambda \in \text{Par}} h_\lambda(x) m_\lambda(y) \). Thus, Exercise 2.47(a) (applied to \( u_\lambda = h_\lambda \) and \( v_\lambda = m_\lambda \)) yields that \( (h_\lambda)_{\lambda \in \text{Par}} \) and \( (m_\lambda)_{\lambda \in \text{Par}} \) are \( k \)-bases of \( \Lambda \), and actually are dual bases with respect to the Hall inner product on \( \Lambda \). In particular, \( (h_\lambda)_{\lambda \in \text{Par}} \) is a \( k \)-basis of \( \Lambda \). This solves Exercise 2.47(b).

Remark. We can use Exercise 2.47(b) to prove that \( (e_\lambda)_{\lambda \in \text{Par}} \) is a \( k \)-basis of \( \Lambda \) as well (in a different way than we have done in the proof of Proposition 2.17). In fact, let us sketch this proof. We are going to proceed similarly to the above proof of Proposition 2.28, but with the roles of the families \( (e_n)_{n \geq 1} \) and \( (h_n)_{n \geq 1} \) switched.

We know (from Exercise 2.47(b)) that \( (h_\lambda)_{\lambda \in \text{Par}} \) is a \( k \)-basis of \( \Lambda \). In other words, the family \( (h_n)_{n \geq 1} \) is algebraically independent and generates the \( k \)-algebra \( \Lambda \). Thus, we can define a \( k \)-algebra homomorphism \( \omega' : \Lambda \to \Lambda \) by setting
\[
(11.191) \quad \omega'(h_n) = e_n \quad \text{for every } n \geq 1.
\]

\(^{353}\)It is here that we are using the assumption that \( u_\lambda \) and \( v_\lambda \) are homogeneous of degree \( |\lambda| \).

\(^{354}\)In fact, it is clearly enough to prove this statement for square matrices of finite size (because block-diagonal matrices can be inverted block-by-block). But for square matrices, this follows from the fact that a surjective endomorphism of a finitely-generated \( k \)-module is an isomorphism (Exercise 2.46(a)).
The identical form of the two recursions in (2.18) shows that this \( \omega' \) also sends \( e_n \) to \( h_n \) for every \( n \geq 1 \). Hence, \( \omega' \) is an involutive automorphism of the \( k \)-algebra \( \Lambda \). In particular, \( \omega' \) is a \( k \)-module isomorphism \( \Lambda \to \Lambda \). By multiplicativity and (11.191), we obtain

\[
(11.192) \quad \omega' (h_\lambda) = e_\lambda \quad \text{for every partition } \lambda.
\]

Thus, the image of the basis \((h_\lambda)_{\lambda \in \mathsf{Par}}\) of \( \Lambda \) under the \( k \)-module isomorphism \( \omega' \) is the family \((e_\lambda)_{\lambda \in \mathsf{Par}}\). Therefore, this latter family \((e_\lambda)_{\lambda \in \mathsf{Par}}\) must be a basis of \( \Lambda \) (being the image of a basis under a \( k \)-module isomorphism).

11.41. **Solution to Exercise 2.48.** Solution to Exercise 2.48. When \( Q \) is a subring of \( k \), the statement of Exercise 2.48 has been proven in (2.30). However, there is no immediate way to reuse the proof of (2.30) in the general case (because this proof made use of logarithms, and these are only defined when \( Q \) is a subring of \( k \)). Nevertheless, it is possible to imitate that proof by defining a notion of "logarithmic derivative" even in the absence of a logarithm. Let us elaborate on this.

Let \( A \) be any commutative ring. Whenever \( Q \in A[[t]] \) is a formal power series with constant term 1, we define the *logarithmic derivative* of \( Q \) to be the power series \( \frac{Q'}{Q} \in A[[t]] \) (where \( Q' \) denotes the derivative of \( Q \), as usual). We denote this logarithmic derivative by \( \text{idr} \). When \( Q \) is a subring of \( A \), we have \( \text{idr} Q = \frac{d}{dt} (\log Q) \), but the concept of \( \text{idr} \) is defined even when \( \log \) is not. Some authors find it instructive to write \( \frac{d}{dt} (\log Q) \) for \( \text{idr} \), but we prefer not to do so, since this might tempt us to write things which make no sense.

It is easy to see that the map

\[
\{ Q \in A[[t]] \mid Q \text{ has constant term } 1 \} \to A[[t]], \quad Q \mapsto \text{idr} Q
\]

is continuous (where we equip \( A[[t]] \) with the usual topology – i.e., the product topology obtained by regarding the set \( A[[t]] \) as a direct product of infinitely many copies of \( A \)). Moreover, whenever \( I \) is a set and \((Q_i)_{i \in I} \in (A[[t]])^I \) is a family of formal power series with constant term 1 such that the product \( \prod_{i \in I} Q_i \) converges (with respect to our topology on \( A[[t]] \)), then

\[
(11.193) \quad \text{idr} \left( \prod_{i \in I} Q_i \right) = \sum_{i \in I} \text{idr} (Q_i).
\]

This equality can be used as a substitute for the famous property of the logarithm to take products into sums in situations where the logarithm is not defined. Furthermore, we have

\[
(11.194) \quad \text{idr} (Q^{-1}) = - \text{idr} Q
\]

---

\(^{355}\)This \( \omega' \) is, of course, precisely the \( \omega \) constructed in the proof of Proposition 2.28. But we cannot use the \( \omega \) constructed in the proof of Proposition 2.28 at this point, because its construction made use of the fact that \((e_\lambda)_{\lambda \in \mathsf{Par}}\) is a \( k \)-basis of \( \Lambda \) (and we want to prove this fact).

\(^{356}\)Proof of (11.193): Let \( I \) be a set, and let \((Q_i)_{i \in I} \in (A[[t]])^I \) be a family of formal power series with constant term 1 such that the product \( \prod_{i \in I} Q_i \) converges (with respect to our topology on \( A[[t]] \)). We need to prove (11.193). Since the map

\[
\{ Q \in A[[t]] \mid Q \text{ has constant term } 1 \} \to A[[t]], \quad Q \mapsto \text{idr} Q
\]

is continuous (and since infinite products are limits of finite products, and infinite sums are limits of finite sums), we can WLOG assume that the set \( I \) is finite. Assuming this, we recall that the product rule for differentiating a product of several power series yields

\[
\left( \prod_{i \in I} Q_i \right)' = \sum_{j \in I} Q_j' \left( \prod_{i \in I \setminus \{j\}} Q_i \right).
\]
whenever $Q \in A[[t]]$ is a formal power series with constant term 1 (where $Q^{-1}$ means the multiplicative inverse of $Q$).

Now, set $A = \Lambda$. We have

$$H(t) = \prod_{i=1}^{\infty} (1 - x_it)^{-1} = \prod_{i \in \{1, 2, 3, \ldots\}} (1 - x_it)^{-1},$$

Now, by the definition of lder $(\prod_{i \in I} Q_i)$, we have

$$\text{lder} \left( \prod_{i \in I} Q_i \right) = \frac{(\prod_{i \in I} Q_i)'}{\prod_{i \in I} Q_i} = \frac{1}{\prod_{i \in I} Q_i} \sum_{j \in I} \frac{Q'_j}{Q_j} \left( \prod_{i \in I \setminus \{j\}} Q_i \right)$$

$$= \sum_{j \in I} Q'_j \frac{1}{Q_j}$$

$$= \sum_{j \in I} \text{lder} (Q_j) = \sum_{i \in I} \text{lder} (Q_i).$$

This proves (11.193).

Proof of (11.194): Let $Q \in A[[t]]$ be a formal power series with constant term 1. Then, the product rule for the derivative of a product of two power series yields $(QQ^{-1})' = Q'Q^{-1} + Q(Q^{-1})'$, so that $Q'Q^{-1} + Q(Q^{-1})' = \left( \frac{QQ^{-1}}{Q^{-1}} \right)' = 1' = 0$

and thus $Q(Q^{-1})' = -Q'Q^{-1}$, so that $\frac{(Q^{-1})'}{Q^{-1}} = -\frac{Q'}{Q}$. But the definition of lder $(Q^{-1})$ yields $\text{lder} (Q^{-1}) = \frac{(Q^{-1})'}{Q^{-1}} = -\frac{Q'}{Q}$.

This proves (11.194).
so that
\[
\text{lder}(H(t)) = \text{lder} \left( \prod_{i \in \{1,2,3,\ldots\}} (1-x_i t)^{-1} \right) = \sum_{i \in \{1,2,3,\ldots\}} \text{lder} \left( (1-x_i t)^{-1} \right).
\]
(by (11.193), applied to \( I = \{1,2,3,\ldots\} \) and \( Q_i = (1-x_i t)^{-1} \))

\[
= \sum_{i \in \{1,2,3,\ldots\}} \left( \frac{1}{1-x_i t} \right)^{m} = \sum_{i \in \{1,2,3,\ldots\}} x_i \cdot \frac{1}{1-x_i t}.
\]

Thus,
\[
\sum_{m \geq 0} p_{m+1} t^m = \text{lder}(H(t)) = \frac{(H(t))^f}{H(t)}
\]
(by the definition of \( \text{lder}(H(t)) \))
\[
= \frac{H'(t)}{H(t)}.
\]

This solves the exercise.

Remark: Another solution of Exercise 2.48 proceeds by noticing that the statement of this exercise can be rewritten in the form \( \left( \sum_{m \geq 0} p_{m+1} t^m \right) \cdot H(t) = H'(t) \), which (by comparison of coefficients) is equivalent to saying that

\[
\sum_{m \geq 0} p_{m+1} t^m = \text{lder}(H(t)) = \frac{(H(t))^f}{H(t)}
\]
(by the definition of \( \text{lder}(H(t)) \))
\[
= \frac{H'(t)}{H(t)}.
\]

But it is clear that once (11.195) is proven for \( k = \mathbb{Z} \), it immediately follows that (11.195) also holds for all \( k \) (since \( \Lambda_k = k \otimes \mathbb{Z} \Lambda_{\mathbb{Z}} \)). Proving (11.195) for \( k = \mathbb{Z} \), in turn, boils down to proving (11.195) for \( k = \mathbb{Q} \), which we already know how to do (from the proof of (2.30)).

11.42. Solution to Exercise 2.63. Solution to Exercise 2.63. We recall that the cells of \( \lambda/\mu \) are the cells \((p,q) \in \{1,2,3,\ldots\}^2 \) satisfying \( \mu_p < q \leq \lambda_p \).

(a) In order to solve Exercise 2.63(a), we need to prove the following two claims:

(11.196) (if \( \lambda/\mu \) is a horizontal strip, then every \( i \in \{1,2,3,\ldots\} \) satisfies \( \mu_i \geq \lambda_{i+1} \))

and

(11.197) (if every \( i \in \{1,2,3,\ldots\} \) satisfies \( \mu_i \geq \lambda_{i+1} \), then \( \lambda/\mu \) is a horizontal strip).
Proof of (11.196): Assume that $\lambda/\mu$ is a horizontal strip. In other words, no two cells of $\lambda/\mu$ lie in the same column.

Let $i \in \{1, 2, 3, \ldots\}$. Assume (for the sake of contradiction) that $\mu_i < \lambda_i + 1$. Hence, $\lambda_i + 1 > \mu_i \geq 0$. Also, $\mu \subseteq \lambda$ yields $\mu_i \leq \lambda_i$. Hence, $\mu_i + 1 \leq \mu_i < \lambda_i + 1$. Now, $(i, \lambda_{i+1})$ is a cell of $\lambda/\mu$ (since $\mu_i < \lambda_i + 1 \leq \lambda_i$). Also, $(i + 1, \lambda_{i+1})$ is a cell of $\lambda/\mu$ (since $\mu_{i+1} < \lambda_{i+1} \leq \lambda_i + 1$). Thus, $(i, \lambda_{i+1})$ and $(i + 1, \lambda_{i+1})$ are two cells of $\lambda/\mu$ lying in the same column. This contradicts the fact that no two cells of $\lambda/\mu$ lie in the same column. This contradiction shows that our assumption (that $\mu_i < \lambda_i + 1$) was wrong. Hence, $\mu_i \geq \lambda_i + 1$. This proves (11.196).

Proof of (11.197): Assume that every $i \in \{1, 2, 3, \ldots\}$ satisfies $\mu_i \geq \lambda_i + 1$.

Now, let us (for the sake of contradiction) assume that there exist two distinct cells of $\lambda/\mu$ which lie in the same column. Let $c$ and $d$ be two such cells. Thus, $c$ and $d$ are two distinct cells of $\lambda/\mu$ which lie in the same column.

Write $c$ and $d$ in the forms $c = (p_c, q_c)$ and $d = (p_d, q_d)$ for some positive integers $p_c, q_c, p_d, q_d$. We have $\mu_{p_c} < q_c \leq \lambda_{p_c}$ (since $(p_c, q_c) = c$ is a cell of $\lambda/\mu$) and $\mu_{p_d} < q_d \leq \lambda_{p_d}$ (similarly). The cells $c$ and $d$ lie in the same column; in other words, $q_c = q_d$ (because the cell $c = (p_c, q_c)$ lies in column $q_c$, and the cell $d = (p_d, q_d)$ lies in column $q_d$).

Our situation so far is symmetric with respect to interchanging $c$ with $d$. Hence, we can WLOG assume that $p_c \leq p_d$ (because otherwise, we can switch $c$ with $d$). Assume this.

If we had $p_c = p_d$, then we would have $c = \left(\begin{array}{c} p_c \\ q_c \end{array}\right) = \left(\begin{array}{c} p_d \\ q_d \end{array}\right) = d$, which would contradict the fact that $c$ and $d$ are distinct. Hence, we cannot have $p_c = p_d$. Thus, we have $p_c \neq p_d$. Combined with $p_c \leq p_d$, this yields $p_c < p_d$. Thus, $p_c \leq p_d - 1$ (since $p_c$ and $p_d$ are integers), so that $p_c + 1 \leq p_d$.

Since $\lambda$ is a partition, we have $\lambda_u \geq \lambda_i$ for any positive integers $u$ and $v$ satisfying $u \leq v$. Applying this to $u = p_c + 1$ and $v = p_d$, we obtain $\lambda_{p_c + 1} \geq \lambda_{p_d}$ (since $p_c + 1 \leq p_d$). But recall that every $i \in \{1, 2, 3, \ldots\}$ satisfies $\mu_i \geq \lambda_i + 1$. Applying this to $i = p_c$, we obtain $\mu_{p_c} \geq \lambda_{p_c + 1}$. Now, recall that $\mu_{p_c} < q_c$, so that $q_c > \mu_{p_c} \geq \lambda_{p_c + 1} \geq \lambda_{p_d} \geq q_d$ (since $q_d \leq \lambda_{p_d}$). This contradicts $q_c = q_d$. This contradiction shows that our assumption (that there exist two distinct cells of $\lambda/\mu$ which lie in the same column) was wrong. Hence, no two cells of $\lambda/\mu$ lie in the same column. In other words, $\lambda/\mu$ is a horizontal strip. This proves (11.197).

Now, both (11.196) and (11.197) are proven. Thus, Exercise 2.63(a) is solved.

(b) We have the following equivalence of statements:

$$\begin{align*}
(\lambda/\mu \text{ is a vertical strip}) & \iff \text{(no two cells of } \lambda/\mu \text{ lie in the same row)} \\
& \iff \text{(by the definition of a "vertical strip")} \\
& \iff \text{(for every } i \in \{1, 2, 3, \ldots\}, \text{ no two cells of } \lambda/\mu \text{ lie in row } i) \\
& \iff \left(\text{for every } i \in \{1, 2, 3, \ldots\}, \text{ the number of cells of } \lambda/\mu \text{ in row } i \text{ is } \leq 1\right) \\
& \iff \text{(for every } i \in \{1, 2, 3, \ldots\}, \text{ we have } \lambda_i - \mu_i \leq 1) \\
& \iff \text{(for every } i \in \{1, 2, 3, \ldots\}, \text{ we have } \lambda_i \leq \mu_i + 1). 
\end{align*}$$

This solves Exercise 2.63(b).

11.43. **Solution to Exercise 2.64.** **Solution to Exercise 2.64.** (a) We use the notation $f(a_1, a_2, \ldots, a_k)$ defined in Exercise 2.2 whenever $a_1, a_2, \ldots, a_k$ are elements of a commutative $k$-algebra $A$ and $f \in R(x)$. In particular, $f(1)$ is a well-defined element of $k$ for every $f \in R(x)$. Every $f \in R(x)$ satisfies

\[
(11.198) \quad f(1) = \text{(the result of substituting } 1, 0, 0, 0, \ldots \text{ for } x_1, x_2, x_3, \ldots \text{ in } f). 
\]
We have $s_{\lambda/\mu} = \sum_T x^{\text{cont}(T)}$, where $T$ runs through all column-strict tableaux of shape $\lambda/\mu$. In other words, $s_{\lambda/\mu} = \sum_T$ is a column-strict $x^{\text{cont}(T)}$. Hence,

$$s_{\lambda/\mu}(1) = \left( \sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu} x^{\text{cont}(T)} \right)(1)$$

(11.199) \hspace{1cm} = \left( \text{the result of substituting } 1, 0, 0, 0, \ldots \text{ for } x_1, x_2, x_3, \ldots \text{ in } \sum_T x^{\text{cont}(T)} \right) \hspace{1cm} \text{(by (11.198)).}

But the substitution $1, 0, 0, 0, \ldots$ for $x_1, x_2, x_3, \ldots$ has the following effect on any given monomial $x^\alpha$:

- if none of the indeterminates $x_2, x_3, x_4, \ldots$ occur in this monomial $x^\alpha$, then the monomial $x^\alpha$ goes to 1;
- otherwise, the monomial $x^\alpha$ goes to 0.

Hence, applying this substitution to $\sum_T x^{\text{cont}(T)}$ yields a sum of 1’s over all column-strict tableaux $T$ of shape $\lambda/\mu$ having the property that none of the indeterminates $x_2, x_3, x_4, \ldots$ occur in this monomial $x^{\text{cont}(T)}$. In other words,

$$\left( \text{the result of substituting } 1, 0, 0, 0, \ldots \text{ for } x_1, x_2, x_3, \ldots \text{ in } \sum_T x^{\text{cont}(T)} \right)$$

(11.199) \hspace{1cm} = \sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu; \text{none of the indeterminates } x_2, x_3, x_4, \ldots \text{ occur in the monomial } x^{\text{cont}(T)}} 1.$$

Therefore, (11.199) rewrites as

$$s_{\lambda/\mu}(1) = \sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu; \text{none of the indeterminates } x_2, x_3, x_4, \ldots \text{ occur in the monomial } x^{\text{cont}(T)}} 1.$$ (11.200)

This rewrites as

$$s_{\lambda/\mu}(1) = \sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu; \text{all entries of } T \text{ are 1}} 1$$

(11.200) \hspace{1cm} \text{ (because for a column-strict tableau } T, \text{ saying that none of the indeterminates } x_2, x_3, x_4, \ldots \text{ occur in the monomial } x^{\text{cont}(T)} \text{ is equivalent to saying that all entries of } T \text{ are 1).}

In order to solve Exercise 2.64(a), we need to prove the following two statements:

(11.201) \hspace{1cm} (\text{if } \lambda/\mu \text{ is a horizontal } n\text{-strip, then } (h_n, s_{\lambda/\mu}) = 1)\)

and

(11.202) \hspace{1cm} (\text{if } \lambda/\mu \text{ is not a horizontal } n\text{-strip, then } (h_n, s_{\lambda/\mu}) = 0)\).

Proof of (11.201): Assume that $\lambda/\mu$ is a horizontal $n$-strip. Thus, in particular, we have $|\lambda/\mu| = n$. Hence, $s_{\lambda/\mu} \in \Lambda_{|\lambda/\mu|} = \Lambda_n$ (since $|\lambda/\mu| = n$). Thus, Exercise 2.41(b) (applied to $f = s_{\lambda/\mu}$) yields

$$s_{\lambda/\mu}(1) = \sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu; \text{all entries of } T \text{ are 1}} 1 \hspace{1cm} \text{(by (11.200)).}$$

(11.203)

Recall that $\lambda/\mu$ is a horizontal $n$-strip, therefore a horizontal strip. In other words, no two cells of $\lambda/\mu$ lie in the same column. Hence, if we fill every cell of the Ferrers diagram of $\lambda/\mu$ with a 1, we obtain a column-strict tableau $T$ of shape $\lambda/\mu$ having the property that all entries of $T$ are 1. Therefore, such a
tableau exists. It is also clearly unique (because requiring that all entries be 1 does not leave any freedom in choosing the entries). Therefore, there exists exactly one such tableau. This shows that the sum on the right hand side of (11.203) has exactly one addend, and therefore equals 1 (since the addend is 1). Hence, (11.203) rewrites as \((h_n, s_{\lambda/\mu}) = 1\), and thus (11.201) is proven.

**Proof of (11.202):** Assume that \(\lambda/\mu\) is not a horizontal \(n\)-strip. We need to show that \((h_n, s_{\lambda/\mu}) = 0\).

If \(|\lambda/\mu| \neq n\), then Exercise 2.41(a) (applied to \(m = |\lambda/\mu|, f = h_n\) and \(g = s_{\lambda/\mu}\)) yields \((h_n, s_{\lambda/\mu}) = 0\) (since \(h_n \in \Lambda_n\) and \(s_{\lambda/\mu} \in \Lambda_{|\lambda/\mu|}\)). Thus, \((h_n, s_{\lambda/\mu}) = 0\) is proven if \(|\lambda/\mu| \neq n\). Hence, for the rest of the proof of (11.203) we WLOG assume that we don’t have \(|\lambda/\mu| \neq n\). Thus, \(|\lambda/\mu| = n\). We can thus prove (11.203) just as we did before (in the proof of (11.201)).

But if \(\lambda/\mu\) were a horizontal strip, then \(\lambda/\mu\) would be a horizontal \(n\)-strip (since \(|\lambda/\mu| = n\)), which would contradict our assumption that \(\lambda/\mu\) is not a horizontal \(n\)-strip. Hence, \(\lambda/\mu\) cannot be a horizontal strip. As a consequence, there must be two cells of \(\lambda/\mu\) which lie in the same column. If \(T\) is a column-strict tableau of shape \(\lambda/\mu\), then these two cells must be filled with two different entries in \(T\) (because the entries of a column-strict tableau are strictly increasing top-to-bottom down columns, and hence all entries in any given column must be pairwise distinct), which is impossible if all entries of \(T\) are to be 1. Therefore, there exists no column-strict tableau \(T\) of shape \(\lambda/\mu\) such that all entries of \(T\) are 1. Hence, the sum on the right hand side of (11.203) is empty, and therefore equals 0. So (11.203) rewrites as \((h_n, s_{\lambda/\mu}) = 0\), and thus (11.202) is proven.

Now that both (11.201) and (11.202) are proven, Exercise 2.64(a) is solved.

(b) Notice first that

\[
(11.204) \quad (h_n, s_{\lambda/\mu}) = c^\lambda_{\mu,(n)} \quad \text{for every } n \in \mathbb{N}, \lambda \in \text{Par and } \mu \in \text{Par}.
\]

Recall that any two partitions \(\mu\) and \(\nu\) satisfy

\[
(11.205) \quad s_\mu s_\nu = \sum_{\lambda \in \text{Par}} c^\lambda_{\mu,\nu} s_\lambda = \sum_{\tau \in \text{Par}} c^\tau_{\mu,\nu} s_\tau
\]

(here, we renamed the summation index \(\lambda\) as \(\tau\)).

---

Proof: Let \(n \in \mathbb{N}, \lambda \in \text{Par and } \mu \in \text{Par}.\) From Remark 2.37, we know that \(s_{\lambda/\mu} = \sum_{\nu} c^\lambda_{\mu,\nu} s_\nu\), where the sum is over all \(\nu \in \text{Par}\). Thus, for every given partition \(\tau\), the \(s_\tau\)-coordinate of \(s_{\lambda/\mu}\) in the basis \((s_\nu)_{\nu \in \text{Par}}\) of \(\Lambda\) equals \(c^\lambda_{\mu,\tau}\). But since \((s_\nu)_{\nu \in \text{Par}}\) is an orthonormal basis of the \(k\)-module \(\Lambda\) with respect to the Hall inner product, this \(s_\tau\)-coordinate also equals \((s_\tau, s_{\lambda/\mu})\). Comparing these two expressions for this \(s_\tau\)-coordinate, we obtain \(c^\lambda_{\mu,\tau} = (s_\tau, s_{\lambda/\mu})\). Applying this to \(\tau = (n)\), we obtain \(c^\lambda_{\mu,(n)} = \left(\frac{a(n), s_{\lambda/\mu}}{h_n} = (h_n, s_{\lambda/\mu})\right)\), which proves (11.204).
Let $n \in \mathbb{N}$. Let $\lambda$ be a partition. We have $s_{(n)} = h_n$, so that $h_n = s_{(n)}$. Now,

$$s_{\lambda} h_n = s_{\lambda} s_{(n)} = \sum_{\tau \in \text{Par}} c_{\lambda, (n)}^{\tau} s_{\tau} \quad (\text{by (11.205), applied to } \mu = \lambda \text{ and } \nu = (n))$$

$$= \sum_{\tau \in \text{Par}} (h_n, s_{\tau/\lambda}) s_{\tau} = \sum_{\tau \in \text{Par}; \ \lambda \subseteq \tau} (h_n, s_{\tau/\lambda}) s_{\tau}$$

where, we have ridden the sum of all its addends in which $\lambda \notin \tau$:

$$\left(\begin{array}{l}
\text{these addends were zero (because if } \lambda \notin \tau, \text{ then } (h_n, s_{\tau/\lambda})_{\tau/\lambda = 0} = 0)
\end{array}\right)$$

$$= \sum_{\tau \in \text{Par}; \ \lambda \subseteq \tau; \ \tau / \lambda \text{ is a horizontal } n\text{-strip}} (h_n, s_{\tau/\lambda}) s_{\tau} + \sum_{\tau \in \text{Par}; \ \lambda \subseteq \tau; \ \tau / \lambda \text{ is not a horizontal } n\text{-strip}} (h_n, s_{\tau/\lambda}) s_{\tau}$$

$$= \sum_{\tau \in \text{Par}; \ \lambda \subseteq \tau; \ \tau / \lambda \text{ is a horizontal } n\text{-strip}} s_{\tau} + \sum_{\tau \in \text{Par}; \ \lambda \subseteq \tau; \ \tau / \lambda \text{ is not a horizontal } n\text{-strip}} 0 s_{\tau}$$

$$= \sum_{\tau \in \text{Par}; \ \lambda \subseteq \tau; \ \tau / \lambda \text{ is a horizontal } n\text{-strip}} s_{\tau} = \sum_{\lambda^+ \subseteq \tau; \ \lambda^+ / \lambda \text{ is a horizontal } n\text{-strip}} s_{\lambda^+}$$

(here, we renamed the summation index $\tau$ as $\lambda^+$)

$$= \sum_{\lambda^+; \lambda^+ / \lambda \text{ is a horizontal } n\text{-strip}} s_{\lambda^+}.$$

This proves (2.36). Thus, Exercise 2.64(b) is solved.

11.44. Solution to Exercise 2.65. Solution to Exercise 2.65. Let us use the notations introduced in the proof of Theorem 2.31. In particular, we use the words “letter” and “positive integer” as synonyms.

We shall use the following lemma:

**Lemma 11.3.** Let $P$ be a column-strict tableau, and let $j$ and $j'$ be two letters. Applying RS-insertion to the tableau $P$ and the letter $j$ yields a new column-strict tableau $P'$ and a corner cell $c$. Applying RS-insertion to the tableau $P'$ and the letter $j'$ yields a new column-strict tableau $P''$ and a corner cell $c'$.

(a) Assume that $j \leq j'$. Then, the cell $c'$ is in the same row as the cell $c$ or in a row further up; it is also in a column further right than $c$.

(b) Assume instead that $j > j'$. Then, the cell $c'$ is in a row further down than the cell $c$; it is also in the same column as $c$ or in a column further left.

Lemma 11.3 is part of the Row bumping lemma that appeared in our proof of Theorem 2.31. We shall first concentrate on proving (2.36).

**Alternative proof of (2.36).** Let $\lambda$ be a partition, and let $n \in \mathbb{N}$. The definition of $s_{\lambda}$ yields

$$s_{\lambda} = \sum_{T \text{ is a column-strict tableau of shape } \lambda} x^{|\text{cont}(T)|}.$$
The definition of $h_n$ yields
\begin{equation}
(11.206) \quad h_n = \sum_{i_1 \leq i_2 \leq \cdots \leq i_n} x_{i_1} x_{i_2} \cdots x_{i_n}.
\end{equation}
Multiplying these two identities, we obtain
\begin{equation}
(11.207) \quad s_{\lambda} h_n = \sum_{T \text{ is a column-strict tableau of shape } \lambda} x^{\cont(T)} \sum_{i_1 \leq i_2 \leq \cdots \leq i_n} x_{i_1} x_{i_2} \cdots x_{i_n},
\end{equation}
where $A$ is the set of all pairs $(T, (i_1, i_2, \ldots, i_n))$ of a column-strict tableau $T$ of shape $\lambda$ and an $n$-tuple $(i_1, i_2, \ldots, i_n)$ of positive integers satisfying $i_1 \leq i_2 \leq \cdots \leq i_n$. Consider this set $A$.

On the other hand, every partition $\lambda^+$ satisfies
\begin{equation}
(11.208) \quad s_{\lambda^+} = \sum_{S \text{ is a column-strict tableau of shape } \lambda^+} x^{\cont(S)} \quad \text{(by the definition of } s_{\lambda^+} \text{)}
\end{equation}
Hence,
\begin{equation}
(11.209) \quad \sum_{\lambda^+: \lambda^+/\lambda \text{ is a horizontal } n\text{-strip}} \sum_{S \text{ is a column-strict tableau of shape } \lambda^+} x^{\cont(S)} = \sum_{(\lambda^+, S) \in B} x^{\cont(S)},
\end{equation}
where $B$ is the set of all pairs $(\lambda^+, S)$ of a partition $\lambda^+$ and a column-strict tableau $S$ of shape $\lambda^+$ such that $\lambda^+/\lambda$ is a horizontal $n$-strip. Consider this set $B$.

We shall now prove that there exists a bijection $i : A \rightarrow B$ which has the property that
\begin{equation}
(11.210) \quad x^{\cont(T)} x_{i_1} x_{i_2} \cdots x_{i_n} = x^{\cont(S)}
\end{equation}
whenever some $(T, (i_1, i_2, \ldots, i_n)) \in A$ and $(\lambda^+, S) \in B$ satisfy $i ((T, (i_1, i_2, \ldots, i_n))) = (\lambda^+, S)$. Once this will be proven, it will immediately follow that
\begin{equation}
(11.207) \quad = \sum_{(\lambda^+, S) \in B} x^{\cont(S)},
\end{equation}
and therefore $s_{\lambda} h_n = \sum_{(T, (i_1, i_2, \ldots, i_n)) \in A} x^{\cont(T)} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{(\lambda^+, S) \in B} x^{\cont(S)} = \sum_{\lambda^+: \lambda^+/\lambda \text{ is a horizontal } n\text{-strip}} x^{\cont(S)} \quad \text{(by } (11.209) \text{)},
\end{equation}
and thus (2.36) will be proven. Hence, in order to complete the proof of (2.36), it is enough to prove that there exists a bijection $i : A \rightarrow B$ which has the property (11.210).

We construct such a bijection $i : A \rightarrow B$ explicitly. Namely, for every $(T, (i_1, i_2, \ldots, i_n)) \in A$, we define $i ((T, (i_1, i_2, \ldots, i_n)))$ as follows: Construct a sequence $(T_0, T_1, \ldots, T_n)$ of column-strict tableaux recursively: We set $T_0 = T$. For every $k \in \{1, 2, \ldots, n\}$, if $T_{k-1}$ is already defined, we let $T_k$ be the column-strict tableau obtained by applying RS-insertion to the tableau $T_{k-1}$ and the letter $i_k$. (This RS-insertion also returns a corner cell, but we do not care about it.) Thus, a sequence $(T_0, T_1, \ldots, T_n)$ is defined. We now set $S = T_n$, and let $\lambda^+$ be the shape of $S$. It is easy to see that $\lambda^+/\lambda$ is a horizontal $n$-strip\footnote{Proof. Notice that $i_1 \leq i_2 \leq \cdots \leq i_n$ (since $(T, (i_1, i_2, \ldots, i_n)) \in A$). The tableau $S$ has been obtained from $T$ by applying RS-insertion $n$ times, using the letters $i_1, i_2, \ldots, i_n$ in this order. Each time that we have applied RS-insertion, the shape of our tableau has grown by a new cell. According to Lemma 11.3(a), each of these cells (except for the first one) lies in a column further right than the previous one (because the letters $i_1, i_2, \ldots, i_n$ that we inserted satisfy $i_1 \leq i_2 \leq \cdots \leq i_n$). Therefore, no two of these cells lie in the same column. Since these cells are precisely the cells of $\lambda^+/\lambda$ (because $\lambda^+$ is the shape of $S$, while $\lambda$ is the shape of $T$), this means that no two cells of $\lambda^+/\lambda$ lie in the same column. In other words, $\lambda^+/\lambda$ is a }. Thus, $(\lambda^+, S) \in B$. Now,
set \(i((T, (i_1, i_2, \ldots, i_n))) = (\lambda^+, S)\). We have therefore defined a map \(i : A \to B\). It remains to prove that this map \(i\) is a bijection and satisfies (11.210).

Proving that the map \(i\) satisfies (11.210) is easy\(^{360}\). It remains to show that \(i\) is a bijection. We will achieve this by constructing an inverse map.

Indeed, let us define a map \(r : B \to A\). For every \((\lambda^+, S) \in B\), we define \(r((\lambda^+, S))\) as follows: We know that \(\lambda^+ / \lambda\) is a horizontal n-strip (since \((\lambda^+, S) \in B\)). We can thus uniquely label the \(n\) cells of \(\lambda^+ / \lambda\) by \(c_1, c_2, \ldots, c_n\) from right to left. Consider these cells \(c_1, c_2, \ldots, c_n\). Construct a sequence \((S_0, S_1, \ldots, S_n)\) of column-strict tableaux and a sequence \((j_1, j_2, \ldots, j_n)\) of positive integers recursively: We set \(S_0 = S\). For every \(k \in \{1, 2, \ldots, n\}\), if \(S_{k-1}\) is already defined, we apply reverse bumping to the tableau \(S_{k-1}\) and its corner cell \(c_k\). We denote the resulting tableau by \(S_k\), and the resulting letter by \(j_k\). \(^{361}\) Thus, two sequences \((S_0, S_1, \ldots, S_n)\) and \((j_1, j_2, \ldots, j_n)\) are defined. We now set \(T = S_n\) and \((i_1, i_2, \ldots, i_n) = (j_n, j_{n-1}, \ldots, j_1)\).

So the tableau \(T\) is obtained from \(S\) by successively applying reverse bumping using the corner cells \(c_1, c_2, \ldots, c_n\) (in this order), and \(j_1, j_2, \ldots, j_n\) are the letters that are obtained from the reverse-bumping procedure. Since reverse bumping is the inverse map to RS-insertion, this yields that we can obtain \(S\) back from \(T\) by successively applying RS-insertion using the letters \(j_n, j_{n-1}, \ldots, j_1\) (that is, the letters \(i_1, i_2, \ldots, i_n\)), and that these successive RS-insertion steps recover the corner cells \(c_n, c_{n-1}, \ldots, c_1\) in this order. Now, it is easy to see that the tableau \(T\) has shape \(\lambda\) (because in passing from \(S\) to \(T\), we lost the corner cells \(c_n, c_{n-1}, \ldots, c_1\), which are exactly the cells of \(\lambda^+ / \lambda\)) and the letters \(i_1, i_2, \ldots, i_n\) satisfy horizontal strip. Since \(\lambda^+ / \lambda\) has precisely \(n\) cells (because we have applied RS-insertion exactly \(n\) times, gaining precisely one cell every time), this yields that \(\lambda^+ / \lambda\) is a horizontal n-strip, qed.

\(^{360}\)Proof. Assume that some \((T, (i_1, i_2, \ldots, i_n)) \in A\) and \((\lambda^+, S) \in B\) satisfy \(i((T, (i_1, i_2, \ldots, i_n))) = (\lambda^+, S)\). We need to show that (11.210) holds.

According to the definition of \(i((T, (i_1, i_2, \ldots, i_n)))\), the tableau \(S\) is obtained by successively applying RS-insertion to the tableau \(T\) using the letters \(i_1, i_2, \ldots, i_n\). But whenever a tableau \(V\) results from applying RS-insertion to a column-strict tableau \(U\) and a letter \(j\), the multiset of entries of \(V\) is obtained from the multiset of entries of \(U\) by tossing in the letter \(j\). Thus, the multiset of entries of \(S\) is obtained from the multiset of entries of \(T\) by tossing in the \(n\) letters \(i_1, i_2, \ldots, i_n\) (because \(S\) is obtained by successively applying RS-insertion to the tableau \(T\) using the letters \(i_1, i_2, \ldots, i_n\)). Hence, \(x^\text{out}(S) = x^\text{out}(T) x_{i_1} x_{i_2} \ldots x_{i_n}\). Thus, (11.210) is proven.

\(^{361}\)In order to verify that this definition makes sense, we need to check that \(c_k\) is a corner cell of \(S_{k-1}\) for each \(k \in \{1, 2, \ldots, n\}\). Let us sketch a proof of this fact.

In fact, we shall show a stronger claim:

Claim CC: For each \(k \in \{1, 2, \ldots, n\}\), the shape of the tableau \(S_{k-1}\) is obtained from \(\lambda^+\) by removing the corner cells \(c_1, c_2, \ldots, c_{k-1}\), and the cell \(c_k\) is a corner cell of \(S_{k-1}\).

Proof. We use induction over \(k\).

Induction base: The shape of the tableau \(S_0\) is \(\lambda^+\) (since \(S_0 = S\) has shape \(\lambda^+\)), which is clearly the shape obtained from \(\lambda^+\) by removing the corner cells \(c_1, c_2, \ldots, c_0\). Moreover, the cell \(c_1\) is the rightmost cell of \(\lambda^+ / \lambda\) (since we labelled the cells of \(\lambda^+ / \lambda\) by \(c_1, c_2, \ldots, c_0\) from right to left), and thus \(c_1\) has no cells of \(\lambda^+\) to its right; but it also has no cells of \(\lambda^+\) below it (since any such cell would have to belong to \(\lambda^+ / \lambda\), but \(\lambda^+ / \lambda\) is a horizontal strip and therefore cannot have two cells one under the other). Thus, \(c_1\) is a corner cell of \(S = S_0\). Therefore, Claim CC is proven for \(k = 1\). This completes the induction base.

Induction step: Fix some \(K \in \{1, 2, \ldots, n - 1\}\). Assume that Claim CC is proven for \(k = K\). We need to show that Claim CC holds for \(k = K + 1\).

We know that Claim CC is proven for \(k = K\). In other words, the shape of the tableau \(S_{K-1}\) is obtained from \(\lambda^+\) by removing the corner cells \(c_1, c_2, \ldots, c_{K-1}\), and the cell \(c_K\) is a corner cell of \(S_{K-1}\). Now, the shape of the tableau \(S_K\) is obtained from the shape of \(S_{K-1}\) by removing the corner cell \(c_K\) (because \(S_K\) is obtained by applying reverse bumping to the tableau \(S_{K-1}\) and its corner cell \(c_K\)), and thus is obtained from \(\lambda^+\) by removing the corner cells \(c_1, c_2, \ldots, c_K\) (since the shape of the tableau \(S_{K-1}\) is obtained from \(\lambda^+\) by removing the corner cells \(c_1, c_2, \ldots, c_K\)).

Recall that we labelled the cells of \(\lambda^+ / \lambda\) by \(c_1, c_2, \ldots, c_n\) from right to left. Therefore, the cell \(c_{K+1}\) belongs to \(\lambda^+ / \lambda\), and all cells of \(\lambda^+ / \lambda\) that lie to the right of \(c_{K+1}\) are \(c_1, c_2, \ldots, c_K\). Thus, all cells of \(\lambda^+\) that lie to the right of \(c_{K+1}\) are \(c_1, c_2, \ldots, c_K\) (because all cells of \(\lambda^+\) that lie to the right of \(c_{K+1}\) must belong to the skew shape \(\lambda^+ / \lambda\) (since \(c_{K+1}\) itself belongs to this skew shape), and so they must be cells of \(\lambda^+ / \lambda\) that lie to the right of \(c_{K+1}\)). In other words, none of the cells of the shape \(\lambda^+\) lies to the right of \(c_{K+1}\) except for \(c_1, c_2, \ldots, c_K\). Therefore, none of the cells of the shape of \(S_K\) lies to the right of \(c_{K+1}\) (because the shape of the tableau \(S_K\) is obtained from \(\lambda^+\) by removing the corner cells \(c_1, c_2, \ldots, c_K\), and thus its cells are precisely the cells of the shape \(\lambda^+\) other than \(c_1, c_2, \ldots, c_K\)). But none of the cells of the shape of \(S_K\) lies underneath \(c_{K+1}\) either (because any such cell would have to belong to \(\lambda^+ / \lambda\), but \(\lambda^+ / \lambda\) is a horizontal strip and thus cannot have two cells underneath one another). Therefore, the cell \(c_{K+1}\) has neither a bottom neighbor nor a right neighbor in the shape of \(S_K\); in other words, it is a corner cell of \(S_K\).

Thus we have shown that the shape of the tableau \(S_K\) is obtained from \(\lambda^+\) by removing the corner cells \(c_1, c_2, \ldots, c_K\), and the cell \(c_{K+1}\) is a corner cell of \(S_K\). This proves Claim CC for \(k = K + 1\). Thus, the induction step is proven. Claim CC is thus proven. We hence conclude that our definition makes sense.
were labelled by \( c_j \) in a column further left. But this contradicts the fact that the tableau \( S \) (since (2.37)). The definition of \( h_r \) is thus defined. It is now easy to prove that the maps \( i \) and \( r \) are mutually inverse (since RS-insertion and reverse bumping are inverse maps), and thus \( i \) is a bijection. Thus, there exists a bijection \( i : A \rightarrow B \) which has the property (11.210). This completes our proof of (2.36).

The proof of (2.37) is almost entirely analogous. We give it for the sake of completeness (but most of it is copypasted material from the proof above).

**Alternative proof of (2.37).** Let \( \lambda \) be a partition, and let \( n \in \mathbb{N} \). The definition of \( s_\lambda \) yields

\[
 s_\lambda = \sum_{T \text{ is a column-strict tableau of shape } \lambda} x^{\text{cont}(T)}.
\]

The definition of \( h_n \) yields

\[
 e_n = \sum_{i_1 < i_2 < \cdots < i_n} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{i_1 > i_2 > \cdots > i_n} x_{i_n} x_{i_{n-1}} \cdots x_{i_1} = \sum_{i_1 > i_2 > \cdots > i_n} x_{i_1} x_{i_2} \cdots x_{i_n}
\]

(here, we substituted \((i_1, i_2, \ldots, i_n)\) for \((i_n, i_{n-1}, \ldots, i_1)\) in the sum). Multiplying these two identities, we obtain

\[
 s_\lambda e_n = \sum_{T \text{ is a column-strict tableau of shape } \lambda} x^{\text{cont}(T)} \sum_{i_1 > i_2 > \cdots > i_n} x_{i_1} x_{i_2} \cdots x_{i_n} \quad \quad \quad (11.211)
\]

\[
 = \sum_{(T, (i_1, i_2, \ldots, i_n)) \in A} x^{\text{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n},
\]

where \( A \) is the set of all pairs \((T, (i_1, i_2, \ldots, i_n))\) of a column-strict tableau \( T \) of shape \( \lambda \) and an \( n \)-tuple \((i_1, i_2, \ldots, i_n)\) of positive integers satisfying \( i_1 > i_2 > \cdots > i_n \). Consider this set \( A \).

On the other hand, every partition \( \lambda^+ \) satisfies

\[
 s_{\lambda^+} = \sum_{T \text{ is a column-strict tableau of shape } \lambda^+} x^{\text{cont}(T)}
\]

(by the definition of \( s_{\lambda^+} \))

\[
 (11.212)
\]

\[
 = \sum_{S \text{ is a column-strict tableau of shape } \lambda^+} x^{\text{cont}(S)}. \quad \quad \quad \text{(here, we renamed the summation index } T \text{ as } S). \quad \quad \quad (11.213)
\]

Hence,

\[
 \sum_{\lambda^+: \lambda^+/\lambda \text{ is a vertical } n\text{-strip}} s_{\lambda^+} = \sum_{S \text{ is a column-strict tableau of shape } \lambda^+} x^{\text{cont}(S)} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{\lambda^+: \lambda^+/\lambda \text{ is a vertical } n\text{-strip tableau of shape } \lambda^+} x^{\text{cont}(S)}
\]

\[
 (11.214)
\]

where \( B \) is the set of all pairs \((\lambda^+, S)\) of a partition \( \lambda^+ \) and a column-strict tableau \( S \) of shape \( \lambda^+ \) such that \( \lambda^+/\lambda \) is a vertical \( n \)-strip. Consider this set \( B \).

We shall now prove that there exists a bijection \( i : A \rightarrow B \) which has the property that

\[
 x^{\text{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n} = x^{\text{cont}(S)}
\]

\[
 (11.215)
\]

\footnote{Proof. Assume the contrary. Then, we don’t have \( i_1 \leq i_2 \leq \cdots \leq i_n \). In other words, we don’t have \( j_n \leq j_{n-1} \leq \cdots \leq j_1 \) (since \((i_1, i_2, \ldots, i_n) = (j_n, j_{n-1}, \ldots, j_1)\)). Hence, there exists a \( k \in \{2, 3, \ldots, n\} \) such that \( j_k > j_{k-1} \). Consider this \( k \). Recall that the tableau \( S_k \) and the letter \( j_k \) were obtained by applying reverse bumping to the tableau \( S_{k-1} \) and its corner cell \( c_k \), while the tableau \( S_{k-1} \) and the letter \( j_{k-1} \) were obtained by applying reverse bumping to the tableau \( S_{k-2} \) and its corner cell \( c_{k-1} \). Since reverse bumping is the inverse map to RS-insertion, this entails that conversely, the tableau \( S_{k-1} \) and its corner cell \( c_k \) are obtained by applying RS-insertion to the tableau \( S_k \) and the letter \( j_k \), and the tableau \( S_{k-2} \) and its corner cell \( c_{k-1} \) are obtained by applying RS-insertion to the tableau \( S_{k-1} \) and the letter \( j_{k-1} \). Hence, Lemma 11.13(b) (applied to \( S_k \), \( j_k \), \( j_{k-1} \), \( S_{k-1} \), \( c_k \), \( S_{k-2} \) and \( c_{k-1} \) instead of \( P, j, j', P', c, P', c' \) and \( c' \)) yields that the cell \( c_{k-1} \) is in the same column as \( c_k \) or in a column further left. But this contradicts the fact that \( c_{k-1} \) lies in a column further right than \( c_k \) (since the cells of \( \lambda^+/\lambda \) were labelled by \( c_1, c_2, \ldots, c_n \) from right to left, and lie in different columns). This contradiction completes our proof.}
whenever some $(T, (i_1, i_2, \ldots, i_n)) \in \mathcal{A}$ and $(\lambda^+, S) \in \mathcal{B}$ satisfy $i((T, (i_1, i_2, \ldots, i_n))) = (\lambda^+, S)$. Once this will be proven, it will immediately follow that $\sum_{(T, (i_1, i_2, \ldots, i_n)) \in \mathcal{A}} x^{\text{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{(\lambda^+, S) \in \mathcal{B}} x^{\text{cont}(S)}$, and therefore (11.212) will become

$$s_\lambda c_n = \sum_{(T, (i_1, i_2, \ldots, i_n)) \in \mathcal{A}} x^{\text{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n} = \sum_{(\lambda^+, S) \in \mathcal{B}} x^{\text{cont}(S)} = \sum_{\lambda^+, \lambda^+/\lambda \text{ is a vertical } n\text{-strip}} s_\lambda$$

(by (11.214)),

and thus (2.37) will be proven. Hence, in order to complete the proof of (2.37), it is enough to prove that there exists a bijection $i : \mathcal{A} \to \mathcal{B}$ which has the property (11.215).

We construct such a bijection $i : \mathcal{A} \to \mathcal{B}$ explicitly. Namely, for every $(T, (i_1, i_2, \ldots, i_n)) \in \mathcal{A}$, we define $i((T, (i_1, i_2, \ldots, i_n)))$ as follows: Construct a sequence $(T_0, T_1, \ldots, T_n)$ of column-strict tableaux recursively: We set $T_0 = T$. For every $k \in \{1, 2, \ldots, n\}$, if $T_{k-1}$ is already defined, we let $T_k$ be the column-strict tableau obtained by applying RS-insertion to the tableau $T_{k-1}$ and the letter $i_k$. (This RS-insertion also returns a corner cell, but we do not care about it.) Thus, a sequence $(T_0, T_1, \ldots, T_n)$ is defined. We now set $S = T_n$, and let $\lambda^+$ be the shape of $S$. It is easy to see that $\lambda^+/\lambda$ is a vertical $n$-strip. Thus, $(\lambda^+, S) \in \mathcal{B}$. Now, set $i((T, (i_1, i_2, \ldots, i_n))) = (\lambda^+, S)$. We have therefore defined a map $i : \mathcal{A} \to \mathcal{B}$. It remains to prove that this map $i$ is a bijection and satisfies (11.215).

Proving that the map $i$ satisfies (11.215) is easy. It remains to show that $i$ is a bijection. We will achieve this by constructing an inverse map.

Indeed, let us define a map $r : \mathcal{B} \to \mathcal{A}$. For every $(\lambda^+, S) \in \mathcal{B}$, we define $r((\lambda^+, S))$ as follows: We know that $\lambda^+/\lambda$ is a vertical $n$-strip (since $(\lambda^+, S) \in \mathcal{B}$). We can thus uniquely label the $n$ cells of $\lambda^+/\lambda$ by $c_1, c_2, \ldots, c_n$ from bottom to top. Consider these cells $c_1, c_2, \ldots, c_n$. Construct a sequence $(S_0, S_1, \ldots, S_n)$ of column-strict tableaux and a sequence $(j_1, j_2, \ldots, j_n)$ of positive integers recursively: We set $S_0 = S$. For every $k \in \{1, 2, \ldots, n\}$, if $S_{k-1}$ is already defined, we apply reverse bumping to the tableau $S_{k-1}$ and its corner cell $c_k$. We denote the resulting tableau by $S_k$, and the resulting letter by $j_k$. Thus, two sequences $(S_0, S_1, \ldots, S_n)$ and $(j_1, j_2, \ldots, j_n)$ are defined. We now set $T = S_n$ and $(i_1, i_2, \ldots, i_n) = (j_n, j_{n-1}, \ldots, j_1)$. So the tableau $T$ is obtained from $S$ by successively applying reverse bumping using the corner cells $c_1, c_2, \ldots, c_n$ (in this order), and $j_1, j_2, \ldots, j_n$ are the letters that are obtained from the reverse-bumping procedure. Since reverse bumping is the inverse map to RS-insertion, this yields that we can obtain $S$ back from $T$ by successively applying RS-insertion using the letters $j_n, j_{n-1}, \ldots, j_1$ (that is, the letters $i_1, i_2, \ldots, i_n$), and that these successive RS-insertion steps recover the corner cells $c_n, c_{n-1}, \ldots, c_1$ in this order. Now, it is easy to see that the tableau $T$ has shape $\lambda$ (because in passing from $S$ to $T$, we lost the corner cells $c_n, c_{n-1}, \ldots, c_1$, which are exactly the cells of $\lambda^+/\lambda$) and the letters $i_1, i_2, \ldots, i_n$ satisfy $i_1 > i_2 > \cdots > i_n$. Hence, $(T, (i_1, i_2, \ldots, i_n)) \in \mathcal{A}$. We set $r((\lambda^+, S)) = (T, (i_1, i_2, \ldots, i_n))$. The

\[363\] \text{Proof.} Notice that $i_1 > i_2 > \cdots > i_n$ (since $(T, (i_1, i_2, \ldots, i_n)) \in \mathcal{A}$). The tableau $S$ has been obtained from $T$ by applying RS-insertion $n$ times, using the letters $i_1, i_2, \ldots, i_n$ in this order. Each time that we have applied RS-insertion, the shape of our tableau has grown by a new cell. According to Lemma 11.3(b), each of these cells (except for the first one) lies in a row further down than the previous one (because the letters $i_1, i_2, \ldots, i_n$ that we inserted satisfy $i_1 > i_2 > \cdots > i_n$). Therefore, no two of these cells lie in the same row. Since these cells are precisely the cells of $\lambda^+/\lambda$ (because $\lambda^+$ is the shape of $S$, while $\lambda$ is the shape of $T$), this means that no two cells of $\lambda^+/\lambda$ lie in the same row. In other words, $\lambda^+/\lambda$ is a vertical strip. Since $\lambda^+/\lambda$ has precisely $n$ cells (because we have applied RS-insertion exactly $n$ times, gaining precisely one cell every time), this yields that $\lambda^+/\lambda$ is a vertical $n$-strip, qed.

\[364\] \text{Proof.} Assume that some $(T, (i_1, i_2, \ldots, i_n)) \in \mathcal{A}$ and $(\lambda^+, S) \in \mathcal{B}$ satisfy $i((T, (i_1, i_2, \ldots, i_n))) = (\lambda^+, S)$. We need to show that (11.215) holds.

According to the definition of $i((T, (i_1, i_2, \ldots, i_n)))$, the tableau $S$ is obtained by successively applying RS-insertion to the tableau $T$ using the letters $i_1, i_2, \ldots, i_n$. But whenever a tableau $V$ results from applying RS-insertion to a column-strict tableau $U$ and a letter $j$, the multiset of entries of $V$ is obtained from the multiset of entries of $U$ by tossing in the letter $j$. Thus, the multiset of entries of $S$ is obtained from the multiset of entries of $T$ by tossing in the $n$ letters $i_1, i_2, \ldots, i_n$ (because $S$ is obtained by successively applying RS-insertion to the tableau $T$ using the letters $i_1, i_2, \ldots, i_n$). Hence, $x^{\text{cont}(S)} = x^{\text{cont}(T)} x_{i_1} x_{i_2} \cdots x_{i_n}$. Thus, (11.215) is proven.

\[365\] One again has to check that this is well-defined. The proof is very similar to the analogous argument in our proof of (2.36), and is left to the reader.
map \( r : B \to A \) is thus defined. It is now easy to prove that the maps \( i \) and \( r \) are mutually inverse (since RS-insertion and reverse bumping are inverse maps), and thus \( i \) is a bijection. Thus, there exists a bijection \( i : A \to B \) which has the property (11.215). This completes our proof of (2.37). \( \square \)

Now, both (2.36) and (2.37) are proven. Hence, Theorem 2.59 is proven again.

11.45. Solution to Exercise 2.66. Solution to Exercise 2.66. Before we start solving any specific part of this exercise, let us state some general properties of determinants (and prove some of them):

- Every \( m \in \mathbb{N} \) and every matrix \( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \in A^{m \times m} \) satisfy

\[
(11.216) \quad \det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \right) = \sum_{\sigma \in S_m} (-1)^\sigma \prod_{i=1}^m \alpha_{i,\sigma(i)}.
\]

(This is simply the explicit formula for the determinant of a matrix as a sum over permutations.)

- If a positive integer \( m \) and a matrix \( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \in A^{m \times m} \) are such that every \( j \in \{1, 2, \ldots, m-1\} \) satisfies \( \alpha_{m,j} = 0 \), then

\[
(11.217) \quad \det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \right) = \alpha_{m,m} \cdot \det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,m-1} \right).
\]

- Every positive integer \( m \) and every matrix \( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \in A^{m \times m} \) satisfy

\[
(11.218) \quad \det \left( (\alpha_{i,j} \alpha_{m,m} - \alpha_{i,m} \alpha_{m,j})_{i,j=1,2,\ldots,m-1} \right) = \alpha_{m,m}^{m-2} \cdot \det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \right)
\]

if \( \alpha_{m,m} \) is an invertible element of \( A \).

**Proof of (11.218):** Let \( m \) be a positive integer. Let \( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \in A^{m \times m} \) be a matrix such that \( \alpha_{m,m} \) is an invertible element of \( A \).

For every \( (i,j) \in \{1, 2, \ldots, m\}^2 \), define an element \( \beta_{i,j} \) of \( A \) by

\[
(11.219) \quad \beta_{i,j} = \begin{cases} 
\alpha_{m,m}, & \text{if } i = j; \\
-\alpha_{m,j}, & \text{if } i = m \text{ and } j \neq m.
\end{cases}
\]

Then, the matrix \( (\beta_{i,j})_{i,j=1,2,\ldots,m} \) is lower-triangular. Since the determinant of a lower-triangular matrix equals the product of its diagonal entries, we thus have

\[
(11.220) \quad \det \left( (\beta_{i,j})_{i,j=1,2,\ldots,m} \right) = \prod_{i=1}^m \beta_{i,i} = \prod_{i=1}^m \alpha_{m,m} = \alpha_{m,m}^m.
\]

Cell \( c_k \) are obtained by applying RS-insertion to the tableau \( S_k \) and the letter \( j_k \), and the tableau \( S_{k-1} \) and its corner cell \( c_{k-1} \) are obtained by applying RS-insertion to the tableau \( S_{k-1} \) and the letter \( j_{k-1} \). Hence, Lemma 11.3(a) (applied to \( S_k, j_k, j_{k-1}, S_{k-1}, c_k, S_{k-2} \) and \( c_{k-1} \) instead of \( P, j, j', P', c, P'' \) and \( c' \)) yields that the cell \( c_{k-1} \) is in the same row as \( c_k \) or in a row further up. But this contradicts the fact that \( c_{k-1} \) lies in a row further down than \( c_k \) (since the cells of \( \lambda^+ / \lambda \) were labelled by \( c_1, c_2, \ldots, c_n \) from bottom to top, and lie in different rows). This contradiction completes our proof.

367In fact, the condition that every \( j \in \{1, 2, \ldots, m-1\} \) satisfies \( \alpha_{m,j} = 0 \) means that all entries of the \( m \)-th row of the \( m \times m \)-matrix \( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \) are zeroes apart from (possibly) the last entry. Hence, applying Laplace expansion to the determinant \( \det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \right) \) of this matrix yields a sum of products, all of which are zero apart from

\[
\alpha_{m,m} \cdot \left( \text{the } (m, m)\text{-th cofactor of the matrix } (\alpha_{i,j})_{i,j=1,2,\ldots,m} \right) = \alpha_{m,m} \cdot \det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,m-1} \right).
\]

367
But since the determinant of a product of two square matrices equals the product of their determinants, we have

$$\det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \cdot (\beta_{i,j})_{i,j=1,2,\ldots,m} \right) = \det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \right) \cdot \det \left( (\beta_{i,j})_{i,j=1,2,\ldots,m} \right)$$

(by (11.220))

$$= \det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \right) \cdot \alpha_{m,m}^m.$$  

Since $$(\alpha_{i,j})_{i,j=1,2,\ldots,m} \cdot (\beta_{i,j})_{i,j=1,2,\ldots,m} = (\sum_{k=1}^m \alpha_{i,k} \beta_{k,j})_{i,j=1,2,\ldots,m},$$ this rewrites as

$$(11.221) \quad \det \left( \sum_{k=1}^m \alpha_{i,k} \beta_{k,j} \right)_{i,j=1,2,\ldots,m} = \det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \right) \cdot \alpha_{m,m}^m.$$  

Now, for every $(i,j) \in \{1, 2, \ldots, m\}^2,$ define an element $\gamma_{i,j}$ of $A$ by

$$\gamma_{i,j} = \sum_{k=1}^m \alpha_{i,k} \beta_{k,j}. $$

Then,

$$(11.222) \quad \det \left( \gamma_{i,j} \right)_{i,j=1,2,\ldots,m} = \det \left( \sum_{k=1}^m \alpha_{i,k} \beta_{k,j} \right)_{i,j=1,2,\ldots,m} = \det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \right) \cdot \alpha_{m,m}^m$$

(by (11.221)).

But for every $(i,j),$ we can simplify the expression for $\gamma_{i,j}$ given by (11.222) by plugging in the definition of $\beta_{k,j}$ (which guarantees that no more than two of the terms of the sum will be nonzero). We obtain

$$(11.224) \quad \gamma_{i,j} = \begin{cases} \alpha_{i,j} \alpha_{m,m} - \alpha_{i,m} \alpha_{m,j}, & \text{if } i \neq m \text{ and } j \neq m; \\ 0, & \text{if } i = m \text{ and } j \neq m; \\ \alpha_{i,m} \alpha_{m,m}, & \text{if } j = m \end{cases}. $$

In particular, this shows that every $j \in \{1, 2, \ldots, m - 1\}$ satisfies $\gamma_{m,j} = 0.$ Hence, (11.217) (applied to $\gamma_{i,j}$ instead of $\alpha_{i,j}$) yields

$$\det \left( \gamma_{i,j} \right)_{i,j=1,2,\ldots,m} = \frac{\gamma_{m,m}}{\alpha_{m,m} \alpha_{m,m}} \cdot \det \left( \gamma_{i,j} \right)_{i,j=1,2,\ldots,m-1} \quad \text{(by (11.224), since } m = m)$$

$$= \alpha_{m,m} \alpha_{m,m} \cdot \det \left( (\alpha_{i,j} \alpha_{m,m} - \alpha_{i,m} \alpha_{m,j})_{i,j=1,2,\ldots,m-1} \right)$$

$$= \alpha_{m,m}^2 \cdot \det \left( (\alpha_{i,j} \alpha_{m,m} - \alpha_{i,m} \alpha_{m,j})_{i,j=1,2,\ldots,m-1} \right).$$

Compared with (11.223), this yields

$$\alpha_{m,m}^2 \cdot \det \left( (\alpha_{i,j} \alpha_{m,m} - \alpha_{i,m} \alpha_{m,j})_{i,j=1,2,\ldots,m-1} \right) = \det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \right) \cdot \alpha_{m,m}^m.$$  

We can divide both sides of this equality by $\alpha_{m,m}^2$ (since $\alpha_{m,m}$ is invertible in $A$), and thus obtain

$$\det \left( (\alpha_{i,j} \alpha_{m,m} - \alpha_{i,m} \alpha_{m,j})_{i,j=1,2,\ldots,m-1} \right) = \alpha_{m,m}^{m-2} \cdot \det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,m} \right).$$

This proves (11.218).

[Remark: The equality (11.218) holds even without requiring that $\alpha_{m,m}$ be invertible, if we have $m \geq 2$ (of course, if $m$ is not $\geq 2$, then the $\alpha_{m,m}^{m-2}$ on the right hand side of (11.218) does not make sense unless $\alpha_{m,m}$ is invertible). There are several ways to see why this is so. One of these ways proceeds as follows: First of all, one should notice that our above proof of (11.218) works
without requiring that $\alpha_{m,m}$ be invertible, as long as $\alpha_{m,m}$ is a non-zero-divisor in $A$ and we have $m \geq 2$. But for any fixed $m \geq 2$, the equality (11.218) is a polynomial identity in the elements $\alpha_{i,j}$ of $A$; thus, it suffices to prove it when $\alpha_{i,j}$ are distinct indeterminates $X_{i,j}$ in the polynomial ring $\mathbb{Z}[X_{i,j} \mid (i,j) \in \{1,2,\ldots,m\}]^2$. But in this case, $\alpha_{m,m}$ is clearly a non-zero-divisor, and so our above proof applies. (An alternative approach would be to replace $\alpha_{m,m}$ by $X + \alpha_{m,m}$ in the polynomial ring $A[X]$; again, $X + \alpha_{m,m}$ is a non-zero-divisor even if $\alpha_{m,m}$ is not.)

Now, rather than solve parts (a) and (b) of the exercise separately, we are going to prove a result from which both of these parts will easily follow:

**Proposition 11.4.** Let $A$ be a commutative ring. Let $n \in \mathbb{N}$. For every $i \in \{1,2,\ldots,n\}$, let $a_i$, $b_i$, $c_i$ and $d_i$ be four elements of $A$. Assume that $a_i d_j - b_i c_j$ is an invertible element of $A$ for every $i \in \{1,2,\ldots,n\}$ and $j \in \{1,2,\ldots,n\}$. Then,

$$
\det \left( \begin{array}{c}
1 \\
(a_i d_j - b_i c_j)_{i,j=1,2,\ldots,m}
\end{array} \right)_{i,j=1,2,\ldots,m} = \prod_{1 \leq j < i \leq n} \frac{((a_i b_j - a_j b_i) (c_j d_i - c_i d_j))}{\prod_{(i,j) \in \{1,2,\ldots,n\}^2} (a_i d_j - b_i c_j)}
$$

*Proof of Proposition 11.4.* We prove this by induction over $n$. The base case ($n = 0$) is obvious, as it claims an equality between the determinant of a $0 \times 0$-matrix (defined to be 1) and the ratio of two empty products (thus $\frac{1}{1} = 1$). For the induction step, we fix some positive integer $m$, and we set out to prove the equality

(11.225) $$
\det \left( \begin{array}{c}
1 \\
(a_i d_j - b_i c_j)_{i,j=1,2,\ldots,m}
\end{array} \right)_{i,j=1,2,\ldots,m} = \prod_{1 \leq j < i \leq m} \frac{((a_i b_j - a_j b_i) (c_j d_i - c_i d_j))}{\prod_{(i,j) \in \{1,2,\ldots,m\}^2} (a_i d_j - b_i c_j)}
$$

assuming that we already know

(11.226) $$
\det \left( \begin{array}{c}
1 \\
(a_i d_j - b_i c_j)_{i,j=1,2,\ldots,m-1}
\end{array} \right)_{i,j=1,2,\ldots,m-1} = \prod_{1 \leq j < i \leq m-1} \frac{((a_i b_j - a_j b_i) (c_j d_i - c_i d_j))}{\prod_{(i,j) \in \{1,2,\ldots,m-1\}^2} (a_i d_j - b_i c_j)}
$$

to be true.

Now, we know that $\frac{1}{a_m d_m - b_m c_m}$ is an invertible element of $A$ (because it is the inverse of $a_m d_m - b_m c_m$).

Thus, (11.218) (applied to $\alpha_{i,j} = \frac{1}{a_i d_j - b_i c_j}$) yields

$$
\det \left( \begin{array}{c}
1 \\
(a_i d_j - b_i c_j)_{i,j=1,2,\ldots,m-1}
\end{array} \right)_{i,j=1,2,\ldots,m-1} = \prod_{1 \leq j < i \leq m-1} \frac{((a_i b_j - a_j b_i) (c_j d_i - c_i d_j))}{\prod_{(i,j) \in \{1,2,\ldots,m-1\}^2} (a_i d_j - b_i c_j)}
$$

Solving this for $\det \left( \begin{array}{c}
1 \\
(a_i d_j - b_i c_j)_{i,j=1,2,\ldots,m-1}
\end{array} \right)_{i,j=1,2,\ldots,m}$, we obtain

(11.227) $$
\det \left( \begin{array}{c}
1 \\
(a_i d_j - b_i c_j)_{i,j=1,2,\ldots,m}
\end{array} \right)_{i,j=1,2,\ldots,m} = (a_m d_m - b_m c_m)^{m-2} \cdot \det \left( \begin{array}{c}
1 \\
(a_i d_j - b_i c_j)_{i,j=1,2,\ldots,m-1}
\end{array} \right)_{i,j=1,2,\ldots,m-1}
$$

---

[368] A *non-zero-divisor* in a commutative ring $B$ means an element $b \in B$ such that every element $c \in B$ satisfying $bc = 0$ must satisfy $c = 0$. 
But straightforward computations show that every \((i, j) \in \{1, 2, \ldots, m\}^2\) satisfy
\[
\frac{1}{a_i d_j - b_i c_j} \cdot \frac{1}{a_m d_m - b_m c_m} - \frac{1}{a_i d_m - b_i c_m} \cdot \frac{1}{a_m d_j - b_m c_j} = \frac{1}{(a_i d_m - b_i c_m)(a_m d_j - b_m c_j)} \cdot \frac{1}{a_i d_m - b_i c_m} \cdot \frac{1}{a_m d_j - b_m c_j}.
\]
Hence, the matrix
\[
\left( \frac{1}{a_i d_j - b_i c_j} \cdot \frac{1}{a_m d_m - b_m c_m} - \frac{1}{a_i d_m - b_i c_m} \cdot \frac{1}{a_m d_j - b_m c_j} \right)_{i,j=1,2,\ldots,m-1}
\]
can be rewritten as
\[
\left( \frac{c_j d_m - c_m d_j}{(a_m d_m - b_m c_m)(a_m d_j - b_m c_j)} \cdot \frac{a_m b_i - a_i b_m}{a_i d_m - b_i c_m} \cdot \frac{1}{a_i d_j - b_i c_j} \right)_{i,j=1,2,\ldots,m-1}.
\]
This means that this matrix can be obtained from the matrix \(\left( \frac{1}{a_i d_j - b_i c_j} \right)_{i,j=1,2,\ldots,m-1}\) by multiplying every row with \(\frac{a_m b_i - a_i b_m}{a_i d_m - b_i c_m}\), where \(i\) is the index of this row, and then multiplying every column with \(\frac{c_j d_m - c_m d_j}{(a_m d_m - b_m c_m)(a_m d_j - b_m c_j)}\), where \(j\) is the index of the column. As a consequence, the determinant of this matrix is
\[
\left( \prod_{j=1}^{m-1} \left( \frac{1}{a_m d_m - b_m c_m} \cdot \frac{c_j d_m - c_m d_j}{a_m d_j - b_m c_j} \right) \right) \cdot \left( \prod_{i=1}^{m-1} \frac{a_m b_i - a_i b_m}{a_i d_m - b_i c_m} \right) \cdot \det \left( \left( \frac{1}{a_i d_j - b_i c_j} \right)_{i,j=1,2,\ldots,m-1} \right)
\]
(because when a row of a matrix is multiplied by a scalar, the determinant of the matrix gets multiplied by the same scalar, and the same rule holds for columns). Hence, we have shown that
\[
\det \left( \left( \frac{1}{a_i d_j - b_i c_j} \cdot \frac{1}{a_m d_m - b_m c_m} - \frac{1}{a_i d_m - b_i c_m} \cdot \frac{1}{a_m d_j - b_m c_j} \right)_{i,j=1,2,\ldots,m-1} \right)
\]
\[
= \prod_{j=1}^{m-1} \left( \frac{1}{a_m d_m - b_m c_m} \cdot \frac{c_j d_m - c_m d_j}{a_m d_j - b_m c_j} \right) \cdot \left( \prod_{i=1}^{m-1} \frac{a_m b_i - a_i b_m}{a_i d_m - b_i c_m} \right) \cdot \det \left( \left( \frac{1}{a_i d_j - b_i c_j} \right)_{i,j=1,2,\ldots,m-1} \right)
\]
\[
= \prod_{i,j=1}^{m-1} \frac{1}{a_i d_j - b_i c_j} \cdot \frac{1}{a_m d_m - b_m c_m} - \frac{1}{a_i d_m - b_i c_m} \cdot \frac{1}{a_m d_j - b_m c_j}
\]
\[
= \prod_{i,j=1}^{m-1} \left( \frac{a_i b_j - a_j b_i}{a_i d_j - b_i c_j} \right) \cdot \det \left( \left( \frac{1}{a_i d_j - b_i c_j} \right)_{i,j=1,2,\ldots,m-1} \right)
\]
\[
= \prod_{1 \leq j < i \leq m-1} \left( \frac{a_i b_j - a_j b_i}{a_i d_j - b_i c_j}(c_j d_i - c_i d_j) \right) \cdot \prod_{(i,j) \in \{1,2,\ldots,m-1\}^2} (a_i d_j - b_j c_j)
\]
We can now plug this into (11.227) and make straightforward simplifications (splitting apart and pulling together products), and obtain (11.225). Thus, the induction step is complete, and Proposition 11.4 is proven.

Now, let us solve the exercise.

(a) First solution of Exercise 2.66(a): Exercise 2.66(a) follows from Proposition 11.4 (applied to \(a_i, 1, b_i\) and 1 instead of \(a_i, b_i, c_i\) and \(d_i\)).

Second solution of Exercise 2.66(a) (sketched): The statement of Exercise 2.66(a) is an identity between two rational functions in the variables \(a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n\), and is easily seen to be equivalent to a polynomial identity in these variables (by multiplying both sides through with the common denominator \(\prod_{(i,j) \in \{1,2,\ldots,n\}^2} (a_i - b_j)\)). It is well-known that such identities need only be checked on complex numbers to ensure that they hold for any elements of any ring. So we only need to prove the statement of Exercise 2.66(a) in the case when \(A = \mathbb{C}\). But the statement is well-known in this case (see, e.g., [50, Lemma 5.15.3 (Lemma 4.48 in the arXiv version)] and various other sources for the proof).
(b) First solution of Exercise 2.66(b): Exercise 2.66(b) follows from Proposition 11.4 (applied to 1, $a_i$, $b_i$ and 1 instead of $a_i$, $b_i$, $c_i$ and $d_i$).

Second solution of Exercise 2.66(b): Just as in the Second solution of Exercise 2.66(a), we can see that it is enough to solve Exercise 2.66(b) in the case when $A = \mathbb{C}$. But in this case, the statement of this exercise is well-known, and proven, e.g., in [50, Corollary 5.15.4 (Corollary 4.49 in the arXiv version)] and [37, Cauchy’s lemma, p. 18].

Remark: One could also derive the statement of Exercise 2.66(b) from Exercise 2.66(a) by applying the latter to $1/a_i$ instead of $a_i$, after first WLOG assuming that the $a_i$ are invertible (but one needs to justify this WLOG assumption).

(c) Alternative proof of Theorem 2.31. We are going to show that, for every $n \in \mathbb{N}$, we have

\begin{equation}
\prod_{i,j=1}^{n} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \Par} s_{\lambda}(x_1, x_2, ..., x_n) s_{\lambda}(y_1, y_2, ..., y_n)
\end{equation}

in the ring $k[[x_1, x_2, ..., x_n, y_1, y_2, ..., y_n]]$. Once this is proven, Theorem 2.31 will easily follow.\footnote{Proof. Assume that (11.228) is proven for all $n \in \mathbb{N}$. We need to show that

\begin{equation}
\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \Par} s_{\lambda}(x) s_{\lambda}(y)
\end{equation}

in the ring $k[[x, y]] = k[[x_1, x_2, x_3, ..., y_1, y_2, y_3, ...]]$. In order to do so, it is clearly enough to prove that for any two weak compositions $\alpha$ and $\beta$, the coefficient of $x^\alpha y^\beta$ in $\prod_{i,j=1}^{n} (1 - x_i y_j)^{-1}$ equals the coefficient of $x^\alpha y^\beta$ in $\sum_{\lambda \in \Par} s_{\lambda}(x) s_{\lambda}(y)$. So fix two weak compositions $\alpha$ and $\beta$. Write $\alpha$ and $\beta$ as $\alpha = (\alpha_1, \alpha_2, \alpha_3, ...)$ and $\beta = (\beta_1, \beta_2, \beta_3, ...)$. Choose some $n \in \mathbb{N}$ such that every integer $m > n$ satisfies $\alpha_m = \beta_m = 0$. (Such an $m$ clearly exists, since $\alpha$ and $\beta$ are finitely supported.) Then, the coefficient of $x^\alpha y^\beta$ in $\prod_{i,j=1}^{n} (1 - x_i y_j)^{-1}$ equals the coefficient of $x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} y_1^{\beta_1} y_2^{\beta_2} ... y_n^{\beta_n}$ in $\prod_{i,j=1}^{n} (1 - x_i y_j)^{-1}$, whereas the coefficient of $x^\alpha y^\beta$ in $\sum_{\lambda \in \Par} s_{\lambda}(x) s_{\lambda}(y)$ equals the coefficient of $x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} y_1^{\beta_1} y_2^{\beta_2} ... y_n^{\beta_n}$ in $\sum_{\lambda \in \Par} s_{\lambda}(x_1, x_2, ..., x_n) s_{\lambda}(y_1, y_2, ..., y_n)$. Since the coefficients of $x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} y_1^{\beta_1} y_2^{\beta_2} ... y_n^{\beta_n}$ in $\prod_{i,j=1}^{n} (1 - x_i y_j)^{-1}$ and in $\sum_{\lambda \in \Par} s_{\lambda}(x_1, x_2, ..., x_n) s_{\lambda}(y_1, y_2, ..., y_n)$ are equal (because of (11.228)), this shows that the coefficients of $x^\alpha y^\beta$ in $\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}$ and in $\sum_{\lambda \in \Par} s_{\lambda}(x) s_{\lambda}(y)$ are equal; but this is exactly what we need to prove. Thus, Theorem 2.31 follows if (11.228) is proven.}

So let $n \in \mathbb{N}$. For every $i \in \{1, 2, ..., n\}$ and $j \in \{1, 2, ..., n\}$, the element $1 - x_i y_j$ of the ring $k[[x_1, x_2, ..., x_n, y_1, y_2, ..., y_n]]$ is invertible. Hence, Exercise 2.66(b) (applied to $A = k[[x_1, x_2, ..., x_n, y_1, y_2, ..., y_n]]$, $a_i = x_i$ and $b_j = y_j$) yields

\begin{equation}
\det \left( \frac{1}{1 - x_i y_j} \right)_{i,j=1,2,...,n} = \prod_{1 \leq i < j \leq n} \frac{((x_i - x_j) (y_i - y_j))}{(1 - x_i y_j)} = \prod_{1 \leq i < j \leq n} \frac{((x_i - x_j) (y_i - y_j))}{(1 - x_i y_j)}
\end{equation}

in the ring $k[[x_1, x_2, ..., x_n, y_1, y_2, ..., y_n]]$.

We will use the notations of Definition 2.50 and Proposition 2.52 (but we do not require $k$ to be $\mathbb{Z}$ or a field of characteristic not equal to 2 as was done in Proposition 2.52). We have $a_\rho = \prod_{1 \leq i < j \leq n} (x_i - x_j)$ (as proven in the proof of Proposition 2.52). Thus, $a_\rho (x_1, x_2, ..., x_n) = \prod_{1 \leq i < j \leq n} (x_i - x_j)$ and $a_\rho (y_1, y_2, ..., y_n) = \prod_{1 \leq i < j \leq n} (y_i - y_j)$.\footnote{Proof. Assume that (11.228) is proven for all $n \in \mathbb{N}$. We need to show that

\begin{equation}
\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \Par} s_{\lambda}(x) s_{\lambda}(y)
\end{equation}

in the ring $k[[x, y]] = k[[x_1, x_2, x_3, ..., y_1, y_2, y_3, ...]]$. In order to do so, it is clearly enough to prove that for any two weak compositions $\alpha$ and $\beta$, the coefficient of $x^\alpha y^\beta$ in $\prod_{i,j=1}^{n} (1 - x_i y_j)^{-1}$ equals the coefficient of $x^\alpha y^\beta$ in $\sum_{\lambda \in \Par} s_{\lambda}(x) s_{\lambda}(y)$. So fix two weak compositions $\alpha$ and $\beta$. Write $\alpha$ and $\beta$ as $\alpha = (\alpha_1, \alpha_2, \alpha_3, ...) = (1, 0, 1, 0, ...)$ and $\beta = (\beta_1, \beta_2, \beta_3, ...)$. Choose some $n \in \mathbb{N}$ such that every integer $m > n$ satisfies $\alpha_m = \beta_m = 0$. (Such an $m$ clearly exists, since $\alpha$ and $\beta$ are finitely supported.) Then, the coefficient of $x^\alpha y^\beta$ in $\prod_{i,j=1}^{n} (1 - x_i y_j)^{-1}$ equals the coefficient of $x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} y_1^{\beta_1} y_2^{\beta_2} ... y_n^{\beta_n}$ in $\prod_{i,j=1}^{n} (1 - x_i y_j)^{-1}$, whereas the coefficient of $x^\alpha y^\beta$ in $\sum_{\lambda \in \Par} s_{\lambda}(x) s_{\lambda}(y)$ equals the coefficient of $x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} y_1^{\beta_1} y_2^{\beta_2} ... y_n^{\beta_n}$ in $\sum_{\lambda \in \Par} s_{\lambda}(x_1, x_2, ..., x_n) s_{\lambda}(y_1, y_2, ..., y_n)$. Since the coefficients of $x_1^{\alpha_1} x_2^{\alpha_2} ... x_n^{\alpha_n} y_1^{\beta_1} y_2^{\beta_2} ... y_n^{\beta_n}$ in $\prod_{i,j=1}^{n} (1 - x_i y_j)^{-1}$ and in $\sum_{\lambda \in \Par} s_{\lambda}(x_1, x_2, ..., x_n) s_{\lambda}(y_1, y_2, ..., y_n)$ are equal (because of (11.228)), this shows that the coefficients of $x^\alpha y^\beta$ in $\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}$ and in $\sum_{\lambda \in \Par} s_{\lambda}(x) s_{\lambda}(y)$ are equal; but this is exactly what we need to prove. Thus, Theorem 2.31 follows if (11.228) is proven.}
\[
\prod_{1 \leq i < j \leq n} (y_i - y_j). \text{ Thus, (11.229) becomes }
\]
\[
\det \left( \frac{1}{1 - x_i y_j} \right)_{i,j=1,2,...,n}
= \left( \prod_{1 \leq i < j \leq n} (x_i - x_j) \right) \cdot \left( \prod_{1 \leq i < j \leq n} (y_i - y_j) \right) \cdot \prod_{i,j=1}^{n} \frac{1}{1 - x_i y_j}^{-1}
= a_\rho(x_1, x_2, ..., x_n) \cdot a_\rho(y_1, y_2, ..., y_n) \cdot \prod_{i,j=1}^{n} (1 - x_i y_j)^{-1}.
\]

On the other hand, (11.216) (applied to \( m = n \) and \( \alpha_{i,j} = \frac{1}{1 - x_i y_j} \)) yields
\[
\det \left( \frac{1}{1 - x_i y_j} \right)_{i,j=1,2,...,n} = \sum_{\sigma \in S_n} (-1)^\sigma \prod_{i=1}^{n} \frac{1}{1 - x_i y_{\sigma(i)}} = \sum_{\gamma \in \mathbb{N}} \sum_{k \in \mathbb{N}} (x_1 y_{\sigma(1)})^k \cdot (x_2 y_{\sigma(2)})^k \cdot ... \cdot (x_n y_{\sigma(n)})^k
= \sum_{\sigma \in S_n} \sum_{(k_1, k_2, ..., k_n) \in \mathbb{N}^n} \left( x_1 y_{\sigma(1)} \right)^{k_1} \left( x_2 y_{\sigma(2)} \right)^{k_2} \cdot ... \cdot \left( x_n y_{\sigma(n)} \right)^{k_n}
= \sum_{\sigma \in S_n} \sum_{(k_1, k_2, ..., k_n) \in \mathbb{N}^n} \left( x_1^{k_1} x_2^{k_2} \cdot ... \cdot x_n^{k_n} \right) \cdot \left( y_{\sigma(1)}^{k_1} y_{\sigma(2)}^{k_2} \cdot ... \cdot y_{\sigma(n)}^{k_n} \right)
= \sum_{(k_1, k_2, ..., k_n) \in \mathbb{N}^n} x_1^{k_1} x_2^{k_2} \cdot ... \cdot x_n^{k_n} \sum_{\sigma \in S_n} (-1)^\sigma y_{\sigma(1)}^{k_1} y_{\sigma(2)}^{k_2} \cdot ... \cdot y_{\sigma(n)}^{k_n}.
\]

But every \((k_1, k_2, ..., k_n) \in \mathbb{N}^n\) satisfies
\[
\det \left( y_j^{k_i} \right)_{i,j=1,2,...,n} = \sum_{\sigma \in S_n} (-1)^\sigma \prod_{i=1}^{n} y_{\sigma(i)}^{k_i}
= \sum_{\sigma \in S_n} (-1)^\sigma y_{\sigma(1)}^{k_1} y_{\sigma(2)}^{k_2} \cdot ... \cdot y_{\sigma(n)}^{k_n},
\]
(by (11.216), applied to \( m = n \) and \( \alpha_{i,j} = y_j \)).
Hence, (11.231) becomes

\[
\det \left( \frac{1}{1 - x_i y_j} \right)_{i,j=1,2,...,n} = \sum_{(k_1, k_2, ..., k_n) \in \mathbb{N}^n} x_1^{k_1} x_2^{k_2} ... x_n^{k_n} \sum_{\sigma \in S_n} (-1)^{\sigma} y_{\sigma(1)}^{k_1} y_{\sigma(2)}^{k_2} ... y_{\sigma(n)}^{k_n}
\]

(by (11.232))

\[
= \sum_{(k_1, k_2, ..., k_n) \in \mathbb{N}^n} x_1^{k_1} x_2^{k_2} ... x_n^{k_n} \det \left( y_i^{k_j} \right)_{i,j=1,2,...,n}
\]

(11.233)

\[
= \sum_{(k_1, k_2, ..., k_n) \in \mathbb{N}^n, \text{the integers } k_1, k_2, ..., k_n \text{ are pairwise distinct}} x_1^{k_1} x_2^{k_2} ... x_n^{k_n} \det \left( y_i^{k_j} \right)_{i,j=1,2,...,n}.
\]

(Here, we have removed from our sum all addends in which the integers \(k_1, k_2, ..., k_n\) are not pairwise distinct. This did not change the value of the sum, because all these addends are zero\(^{370}\).)

On the other hand, every \((k_1, k_2, ..., k_n) \in \mathbb{N}^n\) satisfies

\[
\det \left( x_i^{k_j} \right)_{i,j=1,2,...,n} = \sum_{\sigma \in S_n} (-1)^{\sigma} \prod_{i=1}^{n} x_{\sigma(i)}^{k_i}
\]

(by (11.216), applied to \(m = n\) and \(\alpha_{i,j} = x_i^{k_j}\))

\[
= \sum_{\sigma \in S_n} (-1)^{\sigma} x_{\sigma(1)}^{k_1} x_{\sigma(2)}^{k_2} ... x_{\sigma(n)}^{k_n}.
\]

\(^{370}\)In fact, if \((k_1, k_2, ..., k_n) \in \mathbb{N}^n\) is such that the integers \(k_1, k_2, ..., k_n\) are not pairwise distinct, then the matrix \(\left( y_i^{k_j} \right)_{i,j=1,2,...,n}\) has two equal rows, which causes its determinant \(\det \left( y_i^{k_j} \right)_{i,j=1,2,...,n}\) to be 0, and thus the addend corresponding to this \((k_1, k_2, ..., k_n) \in \mathbb{N}^n\) is

\[
x_1^{k_1} x_2^{k_2} ... x_n^{k_n} \det \left( y_i^{k_j} \right)_{i,j=1,2,...,n} = 0.
\]
Thus,

\[
\sum_{(k_1, k_2, \ldots, k_n) \in \mathbb{N}^n; \; k_1 > k_2 > \cdots > k_n} \det \left( x_{i,j}^{k_i} \right)_{i,j=1,2,\ldots,n} \det \left( y_{i,j}^{k_i} \right)_{i,j=1,2,\ldots,n} = \sum_{\sigma \in S_n} \left( -1 \right)^{\sigma(1)} x_{\sigma(1)}^{k_1} x_{\sigma(2)}^{k_2} \cdots x_{\sigma(n)}^{k_n} \det \left( \left( y_{i,j}^{k_i} \right)_{i,j=1,2,\ldots,n} \right)
\]

\[
= \sum_{(k_1, k_2, \ldots, k_n) \in \mathbb{N}^n; \; \sigma \in S_n} \left( -1 \right)^{\sigma(1)} x_{\sigma(1)}^{k_1} x_{\sigma(2)}^{k_2} \cdots x_{\sigma(n)}^{k_n} \det \left( \left( y_{i,j}^{k_i} \right)_{i,j=1,2,\ldots,n} \right)
\]

\[
= \sum_{(k_1, k_2, \ldots, k_n) \in \mathbb{N}^n; \; \sigma \in S_n} \left( -1 \right)^{\sigma(1)} \frac{k_{\sigma(1)}}{x_{\sigma(1)}} \frac{k_{\sigma(2)}}{x_{\sigma(2)}} \cdots \frac{k_{\sigma(n)}}{x_{\sigma(n)}} \det \left( \left( y_{i,j}^{k_i} \right)_{i,j=1,2,\ldots,n} \right)
\]

\[
= \sum_{(k_1, k_2, \ldots, k_n) \in \mathbb{N}^n; \; k_{\sigma(1)} > k_{\sigma(2)} > \cdots > k_{\sigma(n)}} \frac{k_{\sigma(1)}}{x_{\sigma(1)}} \frac{k_{\sigma(2)}}{x_{\sigma(2)}} \cdots \frac{k_{\sigma(n)}}{x_{\sigma(n)}} \det \left( \left( y_{i,j}^{k_i} \right)_{i,j=1,2,\ldots,n} \right) = \left( -1 \right)^{\sigma} \det \left( \left( y_{i,j}^{k_i} \right)_{i,j=1,2,\ldots,n} \right)
\]

(here, we substituted \((k_{\sigma(1)}, k_{\sigma(2)}, \ldots, k_{\sigma(n)})\) for \((k_1, k_2, \ldots, k_n)\) in the sum

\[\text{(this is allowed, since } \sigma \text{ is a permutation)}\]

\[
= \sum_{(k_1, k_2, \ldots, k_n) \in \mathbb{N}^n; \; \sigma \in S_n} \left( -1 \right)^{\sigma(1)} \frac{k_{\sigma(1)}}{x_{\sigma(1)}} \frac{k_{\sigma(2)}}{x_{\sigma(2)}} \cdots \frac{k_{\sigma(n)}}{x_{\sigma(n)}} \det \left( \left( y_{i,j}^{k_i} \right)_{i,j=1,2,\ldots,n} \right)
\]

\[
= \sum_{(k_1, k_2, \ldots, k_n) \in \mathbb{N}^n; \; k_{\sigma(1)} > k_{\sigma(2)} > \cdots > k_{\sigma(n)}} \frac{k_{\sigma(1)}}{x_{\sigma(1)}} \frac{k_{\sigma(2)}}{x_{\sigma(2)}} \cdots \frac{k_{\sigma(n)}}{x_{\sigma(n)}} \det \left( \left( y_{i,j}^{k_i} \right)_{i,j=1,2,\ldots,n} \right)
\]

\[\text{(since permuting the rows of a matrix multiplies its determinant by the sign of the permutation)}\]

\[
= \sum_{(k_1, k_2, \ldots, k_n) \in \mathbb{N}^n; \; \sigma \in S_n} \left( -1 \right)^{\sigma(1)} \frac{k_{\sigma(1)}}{x_{\sigma(1)}} \frac{k_{\sigma(2)}}{x_{\sigma(2)}} \cdots \frac{k_{\sigma(n)}}{x_{\sigma(n)}} \det \left( \left( y_{i,j}^{k_i} \right)_{i,j=1,2,\ldots,n} \right)
\]

\[
= \sum_{(k_1, k_2, \ldots, k_n) \in \mathbb{N}^n; \; k_{\sigma(1)} > k_{\sigma(2)} > \cdots > k_{\sigma(n)}} \frac{k_{\sigma(1)}}{x_{\sigma(1)}} \frac{k_{\sigma(2)}}{x_{\sigma(2)}} \cdots \frac{k_{\sigma(n)}}{x_{\sigma(n)}} \det \left( \left( y_{i,j}^{k_i} \right)_{i,j=1,2,\ldots,n} \right)
\]

\[\text{(since there exists exactly one } \sigma \in S_n \text{ such that } k_{\sigma(1)} > k_{\sigma(2)} > \cdots > k_{\sigma(n)} \text{)}\]

\[\text{(since the integers } k_1, k_2, \ldots, k_n \text{ are pairwise distinct, and thus there is exactly one permutation which arranges these integers in decreasing order)}\]

\[\left[ \begin{array}{c}
\sum_{(k_1, k_2, \ldots, k_n) \in \mathbb{N}^n; \; \text{the integers } k_1, k_2, \ldots, k_n \text{ are pairwise distinct}} \sum_{\sigma \in S_n; \; k_{\sigma(1)} > k_{\sigma(2)} > \cdots > k_{\sigma(n)}} \frac{k_{\sigma(1)}}{x_{\sigma(1)}} \frac{k_{\sigma(2)}}{x_{\sigma(2)}} \cdots \frac{k_{\sigma(n)}}{x_{\sigma(n)}}
\end{array} \right]
\]

\[\text{= } x_{1}^{k_1} x_{2}^{k_2} \cdots x_{n}^{k_n} \det \left( \left( y_{i,j}^{k_i} \right)_{i,j=1,2,\ldots,n} \right)
\]

\[\text{(because there exists exactly one } \sigma \in S_n \text{ such that } k_{\sigma(1)} > k_{\sigma(2)} > \cdots > k_{\sigma(n)} \text{)}\]

\[\text{(since the integers } k_1, k_2, \ldots, k_n \text{ are pairwise distinct, and thus there is exactly one permutation which arranges these integers in decreasing order)}\]

\[+ \sum_{(k_1, k_2, \ldots, k_n) \in \mathbb{N}^n; \; \text{the integers } k_1, k_2, \ldots, k_n \text{ are not pairwise distinct}} \sum_{\sigma \in S_n; \; k_{\sigma(1)} > k_{\sigma(2)} > \cdots > k_{\sigma(n)}} \frac{k_{\sigma(1)}}{x_{\sigma(1)}} \frac{k_{\sigma(2)}}{x_{\sigma(2)}} \cdots \frac{k_{\sigma(n)}}{x_{\sigma(n)}}
\]

\[\text{= } 0
\]

\[\text{(because the matrix } \left( y_{i,j}^{k_i} \right)_{i,j=1,2,\ldots,n} \text{ has two equal rows (since the integers } k_1, k_2, \ldots, k_n \text{ are not pairwise distinct, i.e., there are two equal among them))}\]
Compared with (11.233), this yields

$$\det \left( \frac{1}{1 - x_i y_j} \right)_{i,j=1,2,...,n} = \sum_{(k_1, k_2, ..., k_n) \in \mathbb{N}^n; \ k_1 > k_2 > ... > k_n} \det \left( x_j^{k_i} \right)_{i,j=1,2,...,n} \det \left( y_j^{k_i} \right)_{i,j=1,2,...,n} = \sum_{(\lambda_1, \lambda_2, ..., \lambda_n) \in \mathbb{N}^n; \ \lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n} \det \left( x_j^{\lambda_i + n - i} \right)_{i,j=1,2,...,n} \det \left( y_j^{\lambda_i + n - i} \right)_{i,j=1,2,...,n}$$

where, we substituted \((\lambda_1 + n - 1, \lambda_2 + n - 2, ..., \lambda_n + n - n)\) for \((k_1, k_2, ..., k_n)\) in the sum (since the map

\[
\{(\lambda_1, \lambda_2, ..., \lambda_n) \in \mathbb{N}^n \mid \lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n\} \rightarrow \{(k_1, k_2, ..., k_n) \in \mathbb{N}^n \mid k_1 > k_2 > ... > k_n\},
\]

\((\lambda_1, \lambda_2, ..., \lambda_n) \mapsto (\lambda_1 + n - 1, \lambda_2 + n - 2, ..., \lambda_n + n - n)\) is a bijection)

(11.234)

$$= \sum_{\lambda=(\lambda_1, \lambda_2, ..., \lambda_n) \in \mathbb{N}^n; \ \lambda \text{ is a partition with at most } n \text{ parts}} \det \left( x_j^{\lambda_i + n - i} \right)_{i,j=1,2,...,n} \det \left( y_j^{\lambda_i + n - i} \right)_{i,j=1,2,...,n}.$$
and thus

\[
a_{\lambda+\rho} = a_{(\lambda_1+n-1,\lambda_2+n-2,\ldots,\lambda_n+n-n)} = \det \begin{pmatrix} (x_j^{\lambda_j+n-i})_{i,j=1,2,\ldots,n} \\ (x_i^{\lambda_i+n-i})_{i,j=1,2,\ldots,n} \end{pmatrix}
\]

(by the definition of \(a_{(\alpha_1,\alpha_2,\ldots,\alpha_n)}\))

\[
= \det \left( \begin{pmatrix} (x_i^{\lambda_i+n-i})_{i,j=1,2,\ldots,n} \end{pmatrix}^T \right) = \det \left( (x_j^{\lambda_j+n-i})_{i,j=1,2,\ldots,n} \right)
\]

(since the determinant of a matrix is preserved under transposition).

Hence, we get the two equalities \(a_{\lambda+\rho} (x_1, x_2, \ldots, x_n) = \det \left( (x_i^{\lambda_i+n-i})_{i,j=1,2,\ldots,n} \right)\) and \(a_{\lambda+\rho} (y_1, y_2, \ldots, y_n) = \det \left( (y_j^{\lambda_j+n-i})_{i,j=1,2,\ldots,n} \right)\) for any partition \(\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)\) with at most \(n\) parts. Thus, (11.234) becomes

\[
\det \left( \frac{1}{1-x_i y_j} \right)_{i,j=1,2,\ldots,n} = \sum_{\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_n)\in\mathbb{N}^n \text{ is a partition with at most } n \text{ parts}} a_{\lambda+\rho} (x_1, x_2, \ldots, x_n) a_{\lambda+\rho} (y_1, y_2, \ldots, y_n)
\]

\[
= \sum_{\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_n)\in\mathbb{N}^n \text{ is a partition with at most } n \text{ parts}} a_{\lambda+\rho} (x_1, x_2, \ldots, x_n) a_{\lambda+\rho} (y_1, y_2, \ldots, y_n).
\]

Compared to (11.229), this yields

\[
\sum_{\lambda \text{ is a partition with at most } n \text{ parts}} a_{\lambda+\rho} (x_1, x_2, \ldots, x_n) a_{\lambda+\rho} (y_1, y_2, \ldots, y_n)
\]

\[
= \left( \prod_{1 \leq i < j \leq n} \frac{1}{x_i - x_j} \right) \cdot \left( \prod_{1 \leq i < j \leq n} \frac{1}{y_i - y_j} \right) \cdot \prod_{i,j=1}^{n} \frac{1}{1-x_i y_j}
\]

\[
= a_{\rho} (x_1, x_2, \ldots, x_n) \cdot a_{\rho} (y_1, y_2, \ldots, y_n) \cdot \prod_{i,j=1}^{n} (1-x_i y_j)^{-1}.
\]

(11.235)

Now, Corollary 2.54 says that \(s_\lambda (x_1, x_2, \ldots, x_n) = \frac{a_{\lambda+\rho}}{a_\rho} \) for every partition \(\lambda\) with at most \(n\) parts. Thus,

\[
s_\lambda (x_1, x_2, \ldots, x_n) = \frac{a_{\lambda+\rho} (x_1, x_2, \ldots, x_n)}{a_\rho (x_1, x_2, \ldots, x_n)} \quad \text{and} \quad s_\lambda (y_1, y_2, \ldots, y_n) = \frac{a_{\lambda+\rho} (y_1, y_2, \ldots, y_n)}{a_\rho (y_1, y_2, \ldots, y_n)}
\]

for every partition \(\lambda\).
with at most \( n \) parts. Hence,

\[
\sum_{\lambda \text{ is a partition with at most } n \text{ parts}} \frac{s_\lambda(x_1, x_2, \ldots, x_n)}{a_\rho(x_1, x_2, \ldots, x_n)} \cdot \frac{s_\lambda(y_1, y_2, \ldots, y_n)}{a_\rho(y_1, y_2, \ldots, y_n)} = \sum_{\lambda \text{ is a partition with at most } n \text{ parts}} \frac{a_{\lambda+\rho}(x_1, x_2, \ldots, x_n)}{a_\rho(x_1, x_2, \ldots, x_n)} \cdot \frac{a_{\lambda+\rho}(y_1, y_2, \ldots, y_n)}{a_\rho(y_1, y_2, \ldots, y_n)}
\]

\[
= \sum_{\lambda \text{ is a partition with at most } n \text{ parts}} \frac{1}{a_\rho(x_1, x_2, \ldots, x_n) \cdot a_\rho(y_1, y_2, \ldots, y_n)} \cdot \sum_{\lambda \text{ is a partition with at most } n \text{ parts}} a_{\lambda+\rho}(x_1, x_2, \ldots, x_n) a_{\lambda+\rho}(y_1, y_2, \ldots, y_n)
\]

\[
= \frac{1}{a_\rho(x_1, x_2, \ldots, x_n) \cdot a_\rho(y_1, y_2, \ldots, y_n)} \cdot a_\rho(x_1, x_2, \ldots, x_n) \cdot a_\rho(y_1, y_2, \ldots, y_n) \cdot \prod_{i,j=1}^{n} (1 - x_i y_j)^{-1}
\]

(11.236)

\[
= \prod_{i,j=1}^{n} (1 - x_i y_j)^{-1}.
\]

We are almost done. Let us now notice that every partition \( \lambda \) with more than \( n \) parts satisfies

\[
(11.237) \quad s_\lambda(x_1, x_2, \ldots, x_n) = 0
\]

(according to Exercise 2.26(b)). Thus, in the sum \( \sum_{\lambda \in \text{Par}} s_\lambda(x_1, x_2, \ldots, x_n) s_\lambda(y_1, y_2, \ldots, y_n) \), all addends corresponding to partitions \( \lambda \) having more than \( n \) parts are 0. We can therefore remove these addends from the sum. Hence,

\[
\sum_{\lambda \in \text{Par}} s_\lambda(x_1, x_2, \ldots, x_n) s_\lambda(y_1, y_2, \ldots, y_n)
\]

\[
= \sum_{\lambda \text{ is a partition with at most } n \text{ parts}} s_\lambda(x_1, x_2, \ldots, x_n) s_\lambda(y_1, y_2, \ldots, y_n)
\]

\[
= \prod_{i,j=1}^{n} (1 - x_i y_j)^{-1} \quad (\text{by (11.236)}).
\]

Thus, we have proven that (11.228) holds for every \( n \in \mathbb{N} \). As we explained above, this yields that Theorem 2.31 is true.

Remark: In our above solution, we solved Exercise 2.66(b) first, and then used it to prove Theorem 2.31. It is also possible (more or less by treading the above proof backwards) to conversely derive the statement of Exercise 2.66(b) from Theorem 2.31 instead (though Exercise 2.66(b) is usually considered a more elementary fact than Theorem 2.31).

11.46. Solution to Exercise 2.67. Solution to Exercise 2.67. Let us first check that we have

\[
(11.238) \quad h_u h_v = \sum_{i=0}^{v} s_{(u+i,v-i)}
\]

for any two nonnegative integers \( u \) and \( v \) satisfying \( u \geq v \).

Proof of (11.238): Let \( u \) and \( v \) be nonnegative integers satisfying \( u \geq v \). We WLOG assume that \( u > 0 \) (otherwise, \( u = 0 \), and \( u \geq v \) forces \( v = 0 \), so that (11.238) can be checked immediately). We have \( h_u = s_{(u)} \)
and thus

\[(11.239) \quad h_u h_v = s_{(u)} h_v = \sum_{\lambda^+ : \lambda^+/u \text{ is a horizontal } v\text{-strip}} s_{\lambda^+}\]

(by (2.36), applied to \(n = v\) and \(\lambda = (u)\)).

Now, fix a partition \(\lambda^+\) such that \(\lambda^+/u\) is a horizontal \(v\)-strip. Then, the skew diagram \(\lambda^+/u\) has at most one cell in column 1 (since \(\lambda^+/u\) is a horizontal \(v\)-strip and thus contains no two cells in the same column). Hence, the partition \(\lambda^+\) has at most 2 rows (because otherwise, the skew diagram \(\lambda^+/u\) would contain at least two cells in column 1, which would contradict the fact that the skew diagram \(\lambda^+/u\) has at most one cell in column 1). Thus, the partition \(\lambda^+\) has the form \((p, q)\) for two nonnegative integers \(p\) and \(q\) satisfying \(p \geq q\). Consider these two integers \(p\) and \(q\). Since \(\lambda^+/u\) is a horizontal \(v\)-strip, we have

\[|\lambda^+/u| = v \quad \text{and thus} \quad v = |\lambda^+/u| = \left|\sum_{\lambda^+ : \lambda^+/u \text{ is a horizontal } v\text{-strip}} \lambda^+\right| - |(p, q)| = p + q - u.\]

Moreover, \(\lambda^+/u\) is a horizontal \(v\)-strip, so that \((u) \subseteq \lambda^+ = (p, q)\). Thus, \(u \leq p\). Hence, there exists an \(i \in \mathbb{N}\) such that \(p = u + i\).

Consider this \(i\). Now, \(v = p + q - u = u + i + q - u = q + i\), so that \(q = v - i\) and thus \(v - i = q \geq 0\), so that \(i \leq v\) and thus \(i \in \{0, 1, \ldots, v\}\) (since \(i \in \mathbb{N}\)). Hence, \(\lambda^+ = \left(\begin{array}{c} p \\ q \\ \hline u+i \\ v-i \end{array}\right) = (u + i, v - i)\). We have thus shown that the partition \(\lambda^+\) has the form \((u + i, v - i)\) for some \(i \in \{0, 1, \ldots, v\}\).

Now forget that we fixed \(\lambda^+\). We thus have proven that every partition \(\lambda^+\) such that \(\lambda^+/u\) is a horizontal \(v\)-strip has the form \(\lambda^+ = (u + i, v - i)\) for some \(i \in \{0, 1, \ldots, v\}\). Conversely, it is clear that for every \(i \in \{0, 1, \ldots, v\}\), the weak composition \((u + i, v - i)\) is a partition \(\lambda^+\) such that \(\lambda^+/u\) is a horizontal \(v\)-strip. Combining the previous two sentences, we conclude that the partitions \(\lambda^+\) such that \(\lambda^+/u\) is a horizontal \(v\)-strip are precisely the weak compositions of the form \((u + i, v - i)\) for \(i \in \{0, 1, \ldots, v\}\). Therefore,

\[\sum_{\lambda^+ : \lambda^+/u \text{ is a horizontal } v\text{-strip}} s_{\lambda^+} = \sum_{i \in \{0, 1, \ldots, v\}} s_{(u+i,v-i)} = \sum_{i=0}^{v} s_{(u+i,v-i)}.\]

Now, (11.239) becomes

\[h_u h_v = \sum_{\lambda^+ : \lambda^+/u \text{ is a horizontal } v\text{-strip}} s_{\lambda^+} = \sum_{i=0}^{v} s_{(u+i,v-i)}.\]

This proves (11.238).

Now, fix two integers \(a\) and \(b\) satisfying \(a \geq b \geq 0\). We need to prove that \(s_{(a,b)} = h_a h_b - h_{a+1} h_{b-1}\).

If \(b = 0\), then proving \(s_{(a,b)} = h_a h_b - h_{a+1} h_{b-1}\) is very easy.\(^{371}\) Hence, for the rest of the proof of \(s_{(a,b)} = h_a h_b - h_{a+1} h_{b-1}\), we assume WLOG that we don’t have \(b = 0\). Thus, \(b \geq 1\), so that \(b - 1 \geq 0\). Clearly, \(a \geq b\) shows that \(a + 1 \geq b + 1 \geq b - 1\), so that we can apply (11.238) to \(u = a + 1\) and \(v = b - 1\).

\(^{371}\)In fact, assume that \(b = 0\). Then,

\[s_{(a,b)} = s_{(a,0)} = s_{(a)} = h_a\]

and

\[h_a h_b = h_a h_b = h_{a+1} h_{b-1} = h_a 1 - h_{a+1} 0 = h_a.\]

Hence, \(s_{(a,b)} = h_a = h_a h_b - h_{a+1} h_{b-1}\), qed.
Now,
\[
\begin{align*}
  h_a h_b &= h_{a+1} h_{b-1} \\
  &= \sum_{i=0}^{b-1} s(a+i, b-1) \\
  &= \sum_{i=0}^{b-1} s(a+i, b-1) \\
  &= s(a+0, b-0) + \sum_{i=1}^{b} s(a+i, b-1) - \sum_{i=0}^{b-1} s(a+(i+1), b-(i+1)) \\
  &= s(a, b) - s(a+0, b-0)
\end{align*}
\]

(11.240) (by (11.238), applied to \(u=a\) and \(v=b\))

\[= s(a+0, b-0) = s(a, b).\]

We thus have proven \(s(a, b) = h_a h_b - h_{a+1} h_{b-1}\). The exercise is solved.

11.47. **Solution to Exercise 2.68.** **Solution to Exercise 2.68.** (a) We shall prove the statement of Exercise 2.68(a) by induction over the length \(\ell(\mu)\) of \(\mu\).

The induction base (i.e., the case \(\ell(\mu) = 0\)) is trivial. For the induction step, we fix a positive integer \(L\) and assume (as the induction hypothesis) that Exercise 2.68(a) has been solved for all \(\mu\) satisfying \(\ell(\mu) = L - 1\). We now have to solve Exercise 2.68(a) for every partition \(\mu\) satisfying \(\ell(\mu) = L\).

So let \(\mu\) be a partition satisfying \(\ell(\mu) = L\). Write \(\mu\) in the form \((\mu_1, \mu_2, \ldots, \mu_L)\). Let \(\overline{\mu}\) be the partition \((\mu_1, \mu_2, \ldots, \mu_{L-1})\) (this is well-defined since \(L\) is positive). Then, \(\ell(\overline{\mu}) = L - 1\), and hence (by the induction hypothesis) we can apply Exercise 2.68(a) to \(\overline{\mu}\) instead of \(\mu\). As a result, we obtain
\[
  h_{\overline{\mu}} = \sum_{\lambda} K_{\lambda, \overline{\mu}} s_{\lambda},
\]

where the sum ranges over all partitions \(\lambda\).

But the definition of \(h_\mu\) yields \(h_\mu = h_{\mu_1} h_{\mu_2} \ldots h_{\mu_L}\); similarly, \(h_{\overline{\mu}} = h_{\mu_1} h_{\mu_2} \ldots h_{\mu_{L-1}}\). Hence,
\[
  h_\mu = h_{\mu_1} h_{\mu_2} \ldots h_{\mu_L} = h_{\mu_1} h_{\mu_2} \ldots h_{\mu_{L-1}} h_{\mu_L} = \left(\sum_{\lambda} K_{\lambda, \overline{\mu}} s_{\lambda}\right) h_{\mu_L}
\]

\[
= \sum_{\lambda} K_{\lambda, \overline{\mu}} \sum_{\lambda} s_{\lambda} h_{\mu_L} = \sum_{\lambda} \sum_{\lambda} K_{\lambda, \overline{\mu}} s_{\lambda} + \sum_{\lambda} K_{\lambda, \overline{\mu}} s_{\lambda+} + \sum_{\lambda} K_{\lambda, \overline{\mu}} s_{\lambda-}
\]

(11.240) \(= \sum_{\lambda} \sum_{\lambda} K_{\lambda, \overline{\mu}} s_{\lambda}\)

(here, we renamed the summation indices \(\lambda^+\) and \(\lambda\) as \(\lambda\) and \(\lambda^-\))

where all summation indices are supposed to be partitions.

Now, let us fix a partition \(\lambda\). We shall show that
\[
(11.241) \sum_{\lambda^-: \lambda^-/\lambda} K_{\lambda^-, \overline{\mu}} = K_{\lambda, \mu}.
\]

(11.240) \(= \sum_{\lambda} \sum_{\lambda^-: \lambda^-/\lambda} K_{\lambda^-, \overline{\mu}} \)
Proof of (11.241): For every partition $\lambda^-$, the number $K_{\lambda^-_\pi}$ is the number of column-strict tableaux $S$ of shape $\lambda^-$ having $\cont(S) = \pi$. Hence, the sum $\sum_{\lambda^- : \lambda/\lambda^- \text{ is a horizontal } \mu_L\text{-strip}} K_{\lambda^-_\pi}$ is the number of all pairs $(\lambda^-, S)$, where:

- $\lambda^-$ is a partition such that $\lambda/\lambda^-$ is a horizontal $\mu_L$-strip;
- $S$ is a column-strict tableau of shape $\lambda^-$ having $\cont(S) = \pi$.

We will refer to such pairs $(\lambda^-, S)$ as $(\lambda, \mu)$-last-step pairs. So the sum $\sum_{\lambda^- : \lambda/\lambda^- \text{ is a horizontal } \mu_L\text{-strip}} K_{\lambda^-_\pi}$ is the number of $(\lambda, \mu)$-last-step pairs. On the other hand, $K_{\lambda, \mu}$ is the number of all column-strict tableaux $T$ of shape $\lambda$ having $\cont(T) = \mu$. We now will construct a bijection between the $(\lambda, \mu)$-last-step pairs and the column-strict tableaux $T$ of shape $\lambda$ having $\cont(T) = \mu$.

Indeed, let us first define a map $\Phi$ from the set of all $(\lambda, \mu)$-last-step pairs to the set of all column-strict tableaux $T$ of shape $\lambda$ having $\cont(T) = \mu$. Namely, let $(\lambda^-, S)$ be a $(\lambda, \mu)$-last-step pair. Then, $\lambda^-$ is a partition such that $\lambda/\lambda^-$ is a horizontal $\mu_L$-strip, whereas $S$ is a column-strict tableau of shape $\lambda^-$ having $\cont(S) = \pi$. In particular, all entries of $S$ are $\leq \lambda$ (since $\cont(S) = \pi = (\mu_1, \mu_2, \ldots, \mu_{\mu_L - 1})$). Now, we can construct the column-strict tableau $S$ of shape $\lambda^-$ to a tableau of shape $\lambda$ by filling the number $\lambda$ into all cells of the skew shape $\lambda/\lambda^-$. The resulting tableau is a column-strict tableau $T$ of shape $\lambda$ having $\cont(T) = \mu$ (indeed, its column-strictness follows from the fact that $\lambda/\lambda^-$ is a horizontal $\mu_L$-strip whereas all entries of $S$ are $\leq \lambda$; and the property $\cont(T) = \mu$ follows from the facts that $\cont(S) = \pi$ and $|\lambda/\lambda^-| = \mu_L$). We define $\Phi(\lambda^-, S)$ to be this tableau. Thus we have defined a map $\Phi$.

Conversely, let us define a map $\Psi$ from the set of all column-strict tableaux $T$ of shape $\lambda$ having $\cont(T) = \mu$ to the set of all $(\lambda, \mu)$-last-step pairs. Namely, let $T$ be a column-strict tableau of shape $\lambda$ having $\cont(T) = \mu$. Then, all entries of $T$ are $\leq \lambda$ (since $\cont(T) = \mu = (\mu_1, \mu_2, \ldots, \mu_L)$), and the cells containing the entries $\lambda$ form a horizontal strip (since $T$ is column-strict). Hence, if we remove all entries $\lambda$ from $T$ (along with their cells), the result will be a column-strict tableau $S$ of some shape $\lambda^-$ such that $\lambda^-$ is a partition (because all entries of $T$ were $\leq \lambda$, so the entries we removed were maximal), $\lambda/\lambda^-$ is a horizontal $\mu_L$-strip (since we removed a total of $\mu_L$ entries and they formed a horizontal strip), and $\cont(S) = \pi$. Consider these $S$ and $\lambda^-$, and define $\Psi(T)$ to be the pair $(\lambda^-, S)$; it is clear that this $\Psi(T)$ is a $(\lambda, \mu)$-last-step pair. Hence, we have defined a map $\Psi$.

It is fairly obvious that the maps $\Phi$ and $\Psi$ are mutually inverse. Hence, they are bijections; in particular, $\Phi$ is a bijection. Thus, we have a bijection between the set of all $(\lambda, \mu)$-last-step pairs and the set of all column-strict tableaux $T$ of shape $\lambda$ having $\cont(T) = \mu$. As a consequence, the number of all $(\lambda, \mu)$-last-step pairs equals the number of all column-strict tableaux $T$ of shape $\lambda$ having $\cont(T) = \mu$. In other words, $\sum_{\lambda^- : \lambda/\lambda^- \text{ is a horizontal } \mu_L\text{-strip}} K_{\lambda^-_\pi} = K_{\lambda, \mu}$ (because $\sum_{\lambda^- : \lambda/\lambda^- \text{ is a horizontal } \mu_L\text{-strip}} K_{\lambda^-_\pi}$ is the number of $(\lambda, \mu)$-last-step pairs, whereas $K_{\lambda, \mu}$ is the number of all column-strict tableaux $T$ of shape $\lambda$ having $\cont(T) = \mu$). This proves (11.241).

Now, (11.240) becomes

$$h_\mu = \sum_{\lambda} \sum_{\lambda^- : \lambda/\lambda^- \text{ is a horizontal } \mu_L\text{-strip}} K_{\lambda^-_\pi} s_\lambda = \sum_{\lambda} K_{\lambda, \mu} s_\lambda = K_{\lambda, \mu}(\text{by (11.241)})$$

Hence, Exercise 2.68(a) is solved for our partition $\mu$. Thus, Exercise 2.68(a) is solved for every partition $\mu$ satisfying $\ell(\mu) = \lambda$. This completes the induction step, and thus Exercise 2.68(a) is solved by induction.

(b) Let us work in the power series ring $k[[x, y]] := k[[x_1, x_2, \ldots, y_1, y_2, \ldots]]$. Every partition $\mu$ satisfies $h_\mu(x) = \sum_\lambda K_{\lambda, \mu} s_\lambda(x)$, where the sum ranges over all partitions $\lambda$ (because Exercise 2.68(a) yields $h_\mu = \sum_\lambda K_{\lambda, \mu} s_\lambda$ in $\Lambda$). In other words, every partition $\mu$ satisfies $h_\mu(x) = \sum_{\lambda \in \Par} K_{\lambda, \mu} s_\lambda(x)$. On the other hand, every partition $\lambda$ satisfies $s_\lambda = \sum_\mu K_{\lambda, \mu} m_\mu$, where the sum ranges over all partitions $\mu$ (this was shown in the proof of Proposition 2.17). In other words, every partition $\lambda$ satisfies $s_\lambda = \sum_{\mu \in \Par} K_{\lambda, \mu} m_\mu$, so
that
\( (11.242) \quad s_\lambda (y) = \sum_{\mu \in \text{Par}} K_{\lambda, \mu} m_\mu (y). \)

Now,
\[
\prod_{i,j=1}^{\infty} \left( 1 - x_i y_j \right)^{-1} = \sum_{\lambda \in \text{Par}} h_\lambda (x) m_\lambda (y) \quad \text{(by (2.28))}
\]
\[= \sum_{\mu \in \text{Par}} \sum_{\lambda \in \text{Par}} K_{\lambda, \mu} s_\lambda (x) m_\mu (y) \quad \text{(here, we renamed the summation index } \lambda \text{ as } \mu)\]
\[= \sum_{\mu \in \text{Par}} \sum_{\lambda \in \text{Par}} K_{\lambda, \mu} s_\lambda (x) m_\mu (y) = \sum_{\lambda \in \text{Par}} s_\lambda (x) \sum_{\mu \in \text{Par}} K_{\lambda, \mu} m_\mu (y) \]
\[= \sum_{\lambda \in \text{Par}} s_\lambda (x) s_\lambda (y). \]

This proves Theorem 2.31. Thus, Exercise 2.68(b) is solved.

(c) We start out just as in the proof of Proposition 2.17: We restrict our attention to a given homogeneous component \( A_n \) and partitions of \( n \). We will check that the family \( (h_\lambda)_{\lambda \in \text{Par}_n} \) expands unitriangularly in the basis \( (s_\lambda)_{\lambda \in \text{Par}_n} \), for an appropriate choice of orderings in \( \text{Par}_n \) indexing the rows and columns.

In the proof of Proposition 2.17, we showed that any two partitions \( \lambda \) and \( \mu \) satisfy \( K_{\lambda, \mu} = 0 \) unless \( \lambda \triangleright \mu \). Hence, the identity \( h_\mu = \sum_{\lambda \in \text{Par}} K_{\lambda, \mu} s_\lambda \), which holds for every partition \( \mu \) (according to Exercise 2.68(a)), rewrites as \( h_\mu = \sum_{\lambda \in \text{Par}} K_{\lambda, \mu} s_\lambda \). Therefore, we have obtained a triangular expansion of the family \( (h_\lambda)_{\lambda \in \text{Par}_n} \) in the basis \( (s_\lambda)_{\lambda \in \text{Par}_n} \) of \( A_n \), with respect to the opposite of the dominance order. Moreover, this expansion is actually unitriangular (because \( K_{\lambda, \lambda} = 1 \) for every partition \( \lambda \), as shown in the proof of Proposition 2.17). Hence, the family \( (h_\lambda)_{\lambda \in \text{Par}_n} \) has a unitriangular expansion in the basis \( (s_\lambda)_{\lambda \in \text{Par}_n} \) of \( A_n \). Thus, this family \( (h_\lambda)_{\lambda \in \text{Par}_n} \) must itself be a \( k \)-basis of \( A_n \). Since this has been proven for all \( n \in \mathbb{N} \), we can combine this to conclude that the family \( (h_\lambda)_{\lambda \in \text{Par}} \) is a \( k \)-basis of \( A \). This solves Exercise 2.68(c).

**Remark.** Just as in the Remark after the solution of Exercise 2.47(b), we can use our solution of Exercise 2.68(c) to further prove that \( (e_\lambda)_{\lambda \in \text{Par}} \) is a \( k \)-basis of \( A \) (in a different way than we have done in the proof of Proposition 2.17).

### 11.48. Solution to Exercise 2.69.

**Solution to Exercise 2.69.** Let us first observe that
\[ (11.243) \quad 3 (h_n) = e_n \quad \text{for every positive integer } n. \]

(In fact, \( h_n = s_{(n)} \), and so the map \( 3 \) maps \( h_n \) to \( s_{(n)} = s_{(1^n)} = e_n \).)

(a) Fix \( n \in \mathbb{N} \). We need to prove that
\[ (11.244) \quad 3 (fh_n) = 3 (f) \cdot 3 (h_n) \quad \text{for every } f \in A. \]

Since this equality is linear in \( f \), it is clearly enough to prove it when \( f \) is of the form \( s_\lambda \) for some \( \lambda \in \text{Par} \) (because \( (s_\lambda)_{\lambda \in \text{Par}} \) is a \( k \)-basis of \( A \)). In other words, it is enough to show that every \( \lambda \in \text{Par} \) satisfies
\[ (11.245) \quad 3 (s_\lambda h_n) = 3 (s_\lambda) \cdot 3 (h_n). \]

Thus, we will focus on proving (11.245) now.
We WLOG assume that \( n \neq 0 \), since otherwise (11.245) is obvious. Let \( \lambda \in \text{Par} \). The equality (2.36) yields

\[
\ell \lambda h_n = \sum_{\lambda^+ : \lambda^+ / \lambda \text{ is a horizontal n-strip}} s_{\lambda^+}.
\]

Applying the map \( \mathfrak{z} \) to both sides of this equality, we obtain

(11.246) \[
\mathfrak{z} (\ell \lambda h_n) = \mathfrak{z} \left( \sum_{\lambda^+ : \lambda^+ / \lambda \text{ is a horizontal n-strip}} s_{\lambda^+} \right) = \sum_{\lambda^+ : \lambda^+ / \lambda \text{ is a horizontal n-strip}} \mathfrak{z} (s_{\lambda^+}) = \sum_{\lambda^+ : \lambda^+ / \lambda \text{ is a horizontal n-strip}} s_{(\lambda^+)^t}.
\]

But for any given partition \( \lambda^+ \), the assertion that \( \lambda^+ / \lambda \) be a horizontal n-strip is equivalent to the assertion that \( (\lambda^+)^t / \lambda^t \) be a vertical n-strip\(^{372} \). Hence, we can replace the summation sign “\( \sum_{\lambda^+ : \lambda^+ / \lambda \text{ is a}} \)” in (11.246) by a “\( \sum_{\lambda^+ : (\lambda^+)^t / \lambda^t \text{ is a}} \)” sign. Thus, we obtain

\[
\sum_{\lambda^+ : \lambda^+ / \lambda \text{ is a horizontal n-strip}} s_{(\lambda^+)^t} = \sum_{\lambda^+ : (\lambda^+)^t / \lambda^t \text{ is a vertical n-strip}} s_{(\lambda^+)^t} = \sum_{\lambda^+ : \lambda^+ / \lambda \text{ is a horizontal n-strip}} s_{\lambda^+},
\]

(11.247) \[
\mathfrak{z} (s_{\lambda h_n}) = \sum_{\lambda^+ : \lambda^+ / \lambda \text{ is a horizontal n-strip}} s_{(\lambda^+)^t} = \sum_{\lambda^+ : \lambda^+ / \lambda \text{ is a vertical n-strip}} s_{\lambda^+}.
\]

On the other hand, the equality (2.37) yields

\[
\ell \lambda e_n = \sum_{\lambda^+ : \lambda^+ / \lambda \text{ is a vertical n-strip}} s_{\lambda^+}.
\]

Applying this to \( \lambda^t \) instead of \( \lambda \), we obtain

\[
\ell \lambda^t e_n = \sum_{\lambda^+ : \lambda^+ / \lambda^t \text{ is a vertical n-strip}} s_{\lambda^+}.
\]

Now,

\[
\mathfrak{z} (s_{\lambda}) \cdot \mathfrak{z} (h_n) = s_{\lambda^t} e_n = \sum_{\lambda^+ : \lambda^+ / \lambda^t \text{ is a vertical n-strip}} s_{\lambda^+}.
\]

Compared with (11.247), this yields \( \mathfrak{z} (s_{\lambda h_n}) = \mathfrak{z} (s_{\lambda}) \cdot \mathfrak{z} (h_n) \). This proves (11.245). This concludes the solution of Exercise 2.69(a).

(b) We shall show that

(11.248) \[
\mathfrak{z} (h_n) = \omega (h_\lambda)
\]

for every partition \( \lambda \).

Proof of (11.248): We will prove (11.248) by induction over \( \ell (\lambda) \). The induction base is the case \( \ell (\lambda) = 0 \), which is utterly obvious. For the induction step, we fix a positive integer \( L \) and assume (as the induction hypothesis) that the equality (11.248) holds if \( \ell (\lambda) = L - 1 \). We need to prove that the equality (11.248) holds if \( \ell (\lambda) = L \).

Let \( \lambda \) be a partition satisfying \( \ell (\lambda) = L \). Write \( \lambda \) in the form \( \lambda = (\lambda_1, \lambda_2, ..., \lambda_L) \) with all of \( \lambda_1, \lambda_2, ..., \lambda_L \) being positive integers. Let \( \overline{\lambda} \) be the partition \((\lambda_1, \lambda_2, ..., \lambda_{L-1})\) (this is obviously well-defined); then, \( \ell (\overline{\lambda}) = L - 1 \).

\(^{372}\)This is because the notion of a vertical n-strip is obtained from the notion of a horizontal n-strip by interchanging the roles of rows and columns, and the operation which sends a partition \( \mu \) to its transpose partition \( \mu^t \) interchanges the roles of rows and columns as well.
Recall that \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_L) \). Hence, by the definition of \( h_\lambda \), we have
\[
 h_\lambda \equiv h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_L} = h_{\lambda_1 h_{\lambda_2} \cdots h_{\lambda_{L-1}}} h_{\lambda_L} = h_{\lambda}.
\]

Applying the map \( \mathfrak{Z} \) to both sides of this equality, we obtain
\[
\mathfrak{Z}(h_\lambda) \equiv \mathfrak{Z}(h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_L}) = \mathfrak{Z}(h_{\lambda_1 h_{\lambda_2} \cdots h_{\lambda_{L-1}}} h_{\lambda_L}) = \mathfrak{Z}(h_{\lambda}).
\]

(by Exercise 2.69(a), applied to \( \lambda \) and \( \lambda_L \) instead of \( \lambda \) and \( n \) for every \( n \geq 1 \))

In other words, the equality (11.248) holds if \( \ell(\lambda) = L. \) This completes the induction step. The induction proof of (11.248) is therefore complete.

The equality (11.248) shows that the \( k \)-linear maps \( \mathfrak{Z} : \Lambda \to \Lambda \) and \( \omega : \Lambda \to \Lambda \) are equal to each other on every element of the basis \((h_\lambda)_{\lambda \in \text{Par}}\) of the \( k \)-module \( \Lambda \). Hence, \( \mathfrak{Z} = \omega. \) This solves Exercise 2.69(b).

(c) Fix two partitions \( \mu \) and \( \nu. \) The first identity in (2.26) yields
\[
(11.249) \quad s_\mu s_\nu = \sum_{\lambda \in \text{Par}} c^\lambda_{\mu, \nu} s_\lambda.
\]

Applying this to \( \mu^t \) and \( \nu^t \) instead of \( \mu \) and \( \nu \), we obtain
\[
(11.250) \quad s_\mu s_\nu^t = \sum_{\lambda \in \text{Par}} c^\lambda_{\mu, \nu^t} s_\lambda.
\]

Applying the map \( \mathfrak{Z} \) to the identity (11.249), we obtain
\[
\mathfrak{Z}(s_\mu s_\nu) = \mathfrak{Z} \left( \sum_{\lambda \in \text{Par}} c^\lambda_{\mu, \nu} s_\lambda \right) = \sum_{\lambda \in \text{Par}} \left( \mathfrak{Z}(c^\lambda_{\mu, \nu}) \right) s_\lambda^t = \sum_{\lambda \in \text{Par}} c^\lambda_{\mu, \nu^t} s_\lambda^t.
\]

But recall that \( \omega \) is a \( k \)-algebra homomorphism. Since \( \mathfrak{Z} = \omega \) (by Exercise 2.69(b)), this shows that \( \mathfrak{Z} \) is a \( k \)-algebra homomorphism. Hence,
\[
\mathfrak{Z}(s_\mu s_\nu) = \mathfrak{Z}(s_\mu) \cdot \mathfrak{Z}(s_\nu) = s_\mu^t s_\nu^t = \sum_{\lambda \in \text{Par}} c^\lambda_{\mu^t, \nu^t} s_\lambda \quad \text{by (11.250)).}
\]

Compared with (11.251), this yields \( \sum_{\lambda \in \text{Par}} c^\lambda_{\mu^t, \nu^t} s_\lambda = \sum_{\lambda \in \text{Par}} c^\lambda_{\mu, \nu^t} s_\lambda. \) Since \((s_\lambda)_{\lambda \in \text{Par}}\) is a basis of the \( k \)-module \( \Lambda \), we can compare coefficients in this equality, and obtain
\[
\begin{align*}
(11.252) \quad c^\lambda_{\mu^t, \nu^t} &= c^\lambda_{\mu^t, \nu^t^t} \quad \text{for every } \lambda \in \text{Par}.
\end{align*}
\]

We can substitute \( \lambda^t \) for \( \lambda \) in this result, and conclude that \( c^\lambda_{\mu^t, \nu^t} = c^\lambda_{\mu^t, \nu^t^t} \) for every \( \lambda \in \text{Par}. \) Since \((\lambda^t)^t = \lambda^t \), this rewrites as \( c^\lambda_{\mu, \nu^t} = c^\lambda_{\mu^t, \nu^t}. \) This solves Exercise 2.69(c).
(d) Let $\mu$ and $\lambda$ be two partitions such that $\mu \subseteq \lambda$. In Remark 2.37, it has been shown that $s_{\lambda/\mu} = \sum_{\nu} c_{\mu,\nu}^{\lambda} s_{\nu}$, where the sum ranges over all partitions $\nu$. In other words,

\begin{equation}
(11.253)
\sum_{\nu} c_{\mu,\nu}^{\lambda} s_{\nu}.
\end{equation}

Applying this to $\lambda^t$ and $\mu^t$ instead of $\lambda$ and $\mu$, we obtain

\begin{equation}
(11.254)
\sum_{\nu} c_{\mu^t,\nu}^{\lambda^t} s_{\nu}.
\end{equation}

Applying the map $\mathfrak{f}$ to both sides of the identity (11.253), we obtain

\begin{equation}
\mathfrak{f} \left( \sum_{\nu} c_{\mu,\nu}^{\lambda} s_{\nu} \right) = \sum_{\nu} c_{\mu^t,\nu}^{\lambda^t} s_{\nu}.
\end{equation}

(by Exercise 2.69(c))

\begin{equation}
\left( \text{here, we substituted } \nu \text{ for } \nu^t \text{ in the sum, since the map } \text{Par} \rightarrow \text{Par}, \; \nu \rightarrow \nu^t \text{ is a bijection} \right)
\end{equation}

\begin{equation}
= s_{\lambda^t/\mu^t}.
\end{equation}

Since $\mathfrak{f} = \varpi$ (by Exercise 2.60(b)), this rewrites as $\omega \left( s_{\lambda/\mu} \right) = s_{\lambda^t/\mu^t}$. This proves the first identity of (2.21).

It is clear that the skew Schur function $s_{\lambda/\mu}$ is homogeneous of degree $|\lambda/\mu|$. In other words, $s_{\lambda/\mu} \in \Lambda_{|\lambda/\mu|}$.

Hence, (2.20) (applied to $f = s_{\lambda/\mu}$ and $n = |\lambda/\mu|$) yields

\begin{equation}
S \left( s_{\lambda/\mu} \right) = (-1)^{|\lambda/\mu|} \omega \left( s_{\lambda/\mu} \right) = (-1)^{|\lambda/\mu|} s_{\lambda^t/\mu^t}.
\end{equation}

This proves the second identity of (2.21). The proof of (2.21) is thus complete.

11.49. **Solution to Exercise 2.70.** Solution to Exercise 2.70. (a)

First solution to Exercise 2.70(a): Let us first show that $\prod_{i,j=1}^{\infty} (1 + x_i y_j) = \sum_{\lambda \in \text{Par}} e_{\lambda}(x) m_{\lambda}(y)$.

In fact, in the proof of Proposition 2.43, we showed that

\begin{equation}
\prod_{i=1}^{\infty} (1 + x_i t) = \sum_{n \geq 0} e_n(x) t^n
\end{equation}

in the ring $(k[[x]])[[t]]$. For every $j \in \{1, 2, 3, \ldots\}$, we have

\begin{equation}
\prod_{i=1}^{\infty} (1 + x_i y_j) = \sum_{n \geq 0} e_n(x) y_j^n
\end{equation}

in the ring $(k[[x]])[[y]]$ (in fact, this results from (11.256) by substituting $y_j$ for $t$). Thus, (11.257) holds in $k[[x, y]]$ (since $(k[[x]])[[y]] = k[[x, y]]$ as rings). Now,

\begin{equation}
\prod_{i,j=1}^{\infty} (1 + x_i y_j) = \prod_{j=1}^{\infty} \prod_{i=1}^{\infty} (1 + x_i y_j) = \prod_{j=1}^{\infty} \sum_{n \geq 0} e_n(x) y_j^n
\end{equation}

(by (11.257))

\begin{equation}
= \sum_{\lambda \in \text{Par}} e_{\lambda}(x) m_{\lambda}(y) \quad \text{(by (11.255))}.
\end{equation}
Next, we shall show that \[\sum_{\lambda \in \text{Par}} s_{\lambda}^{x}(x) s_{\lambda'}^{y}(y) = \sum_{\lambda \in \text{Par}} e_{\lambda}(x) m_{\lambda}(y).\]

In fact, consider the \(k\)-algebra \((k[[x]])[[y]] = k[[x,y]].\) This \(k\)-algebra \((k[[x]])[[y]]\) is a \(k[[y]]\)-algebra, and contains \(\Lambda[[y]]\) as a \(k[[y]]\)-subalgebra.

The equality (2.24) yields

\[
\sum_{\lambda \in \text{Par}} s_{\lambda}(x) s_{\lambda}(y) = \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \text{Par}} h_{\lambda}(x) m_{\lambda}(y) \quad \text{(by (2.28)).}
\]

This is an equality in \(\Lambda[[y]]\) (because \(s_{\lambda}(x)\) and \(h_{\lambda}(x)\) belong to \(\Lambda\) for every \(\lambda \in \text{Par}\)).

Recall that \(\omega: \Lambda \rightarrow \Lambda\) is a \(k\)-algebra homomorphism. It thus induces a \((k[[y]])\)-algebra homomorphism \(\omega[[y]]: \Lambda[[y]] \rightarrow \Lambda[[y]]\) which sends every \(q \in \Lambda\) to \(\omega(q)\), and is continuous with respect to the usual topology\(^{373}\) on \(\Lambda[[y]]\). Applying this homomorphism \(\omega[[y]]\) to both sides of (11.259), we obtain

\[
\sum_{\lambda \in \text{Par}} \omega(s_{\lambda}(x)) s_{\lambda}(y) = \sum_{\lambda \in \text{Par}} \omega(h_{\lambda}(x)) m_{\lambda}(y)
\]

(because \(\omega[[y]]\), being a \((k[[y]])\)-algebra homomorphism, leaves the \(s_{\lambda}(y)\) and \(m_{\lambda}(y)\) terms unchanged, while the \(s_{\lambda}(x)\) and \(h_{\lambda}(x)\) terms are elements of \(\Lambda\) and thus are transformed as by \(\omega\)).

But we know that

\[
\omega(h_n) = e_n \quad \text{for every positive integer } n.
\]

Using this fact and the fact that \(\omega\) is a \(k\)-algebra homomorphism, it is easy to see that

\[
\omega(h_{\lambda}) = e_{\lambda} \quad \text{for every partition } \lambda
\]

(because if \(\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{\ell})\) with \(\ell = \ell(\lambda)\), then \(h_{\lambda} = h_{\lambda_1} h_{\lambda_2} \cdots h_{\lambda_{\ell}}\) and \(e_{\lambda} = e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_{\ell}}\)). Furthermore,

\[
\omega(s_{\lambda}) = s_{\lambda'} \quad \text{for every partition } \lambda.
\]

(This follows from the first equality in (2.21) by setting \(\mu = \varnothing\).) Using (11.262) and (11.263), we can rewrite (11.260) as

\[
\sum_{\lambda \in \text{Par}} s_{\lambda}(x) s_{\lambda}(y) = \sum_{\lambda \in \text{Par}} e_{\lambda}(x) m_{\lambda}(y).
\]

Hence,

\[
\sum_{\lambda \in \text{Par}} e_{\lambda}(x) m_{\lambda}(y) = \sum_{\lambda \in \text{Par}} s_{\lambda'}(x) s_{\lambda}(y) = \sum_{\lambda \in \text{Par}} \sum_{\lambda' \in \text{Par}} s_{(\lambda')':(x)} = s_{\lambda}(x) \quad (\text{since } (\lambda')' = \lambda)
\]

\[
= \sum_{\lambda \in \text{Par}} s_{\lambda}(x) s_{\lambda'}(y).
\]

Combined with (11.258), this yields

\[
\prod_{i,j=1}^{\infty} (1 + x_i y_j) = \sum_{\lambda \in \text{Par}} e_{\lambda}(x) m_{\lambda}(y) = \sum_{\lambda \in \text{Par}} s_{\lambda}(x) s_{\lambda'}(y).
\]

This solves Exercise 2.70(a).

**Second solution to Exercise 2.70(a):** We can prove (11.258) as in the First solution to Exercise 2.70(a). Thus, in order to solve Exercise 2.70(a), it remains to verify

\[
\prod_{i,j=1}^{\infty} (1 + x_i y_j) = \sum_{\lambda \in \text{Par}} s_{\lambda}(x) s_{\lambda'}(y).
\]

\(^{373}\)The usual topology on a power series ring \(Z[[y]]\) (where \(Z\) is a commutative ring) is the direct-product topology obtained by viewing the set \(Z[[y]]\) as a direct product of many copies of \(Z\) (this is done by identifying every power series with the family of its coefficients), each of which is endowed with the discrete topology. In this topology, a sequence (or, more generally, a net) \((f_i)\) of power series converges to a power series \(f\) if and only if for every monomial \(m\), the sequence (resp. net) (the coefficient of \(m\) in \(f_i\)) converges to (the coefficient of \(m\) in \(f\)) with respect to the discrete topology.
Instead of proving (11.266), we will verify the identity
\[
(11.267) \quad \prod_{i,j=1}^{\infty} (1 + tx_i y_j) = \sum_{\lambda \in \text{Par}} t^{\lambda} s_\lambda (x) s_{\lambda'} (y)
\]
in the ring \( R(x, y) [[t]] \). This identity will clearly yield (11.266) (by substituting \( t = 1 \)), and thus conclude the solution of Exercise 2.70(a).

We shall prove (11.267) in a similar way to how we proved (2.25), but using a variation on the RSK correspondence. This is not in itself a new idea; indeed, this is how the equivalent identity (11.266) is proven in [157, Theorem 7.14.3], in [82, §7] and in [141, Theorem 4.8.6]. However, instead of using the dual RSK algorithm (also known as the RSK* algorithm; see [157, §7.14], [82, §5] and [141, Theorem 4.8.5] for it) like these proofs do, we introduce a variation of the RSK algorithm that relies on the same RS-insertion operation but changes the order in which the biletters are processed. (We will reprove (11.267) using the dual RSK algorithm in the Third solution further below.)

For a given partition \( \lambda \), let us define a row-strict tableau of shape \( \lambda \) to be an assignment \( T \) of entries in \( \{1, 2, 3, \ldots \} \) to the cells of the Ferrers diagram for \( \lambda \) which is weakly increasing left-to-right in columns, and strictly increasing top-to-bottom in rows. It is clear that if \( \lambda \) is a partition, then the row-strict tableaux of shape \( \lambda \) are in 1-to-1 correspondence with the column-strict tableaux of shape \( \lambda' \), and the correspondence is given by transposing the tableau (i.e., taking whatever entry was assigned to a cell \( (i, j) \) in the input tableau, and reassigning it to the cell \( (j, i) \) in the output tableau). Hence, for every partition \( \lambda \), we have
\[
(11.268) \quad \sum_{Q \text{ is a row-strict tableau of shape } \lambda} x^{\text{cont}(Q)} = \sum_{T \text{ is a column-strict tableau of shape } \lambda'} x^{\text{cont}(T)} = s_{\lambda'} \quad \text{(since this is how } s_{\lambda'} \text{ is defined)}
\]
where \( x^{\text{cont}(Q)} \) is defined for a row-strict tableau \( Q \) in the same way as it is defined for a column-strict tableau \( Q \). Substituting \( y \) for \( x \) in this equality, we obtain
\[
(11.269) \quad \sum_{Q \text{ is a row-strict tableau of shape } \lambda} y^{\text{cont}(Q)} = s_{\lambda'} (y).
\]
We also have
\[
(11.270) \quad \sum_{P \text{ is a column-strict tableau of shape } \lambda} x^{\text{cont}(P)} = s_{\lambda} (x)
\]
(because this is how \( s_{\lambda} = s_{\lambda}(x) \) is defined) for every partition \( \lambda \).

A tableau-cotableau pair will mean a pair \( (P, Q) \) such that \( P \) is a column-strict tableau, \( Q \) is a row-strict tableau, and \( P \) and \( Q \) both have shape \( \lambda \) for one and the same partition \( \lambda \). Multiplying the identities (11.270) and (11.269) and multiplying the result with \( t^{\lambda} \), we obtain
\[
(11.271) \quad \sum_{(P, Q) \text{ is a pair with } P \text{ being a column-strict tableau of shape } \lambda \text{ and } Q \text{ being a row-strict tableau of shape } \lambda} t^{\lambda} x^{\text{cont}(P)} y^{\text{cont}(Q)} = t^{\lambda} s_{\lambda} (x) s_{\lambda'} (y)
\]
for every partition \( \lambda \). Summing this equality over all partitions \( \lambda \), we obtain
\[
(11.271) \quad \sum_{\lambda \in \text{Par}} t^{\lambda} x^{\text{cont}(P)} y^{\text{cont}(Q)} = \sum_{\lambda \in \text{Par}} t^{\lambda} s_{\lambda} (x) s_{\lambda'} (y),
\]
where the sum on the left hand side is over all tableau-cotableau pairs \( (P, Q) \) and where \( \lambda \) denotes the common shape of \( P \) and \( Q \).

We shall use all notations introduced in the proof of Theorem 2.31. Define the antilexicographic order \( \leq_{\text{alex}} \) to be the total order on the set of all biletters which is given by
\[
\left( \begin{array}{c} i_1 \\ j_1 \end{array} \right) \leq_{\text{alex}} \left( \begin{array}{c} i_2 \\ j_2 \end{array} \right) \iff \text{we have } i_1 \leq i_2, \text{ and if } i_1 = i_2, \text{ then } j_1 \geq j_2.
\]
We denote by \(<_{\text{alex}}\) the (strict) smaller relation of this order. A strict cobiword will mean an array \((i^1_j^1) \cdots (i^i_j^i)\) in which the biletters satisfy \((i^1_j^1) <_{\text{alex}} \cdots <_{\text{alex}} (i^i_j^i)\) (that is, the biletters are distinct and ordered). Strict cobiwords are clearly in 1-to-1 correspondence with sets (not multisets!) of biletters. Now, the left hand side of (11.267) is \(\prod_{i,j=1}^{\infty} (1 + tx_i y_j) = \prod_{i,j=1}^{\infty} (1 + tx_j y_i)\) (here, we substituted \((j,i)\) for the summation index \((i,j)\)), and thus can be rewritten as the sum of \(t^{\ell} x_{j^1} y_{i^1} x_{j^2} y_{i^2} \cdots x_{j^\ell} y_{i^\ell}\) over all sets \(\{(i^1_j^1), \ldots, (i^\ell_j^\ell)\}\) of biletters. Thus, the left hand side of (11.267) is the sum \(\sum t^{\ell} x_{\text{cont}(j)} y_{\text{cont}(i)}\) over all strict cobiwords \((i_j)\), where \(\ell\) stands for the number of biletters in the biword. Meanwhile, we know that the right hand side of (11.267) is the sum \(\sum |\lambda| x_{\text{cont}(P)} y_{\text{cont}(Q)}\) over all tableau-cotableau pairs \((P,Q)\) (because of (11.271)). Thus, in order to prove (11.267), we only need to construct a bijection between the strict cobiwords \((i_j)\) and the tableau-cotableau pairs \((P,Q)\), which has the property that

\[
\text{cont}(i) = \text{cont}(Q);
\]
\[
\text{cont}(j) = \text{cont}(P).
\]

This bijection is the coRSK algorithm, which we shall now define.\(^{374}\)

Let \((i_j)\) be a strict cobiword. Starting with the pair \((P_0, Q_0) = (\emptyset, \emptyset)\) and \(m = 0\), the algorithm applies the following steps (see Example 11.5 below):

- If \(i_{m+1}\) does not exist (that is, \(m\) is the length of \(i\)), stop.
- Apply RS-insertion to the column-strict tableau \(P_m\) and the letter \(j_{m+1}\) (the bottom letter of \((i^m_{j_{m+1}})\)). Let \(P_{m+1}\) be the resulting column-strict tableau, and let \(c_{m+1}\) be the resulting corner cell.
- Create \(Q_{m+1}\) from \(Q_m\) by adding the top letter \(i_{m+1}\) of \((i^m_{j_{m+1}})\) to \(Q_m\) in the cell \(c_{m+1}\) (which, as we recall, is the extra corner cell of \(P_{m+1}\) not present in \(P_m\)).
- Set \(m\) to \(m + 1\).

After all of the biletters have been thus processed, the result of the coRSK algorithm is \((P_\ell, Q_\ell) =: (P,Q)\).

**Example 11.5.** The term in the expansion of the left side of (11.266) corresponding to

\[
(x_4y_1)(x_2y_1)(x_1y_2)(x_3y_4)(x_1y_4)(x_2y_5)
\]

\(^{374}\)Such a bijection will then automatically satisfy \(|\lambda| = \ell\), where \(\lambda\) is the (common) shape of \(P\) and \(Q\), and where \(\ell\) is the length of the strict cobiword \((i_j)\). This is because \(|\lambda| = |\text{cont}(Q)|\) and \(\ell = |\text{cont}(i)|\).

\(^{375}\)This algorithm appears in Fulton’s [54, §A.4] as construction (1d).
is the strict cobiword \( \binom{ij}{1} = \binom{112445}{421312} \), and the coRSK algorithm applied to this cobiword proceeds as follows:

\[
\begin{align*}
P_0 &= \emptyset & Q_0 &= \emptyset \\
P_1 &= 4 & Q_1 &= 1 \\
P_2 &= 2 & Q_2 &= 1 \\
P_3 &= 2 & Q_3 &= 1 \\
P_4 &= 2 & Q_4 &= 1 \\
P_5 &= 2 & Q_5 &= 1 \\
P_6 &= 1 & Q_6 &= 1 \\
\end{align*}
\]

It is clear that \( P_m \) remains a column-strict tableau of some Ferrers shape throughout the execution of the coRSK algorithm, and that \( Q_m \) remains a filling of the same shape as \( P_m \) which is (at least) weakly increasing left-to-right along rows and weakly increasing top-to-bottom in columns. But we can also see that \( Q_m \) is strictly increasing left-to-right along rows\(^{376}\), so that \( Q_m \) is a row-strict tableau. Thus, the result \((P, Q)\) of the coRSK algorithm is a tableau-cotableau pair.

To see that the coRSK algorithm is a bijection, we show how to recover \( \binom{ij}{1} \) from \((P, Q)\). This is done by reverse bumping in the same way as for the usual RSK algorithm, with the only difference that now \( Q_m \) is obtained by removing the bottommost (rather than the rightmost) occurrence of the letter \( i_{m+1} \) from \( Q_{m+1} \).

Finally, to see that the coRSK map is surjective, one needs to show that the reverse bumping procedure can be applied to any tableau-cotableau pair \((P, Q)\), and will result in a strict cobiword \( \binom{ij}{1} \). We leave this verification to the reader.

So the coRSK map is a bijection having the required properties. As we have said, this proves (11.267). This solves Exercise 2.70(a) again.

\(^{376}\)Indeed, this follows from the observation that when one has a string of equal letters \( i_m = i_{m+1} = \cdots = i_{m+r} \) on top of the strict cobiword, then the bottom letters bumped in are \( j_m > j_{m+1} > \cdots > j_{m+r} \), and therefore (as a consequence of the last claim of part (b) of the row bumping lemma) the new cells form a vertical strip, that is, no two of these cells lie in the same row. Actually, more can be said: Each of these new cells (except for the first one) is in a row further down than the previous one. We will use this stronger fact further below.

\(^{377}\)It necessarily has to be the bottommost occurrence, since (according to the previous footnote) the cell into which \( i_{m+1} \) was filled at the step from \( Q_m \) to \( Q_{m+1} \) lies further down than any existing cell of \( Q_m \) containing the letter \( i_{m+1} \).

\(^{378}\)It is easy to see that repeatedly applying reverse bumping to \((P, Q)\) will result in a sequence \( \binom{i_1}{j_1} <_{\text{alex}} \binom{i_2}{j_2} <_{\text{alex}} \cdots <_{\text{alex}} \binom{i_{r+1}}{j_{r+1}} \). Since the chain of inequalities \( i_1 \leq i_2 \leq \cdots \leq i_r \) is clear from the choice of entry to reverse-bump, it only remains to show that for every string \( i_m = i_{m+1} = \cdots = i_{m+r} \) of equal top letters, the corresponding bottom letters strictly decrease (that is, \( j_m > j_{m+1} > \cdots > j_{m+r} \)). One way to see it is to apply part (a) of the row bumping lemma.
Indeed, using (11.262), we can rewrite (11.260) as
\[ \sum_{\lambda \in \text{Par}} \omega(s_\lambda(x)) s_\lambda(y) = \sum_{\lambda \in \text{Par}} e_\lambda(x) m_\lambda(y). \]

Compared with (11.258), this yields
\[
\sum_{\lambda \in \text{Par}} \omega(s_\lambda(x)) s_\lambda(y) = \prod_{i,j=1}^{\infty} (1 + x_i y_j) = \prod_{i,j=1}^{\infty} (1 + x_j y_i) \quad \text{(here, we substituted \((j, i)\) for \((i, j)\) in the sum)}
\]
\[ = \prod_{i,j=1}^{\infty} (1 + y_i x_j) = \sum_{\lambda \in \text{Par}} s_\lambda(y) s_{\lambda'}(x) \quad \text{(by (11.266), with \(x\) and \(y\) substituted for \(y\) and \(x\))}
\]
\[ = \sum_{\lambda \in \text{Par}} s_{\lambda'}(x) s_\lambda(y). \]

Since the \(s_\lambda(y)\) for \(\lambda \in \text{Par}\) are linearly independent over \(k[|x|]\), we can compare coefficients before \(s_\lambda(y)\) in this equality. As a result, we obtain \(\omega(s_\lambda(x)) = s_{\lambda'}(x)\) for every \(\lambda \in \text{Par}\). Thus, (11.263) is proven again.

**Remark:** By combining the two solutions of Exercise 2.70(a) given above, one can obtain a new proof of the equality (11.263) (a proof which does not use (2.21)). Indeed, using (11.262), we can rewrite (11.260) as
\[ \sum_{\lambda \in \text{Par}} \omega(s_\lambda(x)) s_\lambda(y) = \sum_{\lambda \in \text{Par}} e_\lambda(x) m_\lambda(y). \]

Third solution to Exercise 2.70(a): In order to solve Exercise 2.70(a), it suffices to verify the identity (11.267). (This can be proven by the same argument as in the Second solution to Exercise 2.70(a)).

We shall now verify the identity (11.267) using the so-called dual RSK algorithm (also known as RSK* algorithm)\(^{379}\). This will be fairly similar to the Second solution to Exercise 2.70(a) given above, but not identical to it; in particular, the dual RSK algorithm that we will introduce below will (unlike the coRSK algorithm from the Second solution) not rely on the same row bumping operation as the usual RSK algorithm, but on a somewhat modified version of it.

For a given partition \(\lambda\), let us define a row-strict tableau of shape \(\lambda\) to be an assignment \(T\) of entries in \(\{1, 2, 3, \ldots\}\) to the cells of the Ferrers diagram for \(\lambda\) which is weakly increasing left-to-right in columns, and strictly increasing top-to-bottom in rows. The equality (11.268) holds\(^{380}\), thus, we have
\[ \sum_{P\text{ a row-strict tableau of shape }\lambda} x^{\text{cont}(P)} = \sum_{Q\text{ a row-strict tableau of shape }\lambda} x^{\text{cont}(Q)} = s_{\lambda'}(x). \quad \text{(by (11.268))} \]
(11.272)

We also have
\[ \sum_{Q\text{ a column-strict tableau of shape }\lambda} x^{\text{cont}(Q)} = \sum_{P\text{ a column-strict tableau of shape }\lambda} x^{\text{cont}(P)} = s_{\lambda}(x). \]
(because this is how \(s_{\lambda} = s_{\lambda}(x)\) is defined) for every partition \(\lambda\). Substituting \(y\) for \(x\) in this equality, we obtain
\[ \sum_{Q\text{ a column-strict tableau of shape }\lambda} y^{\text{cont}(Q)} = s_{\lambda}(y). \]
(11.273)

A cotableau-tableau pair will mean a pair \((P, Q)\) such that \(P\) is a row-strict tableau, \(Q\) is a column-strict tableau, and \(P\) and \(Q\) both have shape \(\lambda\) for one and the same partition \(\lambda\). Multiplying the identities (11.272) and (11.273) and multiplying the result with \(t^{[\lambda]}\), we obtain
\[ t^{[\lambda]} \sum_{(P, Q)\text{ is a pair with }\text{P being a row-strict tableau of shape }\lambda, \text{Q being a column-strict tableau of shape }\lambda} x^{\text{cont}(P)} y^{\text{cont}(Q)} = t^{[\lambda]} s_{\lambda'}(x) s_{\lambda}(y). \]

\(^{379}\)We will define this algorithm further below. It also frequently appears in literature: see, e.g., [157, §7.14], [82, §5], [141, Theorem 4.8.5] and (with different conventions) [54, §A.4.3, Prop. 3].

\(^{380}\)Indeed, it can be proven as in the Second solution to Exercise 2.70(a).
for every partition $\lambda$. Summing this equality over all partitions $\lambda$, we obtain
\[
\sum_{\lambda \in \text{Par}} t^{\lambda|} x^{\text{cont}(P)} y^{\text{cont}(Q)} = \sum_{\lambda \in \text{Par}} t^{\lambda|} s_{\lambda'}(x) s_{\lambda}(y),
\]
where the sum on the left hand side is over all cotableau-tableau pairs $(P, Q)$ and where $\lambda$ denotes the common shape of $P$ and $Q$. This becomes
\[
\sum_{\lambda \in \text{Par}} t^{\lambda|} x^{\text{cont}(P)} y^{\text{cont}(Q)} = \sum_{\lambda \in \text{Par}} t^{\lambda|} s_{\lambda'}(x) s_{\lambda}(y) = \sum_{\lambda \in \text{Par}} t^{\lambda|} \frac{s_{(\lambda')}^t(x)}{s_{\lambda}(x)} = \sum_{\lambda \in \text{Par}} t^{\lambda|} s_{\lambda'}(x) s_{\lambda}(y) = \sum_{\lambda \in \text{Par}} t^{\lambda|} s_{\lambda'}(x) s_{\lambda}(y).
\]

We shall use all notations introduced in the proof of Theorem 2.31. A strict biword will mean an array \(\binom{j}{i} = (i_1^{(j_1)} \cdots i_j^{(j_j)})\) in which the biletters satisfy \(i_1^{(j_1)} <_\text{lex} \cdots <_\text{lex} i_j^{(j_j)}\) (that is, the biletters are distinct and ordered with respect to the lexicographic order). Strict biwords are clearly in 1-to-1 correspondence with sets (not multisets!) of biletters. Now, the left hand side of (11.267) is \(\prod_{i,j=1}^{\infty} (1 + tx_jy_i) = \prod_{i,j=1}^{\infty} (1 + tx_jy_i)\) (here, we substituted \((j, i)\) for the summation index \((i, j)\), and thus can be rewritten as the sum of \(t^\ell(x_1y_1) (x_2y_2) \cdots (x_jy_i)\) over all sets \(\binom{i}{j_1}, \ldots, \binom{i}{j_j}\) of biletters. Thus, the left hand side of (11.267) is the sum \(\sum t^\ell x^{\text{cont}(j)} y^{\text{cont}(i)}\) over all strict biwords \(\binom{j}{i}\), where \(\ell\) stands for the number of biletters in the biword. Meanwhile, we know that the right hand side of (11.267) is the sum \(\sum t^{\lambda|} x^{\text{cont}(P)} y^{\text{cont}(Q)}\) over all cotableau-tableau pairs $(P, Q)$ (because of (11.274)). Thus, in order to prove (11.267), we only need to construct a bijection between the strict biwords \(\binom{j}{i}\) and the cotableau-tableau pairs $(P, Q)$, which has the property that
\[
\text{cont}(i) = \text{cont}(Q); \quad \text{cont}(j) = \text{cont}(P).
\]

This bijection is the dual RSK algorithm, which we shall define below.

First, we shall introduce a simpler operation which we call dual RS-insertion (and which is similar to RS-insertion, but not identical with it).

Dual RS-insertion takes as input a row-strict tableau $P$ and a letter $j$, and returns a new row-strict tableau $P'$ along with a corner cell $c$ of $P'$, which is constructed as follows: Start out by setting $P' = P$. The letter $j$ tries to insert itself into the first row of $P'$ by either bumping out the leftmost letter in the first row larger or equal to $j$, or else placing itself at the right end of the row if no such larger letter exists. If a letter was bumped from the first row, this letter follows the same rules to insert itself into the second row, and so on. This series of bumps must eventually come to an end. At the end of the bumping, the tableau $P'$ created has an extra corner cell not present in $P$. If we call this corner cell $c$, then $P'$ (in its final form) and $c$ are what the dual RS-insertion operation returns. One says that $P'$ is the result of dually inserting $c$ into the tableau $P$. It is straightforward to see that this resulting filling $P'$ is a row-strict tableau.

**Example 11.6.** To give an example of this operation, let us dually insert the letter $j = 3$ into the row-strict tableau
\[
\begin{array}{cccc}
1 & 2 & 3 & 5 \\
1 & 2 & 4 \\
3 & 4 & 7
\end{array}
\]
we are showing all intermediate states of $P'$; the underlined letter is always the one
\[
\begin{array}{cccc}
1 & 2 & 3 & 5 \\
1 & 2 & 4 & 7
\end{array}
\]

381Such a bijection will then automatically satisfy \(|\lambda| = \ell\), where $\lambda$ is the (common) shape of $P$ and $Q$, and where $\ell$ is the length of the strict biword \(\binom{j}{i}\). This is because \(|\lambda| = |\text{cont}(Q)|\) and $\ell = |\text{cont}(i)|$.

382We will leave many of its properties unproven, because their proofs are analogous (or at least very similar) to the proofs of the corresponding properties of RS-insertion, and thus can be easily reconstructed by the reader.

383Indeed, the reader can check that $P'$ remains a row-strict tableau throughout the algorithm that defines dual RS-insertion.
that is going to be bumped out at the next step):

\[
\begin{array}{c|c|c}
1 & 2 & 3 \\
1 & 2 & 4 \\
3 & 4 & 7 \\
\end{array} \quad \begin{array}{c|c|c}
1 & 2 & 3 \\
1 & 2 & 4 \\
3 & 4 & 7 \\
\end{array} \quad \begin{array}{c|c|c}
1 & 2 & 3 \\
1 & 2 & 3 \\
3 & 4 & 7 \\
\end{array}
\]

The last tableau in this sequence is the row-strict tableau that is returned. The corner cell that is returned
is the second cell of the fifth row (the one containing 6).

Dual RS-insertion will be used as a step in the dual RSK algorithm; the construction will rely on a simple
fact known as the dual row bumping lemma. Let us first define the notion of a dual bumping path (or dual
bumping route): If \( P \) is a row-strict tableau, and \( j \) is a letter, then some letters are inserted into some
cells when dual RS-insertion is applied to \( P \) and \( j \). The sequence of these cells (in the order in which they
see letters inserted into them)\(^{384}\) is called the dual bumping path for \( P \) and \( j \). This dual bumping path
always ends with the corner cell \( c \) which is returned by dual RS-insertion. As an example, when \( j = 3 \) is
dually inserted into the tableau \( P \) shown below, the result \( P' \) is shown with all entries on the bumping path
underlined:

\[
P = \begin{array}{c|c|c}
1 & 2 & 3 \\
1 & 2 & 4 \\
3 & 4 & 7 \\
\end{array} \quad \begin{array}{c|c|c}
1 & 2 & 4 \\
1 & 2 & 4 \\
3 & 4 & 7 \\
\end{array} \quad \begin{array}{c|c|c}
1 & 2 & 3 \\
1 & 2 & 3 \\
3 & 4 & 7 \\
\end{array}
\]

A first simple observation about dual bumping paths is that dual bumping paths trend weakly left (just as
bumping paths for regular RS-insertion do). A subtler property of bumping paths is the following dual row
bumping lemma\(^{385}\):

**Dual row bumping lemma:** Let \( P \) be a row-strict tableau, and let \( j \) and \( j' \) be two letters.
Applying dual RS-insertion to the tableau \( P \) and the letter \( j \) yields a new row-strict tableau
\( P' \) and a corner cell \( c \). Applying dual RS-insertion to the tableau \( P' \) and the letter \( j' \) yields
a new row-strict tableau \( P'' \) and a corner cell \( c' \).

(a) Assume that \( j < j' \). Then, the dual bumping path for \( P' \) and \( j' \) stays strictly to the
right, within each row, of the dual bumping path for \( P \) and \( j \). The cell \( c' \) (in which
the dual bumping path for \( P' \) and \( j' \) ends) is in the same row as the cell \( c \) (in which
the dual bumping path for \( P \) and \( j \) ends) or in a row further up; it is also in a column
further right than \( c \).

(b) Assume instead that \( j \geq j' \). Then, the dual bumping path for \( P' \) and \( j' \) stays weakly to
the left, within each row, of the dual bumping path for \( P \) and \( j \). The cell \( c' \) (in which
the dual bumping path for \( P' \) and \( j' \) ends) is in a row further down than the cell \( c \) (in
which the dual bumping path for \( P \) and \( j \) ends); it is also in the same column as \( c \) or
in a column further left.

This lemma can be easily proven by induction over the row (similarly to the usual row bumping lemma).

We will now define the dual RSK algorithm. Let \( (\emptyset, \emptyset) \) be a strict biword. Starting with the pair
\( (P_0, Q_0) = (\emptyset, \emptyset) \) and \( m = 0 \), the algorithm applies the following steps (see Example 11.7 below):

- If \( i_{m+1} \) does not exist (that is, \( m \) is the length of \( \emptyset \)), stop.
- Apply dual RS-insertion to the row-strict tableau \( P_m \) and the letter \( j_{m+1} \) (the bottom letter of
  \( (i_{m+1}) \)). Let \( P_{m+1} \) be the resulting row-strict tableau, and let \( c_{m+1} \) be the resulting corner cell.

\(^{384}\)In particular, this includes those cells whose entries did not change under the insertion (because the entry inserted was
the same as the entry they contained before the insertion).

\(^{385}\)This lemma is equivalent to the “column bumping lemma” in Fulton [54, p. 187].
• Create $Q_{m+1}$ from $Q_m$ by adding the top letter $i_{m+1}$ of $(i_{m+1}^{m+1})$ to $Q_m$ in the cell $c_{m+1}$ (which, as we recall, is the extra corner cell of $P_{m+1}$ not present in $P_m$).
• Set $m$ to $m + 1$.

After all of the biletters have been thus processed, the result of the dual RSK algorithm is $(P, Q) = (P, Q)$.

**Example 11.7.** The term in the expansion of the left side of (11.266) corresponding to

$$(x_2 y_1)(x_4 y_3)(x_1 y_2)(x_1 y_4)(x_3 y_4)(x_2 y_5)$$

is the strict biword $(1) = (112445)$, and the dual RSK algorithm applied to this biword proceeds as follows:

| $P_0$ = ∅ | $Q_0$ = ∅ |
| $P_1$ = 2 | $Q_1$ = 1 |
| $P_2$ = 2 4 | $Q_2$ = 1 1 |
| $P_3$ = 1 4 2 | $Q_3$ = 1 1 |
| 1 4 | 1 1 |
| $P_4$ = 1 2 | $Q_4$ = 2 |
| 1 3 | $Q_5$ = 2 4 |
| 2 | 4 |
| $P := P_6$ = 1 2 4 | $Q := Q_6$ = 2 4 5 |
| 1 3 | 1 1 |

It is clear that $P_m$ remains a row-strict tableau of some Ferrers shape throughout the execution of the dual RSK algorithm, and that $Q_m$ remains a filling of the same shape as $P_m$ which is (at least) weakly increasing left-to-right along rows and weakly increasing top-to-bottom in columns. But we can also see that $Q_m$ is strictly increasing top-to-bottom along columns\(^{386}\), so that $Q_m$ is a column-strict tableau. Thus, the result $(P, Q)$ of the dual RSK algorithm is a cotableau-tableau pair.

To see that the dual RSK algorithm is a bijection, we show how to recover $(1) = (112445)$ from $(P, Q)$. This is done by *dually reverse bumping* from $(P_{m+1}, Q_{m+1})$ to recover both the biletter $(i_{m+1}^{m+1})$ and the tableaux $(P_m, Q_m)$, as follows. Firstly, $i_{m+1}$ is the maximum entry of $Q_{m+1}$, and $Q_m$ is obtained by removing the bottommost occurrence of this letter $i_{m+1}$ from $Q_{m+1}$.\(^{387}\) To produce $P_m$ and $j_{m+1}$, find the position of the bottommost occurrence of $i_{m+1}$ in $Q_{m+1}$, and start *dually reverse bumping* in $P_{m+1}$ from the entry in this same position, where reverse bumping an entry means inserting it into one row higher by having it bump out the rightmost entry which is smaller or equal to it.\(^{388}\) The entry bumped out of the first row is $j_{m+1}$, and the resulting tableau is $P_m$.

\(^{386}\)Indeed, this follows from the observation that when one has a string of equal letters $i_m = i_{m+1} = \cdots = i_m^r$ on top of the strict biword, then the bottom letters bumped in are $j_m < j_{m+1} < \cdots < j_m^r$, and therefore (as a consequence of the second-to-last claim of part (a) of the dual row bumping lemma) the new cells form a *vertical strip*, that is, no two of these cells lie in the same row. Actually, more can be said: Each of these new cells (except for the first one) is in a row further down than the previous one. We will use this stronger fact further below.

\(^{387}\)It necessarily has to be the bottommost occurrence, since (according to the previous footnote) the cell into which $i_{m+1}$ was filled at the step from $Q_m$ to $Q_{m+1}$ lies further down than any existing cell of $Q_m$ containing the letter $i_{m+1}$.

\(^{388}\)Let us give a few more details on this “dually reverse bumping” procedure. Dually reverse bumping (also known as *dual RS-deletion* or *reverse dual RS-insertion*) is an operation which takes a row-strict tableau $P'$ and a corner cell $c$ of $P'$, and constructs a row-strict tableau $P$ and a letter $j$ such that dual RS-insertion for $P$ and $j$ yields $P'$ and $c$. It starts by setting $P = P'$, and removing the entry in the cell $c$ from $P$. This removed entry is then denoted by $k$, and is inserted into the row of $P$ above $c$, bumping out the rightmost entry which is smaller or equal to $k$. The letter which is bumped out – say, $\ell$ – in turn,
Finally, to see that the dual RSK map is surjective, one needs to show that the dually reverse bumping procedure can be applied to any cotableau-tableau pair \((P, Q)\), and will result in a strict biword \(\{j\}^\ell\). We leave this verification to the reader.\(^{389}\)

So the dual RSK map is a bijection having the required properties. As we have said, this proves (11.267). This solves Exercise 2.70(a) again.

(b) Let us consider the map \(\omega \left[ [y] \right] : \Lambda \left[ [y] \right] \to \Lambda \left[ [y] \right] \) defined as in the solution of Exercise 2.70(a). We have

\[
(11.275) \quad \omega \left( p_n \right) = (-1)^{n-1} p_n \quad \text{for every positive integer } n.
\]

Since \(\omega\) is a \(k\)-algebra homomorphism, this readily entails

\[
(11.276) \quad \omega \left( p_\lambda \right) = (-1)^{|\lambda| - \ell(\lambda)} p_\lambda \quad \text{for every partition } \lambda.
\]

Now, the equality (2.24) yields

\[
(11.277) \quad \sum_{\lambda \in \text{Par}} s_\lambda \left( x \right) s_\lambda \left( y \right) = \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \text{Par}} z_\lambda^{-1} p_\lambda \left( x \right) p_\lambda \left( y \right) \quad \text{(by } (2.28)\text{)}.
\]

This is an equality in \(\Lambda \left[ [y] \right]\) (because \(s_\lambda \left( x \right)\) and \(p_\lambda \left( x \right)\) belong to \(\Lambda\) for every \(\lambda \in \text{Par}\)). Hence, we can apply the map \(\omega \left[ [y] \right] : \Lambda \left[ [y] \right] \to \Lambda \left[ [y] \right] \) to both sides of this equality. As a result, we obtain

\[
\sum_{\lambda \in \text{Par}} \omega \left( s_\lambda \left( x \right) \right) s_\lambda \left( y \right) = \sum_{\lambda \in \text{Par}} z_\lambda^{-1} \omega \left( p_\lambda \left( x \right) \right) p_\lambda \left( y \right)
\]

(because \(\omega \left[ [y] \right]\), being a \(k\left[ [y] \right]\)-algebra homomorphism, leaves the \(s_\lambda \left( y \right)\) and \(p_\lambda \left( y \right)\) terms unchanged, while the \(s_\lambda \left( x \right)\) and \(p_\lambda \left( x \right)\) terms are elements of \(\Lambda\) and thus are transformed as by \(\omega\)). Due to (11.276) and (11.263), this rewrites as

\[
\sum_{\lambda \in \text{Par}} s_\lambda \left( x \right) s_\lambda \left( y \right) = \sum_{\lambda \in \text{Par}} z_\lambda^{-1} \left( -1 \right)^{|\lambda| - \ell(\lambda)} p_\lambda \left( x \right) p_\lambda \left( y \right).
\]

Since \(\sum_{\lambda \in \text{Par}} s_{\lambda^r} \left( x \right) s_\lambda \left( y \right) = \sum_{\lambda \in \text{Par}} s_\lambda \left( x \right) s_{\lambda^r} \left( y \right)\) (this can be proven just as in (11.264)), this rewrites as

\[
\sum_{\lambda \in \text{Par}} s_\lambda \left( x \right) s_{\lambda^r} \left( y \right) = \sum_{\lambda \in \text{Par}} z_\lambda^{-1} \left( -1 \right)^{|\lambda| - \ell(\lambda)} p_\lambda \left( x \right) p_\lambda \left( y \right) = \sum_{\lambda \in \text{Par}} \left( -1 \right)^{|\lambda| - \ell(\lambda)} z_\lambda^{-1} p_\lambda \left( x \right) p_\lambda \left( y \right).
\]

is inserted into the row above it, bumping out the rightmost entry which is smaller or equal to \(\ell\). This procedure continues in the same way until an entry is bumped out of the first row (which will eventually happen). The dually reverse bumping operation returns the resulting tableau \(P\) and the entry which is bumped out of the first row.

It is straightforward to check that the dually reverse bumping operation is well-defined (i.e., \(P\) does stay a row-strict tableau throughout the procedure) and is the inverse of the dual RS-insertion operation. (In fact, these two operations undo each other step by step.)\(^{389}\)

Since the chain of inequalities \(i_1 \leq i_2 \leq \cdots \leq i_\ell\) is clear from the choice of entry to dually reverse-bump, it only remains to show that for every string \(i_m \geq i_{m+1} = \cdots = i_{m+r}\) of equal top letters, the corresponding bottom letters strictly increase (that is, \(j_{m} < j_{m+1} < \cdots < j_{m+r}\)). One way to see it is to apply part (b) of the dual row bumping lemma.\(^{390}\)

Proof of (11.275): Let \(n\) be a positive integer. Proposition 2.28(iii) yields \(S \left( p_n \right) = -p_n\). But (2.20) (applied to \(f = p_n\)) yields \(S \left( p_n \right) = \left( -1 \right)^n \omega \left( p_n \right)\) (since \(p_n \in \Lambda_n\)). Comparing these two equalities yields \(\left( -1 \right)^n \omega \left( p_n \right) = -p_n\), and this quickly rewrites as (11.275).\(^{391}\)

In fact, denoting by \(\ell(\lambda)\) the length \(\ell\) of the partition \(\lambda\), we recall that \(p_\lambda\) is defined as \(p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell}\), and thus

\[
\omega \left( p_\lambda \right) = \omega \left( p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell} \right) = \omega \left( p_{\lambda_1} \right) \omega \left( p_{\lambda_2} \right) \cdots \omega \left( p_{\lambda_\ell} \right) \quad \text{(since } \omega \text{ is a } k\text{-algebra homomorphism)}
\]

\[
= \left( -1 \right)^{\ell(\lambda)-1} p_{\lambda_1} \cdot \left( -1 \right)^{\ell(\lambda)-2} p_{\lambda_2} \cdots \left( -1 \right)^{\ell(\lambda)-1} p_{\lambda_\ell} \quad \text{(by } (11.275)\text{)}
\]

\[
= \left( -1 \right)^{|\lambda_1| + |\lambda_2| + \cdots + |\lambda_\ell|} p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_\ell} = \left( -1 \right)^{|\lambda| - \ell(\lambda)} p_\lambda.
\]
Hence,
\[
\sum_{\lambda \in \text{Par}} (-1)^{|\lambda| - \ell(\lambda)} z^{|\lambda| - 1} p_{\lambda}(x) p_{\lambda}(y) = \sum_{\lambda \in \text{Par}} s_{\lambda}(x) s_{\lambda}(y) = \prod_{i,j=1}^{\infty} (1 + x_i y_j)
\]
(by (11.265)). This solves Exercise 2.70(b).

Let \( n \in \mathbb{N} \). Let \( \mu \) be a partition having at most \( n \) parts. Exercise 2.39(a) yields
\[
\sum_{\lambda \in \text{Par}} s_{\lambda}(x) s_{\lambda/\mu}(y) = s_{\mu}(x) \cdot \prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1}.
\]
in the ring \( k[[x,y]] = k[[x_1, x_2, \ldots, y_1, y_2, y_3, \ldots]] \). Switching the roles of the variables \( x \) and \( y \) in this equality, we obtain
\[
\sum_{\lambda \in \text{Par}} s_{\lambda}(y) s_{\lambda/\mu}(x) = s_{\mu}(y) \cdot \prod_{i,j=1}^{\infty} (1 - y_i x_j)^{-1}.
\]
We can now substitute \( (y_1, y_2, \ldots, y_n, 0, 0, 0, \ldots) \) for \( (y_1, y_2, y_3, \ldots) \) on both sides of this, and obtain the equality
\[
\sum_{\lambda \in \text{Par}} s_{\lambda}(y_1, y_2, \ldots, y_n) s_{\lambda/\mu}(x) = s_{\mu}(y_1, y_2, \ldots, y_n) \cdot \prod_{i=1}^{n} \prod_{j=1}^{\infty} (1 - y_i x_j)^{-1}
\]
in the subring \( (k[[x]])[[y_1, y_2, \ldots, y_n]] \) of \( k[[x,y]] \) (notice that the \( \prod_{i,j=1}^{\infty} (1 - y_i x_j)^{-1} \) on the right hand side became \( \prod_{i,j=1}^{n} \prod_{j=1}^{\infty} (1 - y_i x_j)^{-1} \) because all factors \( 1 - y_i x_j \) with \( i > n \) got sent to \( 1 - 0 x_j = 1 \)). In the sum on the left hand side of this equality, all addends corresponding to partitions \( \lambda \) having more than \( n \) parts are 0 (because Exercise 2.26(b) yields that all such \( \lambda \) satisfy \( s_{\lambda}(x_1, x_2, \ldots, x_n) = 0 \), hence \( s_{\lambda}(y_1, y_2, \ldots, y_n) = 0 \)). Thus, we can remove all these addends, and the equality thus becomes
\[
(11.278) \sum_{\lambda \in \text{Par}; \lambda \text{ has at most } n \text{ parts}} s_{\lambda}(y_1, y_2, \ldots, y_n) s_{\lambda/\mu}(x) = s_{\mu}(y_1, y_2, \ldots, y_n) \cdot \prod_{i=1}^{n} \prod_{j=1}^{\infty} (1 - y_i x_j)^{-1}.
\]
Let \( \rho \) be the \( n \)-tuple \( (n-1, n-2, \ldots, 2, 1, 0) \in \mathbb{N}^n \). We can regard \( \rho \) as a weak composition by padding it with zeroes at the end (i.e., identifying \( \rho \) with the weak composition \( (n-1, n-2, \ldots, 2, 1, 0, 0, 0, 0, \ldots) \)).

For every \( n \)-tuple \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}^n \), we define the alternant \( a_{\alpha} \in k[x_1, x_2, \ldots, x_n] \) as in Definition 2.50. (But other than this, we are not adapting the notations of Section 2.6.)

Now, Corollary 2.54 states that \( s_{\lambda}(x_1, x_2, \ldots, x_n) = \frac{a_{\lambda + \rho}}{a_{\rho}} \) in \( k[x_1, x_2, \ldots, x_n] \) whenever \( \lambda \) is a partition having at most \( n \) parts. Applied to the variables \( y_1, y_2, \ldots, y_n \) instead of \( x_1, x_2, \ldots, x_n \), this yields that \( s_{\lambda}(y_1, y_2, \ldots, y_n) = \frac{a_{\lambda + \rho}(y_1, y_2, \ldots, y_n)}{a_{\rho}(y_1, y_2, \ldots, y_n)} \) whenever \( \lambda \) is a partition having at most \( n \) parts. This equality, applied to \( \lambda = \mu \), yields \( s_{\mu}(y_1, y_2, \ldots, y_n) = \frac{a_{\mu + \rho}(y_1, y_2, \ldots, y_n)}{a_{\rho}(y_1, y_2, \ldots, y_n)} \). Substituting the last two equalities into (11.278), we obtain
\[
\sum_{\lambda \in \text{Par}; \lambda \text{ has at most } n \text{ parts}} \frac{a_{\lambda + \rho}(y_1, y_2, \ldots, y_n)}{a_{\rho}(y_1, y_2, \ldots, y_n)} s_{\lambda/\mu}(x) = \frac{a_{\mu + \rho}(y_1, y_2, \ldots, y_n)}{a_{\rho}(y_1, y_2, \ldots, y_n)} \cdot \prod_{i=1}^{n} \prod_{j=1}^{\infty} (1 - y_i x_j)^{-1}.
\]
Multiplied with \( a_{\rho}(y_1, y_2, \ldots, y_n) \), this becomes
\[
\sum_{\lambda \in \text{Par}; \lambda \text{ has at most } n \text{ parts}} a_{\lambda + \rho}(y_1, y_2, \ldots, y_n) s_{\lambda/\mu}(x) = a_{\mu + \rho}(y_1, y_2, \ldots, y_n) \cdot \prod_{i=1}^{n} \prod_{j=1}^{\infty} (1 - y_i x_j)^{-1}.
\]
Renaming $\lambda$ as $\nu$ on the left hand side of this equality, we obtain

$$(11.279) \quad \sum_{\nu \in \text{Par}; \nu \text{ has at most } n \text{ parts}} a_{\nu + \rho} (y_1, y_2, ..., y_n) s_{\nu/\mu} (x) = a_{\mu + \rho} (y_1, y_2, ..., y_n) \cdot \prod_{i=1}^{n} \prod_{j=1}^{\infty} (1 - y_i x_j)^{-1}.$$ 

Now, let $\lambda$ be a partition having at most $n$ parts. We want to find the coefficient of $y_1^{\lambda_1} y_2^{\lambda_2} ... y_n^{\lambda_n}$ on the left and the right hand sides of (11.279). Here, we regard (11.279) as an equality in the ring $\mathcal{H}$. We first notice that

$$(11.280) \quad \prod_{i=1}^{n} \prod_{j=1}^{\infty} (1 - y_i x_j)^{-1} = \sum_{(q_1, q_2, ..., q_n) \in \mathbb{N}^n} \left( \prod_{j=1}^{n} h_{q_j} (x) \right) \cdot \left( \prod_{j=1}^{n} y_j^{q_j} \right).$$

Let us recall a basic fact from linear algebra, namely the explicit formula for the determinant of a matrix as a sum over permutations: Any matrix $(\alpha_{i,j})_{i,j=1,2,\ldots,\ell}$ over a commutative ring satisfies

$$(11.282) \quad \det \left( (\alpha_{i,j})_{i,j=1,2,\ldots,\ell} \right) = \sum_{\sigma \in \mathfrak{S}_\ell} (-1)^{\sigma} \prod_{i=1}^{\ell} \alpha_{i,\sigma(i)}.$$ 

Applying this to $\ell = n$ and $\alpha_{i,j} = y_i^{(\mu + \rho)_j}$, we obtain

$$(11.283) \quad \det \left( (y_i^{(\mu + \rho)_j})_{i,j=1,2,\ldots,n} \right) = \sum_{\sigma \in \mathfrak{S}_n} (-1)^{\sigma} \prod_{i=1}^{n} y_i^{(\mu + \rho)_{\sigma(i)}} = \sum_{\sigma \in \mathfrak{S}_n} (-1)^{\sigma} \prod_{j=1}^{n} y_j^{(\mu + \rho)_{\sigma(j)}}$$

$392$Proof: The equality (2.14) (with the indices $i$ and $n$ renamed as $j$ and $q$) yields

$$(11.281) \quad \prod_{j=1}^{n} (1 - x_j t)^{-1} = 1 + h_1 (x) t + h_2 (x) t^2 + ... = \sum_{q \geq 0} h_q (x) t^q \quad \text{in} \quad (k[[x]])[[t]].$$

For every $i \in \{1, 2, ..., n\}$, we have

$$\prod_{j=1}^{\infty} (1 - y_i x_j)^{-1} = \sum_{q \geq 0} h_q (x) y_i^q$$

(this follows by substituting $y_i$ for $t$ on both sides of (11.281)). Thus,

$$\prod_{i=1}^{n} \prod_{j=1}^{\infty} (1 - y_i x_j)^{-1} = \prod_{i=1}^{n} \sum_{q \geq 0} h_q (x) y_i^q = \prod_{j=1}^{n} \sum_{q \geq 0} h_q (x) y_j^q$$

$$= \sum_{(q_1, q_2, ..., q_n) \in \mathbb{N}^n} \prod_{j=1}^{n} \left( h_{q_j} (x) \cdot y_j^{q_j} \right) \quad \text{(by the product rule)}$$

$$= \sum_{(q_1, q_2, ..., q_n) \in \mathbb{N}^n} \left( \prod_{j=1}^{n} h_{q_j} (x) \right) \cdot \left( \prod_{j=1}^{n} y_j^{q_j} \right),$$ 

qed.
But for every \( i \) (here, we renamed the index \( i \) as \( j \) in the product). Now, the right hand side of (11.279) becomes

\[
a_{\mu+\rho}(y_1, y_2, \ldots, y_n) = \det \left( \frac{y_j^{n-i}}{y_i^{n-j}} \right)_{i,j=1,2,\ldots,n} = \prod_{j=1}^{\infty} (1 - y_j x_j)^{-1} = \sum_{(q_1, q_2, \ldots, q_n) \in \mathbb{N}^n} \left( \prod_{j=1}^{n} h_{q_j}(x) \right) \left( \prod_{j=1}^{n} y_j^{q_j} \right)
\]

(by the definition of the alternant \( a_{\mu+\rho} \))

\[
= \sum_{\sigma \in \mathfrak{S}_n} (-1)^\sigma \prod_{j=1}^{n} y_j^{(\mu+\rho)_{\sigma(j)}} \prod_{(q_1, q_2, \ldots, q_n) \in \mathbb{N}^n} \left( \prod_{j=1}^{n} h_{q_j}(x) \right) \left( \prod_{j=1}^{n} y_j^{q_j} \right) \left( \prod_{j=1}^{n} y_j^{(\mu+\rho)_{\sigma(j)}} \right)
\]

Hence, the coefficient of \( y_1^{\lambda_1+n-1} y_2^{\lambda_2+n-2} \ldots y_n^{\lambda_n+n-n} \) on the right hand side of (11.279) equals

\[
(11.284) \quad \sum_{\sigma \in \mathfrak{S}_n} (-1)^\sigma \prod_{(q_1, q_2, \ldots, q_n) \in \mathbb{N}^n: q_j+(\mu+\rho)_{\sigma(j)}=\lambda_j+n-j \, \text{for every} \, j \in \{1,2,\ldots,n\}} h_{q_j}(x).
\]

But for every \( \sigma \in \mathfrak{S}_n \), it is easy to see that

\[
(11.285) \quad \prod_{j=1}^{n} h_{q_j}(x) = \prod_{j=1}^{n} h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}}(x).
\]
Hence, the coefficient of \( y_1^{\lambda_1+n-1} y_2^{\lambda_2+n-2} \cdots y_n^{\lambda_n+n-n} \) on the right hand side of (11.279) equals

\[
\sum_{\sigma \in \mathfrak{S}_n} (-1)^{\sigma} \prod_{j=1}^n h_{\lambda_j-j+\sigma(j)}(x) \quad \text{(by (11.284))}
\]

\[
= \sum_{\sigma \in \mathfrak{S}_n} (-1)^{\sigma} \prod_{j=1}^n h_{\lambda_j-n-j-(\mu+\rho)_{\sigma(j)}}(x)
\]

\[
(11.288) = \sum_{\sigma \in \mathfrak{S}_n} (-1)^{\sigma} \prod_{i=1}^n h_{\lambda_j-i+\sigma(i)}(x) = \det \left( \left( h_{\lambda_i-i+j}(x) \right)_{i,j=1,2,\ldots,n} \right)
\]

(because applying (11.282) to \( \ell = n \) and \( \alpha_{i,j} = h_{\lambda_j-i-j}(x) \) yields

\[
(11.287) \quad q_j = \lambda_j + n - j - (\mu + \rho)_{\sigma(j)} \quad \text{for every} \quad j \in \{1,2,\ldots,n\}.
\]

If this n-tuple \((q_1,q_2,\ldots,q_n)\) belongs to \(\mathbb{N}^n\), then the sum (11.286) does have this summand. Thus, if the n-tuple \((q_1,q_2,\ldots,q_n)\) belongs to \(\mathbb{N}^n\), we have

\[
\sum_{(q_1,q_2,\ldots,q_n) \in \mathbb{N}^n} \prod_{j=1}^n h_{q_j}(x) = \prod_{j=1}^n h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}}(x).
\]

Thus, if the n-tuple \((q_1,q_2,\ldots,q_n)\) defined by (11.287) belongs to \(\mathbb{N}^n\), the equality (11.285) is proven. It remains to consider the case when the n-tuple \((q_1,q_2,\ldots,q_n)\) defined by (11.287) does not belong to \(\mathbb{N}^n\). In this case, the sum (11.286) is empty (because the only addend it can have corresponds to the n-tuple \((q_1,q_2,\ldots,q_n)\) defined by (11.287), but this n-tuple does not belong to \(\mathbb{N}^n\) and therefore does not appear in the sum), hence equals 0. But since the n-tuple \((q_1,q_2,\ldots,q_n)\) defined by (11.287) does not belong to \(\mathbb{N}^n\), there must be a \(j \in \{1,2,\ldots,n\}\) satisfying \(\lambda_j + n - j - (\mu + \rho)_{\sigma(j)} \notin \mathbb{N}\). For this \(j\), we have

\[
h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}} = 0.
\]

Hence, the product \(\prod_{j=1}^n h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}}(x)\) has a factor equal to 0; consequently, the product \(\prod_{j=1}^n h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}}(x)\) must be 0. Thus, both the sum (11.286) and the product \(\prod_{j=1}^n h_{\lambda_j+n-j-(\mu+\rho)_{\sigma(j)}}(x)\) are 0, and thus (11.285) is proven.
Hence, the left hand side of (11.279) is

\[
\sum_{\nu \in \text{Par}; \ \nu \text{ has at most } n \text{ parts}} \left( \sum_{\sigma \in \mathfrak{S}_n} (-1)^\sigma \prod_{i=1}^n y_i^{(\nu+\rho)_{\sigma(i)}} \right) s_{\nu/\mu}(x)
\]

\[
= \sum_{\nu \in \text{Par}; \ \nu \text{ has at most } n \text{ parts}} \left( \sum_{\sigma \in \mathfrak{S}_n} (-1)^\sigma \prod_{i=1}^n y_i^{(\nu+\rho)_{\sigma(i)}} \right) s_{\nu/\mu}(x)
\]

\[
= \sum_{\sigma \in \mathfrak{S}_n} (-1)^\sigma \sum_{\nu \in \text{Par}; \ \nu \text{ has at most } n \text{ parts}} \left( \prod_{i=1}^n y_i^{(\nu+\rho)_{\sigma(i)}} \right) s_{\nu/\mu}(x).
\]

Thus, the coefficient of \(y_1^{\lambda_1+n-1}y_2^{\lambda_2+n-2}...y_n^{\lambda_n+n-n}\) on the left hand side of (11.279) equals

(11.289)

\[
\sum_{\sigma \in \mathfrak{S}_n} (-1)^\sigma \sum_{\nu \in \text{Par}; \ \nu \text{ has at most } n \text{ parts}} (\nu+\rho)_{\sigma(j)} = \lambda_j + n - j \text{ for every } j \in \{1,2,...,n\}
\]

Now let us simplify this. First, we claim that every permutation \(\sigma \in \mathfrak{S}_n\) satisfying \(\sigma \neq \text{id}\) satisfies

(11.290)

\[
\sum_{\nu \in \text{Par}; \ \nu \text{ has at most } n \text{ parts}} s_{\nu/\mu}(x) = 0.
\]

Thus, in the outer sum in (11.289), all addends which correspond to permutations \(\sigma \in \mathfrak{S}_n\) satisfying \(\sigma \neq \text{id}\) are 0. We can therefore remove all these addends, leaving only the addend corresponding to \(\sigma = \text{id}\). Thus, the sum simplifies as follows:

\[
\sum_{\sigma \in \mathfrak{S}_n} (-1)^\sigma \sum_{\nu \in \text{Par}; \ \nu \text{ has at most } n \text{ parts}} (\nu+\rho)_{\sigma(j)} = \lambda_j + n - j \text{ for every } j \in \{1,2,...,n\}
\]

(11.290)

\[
\sum_{\nu \in \text{Par}; \ \nu \text{ has at most } n \text{ parts}} s_{\nu/\mu}(x) = 0.
\]

394 Thus, in the outer sum in (11.289), all addends which correspond to permutations \(\sigma \in \mathfrak{S}_n\) satisfying \(\sigma \neq \text{id}\) are 0. We can therefore remove all these addends, leaving only the addend corresponding to \(\sigma = \text{id}\). Thus, the sum simplifies as follows:

(11.290)

\[
\sum_{\nu \in \text{Par}; \ \nu \text{ has at most } n \text{ parts}} s_{\nu/\mu}(x) = 0.
\]

(because there is only one \(\nu \in \text{Par}\) such that \(\nu \text{ has at most } n \text{ parts and satisfies } \nu_j + n - j = \lambda_j + n - j \text{ for every } j \in \{1,2,...,n\}\); namely, this \(\nu \text{ is } \lambda\)). Hence, the coefficient of \(y_1^{\lambda_1+n-1}y_2^{\lambda_2+n-2}...y_n^{\lambda_n+n-n}\) on the left

394 Proof of (11.290): Let \(\sigma \in \mathfrak{S}_n\) be a permutation satisfying \(\sigma \neq \text{id}\). We need to prove (11.290). Of course, it is enough to show that the sum on the left hand side of (11.290) is empty, i.e., that there exists no \(\nu \in \text{Par}\) such that \(\nu \text{ has at most } n \text{ parts and satisfies } (\nu + \rho)_{\sigma(j)} = \lambda_j + n - j \text{ for every } j \in \{1,2,...,n\}\). Assume the contrary. Then, there exists a \(\nu \in \text{Par}\) such that \(\nu \text{ has at most } n \text{ parts and satisfies } (\nu + \rho)_{\sigma(j)} = \lambda_j + n - j \text{ for every } j \in \{1,2,...,n\}\). Consider this \(\nu\). Since \(\nu \text{ is a partition, } \nu + \rho \text{ is a strict partition, i.e., we have } (\nu + \rho)_{\sigma(1)} > (\nu + \rho)_{\sigma(2)} > ... > (\nu + \rho)_{\sigma(n)}\). In other words, the entries of the n-tuple \(\nu + \rho\) are in strictly decreasing order. Since \(\sigma \neq \text{id}\), the permutation \(\sigma\) must mess up the order of the entries of \(\nu + \rho\); thus, we cannot have \((\nu + \rho)_{\sigma(1)} > (\nu + \rho)_{\sigma(2)} > ... > (\nu + \rho)_{\sigma(n)}\). Since \((\nu + \rho)_{\sigma(j)} = \lambda_j + n - j \text{ for every } j \in \{1,2,...,n\}\), this rewrites as follows: We cannot have \(\lambda_1 + n - 1 > \lambda_2 + n - 2 > ... > \lambda_n + n - n\). But since \(\lambda\) is a partition, we have \(\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n\), thus \(\lambda_1 + n - 1 > \lambda_2 + n - 2 > ... > \lambda_n + n - n\). This contradicts the fact that we cannot have \(\lambda_1 + n - 1 > \lambda_2 + n - 2 > ... > \lambda_n + n - n\). This contradiction finishes the proof.
hand side of (11.279) equals

\[
\sum_{\sigma \in S_n} (-1)^\sigma \sum_{\nu \in \text{Par}; \nu \text{ has at most } n \text{ parts}; (\nu+\mu)_{\iota(j)} = \lambda_j + n - j \text{ for every } j \in \{1,2,\ldots,n\}} s_{\nu/\mu} (x) \quad \text{(by (11.289))}
\]

(11.291)

\[= s_{\lambda/\mu} (x). \]

But the coefficients of \(y_1^{\lambda_1+n-1}y_2^{\lambda_2+n-2}\ldots y_n^{\lambda_n+n-n} \) on the left hand side of (11.279) and on the right hand side of (11.279) must be equal. Since the former coefficient is \(s_{\lambda/\mu} (x) \) (by (11.291)), and the latter coefficient is \(\det \left( (h_{\lambda_j-\mu_j-i+j} (x))_{i,j=1,2,\ldots,n} \right) \) (by (11.288)), this shows that

\[s_{\lambda/\mu} (x) = \det \left( (h_{\lambda_j-\mu_j-i+j} (x))_{i,j=1,2,\ldots,n} \right). \]

In other words, \(s_{\lambda/\mu} = \det \left( (h_{\lambda_j-\mu_j-i+j})_{i,j=1,2,\ldots,n} \right). \)

Now, forget that we fixed \(n, \lambda\) and \(\mu\). We thus have proven that if \(n \in \mathbb{N}\), and if \(\lambda\) and \(\mu\) are two partitions having at most \(n\) parts (each), then \(s_{\lambda/\mu} = \det \left( (h_{\lambda_j-\mu_j-i+j})_{i,j=1,2,\ldots,n} \right). \) Renaming \(n\) as \(\ell\) in this claim, we obtain: If \(\ell \in \mathbb{N}\), and if \(\lambda\) and \(\mu\) are two partitions having at most \(\ell\) parts (each), then

(11.292)

\[s_{\lambda/\mu} = \det \left( (h_{\lambda_j-\mu_j-i+j})_{i,j=1,2,\ldots,\ell} \right). \]

This proves (2.22).

Now it remains to prove (2.23). Let \(\ell \in \mathbb{N}\), and let \(\lambda\) and \(\mu\) be two partitions having at most \(\ell\) parts (each).

Let us first notice that every \(m \in \mathbb{Z}\) satisfies

(11.293)

\[\omega (h_m) = e_m. \]

(Indeed, for \(m > 0\) this follows from our knowledge of \(\omega\), while for \(m = 0\) and for \(m < 0\) it is obvious.) But using (2.21), it is easy to see that

(11.294)

\[\omega \left( s_{\lambda/\mu} \right) = s_{\lambda'/\mu'}. \]

Thus,

\[s_{\lambda'/\mu'} = \omega \left( s_{\lambda/\mu} \right) = \omega \left( \det \left( (h_{\lambda_j-\mu_j-i+j})_{i,j=1,2,\ldots,\ell} \right) \right) \quad \text{(by (11.292))} \]

\[= \det \left( \omega \left( h_{\lambda_j-\mu_j-i+j} \right) \right) \quad \text{(by (11.293))} \]

\[= \left( e_{\lambda_j-\mu_j-i+j} \right)_{i,j=1,2,\ldots,\ell} \quad \text{since } \omega \text{ is a } k\text{-algebra homomorphism, and thus} \]

\[= \det \left( e_{\lambda_j-\mu_j-i+j} \right)_{i,j=1,2,\ldots,\ell}. \]

This proves (2.23). Therefore, the proof of Theorem 2.30 is complete.

\[\text{In fact, (11.294) follows immediately from (2.21) in the case when } \mu \subseteq \lambda; \text{ but otherwise it follows from } s_{\lambda/\mu} = 0 \text{ and } s_{\lambda'/\mu'} = 0. \]
11.51. **Solution to Exercise 2.75. Solution to Exercise 2.75.**

Let \((\Lambda_Q)_n\) denote the \(n\)-th graded component of the \(\mathbb{Q}\)-vector space \(\Lambda_Q\). We are going to work in \(\Lambda_Q\) in this solution, making use of the fact that both \((s_\lambda)_{\lambda \in \text{Par}_n}\) and \((p_\lambda)_{\lambda \in \text{Par}_n}\) are bases of the \(\mathbb{Q}\)-vector space \((\Lambda_Q)_n\) (due to Proposition 2.17). Of course, notions such as comultiplication, the antipode, skewing etc. are defined in \(\Lambda_Q\) just in the same way as they have been defined in \(\Lambda\), and their properties are proven analogously.

Our solutions for both parts of the exercise rely on the fact that the trace of an endomorphism of a finite-dimensional vector space can be computed using any basis of the vector space. Specifically, we will be applying this fact to certain endomorphisms of \((\Lambda_Q)_n\), and as bases we will use \((p_\lambda)_{\lambda \in \text{Par}_n}\) and \((s_\lambda)_{\lambda \in \text{Par}_n}\).

(a) The antipode \(S\) of \(\Lambda_Q\) is a graded map, and thus it restricts to an endomorphism of the \(\mathbb{Q}\)-vector space \((\Lambda_Q)_n\). Denote this endomorphism by \(S_n\). We want to compute trace \((S_n)\). (This is well-defined since \((\Lambda_Q)_n\) is a finite-dimensional \(\mathbb{Q}\)-vector space.)

From (2.21), we see that \(S(s_{\lambda/\mu}) = (-1)^{|\lambda/\mu|} s_{\lambda'/\mu'}\) for any partitions \(\lambda\) and \(\mu\) satisfying \(\mu \subseteq \lambda\). In particular, \(S(s_\lambda) = (-1)^{|\lambda|} s_\lambda\) for any partition \(\lambda\). For any partition \(\lambda \in \text{Par}_n\), we now have \(S_n(s_\lambda) = S(s_\lambda) = (-1)^{|\lambda|} s_\lambda\). Thus, the matrix which represents the endomorphism \(S_n\) of \((\Lambda_Q)_n\) with respect to the basis \((s_\lambda)_{\lambda \in \text{Par}_n}\) of \((\Lambda_Q)_n\) is the matrix whose \((\lambda, \mu)\)-th entry (for any \(\lambda \in \text{Par}_n\) and \(\mu \in \text{Par}_n\)) is \((-1)^n\) when \(\mu = \lambda\), and otherwise 0. But the trace trace \((S_n)\) of \(S_n\) (by its definition) is the sum of the diagonal entries of this matrix; hence, this trace equals

\[
\text{trace}(S_n) = \sum_{\substack{\lambda \in \text{Par}_n; \\ \lambda = \lambda'}} (-1)^n \text{ (number of all } \lambda \in \text{Par}_n \text{ satisfying } \lambda = \lambda') = (-1)^n c(n).
\]

On the other hand, every partition \(\lambda\) satisfies \(S(p_{\lambda}) = (-1)^{\ell(\lambda)} p_{\lambda}\) (this follows from Proposition 2.28(i), after recalling that \(S\) is an algebra map and that \(p_{\lambda_1, \lambda_2, \ldots, \lambda_k} = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_k}\)). Hence, the matrix which represents the endomorphism \(S_n\) of \((\Lambda_Q)_n\) with respect to the basis \((p_\lambda)_{\lambda \in \text{Par}_n}\) of \((\Lambda_Q)_n\) is a diagonal matrix, whose \(\lambda\)-th diagonal entry (for any \(\lambda \in \text{Par}_n\)) is \((-1)^n\). But the trace trace \((S_n)\) of \(S_n\) (by its definition) is the sum of the diagonal entries of this matrix; hence, this trace equals

\[
\text{trace}(S_n) = \sum_{\lambda \in \text{Par}_n} (-1)^{\ell(\lambda)} = \sum_{k=0}^n (-1)^k \sum_{\substack{\lambda \in \text{Par}_n; \\ \ell(\lambda) = k}} \text{ (number of all } \lambda \in \text{Par}_n \text{ satisfying } \ell(\lambda) = k) = \sum_{k=0}^n (-1)^k p(n,k).
\]

Compared with trace \((S_n) = (-1)^n c(n)\), this yields \((-1)^n c(n) = \sum_{k=0}^n (-1)^k p(n,k)\). This solves part (a) of the exercise.

(b) Consider the map \(s_1 s_1^\perp : \Lambda \to \Lambda\) which sends every \(f \in \Lambda\) to \(s_1 s_1^\perp f\). This map \(s_1 s_1^\perp\) is graded (because the map \(s_1^\perp : \Lambda \to \Lambda\) lowers the degree of any homogeneous element by 1, while the map \(s_1\) raises it back by 1) and thus restricts to an endomorphism of the \(\mathbb{Q}\)-vector space \((\Lambda_Q)_n\). Denote this endomorphism by \(P_n\). We want to compute trace \((P_n)\). (This is well-defined since \((\Lambda_Q)_n\) is a finite-dimensional \(\mathbb{Q}\)-vector space.)

In the following, if \(A\) is any statement, then \([A]\) will denote the truth value of \(A\) (that is, 1 if \(A\) holds, and 0 if it doesn’t).
We know that the basis \((s_\lambda)_{\lambda \in \text{Par}_n}\) of \((\Lambda_Q)_n\) is orthonormal with respect to the Hall inner product (restricted to \((\Lambda_Q)_n\)). Hence, for any \(\lambda \in \text{Par}_n\), we have

\[
\begin{align*}
\text{(the } s_\lambda\text{-coordinate of } P_n s_\lambda \text{ with respect to this basis)}
= \left( \frac{P_n s_\lambda}{s_1 s_\lambda} , s_\lambda \right) = \left( s_1 s_\lambda^\perp , s_\lambda \right) = \left( s_1 s_\lambda^\perp , s_1 s_\lambda^\perp \right) \\
= \sum_{\mu \in \text{Par}_{n-1}} \left( \frac{s_1 s_\lambda^\perp , s_\mu}{s_1 s_\lambda^\perp} \right)^2 \\
= \sum_{\mu \in \text{Par}_{n-1}} (s_\lambda, s_1 s_\mu)^2.
\end{align*}
\]

Now, every \(\mu \in \text{Par}_{n-1}\) satisfies

\[
s_1 s_\mu = h_1 s_\mu \quad \text{(2.36)}
= \sum_{\mu^+ : \mu^+ / \mu \text{ is a horizontal } 1\text{-strip}} s_{\mu^+} = \sum_{\mu^+ : |\mu^+ / \mu| = 1} s_{\mu^+} = \sum_{\mu^+ \in \text{Par}_n; \mu^+ \subseteq \mu^+} s_{\mu^+}.
\]

Hence, every \(\lambda \in \text{Par}_n\) and \(\mu \in \text{Par}_{n-1}\) satisfy

\[
(s_\lambda, s_1 s_\mu) = \left( s_\lambda, \sum_{\mu^+ \in \text{Par}_n; \mu^+ \subseteq \mu^+} s_{\mu^+} \right) = \sum_{\mu^+ \in \text{Par}_n; \mu^+ \subseteq \mu^+} \left( s_\lambda, s_{\mu^+} \right) = \sum_{\mu^+ \in \text{Par}_n; \mu^+ \subseteq \mu^+} [\lambda = \mu^+] \\
= [\mu \subseteq \lambda].
\]

Now, consider again the basis \((s_\lambda)_{\lambda \in \text{Par}_n}\) of \((\Lambda_Q)_n\). For every \(\lambda \in \text{Par}_n\), we have

\[
\text{(the } s_\lambda\text{-coordinate of } P_n s_\lambda \text{ with respect to this basis)}
= \sum_{\mu \in \text{Par}_{n-1}} \left( \frac{s_\lambda, s_1 s_\mu}{[\mu \subseteq \lambda]} \right)^2 \quad \text{(by (11.295))}
= \sum_{\mu \in \text{Par}_{n-1}} [\mu \subseteq \lambda]^2 = \sum_{\mu \in \text{Par}_{n-1}} [\mu \subseteq \lambda] \\
= \text{(the number of all } \mu \in \text{Par}_{n-1} \text{ such that } \mu \subseteq \lambda) \\
= \text{(the number of all ways to remove a single cell from } \lambda \text{ to obtain a partition)} \\
= \text{(the number of all corners of the Ferrers diagram of } \lambda) \\
= C(\lambda).
\]

Now, consider the matrix which represents the endomorphism \(P_n\) of \((\Lambda_Q)_n\) with respect to the basis \((s_\lambda)_{\lambda \in \text{Par}_n}\) of \((\Lambda_Q)_n\). For any \(\lambda \in \text{Par}_n\) and \(\mu \in \text{Par}_n\), the \((\lambda, \mu)\)-th entry of this matrix is

\[
\text{(the } s_\lambda\text{-coordinate of } P_n s_\mu \text{ with respect to this basis)}.
\]
The trace \( \text{trace} \left( P_n \right) \) of \( P_n \) (by its definition) is the sum of the diagonal entries of this matrix; hence, this trace equals

\[
\text{trace} \left( P_n \right) = \sum_{\lambda \in \text{Par}_n} \left( \text{the } s_{\lambda}-\text{coordinate of } P_n s_{\lambda} \text{ with respect to this basis} \right) = \sum_{\lambda \in \text{Par}_n} C \left( \lambda \right).
\]

(by \( 11.296 \))

On the other hand, it is easy to see that

\[
s_{\lambda} \left( p_n \right) = [n = 1]
\]

for every positive integer \( n \)

(in fact, recall that \( p_n \) is primitive, so that \( \Delta \left( p_n \right) = p_n \otimes 1 + 1 \otimes p_n \) and thus \( s_{\lambda} \left( p_n \right) = \left( s_1, p_n \right) 1 + (s_1, 1) p_n = 0 \) [\( n = 1 \]). But the map \( s_{\lambda} : \Lambda \to \Lambda \) is a derivation (by Proposition 2.73(iii), since \( s_1 \) is primitive). Hence, every partition \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell) \) with \( \ell = \ell(\lambda) \) satisfies

\[
s_{\lambda} \left( p_{\lambda_1 p_{\lambda_2} \ldots p_{\lambda_\ell}} \right) = \sum_{k=1}^{\ell} p_{\lambda_1 p_{\lambda_2} \ldots p_{\lambda_k} 1 \ldots p_{\lambda_{k+1} p_{\lambda_{k+2}} \ldots p_{\lambda_{\ell}}} = [\lambda_k = 1] \quad (\text{by } 11.297)
\]

\[
= \sum_{k=1}^{\ell} p_{\lambda_1 p_{\lambda_2} \ldots p_{\lambda_{k-1} 1} \ldots p_{\lambda_{k+1} p_{\lambda_{k+2}} \ldots p_{\lambda_{\ell}}} = [\lambda_k = 1] \quad (\text{by } 11.298)
\]

and

\[
P_n p_{\lambda} = s_1 s_{\lambda} = p_{\lambda_1 p_{\lambda_2} \ldots p_{\lambda_\ell} = p_{\lambda_1 p_{\lambda_2} \ldots p_{\lambda_\ell}}} = s_{\lambda} = s_{\lambda} \left( p_{\lambda_1 p_{\lambda_2} \ldots p_{\lambda_\ell}} = \sum_{k=1}^{\ell} p_{\lambda_1 p_{\lambda_2} \ldots p_{\lambda_k} 1 \ldots p_{\lambda_{k+1} p_{\lambda_{k+2}} \ldots p_{\lambda_{\ell}}} = [\lambda_k = 1] \right)
\]

\[
+ \sum_{k=1}^{\ell} p_{\lambda_1 p_{\lambda_2} \ldots p_{\lambda_{k-1} 1} \ldots p_{\lambda_{k+1} p_{\lambda_{k+2}} \ldots p_{\lambda_{\ell}}} = [\lambda_k = 1] \quad (\because \text{if } \lambda_k \neq 1, \text{then both sides of this are 0, and otherwise they are clearly equal})
\]

\[
= \sum_{k=1}^{\ell} p_{\lambda_1 p_{\lambda_2} \ldots p_{\lambda_{k-1} 1} \ldots p_{\lambda_{k+1} p_{\lambda_{k+2}} \ldots p_{\lambda_{\ell}}} = [\lambda_k = 1] \quad (\because \text{the number of } k \in \{1, 2, \ldots, \ell \} \text{ such that } \lambda_k = 1 = (\text{the number of parts of } \lambda \text{ equal to 1})}
\]

\[
= \mu_1 \left( \lambda \right) p_{\lambda}.
\]

Hence, the matrix which represents the endomorphism \( P_n \) of \( \left( \Lambda_\mathbb{Q} \right)_n \) with respect to the basis \( \left( p_{\lambda} \right)_{\lambda \in \text{Par}_n} \) of \( \left( \Lambda_\mathbb{Q} \right)_n \) is a diagonal matrix, whose \( \lambda \)-th diagonal entry (for any \( \lambda \in \text{Par}_n \)) is \( \mu_1 \left( \lambda \right) \). But the trace \( \text{trace} \left( P_n \right) \) of \( P_n \) (by its definition) is the sum of the diagonal entries of this matrix; hence, this trace equals

\[
\text{trace} \left( P_n \right) = \sum_{\lambda \in \text{Par}_n} \mu_1 \left( \lambda \right).
\]

Compared with \( \text{trace} \left( P_n \right) = \sum_{\lambda \in \text{Par}_n} C \left( \lambda \right) \), this yields \( \sum_{\lambda \in \text{Par}_n} C \left( \lambda \right) = \sum_{\lambda \in \text{Par}_n} \mu_1 \left( \lambda \right) \). This solves part (b) of the exercise.
11.52. Solution to Exercise 2.76. Solution to Exercise 2.76. We assume WLOG that $k = Z$, since it is enough to prove what we want over $Z$.

(a) Use the primitivity of $p_n$, and Proposition 1.37 to obtain $S(p_n) = -p_n$ for every $n \geq 1$.

But the antipode $S$ of $\Lambda$ is an algebra anti-endomorphism (by Proposition 1.32), therefore an algebra endomorphism (since $\Lambda$ is commutative). Hence, from the fact that $S(p_n) = -p_n$ for every $n \geq 1$, we can deduce by multiplicativity that $S(p_\lambda) = (-1)^{\ell(\lambda)} p_\lambda$ for every partition $\lambda$. Now recall (2.20), and the claim of part (a) follows.

(b) For this part of the exercise, we shall work in $\Lambda_R$ and prove that the endomorphism $\omega$ of $\Lambda_R$ is an isometry. This will clearly yield the analogous statement over $Z$.

Recall that $\left\{ \frac{p_\lambda}{\sqrt{z_\lambda}} \right\}$ is an orthonormal basis of $\Lambda_R$ (by Corollary 2.45(c)). Hence, in order to prove that the endomorphism $\omega$ of $\Lambda_R$ is an isometry, it is enough to show that

$$\left( \omega \left( \frac{p_\lambda}{\sqrt{z_\lambda}} \right), \omega \left( \frac{p_\mu}{\sqrt{z_\mu}} \right) \right) = \left( \frac{p_\lambda}{\sqrt{z_\lambda}}, \frac{p_\mu}{\sqrt{z_\mu}} \right)$$

for any two partitions $\lambda$ and $\mu$. But this follows from part (a) and the fact that the basis $\left\{ \frac{p_\lambda}{\sqrt{z_\lambda}} \right\}$ is orthonormal. Hence, we have shown that the endomorphism $\omega$ of $\Lambda_R$ is an isometry. The same holds therefore for the endomorphism $\omega$ of $\Lambda$, and thus part (b) of the exercise is solved.

(c) To show that $\omega$ is a coalgebra homomorphism, it suffices to check that $(\omega \otimes \omega) \circ \Delta = \Delta \circ \omega$ (the proof of $\epsilon \circ \omega = \epsilon$ is clear). Both sides of this being algebra homomorphisms, this only needs to be checked on the algebra generators $h_n$. On these generators this is very easy to check: Comparing

$$(\omega \otimes \omega) \circ \Delta \left( h_n \right) = (\omega \otimes \omega) \left( \Delta \left( h_n \right) \right) = (\omega \otimes \omega) \left( \sum_{k=0}^n h_k \otimes h_{n-k} \right) = \sum_{k=0}^n \omega \left( h_k \right) \otimes \omega \left( h_{n-k} \right) = \sum_{k=0}^n e_k \otimes e_{n-k}$$

and

$$\Delta \circ \omega \left( h_n \right) = \Delta \left( e_n \right) = \sum_{k=0}^n e_k \otimes e_{n-k}$$

shows $((\omega \otimes \omega) \circ \Delta) \left( h_n \right) = (\Delta \circ \omega) \left( h_n \right)$ for all $n \geq 1$. So $\omega$ is an algebra and coalgebra morphism, thus a Hopf morphism by Proposition 1.46(c).

(d) follows from (b) and (c): Indeed, Definition 2.72 yields

$$(\omega \left( a \right))^\perp \omega \left( b \right) = \sum_{(\omega(b))} \left( \omega \left( a \right), \left( \omega \left( b \right) \right) \right) \omega \left( b \right) = \sum_{(b)} \left( \omega \left( a \right), \omega \left( b_1 \right) \right) \omega \left( b_2 \right)$$

(since $\omega$ is a coalgebra morphism and thus we have $\sum_{(\omega(b))} \left( \omega \left( b_1 \right) \right) \otimes \left( \omega \left( b_2 \right) \right) = \sum_{(b)} \omega \left( b_1 \right) \otimes \omega \left( b_2 \right)$). Since $\omega$ is an isometry, this further simplifies to

$$(\omega \left( a \right))^\perp \omega \left( b \right) = \sum_{(b)} \left( a, b_1 \right) \omega \left( b_2 \right) = \omega \left( \sum_{(b)} \left( a, b_1 \right) b_2 \right) = \omega \left( a^\perp b \right),$$

whence part (d) of the exercise is solved.

(e) and (f) follow from the fact that $s_\lambda^\perp \left( s_\lambda \right) = s_\lambda/\mu$ (applied, respectively, to $\mu = (1^\ell)$ and to $\mu = (\lambda_1)$), and the fact that a parallel translation of a skew Ferrers shape doesn’t change the corresponding skew Schur function.

(g) It should be clear that $S \left( s_\lambda/\mu \right) = (-1)^{\ell(\mu)/\mu} s_\lambda/\mu$ follows from $\omega \left( s_\lambda/\mu \right) = s_\lambda/\mu$, and this, in turn, follows from (d) if one knows that $\omega \left( s_\lambda \right) = s_\lambda$ for every partition $\lambda$.

So it remains to prove that $\omega \left( s_\lambda \right) = s_\lambda$ for every partition $\lambda$. To do so, we use induction over $|\lambda|$, and fix $\lambda$. Part (b) yields that $\omega \left( s_\lambda \right) \omega \left( s_\lambda \right) = 1$. Since $\omega \left( s_\lambda \right) \in \Lambda$, this shows that $\omega \left( s_\lambda \right) = \pm s_\lambda$ for some partition $\nu$ (since a length-1 integral vector cannot have more than one nonzero coordinate). Hence, $\omega \left( s_\lambda \right) = \pm s_\lambda$

(since $\omega$ is an involution). Writing $\lambda$ as $(\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ with all of $\lambda_1, \lambda_2, \ldots, \lambda_\ell$ being positive, and writing
\( \nu \) as \((\nu_1, \nu_2, \nu_3, \ldots)\), we easily see (using (d) and (f)) that \( \omega(e_i^\perp s_\lambda) = \pm h_i^\perp (s_\nu) = \pm s_{(\nu_2, \nu_3, \nu_4, \ldots)} \neq 0 \), so that \( e_i^\perp s_\lambda \neq 0 \), and therefore \( \ell \geq \nu_1 \). On the other hand, (d) and (e) yield \( \omega(h_i^\perp s_\nu) = \pm e_i^\perp (s_\lambda) = \pm s_{(\lambda_1-1, \lambda_2-1, \ldots, \lambda_\ell-1)} \neq 0 \), thus \( h_i^\perp s_\nu \neq 0 \), so that \( \ell \leq \nu_1 \). Combined with \( \ell \geq \nu_1 \), this yields \( \ell = \nu_1 \). Hence, \( h_i^\perp s_\nu = h_i^\perp s_{\nu_1} = s_{(\nu_2, \nu_3, \nu_4, \ldots)} \) (by (f)), so that

\[
\omega(h_i^\perp s_\nu) = \omega(s_{(\nu_2, \nu_3, \nu_4, \ldots)}) = s_{(\nu_2, \nu_3, \nu_4, \ldots)}^t
\]

by the induction assumption. Compared with

\[
\omega(h_i^\perp s_\nu) = \pm e_i^\perp (s_\lambda) = \pm s_{(\lambda_1-1, \lambda_2-1, \ldots, \lambda_\ell-1)}
\]

(a consequence of (d) and (e)), this yields \( s_{(\nu_2, \nu_3, \nu_4, \ldots)} = \pm s_{(\lambda_1-1, \lambda_2-1, \ldots, \lambda_\ell-1)} \), from which it directly follows that \( \pm = + \) and \( (\nu_2, \nu_3, \nu_4, \ldots)^t = (\lambda_1 - 1, \lambda_2 - 1, \ldots, \lambda_\ell - 1) \). This quickly yields \( \nu = \lambda^t \), so that \( \omega(s_\lambda) = \pm s_\nu \) becomes \( \omega(s_\lambda) = s_\lambda^t \) (the \( \pm \) in \( \omega(s_\lambda) = \pm s_\nu \) is the same as in \( s_{(\nu_2, \nu_3, \nu_4, \ldots)} = \pm s_{(\lambda_1-1, \lambda_2-1, \ldots, \lambda_\ell-1)} \)), so we are done.

1153. Solution to Exercise 2.77. Solution to Exercise 2.77. (a) The claim that the sum \( \sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e_i^+ f \) is convergent is very easy to see: Any given \( f \in \Lambda = \bigoplus_{n \in \mathbb{N}} \Lambda_n \) lives in a finite direct sum \( \bigoplus_{n=0}^m \Lambda_n \subset \Lambda \); if we take \( i \in \mathbb{N} \) higher than \( m \), then \( e_i^+ f = 0 \) for degree reasons and therefore \( (-1)^i h_{m+i} e_i^+ f = 0 \).

The claim that the map \( B_m \) is \( k \)-linear is obvious. Exercise 2.77(n) is solved.

(c) Exercise 2.77(c) makes three claims. Let us first prove the first of them: the identity (2.40).

Proof of (2.40): Let \( \lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots) \) be a partition having at most \( n \) parts. Then,

\[
s_\lambda = s_{\lambda/\emptyset} = \det(h_{\lambda_1-i-j})_{i,j=1,2,\ldots,n} = \det(h_{\lambda_i-i-j})_{i,j=1,2,\ldots,n} = s_{(\lambda_1, \lambda_2, \ldots, \lambda_n)}
\]

(by the definition of \( s_{(\lambda_1, \lambda_2, \ldots, \lambda_n)} \)).

This proves (2.40).

Next, we will show that for every \( n \)-tuple \((\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}_n \), the symmetric function \( \overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} \) is either 0 or equals \( \pm s_\nu \) for some partition \( \nu \) having at most \( n \) parts.

Proof. Let \((\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}_n \) be any \( n \)-tuple. The definition of \( \overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} \) yields

\[
\overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = \det(h_{\alpha_i-i-j})_{i,j=1,2,\ldots,n}.
\]

If the \( n \) integers \( \alpha_1 - 1, \alpha_2 - 2, \ldots, \alpha_n - n \) are not pairwise distinct, then the matrix \( (h_{\alpha_i-i-j})_{i,j=1,2,\ldots,n} \) has two equal rows and thus its determinant is 0, so that \( \overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = \det(h_{\alpha_i-i-j})_{i,j=1,2,\ldots,n} = 0 \) in this case. So the claim is proven in this case, and this is a case we do not need to address anymore. Thus, assume WLOG that the \( n \) integers \( \alpha_1 - 1, \alpha_2 - 2, \ldots, \alpha_n - n \) are pairwise distinct. Hence, there exists a (unique) permutation \( \tau \in S_n \) satisfying \( \alpha_\tau(1) > \alpha_\tau(2) > \ldots > \alpha_\tau(n) \). Consider this \( \tau \).

Define an \( n \)-tuple \((\gamma_1, \gamma_2, \ldots, \gamma_n) \in \mathbb{Z}_n \) by setting

\[
\gamma_i = \alpha_\tau(i) - \tau(i) + i \quad \text{for every } i \in \{1, 2, \ldots, n\}.
\]

Then, it is easy to see that \( \gamma_1 \geq \gamma_2 \geq \ldots \geq \gamma_n \). Moreover, every \( i \in \{1, 2, \ldots, n\} \) satisfies \( \gamma_i - i = \alpha_\tau(i) - \tau(i) \) (by the definition of \( \gamma_i \)), and so the matrix \( (h_{\gamma_i-i-j})_{i,j=1,2,\ldots,n} \) is obtained from the matrix \( (h_{\alpha_i-i-j})_{i,j=1,2,\ldots,n} \) by permuting its rows according to the permutation \( \tau \). Since permuting the rows of a matrix multiplies the determinant of the matrix by the sign of the permutation, we thus see

\[
\det(h_{\gamma_i-i-j})_{i,j=1,2,\ldots,n} = (-1)^\tau \det(h_{\alpha_i-i-j})_{i,j=1,2,\ldots,n} = (-1)^\tau \overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)},
\]

so that

\[
(11.299) \quad \overline{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = (-1)^\tau \det(h_{\gamma_i-i-j})_{i,j=1,2,\ldots,n}.
\]

\[396\text{In fact, the map } e_i^+ \text{ lowers degree by } i, \text{ and thus annihilates } \bigoplus_{n=0}^m \Lambda_n \text{ when } i > m.\]

\[397\text{To prove this, it is enough to show that } \gamma_i \geq \gamma_{i+1} \text{ for every } i \in \{1, 2, \ldots, n-1\}. \text{ But this simplifies to } \alpha_\tau(i) - \tau(i) \geq \alpha_\tau(i+1) - \tau(i+1) + 1 \text{ (upon adding } i), \text{ and this follows from } \alpha_\tau(i) - \tau(i) > \alpha_\tau(i+1) - \tau(i+1).\]
If $\gamma_n < 0$, then the matrix $(h_{\gamma_i-i+j})_{i,j=1,2,...,n}$ on the right hand side of this equality has its $n$-th row consist of zeroes only, which implies that the determinant of this matrix is 0, and thus (11.299) becomes $\pi(\alpha_1,\alpha_2,...,\alpha_n) = 0$. Hence, we are done in the case $\gamma_n < 0$. So assume WLOG that we don’t have $\gamma_n < 0$. Thus, $\gamma_n \geq 0$. Combined with $\gamma_1 \geq \gamma_2 \geq ... \geq \gamma_n$, this yields that $(\gamma_1, \gamma_2, ..., \gamma_n)$ is a partition. This partition satisfies

$$s(\gamma_1, \gamma_2, ..., \gamma_n) = s(\gamma_1, \gamma_2, ..., \gamma_n)/\varnothing \ (2.22) \ \det \left( (h_{\gamma_i-i+j})_{i,j=1,2,...,n} \right) = \det \left( (h_{\gamma_i-i+j})_{i,j=1,2,...,n} \right)$$

(since $\varnothing = 0$ for all $j$). Thus, (11.299) becomes

$$\pi(\alpha_1,\alpha_2,...,\alpha_n) = (-1)^{\tau} \det \left( (h_{\gamma_i-i+j})_{i,j=1,2,...,n} \right) = (-1)^{\tau} s(\gamma_1, \gamma_2, ..., \gamma_n),$$

and therefore the symmetric function $\pi(\alpha_1,\alpha_2,...,\alpha_n)$ equals $\pm s_\nu$ for some partition $\nu$ having at most $n$ parts (namely, the partition $\nu$ is $(\gamma_1, \gamma_2, ..., \gamma_n)$, and the $\pm$ sign is $(-1)^{\tau}$).

We thus have shown that for every $n$-tuple $(\alpha_1, \alpha_2, ..., \alpha_n) \in \mathbb{Z}^n$, the symmetric function $\pi(\alpha_1,\alpha_2,...,\alpha_n)$ either is 0 or equals $\pm s_\nu$ for some partition $\nu$ having at most $n$ parts.

In order to complete the solution of Exercise 2.77(c), it now remains to prove (2.41).

For the proof of (2.41): Let $(\alpha_1, \alpha_2, ..., \alpha_n) \in \mathbb{Z}^n$ and $(\beta_1, \beta_2, ..., \beta_n) \in \mathbb{N}^n$ be two $n$-tuples. We need to prove that (2.41) holds.

The definition of $\pi(\alpha_1,\alpha_2,...,\alpha_n)$ yields

$$\pi(\alpha_1,\alpha_2,...,\alpha_n) = \det \left( (h_{\alpha_i-i+j})_{i,j=1,2,...,n} \right).$$

If the $n$ integers $\alpha_1 - 1, \alpha_2 - 2, ..., \alpha_n - n$ are not pairwise distinct, then (2.41) holds (because in this case, the matrices $(h_{\alpha_i-i+j})_{i,j=1,2,...,n}$ and $(h_{\alpha_i-\beta_i-i+j})_{i,j=1,2,...,n}$ have two equal rows each, so their determinants vanish, rendering both sides of (2.41) equal to zero$^{398}$. Thus, we WLOG assume that the $n$ integers $\alpha_1 - 1, \alpha_2 - 2, ..., \alpha_n - n$ are pairwise distinct. Thus, there exists a (unique) permutation $\tau \in S_n$ satisfying $\alpha_{\tau(1)} - 1 > \alpha_{\tau(2)} - 2 > ... > \alpha_{\tau(n)} - n$. Consider this $\tau$, and define an $n$-tuple $(\gamma_1, \gamma_2, ..., \gamma_n) \in \mathbb{Z}^n$ by setting

$$\gamma_i = \alpha_{\tau(i)} - \tau(i) + i \quad \text{for every } i \in \{1,2,...,n\}.$$

Then, it is easy to see that $\gamma_1 \geq \gamma_2 \geq ... \geq \gamma_n$. Moreover, the matrix $(h_{\gamma_i-i+j})_{i,j=1,2,...,n}$ is obtained from the matrix $(h_{\alpha_i-i+j})_{i,j=1,2,...,n}$ by permuting its rows according to the permutation $\tau$ (because every $i \in \{1,2,...,n\}$ satisfies $\gamma_i - i = \alpha_{\tau(i)} - \tau(i)$). This leads to (11.299) again (as in the previous proof), so that

$$\pi(\alpha_1,\alpha_2,...,\alpha_n) = (-1)^{\tau} \det \left( (h_{\gamma_i-i+j})_{i,j=1,2,...,n} \right) = (-1)^{\tau} s(\gamma_1, \gamma_2, ..., \gamma_n).$$

This equality, along with

$$\det \left( (h_{\alpha_i-\beta_j-i+j})_{i,j=1,2,...,n} \right) = (-1)^{\tau} \det \left( (h_{\gamma_i-\beta_j-i+j})_{i,j=1,2,...,n} \right)$$

(which, again, follows from the fact that the matrix on the right hand side is obtained from the matrix on the left hand side by permuting the rows according to $\tau$), shows that in order to prove (2.41), it is enough to show that

$$\pi(1.2.3,...,n)\pi(\gamma_1, \gamma_2, ..., \gamma_n) = \det \left( (h_{\gamma_i-\beta_j-i+j})_{i,j=1,2,...,n} \right).$$

But this is the same equality as (2.41), except with $(\alpha_1, \alpha_2, ..., \alpha_n)$ replaced by $(\gamma_1, \gamma_2, ..., \gamma_n)$. The advantage of $(\gamma_1, \gamma_2, ..., \gamma_n)$ over $(\alpha_1, \alpha_2, ..., \alpha_n)$ is that we know that $\gamma_1 \geq \gamma_2 \geq ... \geq \gamma_n$ (while the analogous chain of inequalities $\alpha_1 \geq \alpha_2 \geq ... \geq \alpha_n$ does not necessarily hold). So, we can **WLOG assume that** $\alpha_1 \geq \alpha_2 \geq ... \geq \alpha_n$ (because otherwise, we can replace $(\alpha_1, \alpha_2, ..., \alpha_n)$ by $(\gamma_1, \gamma_2, ..., \gamma_n)$). Assume this.

If $\alpha_n < 0$, then both matrices $(h_{\alpha_i-i+j})_{i,j=1,2,...,n}$ and $(h_{\alpha_i-\beta_i-i+j})_{i,j=1,2,...,n}$ have their $n$-th row consisting of only zeroes, and therefore the symmetric function $\pi(\alpha_1,\alpha_2,...,\alpha_n)$ equals $\pm s_\nu$ for some partition $\nu$ having at most $n$ parts. So this is a case in which (2.41) trivially holds. We thus assume WLOG that we are not

$^{398}$The left hand side is affected because of $\pi(\alpha_1,\alpha_2,...,\alpha_n) = \det \left( (h_{\alpha_i-i+j})_{i,j=1,2,...,n} \right)$.

$^{399}$For the left hand side, this is because of $\pi(\alpha_1,\alpha_2,...,\alpha_n) = \det \left( (h_{\alpha_i-i+j})_{i,j=1,2,...,n} \right)$.
in this case; hence, we don’t have $\alpha_n < 0$, so we have $\alpha_n \geq 0$. Combined with $\alpha_1 \geq \alpha_2 \geq ... \geq \alpha_n$, this yields that $(\alpha_1, \alpha_2, ..., \alpha_n)$ is a partition.

We have

$$\pi(\beta_1, \beta_2, ..., \beta_n) = \det \left( (h_{\beta_i - i + j})_{i,j=1,2,...,n} \right).$$

If the $n$ integers $\beta_1 - 1$, $\beta_2 - 2$, ..., $\beta_n - n$ are not pairwise distinct, then (2.41) holds (because in this case, the matrix $(h_{\beta_i - i + j})_{i,j=1,2,...,n}$ has two equal rows while the matrix $(h_{\alpha_i - \beta_j - i + j})_{i,j=1,2,...,n}$ has two equal columns; thus, both of these matrices have determinant 0, so that both sides of (2.41) equal to zero\(^{400}\).

Thus, we WLOG assume that the integers $\beta_1 - 1$, $\beta_2 - 2$, ..., $\beta_n - n$ are pairwise distinct. Thus, there exists a (unique) permutation $\zeta \in \mathfrak{S}_n$ satisfying $\beta_{\zeta(1)} - \zeta(1) > \beta_{\zeta(2)} - \zeta(2) > ... > \beta_{\zeta(n)} - \zeta(n)$. Consider this $\zeta$.

Define an $n$-tuple $(\delta_1, \delta_2, ..., \delta_n) \in \mathbb{Z}^n$ by setting

$$\delta_j = \beta_{\zeta(j)} - \zeta(j) + j$$

for every $j \in \{1, 2, ..., n\}$.

Then, it is easy to see that $\delta_1 \geq \delta_2 \geq ... \geq \delta_n \geq 0$\(^{401}\). Thus, $(\delta_1, \delta_2, ..., \delta_n)$ is a partition.

The matrix $(h_{\delta_i - i + j})_{i,j=1,2,...,n}$ is obtained from the matrix $(h_{\beta_i - i + j})_{i,j=1,2,...,n}$ by permuting its rows according to the permutation $\zeta$ (since the definition of $\delta_i$ yields $\delta_i = \beta_{\zeta(i)} - \zeta(i) + i$, thus $\delta_i - i = \beta_{\zeta(i)} - \zeta(i)$ for every $i \in \{1, 2, ..., n\}$). This leads to det $(h_{\delta_i - i + j})_{i,j=1,2,...,n} = (-1)^\zeta \det (h_{\beta_i - i + j})_{i,j=1,2,...,n} = (-1)^\zeta \pi(\beta_1, \beta_2, ..., \beta_n)$ and thus

$$\pi(\beta_1, \beta_2, ..., \beta_n) = (-1)^\zeta \det (h_{\delta_i - i + j})_{i,j=1,2,...,n} = (-1)^\zeta \pi(\delta_1, \delta_2, ..., \delta_n).$$

This equality, and the equality

$$\det (h_{\alpha_i - \beta_j - i + j})_{i,j=1,2,...,n} = (-1)^\zeta \det (h_{\alpha_i - \delta_j - i + j})_{i,j=1,2,...,n}$$

(this time because the matrix on the right hand side is obtained from that on the left hand side by permuting its columns according to $\zeta$) show that in order to prove (2.41), it is enough to show that $\pi(\delta_1, \delta_2, ..., \delta_n) = \det (h_{\alpha_i - \delta_j - i + j})_{i,j=1,2,...,n}.$

This is, of course, the same equality as (2.41), except with $(\beta_1, \beta_2, ..., \beta_n)$ replaced by $(\delta_1, \delta_2, ..., \delta_n)$. The advantage of $(\delta_1, \delta_2, ..., \delta_n)$ over $(\beta_1, \beta_2, ..., \beta_n)$ is that we know that $(\delta_1, \delta_2, ..., \delta_n)$ is a partition. So, we can WLOG assume that $(\beta_1, \beta_2, ..., \beta_n)$ is a partition (because otherwise, we can replace $(\beta_1, \beta_2, ..., \beta_n)$ by $(\delta_1, \delta_2, ..., \delta_n)$). Assume this.

Through a series of WLOG assumptions, we have now ensured that both $(\alpha_1, \alpha_2, ..., \alpha_n)$ and $(\beta_1, \beta_2, ..., \beta_n)$ are partitions. Thus,

$$s(\alpha_1, \alpha_2, ..., \alpha_n) = s(\alpha_1, \alpha_2, ..., \alpha_n)/\varnothing = \det (h_{\alpha_i - \varnothing_j - i + j})_{i,j=1,2,...,n}$$

$$= \det (h_{\alpha_i - i + j})_{i,j=1,2,...,n}$$

(since $\varnothing_j = 0$)

$$= \pi(\alpha_1, \alpha_2, ..., \alpha_n),$$

\(^{400}\)This time, the reason why this causes the left hand side to be zero is the identity $\pi(\beta_1, \beta_2, ..., \beta_n) = \det (h_{\beta_i - i + j})_{i,j=1,2,...,n}$.

\(^{401}\)The $\delta_1 \geq \delta_2 \geq ... \geq \delta_n$ part here is proven like the similar inequalities $\gamma_1 \geq \gamma_2 \geq ... \geq \gamma_n$ shown above; but the $\delta_n \geq 0$ part might require explanation. It stems from the fact that $\delta_n = \beta_{\zeta(n)} - \zeta(n) + n \geq \beta_{\zeta(n)} - n + n = \beta_{\zeta(n)} \geq 0$ (because we have assumed $(\beta_1, \beta_2, ..., \beta_n) \in \mathbb{N}^n$).
so that $\pi_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = s_{(\alpha_1, \alpha_2, \ldots, \alpha_n)}$. Similarly $\pi_{(\beta_1, \beta_2, \ldots, \beta_n)} = s_{(\beta_1, \beta_2, \ldots, \beta_n)}$. Hence,

$$\pi_{(\beta_1, \beta_2, \ldots, \beta_n)} s_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = s_{(\beta_1, \beta_2, \ldots, \beta_n)} s_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = s_{(\alpha_1, \alpha_2, \ldots, \alpha_n)}/(\beta_1, \beta_2, \ldots, \beta_n)$$

(2.22)

This proves (2.41) and thus completes the proof of Exercise 2.77(c).

(d) In the following, we use the so-called Teiser bracket notation: For every assertion $A$, we let $[A]$ denote the integer \( \begin{cases} 1, & \text{if } A \text{ is true;} \\ 0, & \text{if } A \text{ is false} \end{cases} \). (This integer is called the truth value of $A$.)

Before we come to the solution of Exercise 2.77(d), let us recall a general fact about determinants:

- Every commutative ring $A$, every positive integer $N$ and every matrix $(\beta_{i,j})_{i,j=1,2,\ldots,N} \in A^{N \times N}$ satisfy

\[
\det((\beta_{i,j})_{i,j=1,2,\ldots,N}) = \sum_{k=1}^{N} (-1)^{k-1} \beta_{1,k} \det((\beta_{i+1,j+1+(j+1)}_{i,j=1,2,\ldots,N-1}) \quad \text{(by the definition of } \pi_{(\gamma_1, \gamma_2, \ldots, \gamma_{n+1})})
\]

(11.300)

This is just one possible way to write the Laplace expansion formula for the expansion of the determinant of a matrix with respect to its first row.

Let us now solve Exercise 2.77(d). Let $n \in \mathbb{N}$, let $m \in \mathbb{Z}$ and let $(\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Z}^n$. We must prove (2.42).

Define an $(n+1)$-tuple $(\gamma_1, \gamma_2, \ldots, \gamma_{n+1}) \in \mathbb{Z}^{n+1}$ by $(\gamma_1, \gamma_2, \ldots, \gamma_{n+1}) = (m, \alpha_1, \alpha_2, \ldots, \alpha_n)$. Then, $\gamma_i = m$, whereas every $i \in \{1, 2, \ldots, n\}$ satisfies $\gamma_{i+1} = \alpha_i$. Since $(m, \alpha_1, \alpha_2, \ldots, \alpha_n) = (\gamma_1, \gamma_2, \ldots, \gamma_{n+1})$, we have

\[
\pi_{(m, \alpha_1, \alpha_2, \ldots, \alpha_n)} = \pi_{(\gamma_1, \gamma_2, \ldots, \gamma_{n+1})} = \det((h_{\gamma_i, i+j})_{i,j=1,2,\ldots,n+1})
\]

(by definition of $\pi_{(\gamma_1, \gamma_2, \ldots, \gamma_{n+1})}$)

\[
= \sum_{k=1}^{n+1} (-1)^{k-1} h_{\gamma_1-1,k} \det((h_{\gamma_i, i+j})_{i,j=1,2,\ldots,n})
\]

(by (11.300), applied to $A = \Lambda$, $N = n+1$ and $\beta_{i,j} = h_{\gamma_i-i+j}$)

\[
= \sum_{k=0}^{n} (-1)^{(k+1)-1} h_{\gamma_1-1+(k+1)} \det((h_{\gamma_1-1+(i+1)+(j+j+1)})_{i,j=1,2,\ldots,n})
\]

(here, we have substituted $k+1$ for $k$ in the sum)

(11.301)

Now, let us check that

\[
\det((h_{\alpha_i, i+j})_{i,j=1,2,\ldots,n}) = e_k^\uparrow \pi_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} \quad \text{for every } k \in \{0, 1, \ldots, n\}
\]

Proof of (11.302): Let $k \in \{0, 1, \ldots, n\}$. The partition $(1^k)$ has length $k \leq n$, and thus can be identified with the $n$-tuple $(1, 1, \ldots, 1, 0, 0, \ldots, 0)$ of length $k$ times $n$ times. Hence, $\pi_{(1^k)} = s_{(1^k)}$, which is given in (2.40). Hence, $e_k = s_{(1^k)} = \pi_{(1^k)}$ by (2.40), applied to $\lambda = (\beta_1, \beta_2, \ldots, \beta_n)$, so that

\[
e_k^\uparrow \pi_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = \pi_{(\beta_1, \beta_2, \ldots, \beta_n)} \pi_{(\alpha_1, \alpha_2, \ldots, \alpha_n)} = \det((h_{\alpha_i, i+j})_{i,j=1,2,\ldots,n})
\]

(by (2.41))
(since \(\alpha_i - (1 - \lfloor j \geq k + 1 \rfloor) - i + j = \alpha_i - (i + 1) + (j + \lfloor j \geq k + 1 \rfloor)\) for all \((i, j) \in \{1, 2, \ldots, n\}^2\)). This proves (11.302).

Now, (11.301) becomes
\[
\bar{s}(m, \alpha_1, \alpha_2, \ldots, \alpha_n) = \sum_{k=0}^{n} (-1)^k h_{m+k} \det \left( (h_{\gamma_{i+1} - (i+1) + (j+j\geq k+1)})_{i,j=1,2,\ldots, n} \right)
\]
(11.303)
\[
= \sum_{k=0}^{n} (-1)^k h_{m+k} e^+_k \bar{s}(\alpha_1, \alpha_2, \ldots, \alpha_n) = \sum_{i=0}^{n} (-1)^i h_{m+i} e^+_i \bar{s}(\alpha_1, \alpha_2, \ldots, \alpha_n)
\]
(here, we renamed the summation index \(k\) as \(i\)).

But in order to solve Exercise 2.77(d), we need to prove a very similar yet different formula:
\[
\bar{s}(m, \alpha_1, \alpha_2, \ldots, \alpha_n) = \sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e^+_i \bar{s}(\alpha_1, \alpha_2, \ldots, \alpha_n).
\]
These two formulas differ in a minor detail: The sum on the right hand side of (11.304) runs over all \(i \in \mathbb{N}\), whereas the sum on the right hand side of (11.303) only runs over \(i \in \{0, 1, \ldots, n\}\). If we can show that these two sums are equal, then it will follow that the equality (11.304) that we are proving and the equality (11.303) that we have proven are equivalent, and so the former equality must hold, and Exercise 2.77(d) will be solved.

So we need to prove that the sum on the right hand side of (11.304) and the sum on the right hand side of (11.303) are equal. To achieve this, it clearly suffices to show that all addends in which these sums differ are 0. But these addends are the \((-1)^i h_{m+i} e^+_i \bar{s}(\alpha_1, \alpha_2, \ldots, \alpha_n)\) for \(i \in \mathbb{N}\) satisfying \(i > n\). So we need to prove that \((-1)^i h_{m+i} e^+_i \bar{s}(\alpha_1, \alpha_2, \ldots, \alpha_n) = 0\) for all \(i \in \mathbb{N}\) satisfying \(i > n\). Let us do this now.

Let \(i \in \mathbb{N}\) be such that \(i > n\). We need to show that \((-1)^i h_{m+i} e^+_i \bar{s}(\alpha_1, \alpha_2, \ldots, \alpha_n) = 0\).

The second statement of Exercise 2.77(c) says that the symmetric function \(\bar{s}(\alpha_1, \alpha_2, \ldots, \alpha_n)\) either is 0 or equals \(\pm s_\nu\) for some partition \(\nu\) having at most \(n\) parts. If it is 0, then \((-1)^i h_{m+i} e^+_i \bar{s}(\alpha_1, \alpha_2, \ldots, \alpha_n) = 0\) is obvious, and so we are done. Thus, we can WLOG assume that we are in the other case, i.e., the function \(\bar{s}(\alpha_1, \alpha_2, \ldots, \alpha_n)\) equals \(\pm s_\nu\) for some partition \(\nu\) having at most \(n\) parts. Assume that we are in this case, and consider this \(\nu\). Since \(\nu\) has at most \(n\) parts, we have \(\ell(\nu) \leq n\). Now, \(\ell((1^i)) = i > n \geq \ell(\nu)\), and therefore \((1^i) \not\subseteq \nu\). But \(e_i = s_{(1^i)}\) and thus \(e^+_i s_\nu = s^+_{(1^i)} s_\nu = s_\nu / (1^i) = 0\) (because \((1^i) \not\subseteq \nu\)).

Now,
\[
(-1)^i h_{m+i} e^+_i \bar{s}(\alpha_1, \alpha_2, \ldots, \alpha_n) = \pm (-1)^i h_{m+i} e^+_i s_\nu = 0,
\]
which concludes our proof. Exercise 2.77(d) is thus solved.

(b) Let \(n = \ell(\lambda)\). Then, \(\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)\), and the equality (2.40) yields \(s_\lambda = \bar{s}(\lambda_1, \lambda_2, \ldots, \lambda_n)\).

Also, \(\ell((m, \lambda_1, \lambda_2, \lambda_3, \ldots)) = n + 1\), and thus (2.40) (applied to \(n + 1\) and \((m, \lambda_1, \lambda_2, \lambda_3, \ldots)\) instead of \(n\) and \(\lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots)\)) yields
\[
s_{(m, \lambda_1, \lambda_2, \lambda_3, \ldots)} = \bar{s}_{(m, \lambda_1, \lambda_2, \lambda_3, \ldots)} = \bar{s}_{(m, \lambda_1, \lambda_2, \lambda_3, \ldots)}
\]
Compared with
\[
\sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e^+_i \bar{s}_{\lambda} = \sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e^+_i \bar{s}_{(\lambda_1, \lambda_2, \ldots, \lambda_n)}
\]
(11.304)
\[
= \bar{s}_{(m, \lambda_1, \lambda_2, \lambda_3, \ldots)} = \bar{s}_{(m, \lambda_1, \lambda_2, \lambda_3, \ldots)} \quad \text{(by (2.42), applied to } \alpha_j = \lambda_j),
\]
this yields 
\[
\sum_{i \in \mathbb{N}} (-1)^i h_{m+i} e^+_i s_\lambda = s_{(m, \lambda_1, \lambda_2, \lambda_3, \ldots)}.
\]
This solves Exercise 2.77(b).

(c) This follows by induction over \(n\) from Exercise 2.77(d). The induction base (the \(n = 0\) case) is trivial. For the induction step, we need to prove that
\[
\bar{s}_{(\lambda_1, \lambda_2, \ldots, \lambda_{n+1})} = B_{\alpha_1} \circ B_{\alpha_2} \circ \cdots \circ B_{\alpha_{n+1}} (1),
\]
assuming that

\[ (11.305) \quad \mathfrak{s}_{(\alpha_2, \ldots, \alpha_{n+1})} = (B_{\alpha_2} \circ B_{\alpha_3} \circ \cdots \circ B_{\alpha_{n+1}}) (1). \]

But this is fairly straightforward: Applying the map \( B_{\alpha_1} \) to \((11.305)\), we obtain

\[ B_{\alpha_1} \left( \mathfrak{s}_{(\alpha_2, \ldots, \alpha_{n+1})} \right) = B_{\alpha_1} \left( (B_{\alpha_2} \circ B_{\alpha_3} \circ \cdots \circ B_{\alpha_{n+1}}) (1) \right) = (B_{\alpha_1} \circ B_{\alpha_2} \circ \cdots \circ B_{\alpha_{n+1}}) (1). \]

Compared with

\[ B_{\alpha_1} \left( \mathfrak{s}_{(\alpha_2, \ldots, \alpha_{n+1})} \right) = \sum_{i \in \mathbb{N}} (-1)^i h_{\alpha_1 + i} e_i \mathfrak{s}_{(\alpha_2, \ldots, \alpha_{n+1})} \quad \text{(by the definition of the map \( B_{\alpha_1} \))} \]

\[ = \mathfrak{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_{n+1})} \quad \text{(by Exercise 2.77(d), applied to \( \alpha_1 \) and \((\alpha_2, \alpha_3, \ldots, \alpha_{n+1})\) instead of \( m \) and \((\alpha_1, \alpha_2, \ldots, \alpha_n)\))} \]

this yields

\[ \mathfrak{s}_{(\alpha_1, \alpha_2, \ldots, \alpha_{n+1})} = (B_{\alpha_1} \circ B_{\alpha_2} \circ \cdots \circ B_{\alpha_{n+1}}) (1), \]

which completes the induction step and thus the proof.

**Remark:** Our solution to Exercise 2.77(d) was modelled after the rough sketch of a solution to Exercise 2.77(b) given in [173, §4.20]; but it involved many technicalities which are not necessary if one is only interested in a solution to Exercise 2.77(b). (Specifically, if one only wants to solve Exercise 2.77(b), one can avoid the use of Exercise 2.77(c).)

We solved Exercise 2.77(e) using Exercise 2.77(d), but of course one can just as well turn this around and solve Exercise 2.77(d) using Exercise 2.77(e) if one has an independent solution to Exercise 2.77(e). Such an independent solution can be extracted from [16, Corollary 3.30] (using the realization that the immaculate creation operators \( \mathbb{B}_m \) of [16] are lifts of our Bernstein operators \( \mathbb{B}_m \) to \( \text{NSym} \)).

### 11.54. Solution to Exercise 2.79

**Solution to Exercise 2.79.** (a) In fact, for any commutative ring \( A \) and any power series \( f \in A[[t]] \) with constant term 1, there is a unique family \((x_n)_{n \geq 1}\) of elements of \( A \) such that

\[ (11.306) \quad f = \prod_{n=1}^{\infty} (1 - x_n t^n)^{-1}. \]

This family is constructed recursively: If \( x_1, x_2, \ldots, x_{k-1} \) have been determined (for some \( k \geq 1 \)), then \( x_k \) is obtained by comparing coefficients before \( t^k \) in the equation \((11.306)\) (the coefficient before \( t^k \) on the left hand side is a known constant, and the coefficient before \( t^k \) on the right hand side can be written in the form \( x_k + \text{some polynomial in } x_1, x_2, \ldots, x_{k-1} \)), and thus the equality of these coefficients gives a linear equation in \( x_k \) which can be uniquely solved for \( x_k \). The family \((x_n)_{n \geq 1}\) thus constructed satisfies \((11.306)\) (because the constant terms on both sides of \((11.306)\) are 1, whereas for every \( k \geq 1 \), the coefficients before \( t^k \) on both sides of \((11.306)\) are equal due to the construction of \( x_k \)).

(b) Again, this generalizes: Let us say that a power series \( f \in A[[t]] \) over a graded commutative ring \( A \) is **equigraded** if, for every \( n \in \mathbb{N} \), the coefficient of \( f \) before \( t^n \) is homogeneous of degree \( n \). Then, if \( f \in A[[t]] \) is an equigraded power series with constant term 1, then the unique family \((x_n)_{n \geq 1}\) satisfying \((11.306)\) has the property that \( x_n \) is homogeneous of degree \( n \) for every positive \( n \). This is rather easy to see by induction.
(c) By the definition of the \( w_n \), we have \( H(t) = \prod_{n=1}^{\infty} (1 - w_n t^n)^{-1} \). Expanding and comparing coefficients yields precisely \( \sum_{\lambda \in \Par_n} w_\lambda = h_n \).  

(d) From (c), we know that \( h_n = \sum_{\lambda \in \Par_n} w_\lambda \) for every \( n \in \mathbb{N} \). In other words, \( h_n = \sum_{\lambda \in \Par_n} w_\lambda \) for every \( n \in \mathbb{N} \), where we are using the notation \( \lambda \vdash n \) for \( \lambda \in \Par_n \). Multiplying such identities, we conclude that every partition \( \mu = (\mu_1, \mu_2, \ldots, \mu_\ell) \) with \( \ell = \ell(\mu) \) satisfies

\[
\sum_{\lambda(1) \cup \lambda(2) \cup \cdots \cup \lambda(\ell)} w_\lambda = \sum_{\lambda(1) \in \Par; \lambda(2) \in \Par; \cdots; \lambda(\ell) \in \Par} w_\lambda(1) w_\lambda(2) \cdots w_\lambda(\ell),
\]

where \( \lambda(1) \cup \lambda(2) \cup \cdots \cup \lambda(\ell) \) denotes the partition obtained by sorting all the parts of \( \lambda(1), \lambda(2), \ldots, \lambda(\ell) \) in increasing order. This easily shows that the matrix transforming the basis \( \{ h_\lambda \} \) to the basis \( \{ w_\lambda \} \) is unit triangular with respect to the refinement order on \( \Par_n \). Hence, \( \{ w_\lambda \} \) is a basis of \( \Lambda \), and thus \( \{ w_\lambda \} \) is a basis of the whole \( \Lambda \). This proves (d).

An alternative solution of (d) proceeds as follows: From part (c), we see that the \( w_n \) generate \( \Lambda \) as a \( k \)-algebra. In other words, the family \( \{ w_\lambda \} \) spans the \( k \)-module \( \Lambda \). Since everything is graded, this yields that the family \( \{ w_\lambda \} \) spans the \( k \)-module \( \Lambda_n \). Therefore, the \( k \)-linear map

\[
w : \Lambda \to \Lambda_n, \quad h_\lambda \mapsto w_\lambda \quad \text{for every } \lambda \in \Par_n
\]

is surjective. But there is a well-known fact (Exercise 2.46(a)) that a surjective endomorphism of a finitely generated \( A \)-module (where \( A \) is any commutative ring) must be an automorphism. Applying this to the surjective endomorphism \( w \) of the finitely generated \( k \)-module \( \Lambda_n \), we conclude that \( w \) is an automorphism, so that \( \{ w_\lambda \} \in \Par_n \) is a basis of the \( k \)-module \( \Lambda_n \) (being the image of the basis \( \{ h_\lambda \} \)). Hence, \( \{ w_\lambda \} \) is a basis of the whole \( \Lambda \). This proves (d) again.

\[\text{(2.46a)}\]

\[\text{(2.14)}\]

\[\text{(2.15)}\]

\[\text{(2.16)}\]
(e) By the definition of the $w_n$, we have $H(t) = \prod_{n=1}^{\infty} (1 - w_n t^n)^{-1}$. Hence,

$$\frac{d}{dt} \log(H(t)) = \frac{d}{dt} \left( \log \prod_{n=1}^{\infty} (1 - w_n t^n)^{-1} \right) = \frac{d}{dt} \sum_{n=1}^{\infty} \log((1 - w_n t^n)^{-1})$$

$$= \sum_{n=1}^{\infty} \frac{d}{dt} \log((1 - w_n t^n)^{-1}) = \sum_{n=1}^{\infty} \frac{n w_n t^{n-1}}{1 - w_n t^n}.$$ 

Compared with

$$\frac{d}{dt} \log(H(t)) = \frac{H'(t)}{H(t)} = \sum_{m \geq 0} p_{m+1} t^m \quad \text{(by Exercise 2.48),}$$

this yields

$$\sum_{m \geq 0} p_{m+1} t^m = \sum_{n=1}^{\infty} \frac{n w_n t^{n-1}}{1 - w_n t^n}.$$ 

Multiplying this by $t$, we obtain

$$\sum_{m \geq 0} p_{m+1} t^{m+1} = \sum_{n=1}^{\infty} \frac{n w_n t^n}{1 - w_n t^n} = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} n w_n^k t^{nk}$$

$$= \sum_{m=1}^{\infty} \sum_{d|m} d w_d^{m/d} t^m = \sum_{n=1}^{\infty} \sum_{d|n} d w_d^{n/d} t^n,$$

so that $\sum_{n=1}^{\infty} \sum_{d|n} d w_d^{n/d} t^n = \sum_{m \geq 0} p_{m+1} t^{m+1} = \sum_{n \geq 1} p_n t^n$. Comparing coefficients yields the claim of part (e).

---

403 We will be using logarithms here, so prima facie our argument only works when the base ring $k$ is a $\mathbb{Q}$-algebra. However, it is easy to see that our argument can easily be adapted to work in the general case as well. For example, one can argue that even if the power series $\log t$ is not defined if $k$ is not a $\mathbb{Q}$-algebra, the notion of the logarithmic derivative $\frac{d}{dt} (\log Q)$ is well-defined for every $k$ whenever $Q \in k[[t]]$ is a power series with constant term 1 (for example, one could use the formula $\frac{d}{dt} (\log Q) = \frac{Q'(t)}{Q(t)}$ as the definition of the logarithmic derivative) and still has the familiar property of turning products into sums in this generality. See the solution to Exercise 2.48 for details about this.
(f) This is done in [42]. For the sake of completeness: First, let \( n \geq k \geq 2 \). Then, every \( i \in \{2, 3, ..., k-1\} \) satisfies \( f_{i,i} = \sum_{\lambda \in \text{Par}_i, \min \lambda \geq i} w_\lambda = w_{(i)} = w_i \). Thus,

\[
\sum_{i=2}^{k-1} f_{i,i} = \sum_{i=2}^{k-1} w_i = \sum_{\lambda \in \text{Par}_{n-i}, \min \lambda \geq i} w_\lambda = \sum_{\lambda \in \text{Par}_{n-i}, \min \lambda \geq i} w_{\lambda, \min \lambda \geq i} = \sum_{\lambda \in \text{Par}_{n-i}, \min \lambda = i} w_{\lambda} = \sum_{\lambda \in \text{Par}_{n-i}, \min \lambda = i} w_{\lambda} \\
= h_n - w_1 \sum_{\lambda \in \text{Par}_{n-1}} w_{\lambda} = h_n - f_{n,k} \\
= h_n - h_{n-1} - f_{n,k} \\
= h_{(n)} - h_{(n-1)} - f_{n,k} \\
= s_{(n)} - h_{1} s_{(n-1)} - f_{n,k} \\
= s_{(n)} + s_{(n-1,1)} - f_{n,k} \\
= -s_{(n-1,1)} - f_{n,k};
\]

in other words,

\[
-f_{n,k} = s_{(n-1,1)} + \sum_{i=2}^{k-1} f_{i,i} f_{n-i,i}.
\]

From this, we can conclude inductively that \(-f_{n,k}\) is a sum of Schur functions for every \( n \in \mathbb{N} \) and \( k \geq 2 \) (in fact, the trivial cases with \( n < 2 \) have to be taken as induction base). Since \( f_{n,n} = \sum_{\lambda \in \text{Par}_n, \min \lambda \geq n} w_\lambda = w_{(n)} = w_n \), this yields that \(-w_n\) is a sum of Schur functions for every \( n \geq 2 \).
(g) Every partition \( \lambda \) can be uniquely written in the form \( \lambda = (1^{m_1}2^{m_2}3^{m_3} \cdots) \) for some weak composition \((m_1, m_2, m_3, \ldots)\). Thus,

\[
\sum_{\lambda \in \text{Par}} w_{\lambda}(x) r_{\lambda}(y) = \sum_{(m_1, m_2, m_3, \ldots)} w_{(1^{m_1}2^{m_2}3^{m_3} \cdots)}(x) r_{(1^{m_1}2^{m_2}3^{m_3} \cdots)}(y)
\]

(by the definition of \( w_{\lambda} \)) (by the definition of \( r_{\lambda} \))

\[
= \sum_{(m_1, m_2, m_3, \ldots)} \left( \prod_{i \geq 1} (w_i(x))^{m_i} \right) \left( \prod_{i \geq 1} h_{m_i}(y_1^i, y_2^i, y_3^i, \ldots) \right)
\]

\[
= \prod_{i \geq 1} \sum_{m \in \mathbb{N}} h_m(y_1^i, y_2^i, y_3^i, \ldots) (w_i(x))^{m_i}
\]

(by \(2.14\), upon substitution of \((y_1^i, y_2^i, y_3^i, \ldots) \) and \( w_i(x) \) for \( x \) and \( t \))

\[
= \prod_{i \geq 1} \prod_{j \geq 1} (1 - y_j^i w_i(x))^{-1} = \prod_{j \geq 1} \prod_{i \geq 1} (1 - y_j^i w_i(x))^{-1}
\]

\[
= \prod_{j \geq 1} H(y_j) = \prod_{i, j = 1} (1 - x_i y_j)^{-1}.
\]

This proves (g).

(h) For every partition \( \lambda \), both symmetric functions \( w_{\lambda} \) and \( r_{\lambda} \) are homogeneous of degree \( |\lambda| \). (In fact, for \( w_{\lambda} \) this follows from Exercise 2.79(c), whereas for \( r_{\lambda} \) this is easily derived from the definition.)

From Exercise 2.79(g), we obtain \( \sum_{\lambda \in \text{Par}} w_{\lambda}(x) r_{\lambda}(y) = \prod_{j, i = 1} (1 - x_i y_j)^{-1} = \sum_{\lambda \in \text{Par}} s_{\lambda}(x)s_{\lambda}(y) \) (by \(2.24\)). Thus, we can apply Exercise 2.47(a) to \( u_{\lambda} = w_{\lambda} \) and \( v_{\lambda} = r_{\lambda} \). As a result, we obtain that \((w_{\lambda})_{\lambda \in \text{Par}}\) and \((r_{\lambda})_{\lambda \in \text{Par}}\) are \( k \)-bases of \( \Lambda \), and actually are dual bases with respect to the Hall inner product on \( \Lambda \). Thus we have solved Exercise 2.79(h), but also given another proof of Exercise 2.79(d) in the process (because we have shown once again that \((w_{\lambda})_{\lambda \in \text{Par}}\) is a \( k \)-basis of \( \Lambda \).

11.55. Solution to Exercise 2.80. Solution to Exercise 2.80. (a) Let \( f \in A \).

Recall the following fundamental fact from linear algebra: If \( k \) is a commutative ring, if \( A \) is a \( k \)-module, if \((\cdot, \cdot) : A \times A \to k \) is a symmetric \( k \)-bilinear form on \( A \), and if \((u_\lambda)_{\lambda \in L} \) and \((v_\lambda)_{\lambda \in L} \) are two \( k \)-bases of \( A \) which are dual to each other with respect to the form \((\cdot, \cdot) \) (where \( L \) is some indexing set), then every \( a \in A \) satisfies

\[
(1.308) \quad a = \sum_{\lambda \in L} (u_\lambda, a) v_\lambda.
\]

We can apply this fact to \( k = \mathbb{Q} \), \( A = \Lambda \mathbb{Q} \), \( L = \text{Par} \), \((u_\lambda)_{\lambda \in \text{Par}} = (p_\lambda)_{\lambda \in \text{Par}} \), \((v_\lambda)_{\lambda \in \text{Par}} = (z^{-1}_\lambda p_\lambda)_{\lambda \in \text{Par}} \) and \( a = f \) (because the bases \((p_\lambda)_{\lambda \in \text{Par}} \) and \((z^{-1}_\lambda p_\lambda)_{\lambda \in \text{Par}} \) of \( \Lambda \mathbb{Q} \) are dual to each other with respect to the Hall
inner product $\langle \cdot, \cdot \rangle$, as Corollary 2.45(b) shows). As a result, we obtain

$$f = \sum_{\lambda \in \text{Par}} (p_\lambda, f) z_\lambda^{-1} p_\lambda.$$  

Hence,

$$Z(f) = Z \left( \sum_{\lambda \in \text{Par}} (p_\lambda, f) z_\lambda^{-1} p_\lambda \right) = \sum_{\lambda \in \text{Par}} (p_\lambda, f) z_\lambda^{-1} Z (p_\lambda) = \sum_{\lambda \in \text{Par}} (p_\lambda, f) z_\lambda^{-1} z_\lambda p_\lambda$$

$$= \sum_{\lambda \in \text{Par}} (p_\lambda, f) p_\lambda \in \sum_{\lambda \in \text{Par}} z p_\lambda \in \Lambda.$$  

Since we have proven this for every $f \in \Lambda$, we thus have shown that $Z(\Lambda) \subset \Lambda$. This solves Exercise 2.80(a).

(b) First solution of Exercise 2.80(b): Consider two variable sets $x = (x_1, x_2, x_3, \ldots)$ and $y = (y_1, y_2, y_3, \ldots)$. Let $xy$ denote the variable set

$$(x_i y_j)_{(i,j) \in \{1,2,3,\ldots\}^2} = (x_1 y_1, x_1 y_2, x_1 y_3, \ldots, x_2 y_1, x_2 y_2, x_2 y_3, \ldots, x_3 y_1, x_3 y_2, x_3 y_3, \ldots, \ldots).$$

Now, we claim that for every $f \in \Lambda_Q$, there is a well-defined element $f(xy) := f((x_i y_j)_{(i,j) \in \{1,2,3,\ldots\}^2})$ of $\mathbb{Q}[x,y]$.

This claim (11.309) is not obvious! For example, there is generally no well-defined element $f((x_i + y_j)_{(i,j) \in \{1,2,3,\ldots\}^2})$, because e.g. in the case of $f = e_1$ we would have $e_1((x_i + y_j)_{(i,j) \in \{1,2,3,\ldots\}^2}) = \sum_{(i,j) \in \{1,2,3,\ldots\}^2} (x_i + y_j)$, which is a sum containing infinitely many $x_1$’s. So there is some subtlety which allows us to make sense of $f((x_i + y_j)_{(i,j) \in \{1,2,3,\ldots\}^2})$ but not of $f((x_i y_j)_{(i,j) \in \{1,2,3,\ldots\}^2})$.

Here is a sketch of a proof of (11.309): Consider a new variable set $s := (s_{i,j})_{(i,j) \in \{1,2,3,\ldots\}^2}$ whose variables are indexed by pairs of positive integers. This variable set $s$ is still countably infinite, and so there is an isomorphism $\Lambda_Q = \Lambda_Q(x) \to \Lambda_Q(s)$. On the other hand, we can consider the ring $\mathbb{Q}[[s]]$ of formal power series in the variables from the set $s$, and then we have $\Lambda_Q(s) \subset \mathbb{Q}[[s]]$. It is easy to see that $g((x_i y_j)_{(i,j) \in \{1,2,3,\ldots\}^2})$ is a well-defined element of $\mathbb{Q}[[x,y]]$ for every $g \in \mathbb{Q}[[s]]$. Hence, for every $f \in \Lambda_Q$, there is a well-defined element $f(xy) := f((x_i y_j)_{(i,j) \in \{1,2,3,\ldots\}^2})$ of $\mathbb{Q}[[x,y]]$ (because we can regard $f$ as an element of $\Lambda_Q(s) \subset \mathbb{Q}[[s]]$ by means of the isomorphism $\Lambda_Q = \Lambda_Q(x) \to \Lambda_Q(s)$). This proves (11.309).

---

404 Proof. Let $g \in \mathbb{Q}[[s]]$. When $(x_i y_j)_{(i,j) \in \{1,2,3,\ldots\}^2}$ is substituted for $(s_{i,j})_{(i,j) \in \{1,2,3,\ldots\}^2}$ in the power series $g \in \mathbb{Q}[[s]]$, every monomial in the variables $s$ turns into a monomial in the variables $(x,y)$, and any given monomial in $(x,y)$ can only be obtained (this way) from finitely many monomials in $s$ (indeed, a given monomial $x^\alpha y^\beta$ in $(x,y)$ can only be obtained from monomials $\prod_{i,j=1,2,3,\ldots} s_{i,j}^{\gamma_{i,j}}$ whose exponents $\gamma_{i,j}$ satisfy the equations

$$\sum_{i=1}^{\infty} \gamma_{i,j} = \beta_j \quad \text{for all } j \in \{1,2,3,\ldots\};$$

$$\sum_{j=1}^{\infty} \gamma_{i,j} = \alpha_i \quad \text{for all } i \in \{1,2,3,\ldots\};$$

but it is easy to see that these equations leave only finitely many possibilities for the monomial $\prod_{i,j=1,2,3,\ldots} s_{i,j}^{\gamma_{i,j}}$. Hence, the substitution yields an infinite sum of monomials in which every monomial occurs only finitely often, therefore, this sum converges. This shows that $g((x_i y_j)_{(i,j) \in \{1,2,3,\ldots\}^2})$ is a well-defined element of $\mathbb{Q}[[x,y]]$.  

Due to (11.309), we can define a map
\[
\tilde{\Delta}_\pi : \Lambda Q \to \mathbb{Q}[[x,y]],
\]
\[
\tilde{\Delta}_\pi (p_n) = \tilde{\Delta}_\pi (p_n) = \sum_{i=1}^{\infty} x^i = \sum_{j=1}^{\infty} y^j = \sum_{i,j \geq 1} x^i y^j = \sum_{(i,j) \in \{(1,2,3,\ldots)\}} (x_i y_j)^n
\]

This map $\tilde{\Delta}_\pi$ is an evaluation map (in an appropriate sense), and thus a $\mathbb{Q}$-algebra homomorphism. It is also clear that $\tilde{\Delta}_\pi (\Lambda) \subset \mathbb{Z}[[x,y]]$ (where $\mathbb{Z}[[x,y]]$ is regarded as a subring of $\mathbb{Q}[[x,y]]$ in the obvious way).

Now, recall the $\mathbb{Q}$-algebra isomorphism $\Lambda Q \otimes_{\mathbb{Q}} \Lambda Q \to R_Q (x,y)^{E(\infty) \times E(\infty)}$ constructed in (2.2) (applied to $k = \mathbb{Q}$). This entails a $\mathbb{Q}$-algebra injection $\Lambda Q \otimes_{\mathbb{Q}} \Lambda Q \to \mathbb{Q}[[x,y]]$ (since $R_Q (x,y)^{E(\infty) \times E(\infty)} \subset R_Q (x,y) \subset \mathbb{Q}[[x,y]]$). Denote this injection by $\iota$.

We shall now show that
\[
(11.310) \quad \iota \circ \tilde{\Delta}_\pi = \tilde{\Delta}_\pi.
\]

**Proof of (11.310):** Indeed, (11.310) is an equality between $\mathbb{Q}$-algebra homomorphisms (since $\tilde{\Delta}_\pi$, $\iota$ and $\Delta_\pi$ are $\mathbb{Q}$-algebra homomorphisms), and thus, in order to prove it, we only need to check that it holds on a generating set of the $\mathbb{Q}$-algebra $\Lambda Q$. We do this on the generating set $\{p_n\}_{n \geq 1}$, by noticing that every $n \geq 1$ satisfies
\[
(\iota \circ \tilde{\Delta}_\pi) (p_n) = \tilde{\Delta}_\pi (p_n) = \sum_{i=1}^{\infty} x^i = \sum_{j=1}^{\infty} y^j = \sum_{i,j \geq 1} x^i y^j = \sum_{(i,j) \in \{(1,2,3,\ldots)\}} (x_i y_j)^n
\]

So (11.310) is proven.

From (11.310), we obtain $\iota (\tilde{\Delta}_\pi (\Lambda)) = \tilde{\Delta}_\pi (\Lambda) \subset \mathbb{Z}[[x,y]]$, so that $\iota (\tilde{\Delta}_\pi (\Lambda)) = (\iota \circ \tilde{\Delta}_\pi) (\Lambda) \subset \mathbb{Z}[[x,y]]$ and thus $\tilde{\Delta}_\pi (\Lambda) \subset \iota^{-1} (\mathbb{Z}[[x,y]])$.

But it so happens that $\iota^{-1} (\mathbb{Z}[[x,y]]) = \Lambda \otimes_{\mathbb{Z}} \Lambda$. Hence, $\tilde{\Delta}_\pi (\Lambda) \subset \iota^{-1} (\mathbb{Z}[[x,y]]) = \Lambda \otimes_{\mathbb{Z}} \Lambda$. This solves Exercise 2.80(b).

---

405Proof. It is clear that $\iota (\Lambda \otimes_{\mathbb{Z}} \Lambda) \subset \mathbb{Z}[[x,y]]$, so that $\Lambda \otimes_{\mathbb{Z}} \Lambda \subset \iota^{-1} (\mathbb{Z}[[x,y]])$. We are now going to prove the reverse inclusion $\iota^{-1} (\mathbb{Z}[[x,y]]) \subset \Lambda \otimes_{\mathbb{Z}} \Lambda$.

Indeed, let $p \in \iota^{-1} (\mathbb{Z}[[x,y]])$. Then, $p$ is an element of $\Lambda Q \otimes_{\mathbb{Q}} \Lambda Q$ satisfying $\iota (p) \in \mathbb{Z}[[x,y]]$. All coefficients of the power series $\iota (p)$ are integers (since $\iota (p) \in \mathbb{Z}[[x,y]]$).

We can write $p$ in the form $p = \sum_{(\lambda,\mu) \in \text{Par} \times \text{Par}} \rho_{\lambda,\mu} m_\lambda \otimes m_\mu$ with $\rho_{\lambda,\mu}$ being elements of $\mathbb{Q}$ (because $m_\lambda \otimes m_\mu$ is a $\mathbb{Q}$-basis of $\Lambda Q \otimes_{\mathbb{Q}} \Lambda Q$ (since $m_\lambda \lambda \in \text{Par}$ is a $\mathbb{Q}$-basis of $\Lambda Q$)). Consider these elements $\rho_{\lambda,\mu}$. Since $p = \sum_{(\lambda,\mu) \in \text{Par} \times \text{Par}} \rho_{\lambda,\mu} m_\lambda \otimes m_\mu$, we have $\iota (p) = \sum_{(\lambda,\mu) \in \text{Par} \times \text{Par}} \rho_{\lambda,\mu} m_\lambda \otimes m_\mu$ (by the definition of $\iota$). Therefore, for every $(\alpha, \beta) \in \text{Par} \times \text{Par}$, the rational number $\rho_{\alpha,\beta}$ is the coefficient of the power series $\iota (p)$ before the monomial $x^{\alpha} y^{\beta}$ (since there is clearly only one term in the sum $\sum_{(\lambda,\mu) \in \text{Par} \times \text{Par}} \rho_{\lambda,\mu} m_\lambda \otimes m_\mu$ contributing to this coefficient, namely the term $\rho_{\alpha,\beta} m_\lambda \otimes m_\mu$ for $(\lambda, \mu) = (\alpha, \beta)$, and this term contributes $\rho_{\alpha,\beta}$). So it is clear that $\rho_{\alpha,\beta}$ must be an integer (since all coefficients of the power series $\iota (p)$ are integers). So we have shown that $\rho_{\alpha,\beta}$ is an integer for every $(\alpha, \beta) \in \text{Par} \times \text{Par}$. In other words, $\rho_{\alpha,\beta}$ is an integer for every $(\lambda, \mu) \in \text{Par} \times \text{Par}$. Now, $p = \sum_{(\lambda,\mu) \in \text{Par} \times \text{Par}} \rho_{\lambda,\mu} m_\lambda \otimes m_\mu \in \Lambda \otimes_{\mathbb{Z}} \Lambda$.

Forget now that we fixed $p$. We thus have shown that $p \in \Lambda \otimes_{\mathbb{Z}} \Lambda$ for every $p \in \iota^{-1} (\mathbb{Z}[[x,y]])$. In other words, $\iota^{-1} (\mathbb{Z}[[x,y]]) \subset \Lambda \otimes_{\mathbb{Z}} \Lambda$. Combined with $\Lambda \otimes_{\mathbb{Z}} \Lambda \subset \iota^{-1} (\mathbb{Z}[[x,y]])$, this yields $\iota^{-1} (\mathbb{Z}[[x,y]]) = \Lambda \otimes_{\mathbb{Z}} \Lambda$, qed.
Second solution of Exercise 2.80(b): The following solution of Exercise 2.80(b) is a variation on the First one given above. It avoids the use of (11.309) in favor of working with finite variable sets. (This has the advantage that we no longer need to bother about technicalities; but more importantly, this approach will be very useful in solving Exercise 2.80(f) later on.)

Let \( N \in \mathbb{N} \) be arbitrary. We define a \( \mathbb{Q} \)-linear map

\[
\mathcal{E}_N : \Lambda_{\mathbb{Q}} \otimes \Lambda_{\mathbb{Q}} \to \mathbb{Q} \left[ x_1, x_2, ..., x_N, y_1, y_2, ..., y_N \right],
\]

\[
f \otimes g \mapsto f (x_1, x_2, ..., x_N) \ g (y_1, y_2, ..., y_N).
\]

This is well-defined because \( f (x_1, x_2, ..., x_N) \in \mathbb{Q} \left[ x_1, x_2, ..., x_N \right] \) and \( g (y_1, y_2, ..., y_N) \in \mathbb{Q} \left[ y_1, y_2, ..., y_N \right] \) are well-defined polynomials for every \( f \in \Lambda_{\mathbb{Q}} \) and \( g \in \Lambda_{\mathbb{Q}} \) (by Exercise 2.2) and depend linearly on \( f \) and \( g \), respectively. It is easy to see that the map \( \mathcal{E}_N \) is a \( \mathbb{Q} \)-algebra homomorphism.

Next, we define a map

\[
\mathcal{K}_N : \Lambda_{\mathbb{Q}} \to \mathbb{Q} \left[ x_1, x_2, ..., x_N, y_1, y_2, ..., y_N \right],
\]

\[
f \mapsto f \left( \left[ (x_i y_j) \right] \right) \in \mathbb{Q} \left[ x_1, x_2, ..., x_N, y_1, y_2, ..., y_N \right].
\]

Here, \( f \left( \left[ (x_i y_j) \right] \right) \) is defined as follows: Let \((u_1, u_2, ..., u_{N^2})\) be a list of all \( N^2 \) elements of the family \( \left( x_i y_j \right) \) in any arbitrary order, and set \( f \left( \left[ (x_i y_j) \right] \right) = f (u_1, u_2, ..., u_{N^2}). \) (The result does not depend on the order chosen, because \( f \) is symmetric.)

Again, \( \mathcal{K}_N \) is a \( \mathbb{Q} \)-algebra homomorphism (since \( \mathcal{K}_N \) is an evaluation map in an appropriate sense).

We now claim that

\[
(11.311) \quad \mathcal{E}_N \circ \Delta_X = \mathcal{K}_N.
\]

Proof of (11.311): The equality (11.311) is an equality between \( \mathbb{Q} \)-algebra homomorphisms (since \( \mathcal{K}_N \), \( \mathcal{E}_N \) and \( \Delta_X \) are \( \mathbb{Q} \)-algebra homomorphisms), and thus, in order to prove it, we only need to check that it holds on a generating set of the \( \mathbb{Q} \)-algebra \( \Lambda_{\mathbb{Q}} \). We do this on the generating set \( \{ p_n \}_{n \geq 1} \), by noticing that every \( n \geq 1 \) satisfies

\[
(\mathcal{E}_N \circ \Delta_X) (p_n) = \mathcal{E}_N \left( \Delta_X (p_n) \right) = \mathcal{E}_N (p_n \otimes p_n)
\]

\[
= p_n (x_1, x_2, ..., x_N) \ p_n (y_1, y_2, ..., y_N) \quad \text{(by the definition of } \mathcal{E}_N) \]

\[
= \sum_{i=1}^{N} x_i^n \ \sum_{j=1}^{N} y_j^n \quad \text{(by the definition of } \mathcal{E}_N) \]

\[
= \sum_{\left( (i,j) \right) \in \{1,2,\ldots,N\}^2} \ x_i y_j^n \quad \text{(by the definition of } \mathcal{K}_N \text{)}
\]

\[
= p_n \left( \left[ (x_i y_j) \right] \right) \quad \text{(by the definition of } \mathcal{K}_N \text{)}
\]

This proves (11.311).

Now,

\[
\mathcal{E}_N (\Delta_X (\Lambda)) = \mathcal{E}_N \circ \Delta_X (\Lambda) = \mathcal{K}_N (\Lambda) \subset \mathbb{Z} [x_1, x_2, ..., x_N, y_1, y_2, ..., y_N]
\]

(11.312)

(11.312) \( \mathcal{E}_N (p) \in \mathcal{E}_N (\Delta_X (\Lambda)) = \mathbb{Z} [x_1, x_2, ..., x_N, y_1, y_2, ..., y_N] \).
Now, forget that we fixed $N$. We thus have defined a $\mathbb{Q}$-algebra homomorphism $E_N : \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}} \rightarrow \mathbb{Q} [x_1, x_2, ..., x_N, y_1, y_2, ..., y_N]$ for every $N \in \mathbb{N}$, and we have shown that every $p \in \Delta_\times (\Lambda)$ satisfies (11.312) for every $N \in \mathbb{N}$.

Now, we are going to show that

(11.313) \[
\left( \mathcal{E}_N (p) \in \mathbb{Z} [x_1, x_2, ..., x_N, y_1, y_2, ..., y_N] \text{ for every } N \in \mathbb{N} \right),
\]

then $p \in \Lambda \otimes_{\mathbb{Z}} \Lambda$

Proof of (11.313): Let $p \in \Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$ be such that

(11.314) \[
\mathcal{E}_N (p) \in \mathbb{Z} [x_1, x_2, ..., x_N, y_1, y_2, ..., y_N] \text{ for every } N \in \mathbb{N} .
\]

We can write $p$ in the form $p = \sum_{(\lambda, \mu) \in \text{Par} \times \text{Par}} \rho_{\lambda, \mu} m_\lambda \otimes m_\mu$ with $\rho_{\lambda, \mu}$ being elements of $\mathbb{Q}$ (because $(m_\lambda \otimes m_\mu)_{(\lambda, \mu) \in \text{Par} \times \text{Par}}$ is a $\mathbb{Q}$-basis of $\Lambda_{\mathbb{Q}} \otimes_{\mathbb{Q}} \Lambda_{\mathbb{Q}}$ (since $(m_\lambda)_{\lambda \in \text{Par}}$ is a $\mathbb{Q}$-basis of $\Lambda_{\mathbb{Q}}$)). Consider these elements $\rho_{\lambda, \mu}$.

Now, let $(\alpha, \beta) \in \text{Par} \times \text{Par}$ be arbitrary. Choose some $N \in \mathbb{N}$ satisfying $N \geq \ell (\alpha)$ and $N \geq \ell (\beta)$ (such an $N$ clearly exists). Then, $\alpha = (\alpha_1, \alpha_2, ..., \alpha_N)$ and $\beta = (\beta_1, \beta_2, ..., \beta_N)$. Since $p = \sum_{(\lambda, \mu) \in \text{Par} \times \text{Par}} \rho_{\lambda, \mu} m_\lambda \otimes m_\mu$, we have

(11.315) \[
\mathcal{E}_N (p) = \sum_{(\lambda, \mu) \in \text{Par} \times \text{Par}} \rho_{\lambda, \mu} m_\lambda (x_1, x_2, ..., x_N) m_\mu (y_1, y_2, ..., y_N).
\]

The only addend on the right hand side of this equality which has a nonzero coefficient before $x_1^{\alpha_1} x_2^{\alpha_2} ... x_N^{\alpha_N} y_1^{\beta_1} y_2^{\beta_2} ... y_N^{\beta_N}$ is the addend for $(\lambda, \mu) = (\alpha, \beta)$ (since $\alpha$ and $\beta$ are partitions, so the only partition $\lambda$ such that the monomial $x_1^{\alpha_1} x_2^{\alpha_2} ... x_N^{\alpha_N}$ appears in $m_\lambda (x_1, x_2, ..., x_N)$ is $\alpha$, and the only partition $\mu$ such that the monomial $y_1^{\beta_1} y_2^{\beta_2} ... y_N^{\beta_N}$ appears in $m_\mu (y_1, y_2, ..., y_N)$ is $\beta$). This coefficient is $\rho_{\alpha, \beta}$. Hence, (11.315) shows that the coefficient before $x_1^{\alpha_1} x_2^{\alpha_2} ... x_N^{\alpha_N} y_1^{\beta_1} y_2^{\beta_2} ... y_N^{\beta_N}$ in the power series $\mathcal{E}_N (p)$ is $\rho_{\alpha, \beta}$. But this coefficient must be an integer (in fact, (11.314) shows that every coefficient of the power series $\mathcal{E}_N (p)$ is an integer). Thus, $\rho_{\alpha, \beta}$ is an integer.

So we have shown that $\rho_{\alpha, \beta}$ is an integer for every $(\alpha, \beta) \in \text{Par} \times \text{Par}$. In other words, $\rho_{\lambda, \mu}$ is an integer for every $(\lambda, \mu) \in \text{Par} \times \text{Par}$. Now, $p = \sum_{(\lambda, \mu) \in \text{Par} \times \text{Par}} \rho_{\lambda, \mu} m_\lambda \otimes m_\mu \in \Lambda \otimes_{\mathbb{Z}} \Lambda$. This proves (11.313).

Now, we are almost done. Let $p \in \Delta_\times (\Lambda)$. We know that $\mathcal{E}_N (p) \in \mathbb{Z} [x_1, x_2, ..., x_N, y_1, y_2, ..., y_N]$ for every $N \in \mathbb{N}$ (according to (11.312)). Hence, (11.313) yields that $p \in \Lambda \otimes_{\mathbb{Z}} \Lambda$. Since we have proven this for every $p \in \Delta_\times (\Lambda)$, we thus obtain $\Delta_\times (\Lambda) \subseteq \Lambda \otimes_{\mathbb{Z}} \Lambda$. This solves Exercise 2.80(b).

Third solution of Exercise 2.80(b): Here is another solution of Exercise 2.80(b), which entirely gets by without using substitutions. We are going to prove that

(11.316) \[
\Delta_\times (h_n) = \sum_{\lambda \vdash n} s_\lambda \otimes s_\lambda \quad \text{for every } n \in \mathbb{N}.
\]

Once this is proven, it will follow that $\Delta_\times (h_n) \in \Lambda \otimes_{\mathbb{Z}} \Lambda$ for every $n \in \mathbb{N}$, and therefore

\[
\Delta_\times \left( \begin{array}{c}
\frac{h_\lambda}{h_\lambda}
\end{array} \right) = \Delta_\times (h_{\lambda_1} h_{\lambda_2} h_{\lambda_3} ...)
\]

\[
= \Delta_\times (h_{\lambda_1}) \cdot \Delta_\times (h_{\lambda_2}) \cdot \Delta_\times (h_{\lambda_3}) \cdot \cdots \quad \text{(since $\Delta_\times$ is a $\mathbb{Q}$-algebra homomorphism)}
\]

\[
\in \Lambda \otimes_{\mathbb{Z}} \Lambda \cdot \Lambda \otimes_{\mathbb{Z}} \Lambda \cdot \Lambda \otimes_{\mathbb{Z}} \Lambda \cdot \cdots
\]

\[
\subseteq \Lambda \otimes_{\mathbb{Z}} \Lambda
\]

for every partition $\lambda$; and this will immediately yield that $\Delta_\times (\Lambda) \subseteq \Lambda \otimes_{\mathbb{Z}} \Lambda$ (because $\Delta_\times$ is $\mathbb{Z}$-linear, and because $(h_\lambda)_{\lambda \in \text{Par}}$ is a basis of the $\mathbb{Z}$-module $\Lambda$), which will solve Exercise 2.80(b). Hence, in order to solve Exercise 2.80(b), it is enough to prove (11.316).

\[\text{The notation } \lambda \vdash n \text{ here is a synonym for } \lambda \in \text{Par}_n.\]
We need to prove that \((11.316)\). In order to do so, it is clearly enough to show that
\[
\sum_{n \geq 0} \Delta_x (h_n) t^n = \sum_{n \geq 0} \left( \sum_{\lambda \vdash n} s_{\lambda} \otimes s_{\lambda} \right) t^n \quad \text{in} \quad (\Lambda_{\mathbb{Q}} \otimes \Lambda_{\mathbb{Q}}) [[t]]
\]

(because comparing coefficients in \((11.317)\) immediately yields \((11.316)\)).

Recall that for any two \(\mathbb{Q}\)-algebras \(A\) and \(B\), every \(\mathbb{Q}\)-algebra homomorphism \(\varphi : A \to B\) induces a continuous\(^407\) \(\mathbb{Q}[[t]]\)-algebra homomorphism \(\varphi[[t]] : A[[t]] \to B[[t]]\) which is given by
\[
(\varphi[[t]]) \left( \sum_{k \geq 0} a_k t^k \right) = \sum_{k \geq 0} \varphi(a_k) t^k \quad \text{for every} \quad (a_k)_{k \geq 0} \in A^\mathbb{N}.
\]

In particular, the \(\mathbb{Q}\)-algebra homomorphism \(\Delta_x : \Lambda_{\mathbb{Q}} \to \Lambda_{\mathbb{Q}} \otimes \Lambda_{\mathbb{Q}}\) induces a \(\mathbb{Q}[[t]]\)-algebra homomorphism \(\Delta_x[[t]] : \Lambda_{\mathbb{Q}}[[t]] \to (\Lambda_{\mathbb{Q}} \otimes \Lambda_{\mathbb{Q}})[[t]]\). Recall the power series \(H(t) \in \Lambda[[t]]\) defined in \((2.14)\); it satisfies \(H(t) = \sum_{n \geq 0} h_n t^n\). Applying the map \(\Delta_x[[t]]\) to both sides of this equality, we obtain
\[
(\Delta_x[[t]]) (H(t)) = (\Delta_x[[t]]) \left( \sum_{n \geq 0} h_n t^n \right) = \sum_{n \geq 0} \Delta_x(h_n) t^n
\]

(by the definition of \(\Delta_x[[t]]\)).

On the other hand, recall the \(\mathbb{Q}\)-algebra isomorphism \(\Lambda_{\mathbb{Q}} \otimes \Lambda_{\mathbb{Q}} \to R_{\mathbb{Q}}(\mathbb{x}, \mathbb{y})^{S(\infty) \times S(\infty)}\) constructed in \((2.2)\) (applied to \(k = \mathbb{Q}\)). This entails a \(\mathbb{Q}\)-algebra injection \(\Lambda_{\mathbb{Q}} \otimes \Lambda_{\mathbb{Q}} \to \mathbb{Q}[[\mathbb{x}, \mathbb{y}]\) (since \(R_{\mathbb{Q}}(\mathbb{x}, \mathbb{y})^{S(\infty) \times S(\infty)} \subset R_{\mathbb{Q}}(\mathbb{x}, \mathbb{y}) \subset \mathbb{Q}[[\mathbb{x}, \mathbb{y}]\)). Denote this injection by \(\iota\). Then, the \(\mathbb{Q}\)-algebra homomorphism \(\iota : \Lambda_{\mathbb{Q}} \otimes \Lambda_{\mathbb{Q}} \to \mathbb{Q}[[\mathbb{x}, \mathbb{y}]\) induces a \(\mathbb{Q}[[t]]\)-algebra homomorphism \(\iota[[t]] : (\Lambda_{\mathbb{Q}} \otimes \Lambda_{\mathbb{Q}})[[t]] \to \mathbb{Q}[[\mathbb{x}, \mathbb{y}]][[t]]\). This homomorphism \(\iota[[t]]\) is injective (since \(\iota\) is injective).

We need to prove \((11.317)\). For this, it is enough to prove
\[
\left(\iota[[t]]\right) \left( \sum_{n \geq 0} \Delta_x(h_n) t^n \right) = \left(\iota[[t]]\right) \left( \sum_{n \geq 0} \left( \sum_{\lambda \vdash n} s_{\lambda} \otimes s_{\lambda} \right) t^n \right)
\]

(because the injectivity of \(\iota[[t]]\) ensures that \((11.319)\) implies \((11.317)\)).

In \(\mathbb{Q}[[\mathbb{x}, \mathbb{y}]][[t]]\), we have
\[
\left(\iota[[t]]\right) \left( \sum_{n \geq 0} \left( \sum_{\lambda \vdash n} s_{\lambda} \otimes s_{\lambda} \right) t^n \right) = \sum_{n \geq 0} \iota \left( \sum_{\lambda \vdash n} s_{\lambda} \otimes s_{\lambda} \right) t^n = \sum_{n \geq 0} \sum_{\lambda \vdash n} s_{\lambda}(\mathbb{x}) s_{\lambda}(\mathbb{y}) t^n = \sum_{\lambda \in \text{Par}} s_{\lambda}(\mathbb{x}) \sum_{n \geq 0} \iota^{|\lambda|} s_{\lambda}(\mathbb{y}) (t y)^{|\lambda|}
\]

\((11.320)\)

\[
= \sum_{\lambda \in \text{Par}} s_{\lambda}(\mathbb{x}) \left( t y_1, t y_2, t y_3, ... \right) = \prod_{i,j=1}^{\infty} \left( 1 - x_i \cdot t y_j \right)^{-1}
\]

\((11.320)\) (applied to \( (t y_1, t y_2, t y_3, ... \) instead of \(\mathbb{y})\)

\((11.320)\) (applied to \( (t y_1, t y_2, t y_3, ... \) instead of \(\mathbb{y})\)

\[
\text{yields} \quad \prod_{i,j=1}^{\infty} \left( 1 - x_i \cdot t y_j \right)^{-1} = \sum_{\lambda \in \text{Par}} s_{\lambda}(\mathbb{x}) s_{\lambda}(t y_1, t y_2, t y_3, ...)
\]

Meanwhile, exponentiating both sides of the equality \((2.29)\) yields
\[
H(t) = \exp \left( \sum_{m=1}^{\infty} \frac{1}{m} p_m(\mathbb{x}) t^m \right),
\]

\(^{407}\)The word “continuous” refers to the usual topologies on the power series rings \(\mathcal{A}[[t]]\) and \(\mathcal{B}[[t]]\).
Applying the map \( \iota \) to both sides of this equality, we obtain
\[
\exp \left( \sum_{m=1}^{\infty} \frac{1}{m} p_m(x) t^m \right) = \exp \left( \sum_{m=1}^{\infty} \frac{1}{m} p_m(x) t^m \right)
\]
since \( \Delta_x \) is a continuous \( \mathbb{Q} \)-algebra homomorphism, and thus commutes with \( \exp \).

Compared with (11.318), this yields
\[
\sum_{n \geq 0} \Delta_x(h_n) t^n = \exp \left( \sum_{m=1}^{\infty} \frac{1}{m} p_m \otimes p_m t^m \right).
\]

But exponentiating both sides of the equality (2.31) yields
\[
\prod_{i,j=1}^{\infty} (1 - x_i y_j)^{-1} = \exp \left( \sum_{m=1}^{\infty} \frac{1}{m} p_m(x) p_m(y) \right).
\]

Substituting \((ty_1, ty_2, ty_3, \ldots)\) for \(y\) in this equality, we obtain
\[
\prod_{i,j=1}^{\infty} (1 - x_i \cdot ty_j)^{-1} = \exp \left( \sum_{m=1}^{\infty} \frac{1}{m} p_m(x) p_m(ty_1, ty_2, ty_3, \ldots) \right).
\]
Compared with (11.321), this yields
\[ (\iota \, \llbracket [t] \rrbracket) \left( \sum_{n \geq 0} \Delta_x (h_n) \, t^n \right) = \prod_{i,j=1}^{\infty} \left( 1 - x_i \cdot t y_j \right)^{-1}. \]

Compared with (11.320), this yields
\[ (\iota \, \llbracket [t] \rrbracket) \left( \sum_{n \geq 0} \Delta_x (h_n) \, t^n \right) = \left( \iota \, \llbracket [t] \rrbracket \right) \left( \sum_{n \geq 0} \left( \sum_{\lambda \vdash n} s_{\lambda} \otimes s_{\lambda} \right) \, t^n \right). \]

This proves (11.319). Thus, Exercise 2.80(b) is solved.

**Remark:** The three solutions we gave for Exercise 2.80(b) are not that different. The Second solution is a variation on the First solution which trades the use of a substitution of infinitely many variables (with the technical troubles that come along with it) for the inconvenience of having to consider a “sufficiently high $N \in \mathbb{N}$”. The Third solution looks like a different beast, but its main idea – the equality (11.316) – is really just an afterthought of the First solution. Indeed, knowing the equality (11.310) in the First solution, we can easily prove (11.316) as follows: Using the notations of the First solution, we have
\[ \sum_{n \geq 0} h_n \left( (x_i y_j)_{(i,j) \in \{1,2,3,\ldots\}^2} \right) \, t^n = \prod_{i,j=1}^{\infty} \left( 1 - t x_i y_j \right)^{-1} \quad \text{(by (2.14), evaluated on the variable set $xy$)} \]
\[ = \sum_{\lambda \in \text{Par}} t^{\lambda} s_{\lambda} (x) \, s_{\lambda} (y) \quad \text{(by (2.25))} \]
\[ = \sum_{n \in \mathbb{N}} \left( \sum_{\lambda \vdash n} s_{\lambda} (x) \, s_{\lambda} (y) \right) \, t^n \]
in $\mathbb{Q} \llbracket [x,y] \rrbracket \llbracket [t] \rrbracket$. Comparing coefficients before $t^n$ in this equality, we obtain
\[ h_n \left( (x_i y_j)_{(i,j) \in \{1,2,3,\ldots\}^2} \right) = \sum_{\lambda \vdash n} s_{\lambda} (x) \, s_{\lambda} (y) \quad \text{for every } n \in \mathbb{N}. \]

Thus, for every $n \in \mathbb{N}$, we have
\[ \iota \left( \Delta_x (h_n) \right) = \left( \iota \circ \Delta_x \right) (h_n) = \overline{\Delta_x} (h_n) = h_n \left( (x_i y_j)_{(i,j) \in \{1,2,3,\ldots\}^2} \right) = \sum_{\lambda \vdash n} s_{\lambda} (x) \, s_{\lambda} (y) \quad \text{(by (11.310))} \]
\[ = \iota \left( \sum_{\lambda \vdash n} s_{\lambda} \otimes s_{\lambda} \right), \]

which (by the injectivity of $\iota$) yields $\Delta_x (h_n) = \sum_{\lambda \vdash n} s_{\lambda} \otimes s_{\lambda}$. Thus, (11.316) is proven again.

In a similar vein, one can show that
\[ \Delta_x (e_n) = \sum_{\lambda \vdash n} s_{\lambda} \otimes s_{\lambda}, \quad \text{for every } n \in \mathbb{N}. \]

(One would need to use the dual Cauchy identity, i.e., Exercise 2.70(a), instead of (2.25) this time.)

(c) We are going to show that
\[ \epsilon_r (h_n) = (-1)^n \binom{-r}{n} \quad \text{for every } n \in \mathbb{N}. \]

(11.322) This will yield the statement of Exercise 2.80(c) in the same way as (11.316) yielded the statement of Exercise 2.80(b) in the Third solution of Exercise 2.80(b) above. Hence, we only need to prove (11.322) in order to be done with Exercise 2.80(c).
Recall that for any two \( \mathbb{Q} \)-algebras \( \mathcal{A} \) and \( \mathcal{B} \), every \( \mathbb{Q} \)-algebra homomorphism \( \varphi : \mathcal{A} \to \mathcal{B} \) induces a continuous\(^{408}\) \( \mathbb{Q}[[t]] \)-algebra homomorphism \( \varphi[[t]] : \mathcal{A}[[t]] \to \mathcal{B}[[t]] \) which is given by

\[
(\varphi[[t]]) \left( \sum_{k \geq 0} a_k t^k \right) = \sum_{k \geq 0} \varphi(a_k) t^k \quad \text{for every } (a_k)_{k \geq 0} \in \mathcal{A}_t.
\]

Hence, the \( \mathbb{Q} \)-algebra homomorphism \( \epsilon_r : \Lambda \mathbb{Q} \to \mathbb{Q} \) induces a continuous \( \mathbb{Q}[[t]] \)-algebra homomorphism \( \epsilon_r[[t]] : \Lambda \mathbb{Q}[[t]] \to \mathbb{Q}[[t]] \).

Recall the power series \( H(t) \in \Lambda[[t]] \) defined in (2.14); it satisfies \( H(t) = \sum_{n \geq 0} h_n t^n \). Applying the homomorphism \( \epsilon_r[[t]] \) to both sides of this equality, we obtain

\[
(11.323) \quad (\epsilon_r[[t]])(H(t)) = (\epsilon_r[[t]])\left( \sum_{n \geq 0} h_n t^n \right) = \sum_{n \geq 0} \epsilon_r(h_n) t^n.
\]

On the other hand, exponentiating both sides of the equality (2.29) yields

\[
H(t) = \exp \left( \sum_{m=1}^{\infty} \frac{1}{m} p_m(x) t^m \right).
\]

Applying the map \( \epsilon_r[[t]] \) to both sides of this equality, we obtain

\[
(\epsilon_r[[t]])(H(t)) = \left( \exp \left( \sum_{m=1}^{\infty} \frac{1}{m} p_m(x) t^m \right) \right) = \exp \left( \sum_{m=1}^{\infty} \frac{1}{m} \epsilon_r(p_m) t^m \right) \quad \text{(by the definition of } \epsilon_r)\)
\]

\[
= \exp \left( \sum_{m=1}^{\infty} \frac{1}{m} \epsilon_r(p_m) t^m \right) = \exp \left( \sum_{m=1}^{\infty} \frac{1}{m} \epsilon_r(p_m) t^m \right) = \exp \left( \sum_{m=1}^{\infty} \frac{1}{m} \epsilon_r(p_m) t^m \right) \quad \text{(by Newton's binomial formula)}.
\]

Comparing this with (11.323), we obtain

\[
\sum_{n \geq 0} \epsilon_r(h_n) t^n = \sum_{n \geq 0} (-1)^n \binom{-r}{n} t^n.
\]

Comparing coefficients in this equality of power series, we see that \( \epsilon_r(h_n) = (-1)^n \binom{-r}{n} \) for every \( n \in \mathbb{N} \). Thus, (11.322) is proven, and so Exercise 2.80(c) is solved.

\(^{408}\)The word “continuous” refers to the usual topologies on the power series rings \( \mathcal{A}[[t]] \) and \( \mathcal{B}[[t]] \).
(d) Consider the \(\mathbb{Q}\)-algebra homomorphism \(\epsilon_r\) defined in Exercise 2.80(c), and the \(\mathbb{Q}\)-algebra homomorphism \(\Delta_\lambda\) defined in Exercise 2.80(b). We know that both of these maps \(\epsilon_r\) and \(\Delta_\lambda\) are \(\mathbb{Q}\)-algebra homomorphisms. In particular, \(\text{id} : \Lambda_\mathbb{Q} \to \Lambda_\mathbb{Q}\) and \(\epsilon_r : \Lambda_\mathbb{Q} \to \mathbb{Q}\) are \(\mathbb{Q}\)-algebra homomorphisms. Thus, \(\text{id} \otimes \epsilon_r : \Lambda_\mathbb{Q} \otimes \mathbb{Q} \Lambda_\mathbb{Q} \to \Lambda_\mathbb{Q} \otimes \mathbb{Q} \mathbb{Q}\) is a \(\mathbb{Q}\)-algebra homomorphism (by Exercise 1.13(a)). Also, let \(\text{can} : \Lambda_\mathbb{Q} \otimes \mathbb{Q} \mathbb{Q} \to \Lambda_\mathbb{Q}\) be the canonical \(\mathbb{Q}\)-vector space isomorphism sending every \(f \otimes \alpha \in \Lambda_\mathbb{Q} \otimes \mathbb{Q} \mathbb{Q}\) to \(\alpha f \in \Lambda_\mathbb{Q}\).

Now, we claim that

\[
\text{can} \circ (\text{id} \otimes \epsilon_r) \circ \Delta_\lambda = id_r.
\]

Proof of (11.324): The equality (11.324) is an equality between \(\mathbb{Q}\)-algebra homomorphisms (since \(\text{can}, \text{id} \otimes \epsilon_r, \Delta_\lambda\) and \(id_r\) are \(\mathbb{Q}\)-algebra homomorphisms). Consequently, in order to prove it, we only need to check that it holds on a generating set of the \(\mathbb{Q}\)-algebra \(\Lambda_\mathbb{Q}\). We do this on the generating set \((p_n)_{n \geq 1}\), by noticing that every \(n \geq 1\) satisfies

\[
\text{can} \circ (\text{id} \otimes \epsilon_r) \circ \Delta_\lambda(p_n) = \text{can} \left(\left(\text{id} \otimes \epsilon_r\right) \left(\Delta_\lambda(p_n)\right)\right) = \text{can} \left(\left(\text{id} \otimes \epsilon_r\right)(p_n \otimes p_n)\right) = \text{can} (p_n \otimes \epsilon_r(p_n)) = \epsilon_r(p_n) = r p_n = i_r(p_n).
\]

This proves (11.324).

Now, \(\text{can} \circ (\text{id} \otimes \epsilon_r) \circ \Delta_\lambda(\Lambda) = i_r(\Lambda)\), so that

\[
i_r(\Lambda) = (\text{can} \circ (\text{id} \otimes \epsilon_r) \circ \Delta_\lambda)(\Lambda) = \text{can} \left(\left(\text{id} \otimes \epsilon_r\right) \left(\Delta_\lambda(\Lambda)\right)\right) \subset \text{can} \left(\left(\text{id} \otimes \epsilon_r\right)(\Lambda \otimes \mathbb{Z}\Lambda)\right) = \text{can} \left(\Lambda \otimes \mathbb{Z}\epsilon_r(\Lambda)\right) \subset \text{can}(\Lambda \otimes \mathbb{Z}) = \mathbb{Z} \cdot \Lambda \quad \text{(by the definition of can)}
\]

\[
= \Lambda.
\]

This solves Exercise 2.80(d).

(e) Consider the \(\mathbb{Q}\)-algebra homomorphism \(\Delta_\lambda\) defined in Exercise 2.80(b). We have \(\Delta_\lambda(p_n) = p_n \otimes p_n\) for every positive integer \(n\). Now, it is easy to see that

\[
\Delta_\lambda(p_\lambda) = p_\lambda \otimes p_\lambda \quad \text{for every partition } \lambda.
\]

Now, consider the multiplication map \(m_{\Lambda_\mathbb{Q}} : \Lambda_\mathbb{Q} \otimes \Lambda_\mathbb{Q} \to \Lambda_\mathbb{Q}\) of the \(\mathbb{Q}\)-algebra \(\Lambda_\mathbb{Q}\). We claim that

\[
\text{Sq} = m_{\Lambda_\mathbb{Q}} \circ \Delta_\lambda.
\]

\[\text{Proof of (11.325): Let } \lambda \text{ be a partition. Write } \lambda \text{ in the form } \lambda = (\lambda_1, \lambda_2, ..., \lambda_\ell) \text{ with } \ell = \ell (\lambda). \text{ Then, the definition of } p_\lambda \text{ yields } p_\lambda = p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_\ell}. \text{ Applying the map } \Delta_\lambda \text{ to both sides of this equality, we obtain }
\]

\[
\Delta_\lambda (p_\lambda) = \Delta_\lambda (p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_\ell}) = \Delta_\lambda (p_{\lambda_1}) \cdot \Delta_\lambda (p_{\lambda_2}) \cdots \cdot \Delta_\lambda (p_{\lambda_\ell}) \quad \text{(since } \Delta_\lambda \text{ is a } \mathbb{Q}\text{-algebra homomorphism)}
\]

\[
= (p_{\lambda_1} \otimes p_{\lambda_1}) \cdot (p_{\lambda_2} \otimes p_{\lambda_2}) \cdots \cdot (p_{\lambda_\ell} \otimes p_{\lambda_\ell}) \quad \text{(since } \Delta_\lambda (p_n) = p_n \otimes p_n \text{ for every positive integer } n)
\]

\[
= (p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_\ell}) \otimes (p_{\lambda_1} p_{\lambda_2} ... p_{\lambda_\ell}) = p_\lambda \otimes p_\lambda,
\]

which proves (11.325).
Indeed, every partition \( \lambda \) satisfies

\[
\text{Sq} (p\lambda) = p_\lambda^2 = p\lambda p\lambda = m_{\Lambda_Q} \left( p_\lambda \otimes p_\lambda \right) = m_{\Lambda_Q} \left( \Delta_x (p\lambda) \right) = \left( m_{\Lambda_Q} \circ \Delta_x \right) (p\lambda) \quad (\text{by (11.325)}).
\]

Since \( (p_\lambda)_{\lambda \in \text{Par}} \) is a \( \mathbb{Q} \)-basis of \( \Lambda_Q \) (and since \( \text{Sq} \) and \( m_{\Lambda_Q} \circ \Delta_x \) are \( \mathbb{Q} \)-linear maps), this shows that \( \text{Sq} = m_{\Lambda_Q} \circ \Delta_x \). Hence,

\[
\text{Sq} (\Lambda) = \left( m_{\Lambda_Q} \circ \Delta_x \right) (\Lambda) = m_{\Lambda_Q} \left( \Delta_x (\Lambda) \right) \subset m_{\Lambda_Q} (\Lambda \otimes \Lambda) \quad (\text{by Exercise 2.80(b)})
\]

\[
= \Lambda \cdot \Lambda \quad (\text{by the definition of } m_{\Lambda_Q})
\]

\[
= \Lambda.
\]

This solves Exercise 2.80(e).

(f) For every \( N \in \mathbb{N} \), we define a \( \mathbb{Q} \)-algebra homomorphism \( \mathcal{E}_N : \Lambda_Q \otimes \Lambda_Q \rightarrow \mathbb{Q} [x_1, x_2, ..., x_N, y_1, y_2, ..., y_N] \) as it was done in our Second solution of Exercise 2.80(b).

Recall the definition of the maps \( i_r \) in Exercise 2.80(d). We are first going to show that every \( a \in \mathbb{Z} \) and \( b \in \mathbb{Z} \) satisfy

\[
\Delta_a = \Delta_b * \left( \Delta_{\Lambda_Q} \circ i_{a-b} \right) \quad \text{in } \text{Hom} (\Lambda_Q, \Lambda_Q \otimes \Lambda_Q)
\]

(11.326)

(11.326) \( \Delta_a = \Delta_b * (\Delta_{\Lambda_Q} \circ i_{a-b}) \) in \( \text{Hom} (\Lambda_Q, \Lambda_Q \otimes \Lambda_Q) \)

(where \( \Delta_{\Lambda_Q} : \Lambda_Q \rightarrow \Lambda_Q \otimes \Lambda_Q \) is the usual comultiplication of \( \Lambda_Q \)).

\textit{Proof of (11.326):} Let \( a \in \mathbb{Z} \) and \( b \in \mathbb{Z} \). We know that \( \Delta_a \) and \( \Delta_b \) are \( \mathbb{Q} \)-algebra homomorphisms. Also, \( \Delta_{\Lambda_Q} \) is a \( \mathbb{Q} \)-algebra homomorphism (by the axioms of a bialgebra, which we know are satisfied for \( \Lambda_Q \)), and \( i_{a-b} \) is a \( \mathbb{Q} \)-algebra homomorphism. Hence, the composition \( \Delta_{\Lambda_Q} \circ i_{a-b} \) is a \( \mathbb{Q} \)-algebra homomorphism. Thus, the convolution \( \Delta_b * (\Delta_{\Lambda_Q} \circ i_{a-b}) \) is a \( \mathbb{Q} \)-algebra homomorphism (by Exercise 1.63(a), applied to \( \mathbb{Q}, \Lambda_Q, \Lambda_Q \otimes \Lambda_Q, \Delta_b \) and \( \Delta_{\Lambda_Q} \circ i_{a-b} \) instead of \( \mathbb{K}, H, A, f \) and \( g \)). Hence, (11.326) is an equality between \( \mathbb{Q} \)-algebra homomorphisms. Therefore, in order to prove it, we only need to check that it holds on a generating
set of the \( \mathbb{Q} \)-algebra \( \Lambda_Q \). We do this on the generating set \( (p_n)_{n \geq 1} \), by noticing that every \( n \geq 1 \) satisfies

\[
\left( \Delta_b \ast \left( \Delta_{\Lambda_Q} \circ i_{a-b} \right) \right) (p_n)
= m_{\Lambda_Q} \circ (\Delta_b \otimes (\Delta_{\Lambda_Q} \circ i_{a-b}) \circ \Delta_{\Lambda_Q}) (p_n)

(by the definition of convolution)
= m_{\Lambda_Q} \left( \Delta_b \otimes \left( \Delta_{\Lambda_Q} \circ i_{a-b} \right) \right) \left( \frac{\Delta_{\Lambda_Q} (p_n)}{= 1 \otimes p_n + p_n \otimes 1} \right)
= m_{\Lambda_Q} \left( \Delta_b (1) \otimes \left( \Delta_{\Lambda_Q} \circ i_{a-b} \right) (p_n) + \Delta_b (p_n) \otimes \left( \Delta_{\Lambda_Q} \circ i_{a-b} \right) (1) \right)
= \Delta_{\Lambda_Q} (1) \cdot \left( \Delta_{\Lambda_Q} \circ i_{a-b} \right) (p_n) + \Delta_b (p_n) \cdot \left( \Delta_{\Lambda_Q} \circ i_{a-b} \right) (1)

(since \( \Delta_{\Lambda_Q} \circ i_{a-b} \) is a \( \mathbb{Q} \)-algebra homomorphism)

\[
= (1 \otimes 1) \cdot \left( \Delta_{\Lambda_Q} \circ i_{a-b} (p_n) \right) + \Delta_b (p_n) \cdot (1 \otimes 1) = \Delta_{\Lambda_Q} \left( \frac{i_{a-b} (p_n)}{= (a-b)p_n} \right) + \frac{\Delta_b (p_n)}{= \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + b \otimes p_n + p_n \otimes b}
\]

(by the definition of \( i_{a-b} \))

\[
= \Delta_{\Lambda_Q} ((a-b) \cdot p_n) + \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + b \otimes p_n + p_n \otimes b
= (a-b) \cdot \Delta_{\Lambda_Q} (p_n) + \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + b \otimes p_n + p_n \otimes b
\]

(by the definition of \( \Delta_{\Lambda_Q} \))

\[
= (a-b) \cdot (1 \otimes p_n + p_n \otimes 1) + \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + b \otimes p_n + p_n \otimes b
= \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + b \otimes p_n + p_n \otimes b + (a-b) \cdot (1 \otimes p_n + p_n \otimes 1)
\]

\[
= \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + b \otimes p_n + p_n \otimes b + (a-b) \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i}
\]

\[
= \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + a \otimes p_n + p_n \otimes a = \Delta_a (p_n)
\]

(since the definition of \( \Delta_a \) yields \( \Delta_a (p_n) = \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + a \otimes p_n + p_n \otimes a \)).

This proves \((11.326)\).

Before we move on, let us record a simple fact about convolution of maps. Namely, if \( k \) is a commutative ring, and \( C \) is a \( k \)-coalgebra, and \( A \) and \( A' \) are two \( k \)-algebras, and \( f \) and \( g \) are two \( k \)-linear maps \( C \to A \), and \( \alpha : A \to A' \) is a \( k \)-algebra homomorphism, then

\[(11.327) \quad \alpha \circ (f \ast g) = (\alpha \circ f) \ast (\alpha \circ g).\]
This is merely the particular case of (1.17) when $C^\prime = C$ and $\gamma = \text{id}$, and so requires no proof anymore. Now, let $N \in \mathbb{N}$. Our next goal is to show that

\begin{equation}
(11.328) \quad (\mathcal{E}_N \circ \Delta_N)(\Lambda) \subset \mathbb{Z}[x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N].
\end{equation}

**Proof of (11.328):** Let us define a map

$$
\mathcal{L}_N : \mathbb{A}_Q \rightarrow \mathbb{Q}[x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N],
$$

\[ f \mapsto f \left((x_i + y_j)_{(i,j)\in\{1,2,\ldots,N\}^2}\right). \]

Here, $f \left((x_i + y_j)_{(i,j)\in\{1,2,\ldots,N\}^2}\right)$ is defined as follows: Let $(u_1, u_2, \ldots, u_{N^2})$ be a list of all $N^2$ elements of the family $(x_i + y_j)_{(i,j)\in\{1,2,\ldots,N\}^2}$ in any arbitrary order, and set $f \left((x_i + y_j)_{(i,j)\in\{1,2,\ldots,N\}^2}\right) = f(u_1, u_2, \ldots, u_{N^2})$. (The result does not depend on the order chosen, because $f$ is symmetric.) The map $\mathcal{L}_N$ is a $\mathbb{Q}$-algebra homomorphism (since $\mathcal{L}_N$ is an evaluation map in an appropriate sense). For every positive integer $n$, we have

\[
\Delta_N(p_n) = \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + N \otimes p_n + p_n \otimes N \quad \text{(by the definition of } \Delta_N(p_n))
\]

\[
= \sum_{k=1}^{n-1} \binom{n}{k} p_k \otimes p_{n-k} + N \otimes p_n + p_n \otimes N
\]

and thus

\[
(\mathcal{E}_N \circ \Delta_N)(p_n)
\]

\[
= \mathcal{E}_N \left( \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + N \otimes p_n + p_n \otimes N \right)
\]

\[
= \sum_{k=1}^{n-1} \binom{n}{k} p_k \otimes p_{n-k} + N \otimes p_n + p_n \otimes N
\]

(by the definition of $\mathcal{E}_N$)

\begin{equation}
(11.329) \quad = \sum_{k=1}^{n-1} \binom{n}{k} \left( \sum_{i=1}^{N} x_i^k \right) \left( \sum_{j=1}^{N} y_j^{n-k} \right) + N \left( \sum_{j=1}^{N} y_j^n \right) + \left( \sum_{i=1}^{N} x_i^n \right) N,
\end{equation}
while at the same time

\[ L_N(p_n) = p_n \left( (x_i + y_j)_{(i,j) \in \{1,2,\ldots,N\}^2} \right) \]

(by the definition of \( L_N(p_n) \))

\[
= \sum_{(i,j) \in \{1,2,\ldots,N\}^2} (x_i + y_j)^n = \sum_{(i,j) \in \{1,2,\ldots,N\}^2} \binom{n}{k} x_i^k y_j^{n-k}
\]

\[
= \sum_{k=0}^{n} \binom{n}{k} \sum_{(i,j) \in \{1,2,\ldots,N\}^2} x_i^k y_j^{n-k} = \sum_{k=0}^{n} \binom{n}{k} \left( \sum_{i=1}^{N} x_i^k \right) \left( \sum_{j=1}^{N} y_j^{n-k} \right)
\]

\[
= \sum_{k=0}^{n} \binom{n}{k} \left( \sum_{i=1}^{N} x_i^k \right) \left( \sum_{j=1}^{N} y_j^{n-k} \right) + \sum_{k=0}^{n} \binom{n}{k} \left( \sum_{i=1}^{N} x_i^k \right) \left( \sum_{j=1}^{N} y_j^{n-k} \right)
\]

(11.330)

\[
= \sum_{k=1}^{n-1} \binom{n}{k} \left( \sum_{i=1}^{N} x_i^k \right) \left( \sum_{j=1}^{N} y_j^{n-k} \right) + \binom{n}{0} \left( \sum_{i=1}^{N} x_i^0 \right) \left( \sum_{j=1}^{N} y_j^{n} \right) + \binom{n}{0} \left( \sum_{i=1}^{N} x_i^n \right) \left( \sum_{j=1}^{N} y_j^0 \right)
\]

Comparing (11.329) with (11.330) reveals that \((E_N \circ \Delta_N)(p_n) = L_N(p_n)\) for every positive integer \(n\). In other words, the two maps \(E_N \circ \Delta_N\) and \(L_N\) are equal to each other on the generating set \((p_n)_{n \geq 1}\) of the \(\mathbb{Q}\)-algebra \(\Lambda_Q\). Since these two maps \(E_N \circ \Delta_N\) and \(L_N\) are \(\mathbb{Q}\)-algebra homomorphisms (since \(E_N, \Delta_N\) and \(L_N\) are \(\mathbb{Q}\)-algebra homomorphisms), this yields that these two maps must be identical, i.e., we have \(E_N \circ \Delta_N = L_N\). Hence,

\[
(E_N \circ \Delta_N)(\Lambda) = L_N(\Lambda) \subset \mathbb{Z}[x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N]
\]

(because the definition of \(L_N\) immediately shows that \(L_N(f) \in \mathbb{Z}[x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N]\) for every \(f \in \Lambda\). This proves (11.328).

Next, we are going to show that

(11.331)

\[
(E_N \circ \Delta_r)(\Lambda) \subset \mathbb{Z}[x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N].
\]

**Proof of (11.331):** Applying (11.326) to \(a = r\) and \(b = N\), we obtain

\[
\Delta_r = \Delta_N \ast (\Delta_{\Lambda_0} \circ \iota_{r-N}).
\]

Thus,

\[
E_N \circ \Delta_r = E_N \circ (\Delta_N \ast (\Delta_{\Lambda_0} \circ \iota_{r-N})) = (E_N \circ \Delta_N) \ast (E_N \circ \Delta_{\Lambda_0} \circ \iota_{r-N})
\]

(by (11.327), applied to \(\mathbb{Q}, \Lambda_Q, \Lambda_Q \otimes \Lambda_Q, \mathbb{Q}[x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N], \Delta_N, \Delta_{\Lambda_0} \circ \iota_{r-N}\) and \(E_N\) instead of \(k, C, A, A', f, g\) and \(\alpha\)). Thus,

\[
E_N \circ \Delta_r = (E_N \circ \Delta_N) \ast (E_N \circ \Delta_{\Lambda_0} \circ \iota_{r-N}) = m_{\mathbb{Q}[x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N]}((E_N \circ \Delta_N) \otimes (E_N \circ \Delta_{\Lambda_0} \circ \iota_{r-N})) \ast \Delta_{\Lambda_0}
\]
(by the definition of convolution), so that

\[
\frac{(E_N \circ \Delta_r)}{(E_N \circ \Delta_r)} \quad (A)
\]

\[
= m_{Q[x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N]} \circ ((E_N \circ \Delta_N) \circ (E_N \circ \Delta_r \circ i_{r-N})) \circ \Delta_Q
\]

\[
= m_{Q[x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N]} \left( (E_N \circ \Delta_N) \otimes (E_N \circ \Delta_Q \circ i_{r-N}) \right) \circ \Delta_Q \quad (A)
\]

\[
\subseteq m_{Q[x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N]} \left( (E_N \circ \Delta_N) \otimes (E_N \circ \Delta_Q \circ i_{r-N}) \right) \circ \Delta_Q \quad (A)
\]

\[
= (E_N \circ \Delta_N) (\lambda) \cdot \left( E_N \left( \Delta_Q \left( i_{r-N} (\lambda) \right) \right) \right)
\]

\[
\subseteq Z [x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N] \cdot E_N (\lambda) \quad (by (11.328))
\]

\[
\subseteq Z [x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N] \cdot E_N (\lambda) \quad (because the definition of \ E_N \ immediately \ yields \ E_N (f \otimes g) \subseteq Z [x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N] \ for \ any \ f, g \in \Lambda)
\]

This proves (11.331).

Now, forget that we fixed \ N. We thus have proven (11.331) to hold for every \ N \in \mathbb{N}.

Now, let \ p \in \Delta_r (\lambda). Then,

\[
E_N \left( \frac{p}{\in \Delta_r (\lambda)} \right) = (E_N \circ \Delta_r) (\lambda) \in Z [x_1, x_2, \ldots, x_N, y_1, y_2, \ldots, y_N] \quad (by (11.331))
\]

for every \ N \in \mathbb{N}. Hence, (11.313) yields that \ p \in \Lambda \otimes Z \lambda. Since we have shown this for every \ p \in \Delta_r (\lambda), we thus conclude that \ \Delta_r (\lambda) \subseteq \Lambda \otimes Z \lambda. This solves Exercise 2.80(f).

Remark. Here is a rough sketch of an alternative way to conclude this solution of Exercise 2.80(f) after proving (11.328). This is closer to Richard Stanley’s suggested solution than the above.

The family \ (m_\alpha \otimes m_\beta)_{\lambda \in \Lambda} \subseteq \Lambda \times \Lambda \] is a \ Q\)-basis of \ \Lambda \otimes Q \lambda \ (since \ (m_\lambda)_{\lambda \in \Lambda} \ is a \ Q\)-basis of \ \Lambda \). Let us refer to this basis as the \ monomial basis \ of \ \Lambda \otimes Q \lambda. Then, \ \Lambda \otimes \Lambda \ is the subset of \ \Lambda \otimes Q \lambda \ consisting of all elements whose coordinates with respect to the monomial basis all are integers.

We want to prove that \ \Delta_r (\lambda) \subseteq \Lambda \otimes \Lambda \ for every \ r \in \mathbb{Z}. In other words, we want to prove that \ \Delta_r (f) \in \Lambda \otimes \Lambda \ for every \ r \in \mathbb{Z} \ and \ f \in \Lambda. Let us fix \ f \in \Lambda, but not fix \ r. We need to show that \ \Delta_r (f) \in \Lambda \otimes \Lambda \ for every \ r \in \mathbb{Z}; in other words, we need to show that for every \ (\alpha, \beta) \in \Lambda, the \ m_\alpha \otimes m_\beta\-coordinate of \ \Delta_r (f) \ with respect to the monomial basis is an integer for every \ r \in \mathbb{Z}.\]
So let us fix \((\alpha, \beta) \in \text{Par}\). It is not hard to see that the \(m_{\alpha} \otimes m_{\beta}\)-coordinate of \(\Delta_r (f)\) with respect to the monomial basis is a polynomial in \(r\) with rational coefficients.\(^{410}\) We want to prove that this polynomial is integer-valued (i.e., all its values at integer inputs are integers\(^{411}\)). To do so, it suffices to show that its values are integers at all sufficiently high \(r \in \mathbb{Z}\) (because for a polynomial with rational coefficients, the non-integer values appear periodically\(^{412}\), and therefore if no non-integer values appear from a given integer onwards, then the polynomial is integer-valued). In other words, it suffices to show that the \(m_{\alpha} \otimes m_{\beta}\)-coordinate of \(\Delta_r (f)\) with respect to the monomial basis is an integer for all sufficiently high \(r \in \mathbb{Z}\).

Let us prove this now. Our interpretation of “sufficiently high” will be that \(r \geq \ell (\alpha)\) and \(r \geq \ell (\beta)\). So what we need to prove is that the \(m_{\alpha} \otimes m_{\beta}\)-coordinate of \(\Delta_r (f)\) with respect to the monomial basis is an integer whenever \(r\) is an integer satisfying \(r \geq \ell (\alpha)\) and \(r \geq \ell (\beta)\).

Consider such an \(r\). Set \(N = r\) and \(p = \Delta_r (f)\). Then, \(N = r \geq \ell (\alpha)\) and \(N = r \geq \ell (\beta)\) and \(\Delta_N (f) = \Delta_r (f) = p\).

Then,

\[
\mathcal{E}_N \left( \frac{p}{\Delta_N (f)} \right) = \mathcal{E}_N (\Delta_N (f)) = \left( \mathcal{E}_N \circ \Delta_N \right) \left( \frac{f}{\epsilon \in \Lambda} \right) \in \left( \mathcal{E}_N \circ \Delta_N \right) (\Lambda) \subset \mathbb{Z} [x_1, x_2, ..., x_N, y_1, y_2, ..., y_N]
\]

(by \((11.328)\)). Therefore, arguing precisely as in the proof of \((11.313)\) (but using our specific \(\alpha\) and \(\beta\) rather than arbitrary \(\alpha\) and \(\beta\) as in that proof), we can show that \(\rho_{\alpha, \beta}\) is an integer, where \(p\) is written in the form \(p = \sum_{\lambda, \mu} m_{\lambda} \otimes m_{\mu}\) with \(\rho_{\lambda, \mu}\) being elements of \(\mathbb{Q}\). But \(\rho_{\alpha, \beta}\) is precisely the \(m_{\alpha} \otimes m_{\beta}\)-coordinate of \(\Delta_r (f)\) with respect to the monomial basis. Thus, we have shown that the \(m_{\alpha} \otimes m_{\beta}\)-coordinate of \(\Delta_r (f)\) with respect to the monomial basis is an integer (for \(r \in \mathbb{Z}\) satisfying \(r \geq \ell (\alpha)\) and \(r \geq \ell (\beta)\)). As we know, this completes our solution of Exercise 2.80(f) again.

\(g\) Let us first show that the \(\mathbb{Q}\)-module \(\Lambda_{Q}\), endowed with the comultiplication \(\Delta_{\times}\) and the counit \(\epsilon_{1}\), becomes a \(\mathbb{Q}\)-coalgebra.

In order to do so, we must verifying that the diagrams (1.3) and (1.4), with \(C, \Delta, \epsilon\) replaced by \(\Lambda_{Q}, \Delta_{\times}\) and \(\epsilon_{1}\), commute. We will only do this for the diagram (1.3), while leaving the diagram (1.4) to the reader.

So we must check that the diagram (1.3), with \(C, \Delta, \epsilon\) replaced by \(\Lambda_{Q}, \Delta_{\times}\) and \(\epsilon_{1}\), commutes. In other words, we must prove the identity

\[
\left( \Delta_{\times} \otimes \text{id} \right) \circ \Delta_{\times} = \left( \text{id} \otimes \Delta_{\times} \right) \circ \Delta_{\times}.
\]

Proof of \((11.324)\): The equality \((11.324)\) is an equality between \(\mathbb{Q}\)-algebra homomorphisms (since \(\Delta_{\times}\), \(\text{id} \otimes \Delta_{\times}\) and \(\Delta_{\times} \otimes \text{id}\) are \(\mathbb{Q}\)-algebra homomorphisms\(^{413}\)). Consequently, in order to prove it, we only need to check that it holds on a generating set of the \(\mathbb{Q}\)-algebra \(\Lambda_{Q}\).

But every \(n \geq 1\) satisfies

\[
\left( \left( \Delta_{\times} \otimes \text{id} \right) \circ \Delta_{\times} \right) (pn) = \left( \Delta_{\times} \otimes \text{id} \right) \left( \Delta_{\times} (pn) \right) = \left( \Delta_{\times} \otimes \text{id} \right) (p_n) \otimes p_n = p_n \otimes p_n \otimes p_n
\]

and similarly \(\left( \left( \text{id} \otimes \Delta_{\times} \right) \circ \Delta_{\times} \right) (pn) = p_n \otimes p_n \otimes p_n\). Hence, every \(n \geq 1\) satisfies \(\left( \left( \Delta_{\times} \otimes \text{id} \right) \circ \Delta_{\times} \right) (pn) = p_n \otimes p_n \otimes p_n\). Thus, we have checked that the equality \((11.332)\) holds on a generating set of the \(\mathbb{Q}\)-algebra \(\Lambda_{Q}\) (namely, on the generating set \((pn)_{n \geq 1}\)). As a consequence, the proof of \((11.332)\) is complete.

We have thus shown that the \(\mathbb{Q}\)-module \(\Lambda_{Q}\), endowed with the comultiplication \(\Delta_{\times}\) and the counit \(\epsilon_{1}\), becomes a \(\mathbb{Q}\)-coalgebra. This \(\mathbb{Q}\)-coalgebra becomes a \(\mathbb{Q}\)-bialgebra when combined with the existing \(\mathbb{Q}\)-algebra structure on \(\Lambda_{Q}\) (this is because \(\Delta_{\times}\) and \(\epsilon_{1}\) are \(\mathbb{Q}\)-algebra homomorphisms), and this \(\mathbb{Q}\)-bialgebra is cocommutative (this follows from the equality \(T \circ \Delta_{\times} = \Delta_{\times}\), where \(T : \Lambda_{Q} \otimes_{\mathbb{Q}} \Lambda_{Q} \rightarrow \Lambda_{Q} \otimes_{\mathbb{Q}} \Lambda_{Q}\) is the twist map; and this equality can be proven in the same way as we have showed \((11.332)\)). This solves Exercise 2.80(g).

\(^{410}\) This can be proven by noticing that it holds whenever \(f = p_n\) for some \(n \geq 0\) (by inspection of the definition of \(\Delta_{r}\)), and if it holds for two given values of \(f\) then it holds for any of their \(\mathbb{Q}\)-linear combinations and also for their product.

\(^{411}\) This is a weaker statement than saying that it has integer coefficients. (Actually, it does not in general have integer coefficients.)

\(^{412}\) To see this, just work modulo the common denominator of the coefficients of the polynomial.

\(^{413}\) Here, we are using Exercise 1.13(a) to see that \(\text{id} \otimes \Delta_{\times}\) and \(\Delta_{\times} \otimes \text{id}\) are \(\mathbb{Q}\)-algebra homomorphisms.
(h) Corollary 2.45(b) yields that \((p_\lambda)_{\lambda \in \text{Par}}\) and \((z^{-1}_\lambda p_\lambda)_{\lambda \in \text{Par}}\) are dual bases with respect to the Hall inner product on \(\Lambda\). In other words,
\[
(p_\lambda, z^{-1}_\mu p_\mu) = \delta_{\lambda, \mu} \quad \text{for any partitions } \lambda \text{ and } \mu.
\]
Hence,
\[
(11.333) \quad \left( \frac{p_\lambda}{z^{-1}_\mu p_\mu} \right) = \left( \frac{p_\lambda, z^{-1}_\mu p_\mu}{z^{-1}_\mu p_\mu} \right) = z^{-1}_\mu \delta_{\lambda, \mu}
\]
for any partitions \(\lambda\) and \(\mu\).

The Hall inner product \((\cdot, \cdot) : \Lambda Q \times \Lambda Q \to \mathbb{Q}\) is a bilinear form on \(\Lambda Q\). Hence, according to Definition 3.2(b) (below), this inner product induces a \(\mathbb{Q}\)-bilinear form \((\cdot, \cdot)_{\Lambda Q \otimes \Lambda Q} : (\Lambda Q \otimes \Lambda Q) \times (\Lambda Q \otimes \Lambda Q) \to \mathbb{Q}\). Similarly, the Hall inner product \((\cdot, \cdot) : \Lambda \times \Lambda \to \mathbb{Z}\) induces a \(\mathbb{Z}\)-bilinear form \((\cdot, \cdot)_{\Lambda \otimes \Lambda} : (\Lambda \otimes \mathbb{Z}) \times (\Lambda \otimes \mathbb{Z}) \to \mathbb{Z}\), and this latter \(\mathbb{Z}\)-bilinear form is clearly the restriction of the \(\mathbb{Q}\)-bilinear form \((\cdot, \cdot)_{\Lambda Q \otimes \Lambda Q}\) to \(\Lambda \otimes \mathbb{Z}\).

We shall now show that
\[
(11.334) \quad (a \ast b, c) = (a \otimes b, \Delta_X (c))_{\Lambda Q \otimes \Lambda Q} \quad \text{for all } a \in \Lambda Q, \ b \in \Lambda Q \text{ and } c \in \Lambda Q.
\]

**Proof of (11.334):** Let \(a \in \Lambda Q\), \(b \in \Lambda Q\) and \(c \in \Lambda Q\). The equality (11.334) is clearly \(\mathbb{Q}\)-linear in each of \(a\), \(b\) and \(c\). Hence, in proving this equality, we can WLOG assume that \(a\) and \(b\) and \(c\) are elements of the basis \((p_\lambda)_{\lambda \in \text{Par}}\) of the \(\mathbb{Q}\)-module \(\Lambda Q\). Assume this. Thus, \(a = p_\lambda\), \(b = p_\mu\) and \(c = p_\nu\) for some partitions \(\lambda\), \(\mu\) and \(\nu\); consider these partitions.

We have \(\delta_{\lambda, \mu} z_\lambda = \delta_{\lambda, \mu} z_\mu\) (in fact, the two sides of this equality are equal when \(\lambda = \mu\), and both vanish otherwise). Hence, \(a \ast b = p_\lambda \ast p_\mu = \delta_{\lambda, \mu} z_\mu p_\lambda = \delta_{\lambda, \mu} z_\mu p_\lambda\), so that
\[
\left( a \ast b \right) c = \left( \delta_{\lambda, \mu} z_\mu p_\lambda \right) c = \delta_{\lambda, \mu} z_\mu \lambda (p_\lambda, p_\nu) = \delta_{\lambda, \mu} z_\mu \lambda (p_\lambda, p_\nu) = \delta_{\lambda, \mu} z_\nu \lambda (p_\lambda, p_\nu) = \delta_{\lambda, \mu} z_\nu \lambda (p_\lambda, p_\nu).
\]

On the other hand, \(c = p_\nu\), so that
\[
\Delta_X (c) = \Delta_X (p_\nu) = p_\nu \otimes p_\nu \quad \text{(by (11.325), applied to } \nu \text{ instead of } \lambda),
\]
and thus
\[
\left( \frac{a \otimes b}{p_\lambda \otimes p_\mu} \right) \Delta_X (c) = (p_\lambda \otimes p_\mu \otimes p_\nu)_{\Lambda Q \otimes \Lambda Q} = (p_\lambda, p_\nu)_{\Lambda Q \otimes \Lambda Q} = (p_\mu, p_\nu)_{\Lambda Q \otimes \Lambda Q} = z_{\lambda} \delta_{\lambda, \mu} \lambda (p_\lambda, p_\nu) = z_{\lambda} \delta_{\lambda, \mu} \lambda (p_\mu, p_\nu).
\]

The equality in question, (11.334), thus rewrites as \(\delta_{\lambda, \mu} z_\lambda z_\nu \delta_{\lambda, \mu} = z_\nu \delta_{\lambda, \mu} z_\nu \delta_{\lambda, \mu}\) (because \((a \ast b, c) = \delta_{\lambda, \mu} z_\lambda z_\nu \delta_{\lambda, \mu}\) and \((a \otimes b, \Delta_X (c))_{\Lambda Q \otimes \Lambda Q} = z_\nu \delta_{\lambda, \mu} z_\nu \delta_{\lambda, \mu}\)). But the latter equality is obvious (because both of its sides are \(z_\nu^2\) if \(\lambda = \mu = \nu\), and vanish otherwise). Hence, (11.334) must hold as well.

Now that (11.334) is proven, let \(f \in \Lambda\) and \(g \in \Lambda\) be arbitrary. We can apply (11.308) to \(k = \mathbb{Q}\), \(A = \Lambda Q\), \(L = \text{Par}\), \((u_\lambda)_{\lambda \in \text{Par}} = (s_\lambda)_{\lambda \in \text{Par}}\), \((v_\lambda)_{\lambda \in \text{Par}} = (s_\lambda)_{\lambda \in \text{Par}}\) and \(f \ast g\) (because the basis \((s_\lambda)_{\lambda \in \text{Par}}\) of \(\Lambda Q\) is orthonormal with respect to the Hall inner product \((\cdot, \cdot)\), and thus dual to itself with respect to this product). As a result, we obtain
\[
f \ast g = \sum_{\lambda \in \text{Par}} (s_\lambda, f \ast g)_{\lambda \in \text{Par}} = \sum_{\lambda \in \text{Par}} (f \ast g, s_\lambda)_{\lambda \in \text{Par}} = \sum_{\lambda \in \text{Par}} (f \otimes g, \Delta_X (s_\lambda))_{\Lambda Q \otimes \Lambda Q} s_\lambda.
\]
But every $\lambda \in \text{Par}$ satisfies $\Delta_x \left( \sum_{\lambda \in \Lambda} s_{\lambda} \right) \in \Delta_x (\Lambda) \subseteq \Lambda \otimes_{\mathbb{Z}} \Lambda$ (by Exercise 2.80(b)) and thus

$$\left( f \otimes g, \Delta_x (s_{\lambda}) \right)_{\Lambda \otimes_{\mathbb{Z}} \Lambda} = \left( f \otimes g, \Delta_x (s_{\lambda}) \right)_{\Lambda \otimes_{\mathbb{Z}} \Lambda} \in \Lambda.$$ 

Hence,

$$f \ast g = \sum_{\lambda \in \text{Par}} \left( f \otimes g, \Delta_x (s_{\lambda}) \right)_{\Lambda \otimes_{\mathbb{Z}} \Lambda} s_{\lambda} \in \sum_{\lambda \in \text{Par}} \mathbb{Z} s_{\lambda} \subseteq \Lambda.$$

This solves Exercise 2.80(h).

(i) Define a map $U : \Lambda_{\mathbb{Q}} \to \mathbb{Q}$ by

$$U(f) = f(1) \quad \text{for all } f \in \Lambda_{\mathbb{Q}}.$$ 

Notice that $U(f) = f(1)$ is the result of substituting $1, 0, 0, 0, \ldots$ for $x_1, x_2, x_3, \ldots$ in $f$. Hence, $U$ is a $\mathbb{Q}$-algebra homomorphism.

We shall now show that $\epsilon_1 = U$.

For every integer $n \geq 1$, we have

$$U(p_n) = p_n(1) \quad \text{(by the definition of } U)$$

$$= (\text{the result of substituting } 1, 0, 0, 0, \ldots \text{ for } x_1, x_2, x_3, \ldots \text{ in } p_n)$$

$$= (\text{the result of substituting } 1, 0, 0, 0, \ldots \text{ for } x_1, x_2, x_3, \ldots \text{ in } x_1^n + x_2^n + x_3^n + x_4^n + \cdots)$$

$$= 1^n + 0^n + 0^n + 0^n + \cdots = 1$$

$$= \epsilon_1(p_n) \quad \text{(since } \epsilon_1(p_n) = 1 \text{ (by the definition of } \epsilon_1)).$$

In other words, the two $\mathbb{Q}$-algebra homomorphisms $U$ and $\epsilon_1$ are equal to each other on each element of the family $(p_n)_{n \geq 1}$. But since this family $(p_n)_{n \geq 1}$ is a generating set of the $\mathbb{Q}$-algebra $\Lambda_{\mathbb{Q}}$, this yields that the two $\mathbb{Q}$-algebra homomorphisms $U$ and $\epsilon_1$ must be identical (because if two $\mathbb{Q}$-algebra homomorphisms are equal to each other on each element of a generating set of their domain, then they must be identical). That is, $\epsilon_1 = U$. Hence, every $f \in \Lambda_{\mathbb{Q}}$ satisfies $\epsilon_1(f) = U(f) = f(1)$ (by the definition of $U$). This solves Exercise 2.80(i).

11.56. Solution to Exercise 2.82. Solution to Exercise 2.82. We are going to be brief; more detailed proofs for everything except of the (very easy) equivalence $\mathcal{D} \iff \mathcal{J}$ can be found at http://web.mit.edu/~darij/www/algebra/witt5f.pdf (along with generalizations and additional equivalences in the case of $A = \mathbb{Z}$).

As suggested by the hint, we first prove some elementary facts of number theory:

• Every positive integer $n$ satisfies

$$(11.335) \quad \sum_{d \mid n} \phi(d) = n.$$

**Proof:** Let $n$ be a positive integer. For every positive divisor $d$ of $n$, there is a bijection

$$\{ i \in \{1, 2, \ldots, n\} \mid \gcd(i, n) = d \} \to \left\{ j \in \left\{ 1, 2, \ldots, \frac{n}{d} \right\} \mid j \text{ is coprime to } \frac{n}{d} \right\},$$

$$i \mapsto \frac{i}{d},$$

Hence, for every positive divisor $d$ of $n$, we have

$$\left| \{ i \in \{1, 2, \ldots, n\} \mid \gcd(i, n) = d \} \right| = \left| \left\{ j \in \left\{ 1, 2, \ldots, \frac{n}{d} \right\} \mid j \text{ is coprime to } \frac{n}{d} \right\} \right|$$

$$(11.336) \quad = \phi\left(\frac{n}{d}\right) \quad \text{(because this is how } \phi\left(\frac{n}{d}\right) \text{ was defined}).$$
Now,
\[ n = |\{1, 2, \ldots, n\}| = \sum_{d \mid n} |\{i \in \{1, 2, \ldots, n\} \mid \gcd(i, n) = d\}| = \sum_{d \mid n} \phi \left( \frac{n}{d} \right) = \sum_{d \mid n} \phi (d) \]
\[ = \phi \left( \frac{n}{d} \right) \]
(here, we have substituted \( \frac{n}{d} \) in the sum). This proves (11.335).

- Every positive integer \( n \) satisfies
\[ (11.337) \]
\[ \sum_{d \mid n} \mu (d) = \delta_{n,1}. \]

**Proof:** Let \( n \) be a positive integer. Let \( n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k} \) be the prime factorization of \( n \), with all of \( a_1, a_2, \ldots, a_k \) being positive integers. Then, the squarefree positive divisors of \( n \) all have the form \( \prod_{i \in I} p_i \) for some subset \( I \) of \( \{1, 2, \ldots, k\} \). More precisely, there exists a bijection
\[ \{I \subset \{1, 2, \ldots, k\}\} \rightarrow (\text{the set of all squarefree positive divisors of } n), \]
\[ (11.338) \]
\[ I \mapsto \prod_{i \in I} p_i. \]

Now,
\[ \sum_{d \mid n} \mu (d) = \sum_{d \mid n: d \text{ is squarefree}} \mu (d) + \sum_{d \mid n: d \text{ is not squarefree}} \mu (d) = \sum_{d \mid n: d \text{ is squarefree}} \mu (d) \]
\[ = \sum_{I \subset \{1,2,\ldots,k\}} \mu (\prod_{i \in I} p_i) \]
\[ = (-1)^{|I|} \]
\[ = \sum_{I \subset \{1,2,\ldots,k\}} \]
\[ = \left\{ \begin{array}{ll} 1, & \text{if } \{1, 2, \ldots, k\} = \emptyset; \vspace{1mm} \text{if } k = 0; \vspace{1mm} \end{array} \right. \]
\[ = \left\{ \begin{array}{ll} 0, & \text{otherwise} \vspace{1mm} \end{array} \right. \]
\[ = \delta_{n,1}, \]
which proves (11.337).

- Every positive integer \( n \) satisfies
\[ (11.339) \]
\[ \sum_{d \mid n} \mu (d) \frac{n}{d} = \phi (n). \]

**Proof:** This is an elementary fact, but we have a hammer at our disposal, and this looks conspicuously like a nail. Let us define a \( \mathbb{Z} \)-coalgebra. Namely, let \( T \) be the free \( \mathbb{Z} \)-module with basis \( (t_n)_{n \geq 1} \). We define a \( \mathbb{Z} \)-coalgebra structure on \( T \) by setting
\[ \Delta (t_n) = \sum_{d,e \in \{1,2,3,\ldots\}; \quad d+e=n} t_d \otimes t_e = \sum_{d \mid n} t_d \otimes t_{n/d} \]
and \( \epsilon (t_n) = \delta_{n,1} \)
for every \( n \in \{1, 2, 3, \ldots\} \). It is straightforward to check that this makes \( T \) into a cocommutative coalgebra. Hence, \((\text{Hom}_\mathbb{Z}(T, \mathbb{Z}), \star)\) is a \( \mathbb{Z} \)-algebra with unity \( \epsilon \). This algebra \((\text{Hom}_\mathbb{Z}(T, \mathbb{Z}), \star)\) is commutative. We define four elements of \( \text{Hom}_\mathbb{Z}(T, \mathbb{Z}) \):
- a \( \mathbb{Z} \)-linear map \( \phi : T \rightarrow \mathbb{Z} \) which sends \( t_n \) to \( \phi (n) \) for every \( n \geq 1 \);
- a \( \mathbb{Z} \)-linear map \( \mu : T \rightarrow \mathbb{Z} \) which sends \( t_n \) to \( \mu (n) \) for every \( n \geq 1 \);

\[ \text{This is a particular case of the general fact that if } k \text{ is a commutative ring, if } C \text{ is a cocommutative } k \text{-coalgebra, and if } A \text{ is a commutative } k \text{-algebra, then the convolution algebra } (\text{Hom}_k(C,A), \star) \text{ is commutative. This is easy to check.} \]
– a $\mathbb{Z}$-linear map $\tilde{id} : T \to \mathbb{Z}$ which sends $t_n$ to $n$ for every $n \geq 1$;
– a $\mathbb{Z}$-linear map $\tilde{1} : T \to \mathbb{Z}$ which sends $t_n$ to 1 for every $n \geq 1$.

Then, the identity that we want to prove – i.e., the identity (11.339) – is equivalent to the claim that $\bar{\mu} * \tilde{id} = \tilde{\phi}$.\footnote{This is because the left hand side of (11.339) equals $\sum_{d|n} \mu(t_d) \tilde{id}(t_{n/d}) = \left( \bar{\mu} \otimes \tilde{id} \right) \left( \sum_{d|n} t_d \otimes t_{n/d} \right) = m \left( \left( \bar{\mu} \otimes \tilde{id} \right) (\Delta (t_n)) \right) = \bar{\mu} * \tilde{id} (t_n)$, whereas the right hand side equals $\tilde{\phi} (t_n)$.
}

Similarly, the (already proven) identity (11.335) is equivalent to $\tilde{\phi} * \tilde{1} = \tilde{id}$, and the (already proven) identity (11.337) is equivalent to $\bar{\mu} * \tilde{1} = \epsilon$. Thus,

$$\bar{\mu} * \tilde{id} = \mu * \tilde{\phi} * \tilde{1} = \mu * \tilde{1} * \tilde{\phi} \quad \text{(since } \left( \text{Hom}_\mathbb{Z} (T, \mathbb{Z}) , * \right) \text{ is commutative)}$$

As we know, this is equivalent to (11.339), so that (11.339) is proven.

It was not really necessary to phrase this argument in terms of coalgebras; this was only done to illustrate a use of the latter. Our proof can just as well be rewritten as a manipulation of sums, or (as a compromise between concreteness and structure) it can be paraphrased by using the Dirichlet convolution, which is the operation taking two maps $f, g : \{1, 2, 3, \ldots \} \to \mathbb{Z}$ to a third map $h : \{1, 2, 3, \ldots \} \to \mathbb{Z}$ defined by $h(n) = \sum_{d|n} f(d) g\left( \frac{n}{d} \right)$. Of course, this Dirichlet convolution is the same as our convolution $*$ on $\text{Hom}_\mathbb{Z} (T, \mathbb{Z})$, with the only difference that $\mathbb{Z}$-linear maps $T \to \mathbb{Z}$ are replaced by arbitrary maps $\{1, 2, 3, \ldots \} \to \mathbb{Z}$.

• Every positive integer $n$ satisfies

$$(11.340) \quad \sum_{d|n} d\mu(d) \phi\left( \frac{n}{d} \right) = \mu(n).$$

Proof: We use the setup we prepared in the proof of (11.339). Additionally, we define $\tilde{\mu}' : T \to \mathbb{Z}$ as the $\mathbb{Z}$-linear map which sends $t_n$ to $n \mu(n)$ for every $n \geq 1$. Then, $\tilde{\mu}' * \tilde{id} = \epsilon$, because every positive integer $n$ satisfies

$$\left( \tilde{\mu}' * \tilde{id} \right) (t_n) = \sum_{d|n} d\mu(d) \frac{n}{d} = \sum_{d|n} n \mu(d) = n \sum_{d|n} \mu(d) = n \delta_{n,1} = \delta_{n,1} = \epsilon(t_n).$$

But recall that $\tilde{\phi} = \bar{\mu} * \tilde{id} = \tilde{id} * \bar{\mu}$ (since $(\text{Hom}_\mathbb{Z} (T, \mathbb{Z}) , *)$ is commutative), and thus

$$\tilde{\mu}' * \tilde{id} = \tilde{id} * \bar{\mu} = \bar{\mu}$.

This is easily seen to be equivalent to (11.340), and so (11.340) is proven.

• If $k$ is a positive integer, and if $p \in \mathbb{Z}$, $a \in A$ and $b \in A$ are such that $a \equiv b \mod p^k A$, then

$$(11.341) \quad a^p \equiv b^p \mod p^{k+\ell} A \quad \text{for every } \ell \in \mathbb{N}.$$ 

Proof: Let $k$ be a positive integer. Let $p \in \mathbb{Z}$, $a \in A$ and $b \in A$ be such that $a \equiv b \mod p^k A$. We need to prove (11.341). It is clearly enough to show that $a^p \equiv b^p \mod p^{k+1} A$, because then (11.341) will follow by induction over $\ell$. But we have $a \equiv \bar{b} \mod pA$ (since $a \equiv b \mod p^k A$ and since $k$ is positive) and therefore

$$a^{p-1} + a^{p-2} b + \cdots + b^{p-1} \equiv b^{p-1} + b^{p-2} \bar{b} + \cdots + \bar{b}^{p-1} \equiv b^{p-1} + b^{p-2} b + \cdots + b^{p-1} = p b^{p-1} + \cdots + p b^{p-1} = p b^{p-1} \equiv 0 \mod p A,$$

for every $\ell \in \mathbb{N}$.\footnote{This is because the left hand side of (11.339) equals $\sum_{d|n} \bar{\mu}(t_d) \tilde{id}(t_{n/d}) = m \left( \bar{\mu} \otimes \tilde{id} \left( \sum_{d|n} t_d \otimes t_{n/d} \right) \right) = m \left( \left( \bar{\mu} \otimes \tilde{id} \right) (\Delta(t_n)) \right) = \bar{\mu} * \tilde{id}(t_n)$, whereas the right hand side equals $\tilde{\phi}(t_n)$.}
so that $a^{p-1} + \cdots + b^{p-1} \in pA$. Thus,

$$a^p - b^p = \frac{(a-b) \left( a^{p-1} + \cdots + b^{p-1} \right) \in (p^k A) (pA) = p^{k+1} A,}{\text{since } a \equiv b \mod p^k A}$$

so that $a^p \equiv b^p \mod p^{k+1} A$. This proves (11.341).

- Here is a slightly more useful corollary of (11.341): If a prime number $p$ and two elements $a$ and $b$ of $A$ are such that $a \equiv b \mod pA$, then

$$a^n \equiv b^n \mod p^{v_p(N)+1} A \quad \text{for every } n \in \{1,2,3,\ldots\}. \tag{11.342}$$

**Proof:** Let $p$ be a prime number, and let $a$ and $b$ be two elements of $A$ such that $a \equiv b \mod pA$. Let $N \in \{1,2,3,\ldots\}$. Write $N$ in the form $N = p^{v_p(N)} M$ for some positive integer $M$. Then, (11.341) (applied to $\ell = v_p(N)$ and $k = 1$) yields $a^{p^{v_p(N)}} \equiv b^{p^{v_p(N)}} \mod p^{v_p(N)+1} A$, and because of $N = p^{v_p(N)} M$ we have

$$a^N = a^{p^{v_p(N)} M} = \left( a^{p^{v_p(N)}} \right)^M \equiv \left( b^{p^{v_p(N)}} \right)^M \quad \text{(since } a^{p^{v_p(N)}} \equiv b^{p^{v_p(N)}} \mod p^{v_p(N)+1} A)$$

$$= b^{p^{v_p(N)} M} = b^N \mod p^{v_p(N)+1} A \quad \left( \text{since } p^{v_p(N)} M = N \right).$$

This proves (11.342).

We shall also use a fact from commutative algebra – namely, one of the versions of the Chinese Remainder Theorem of ring theory:

**Theorem 11.8.** Let $A$ be a commutative ring. Let $S$ be a finite set. For every $s \in S$, let $I_s$ be an ideal of $A$. Assume that the ideals $I_s$ of $A$ are comaximal 416; this means that every two distinct elements $s$ and $t$ of $S$ satisfy $I_s + I_t = A$. Then:

(a) We have

$$\bigcap_{s \in S} I_s = \prod_{s \in S} I_s.$$

(b) The canonical ring homomorphism

$$A/\left( \bigcap_{s \in S} I_s \right) \to \prod_{s \in S} (A/I_s), \quad a + \bigcap_{s \in S} I_s \mapsto (a + I_s)_{s \in S}$$

is well-defined and a ring isomorphism.

See http://stacks.math.columbia.edu/tag/00DT (Lemma 10.14.3 in the Stacks Project, as of 14 August 2015) or many other sources for a proof of Theorem 11.8. We shall only use part (a) of this theorem.

As a consequence of Theorem 11.8(a), we have the following:

- If $n$ is a positive integer and if $A$ is a commutative ring, then

$$\bigcap_{p \text{ is a prime factor of } n} p^{v_p(n)} A = \prod_{p \text{ is a prime factor of } n} \left( p^{v_p(n)} A \right). \tag{11.343}$$

Indeed, this follows from Theorem 11.8(a), since the ideals $p^{v_p(n)} A$ of $A$ (for varying $p$) are comaximal (because for any two distinct primes $p$ and $q$, we can find integers $x$ and $y$ satisfying $p^{v_p(n)} x + q^{v_q(n)} y = 1$, and therefore we have $p^{v_p(n)} A + q^{v_q(n)} A = A$).

Now, we can step to proving the actual equivalences.

**Proof of the implication $\mathcal{D} \implies \mathcal{F}$:** Assume that Assertion $\mathcal{D}$ holds. That is, there exists a family $(\alpha_n)_{n \geq 1} \in A^{1,2,3,\ldots}$ of elements of $A$ such that every positive integer $n$ satisfies $b_n = \sum_{d|n} d \alpha_d^{n/d}$. Consider this family $(\alpha_n)_{n \geq 1}$. Consider also the family $(w_n)_{n \geq 1} \in A^{1,2,3,\ldots}$ defined in Exercise 2.79(a). This family $(w_n)_{n \geq 1}$ is an algebraically independent generating set of $A^1$ (indeed, this is a restatement of Exercise 2.79(d)). Hence, there exists a unique $A$-algebra homomorphism $f : A \to A$ which satisfies

$$f(w_n) = \alpha_n \quad \text{for every } n \in \{1,2,3,\ldots\}.$$  

416 Some authors use the word “coprime” instead of “comaximal” here.
Consider this \( f \). We have \( p_n = \sum_{d|n} dw_d^{n/d} \) for every positive integer \( n \) (by Exercise 2.79(e)). Thus, for every positive integer \( n \), we have

\[
 f (p_n) = f \left( \sum_{d|n} dw_d^{n/d} \right) = \sum_{d|n} d \left( f (w_d) \right)^{n/d} \quad \text{(since \( f \) is a \( \mathbb{Z} \)-algebra homomorphism)}
\]

\[
 = \sum_{d|n} d \alpha_d^{n/d} = b_n.
\]

Hence, there exists a ring homomorphism \( \Lambda \to A \) which, for every positive integer \( n \), sends \( p_n \) to \( b_n \) (namely, \( f \)). That is, Assertion \( J \) holds, and the implication \( D \implies J \) is proven.

**Proof of the implication** \( J \implies D \): Assume that Assertion \( J \) holds. That is, there exists a ring homomorphism \( \Lambda \to A \) which, for every positive integer \( n \), sends \( p_n \) to \( b_n \). Let \( f \) be such a homomorphism. Consider the family \( (w_n)_{n \geq 1} \in A_{\Lambda}^{(1,2,3,...)} \) defined in Exercise 2.79(a). For every positive integer \( n \), we have \( p_n = \sum_{d|n} dw_d^{n/d} \) (by Exercise 2.79(e)) and thus

\[
 f (p_n) = f \left( \sum_{d|n} dw_d^{n/d} \right) = \sum_{d|n} d (f (w_d))^{n/d} \quad \text{(since \( f \) is a ring homomorphism)}.
\]

Since \( f (p_n) = b_n \) (by the definition of \( f \)), this rewrites as \( b_n = \sum_{d|n} d (f (w_d))^{n/d} \). Thus, there exists a family \( (\alpha_n)_{n \geq 1} \in A^{(1,2,3,...)} \) of elements of \( A \) such that every positive integer \( n \) satisfies \( b_n = \sum_{d|n} d \alpha_d^{n/d} \) (namely, such a family can be defined by \( \alpha_n = f (w_n) \)). Assertion \( D \) thus holds. We have now proven the implication \( J \implies D \).

**Proof of the implication** \( D \implies C \): Assume that Assertion \( D \) holds. That is, there exists a family \( (\alpha_n)_{n \geq 1} \in A^{(1,2,3,...)} \) of elements of \( A \) such that every positive integer \( n \) satisfies \( b_n = \sum_{d|n} d \alpha_d^{n/d} \). Consider this family \( (\alpha_n)_{n \geq 1} \).

We need to prove that Assertion \( C \) holds, i.e., that we have

\[
(11.344) \quad \varphi_p \left( b_{n/p} \right) \equiv b_n \mod p^{v_p(n)} A
\]

for every positive integer \( n \) and every prime factor \( p \) of \( n \). So let us fix a positive integer \( n \) and a prime factor \( p \) of \( n \). We need to prove \( (11.344) \).

The definition of the family \( (\alpha_n)_{n \geq 1} \) shows that \( b_{n/p} = \sum_{d|n/p} d \alpha_d^{(n/p)/d} \), so that

\[
\varphi_p \left( b_{n/p} \right) = \varphi_p \left( \sum_{d|n/p} d \alpha_d^{(n/p)/d} \right) = \sum_{d|n/p} d (\varphi_p (\alpha_d))^{(n/p)/d} \quad \text{(since \( \varphi_p \) is a ring endomorphism)}.
\]

On the other hand,

\[
(11.345) \quad \sum_{d|n} d \alpha_d^{n/d} = \sum_{d|n/p} d \alpha_d^{n/d} + \sum_{d|n/p} \sum_{d|n \mod p^{v_p(n)} \text{ and } d|n/p} d \alpha_d^{n/d} \equiv \sum_{d|n/p} d \alpha_d^{n/d} \mod p^{v_p(n)} A.
\]

Thus, if we succeed to prove that

\[
(11.346) \quad d (\varphi_p (\alpha_d))^{(n/p)/d} \equiv d \alpha_d^{n/d} \mod p^{v_p(n)} A \quad \text{for every } d \mid n/p,
\]

then we will obtain

\[
\varphi_p \left( b_{n/p} \right) = \sum_{d|n/p} d (\varphi_p (\alpha_d))^{(n/p)/d} \equiv \sum_{d|n/p} d \alpha_d^{n/d} \equiv \sum_{d|n} d \alpha_d^{n/d} \quad \text{(by (11.345))}
\]

\[
= b_n \mod p^{v_p(n)} A \quad \text{(since } b_n = \sum_{d|n} d \alpha_d^{n/d} \text{)},
\]

and thus our goal (proving (11.344)) will be achieved. Hence, it remains to prove (11.346).
So let $d$ be any positive divisor of $n/p$. Then, $\varphi_p(\alpha_d) \equiv a_d^p \pmod{pA}$ (because of the axiom $\varphi_p(a) \equiv a^p \pmod{pA}$ for every $a \in A$). Thus, (11.342) (applied to $a = \varphi_p(\alpha_d)$, $b = a_d^p$ and $N = (n/p)/d$) yields $(\varphi_p(\alpha_d))^{(n/p)/d} \equiv (a_d^p)^{(n/p)/d} \pmod{p^\nu_p((n/p)/d)+1}$. Since $(\alpha_d^p)^{(n/p)/d} = a_d^{n/d}$ and $v_p((n/p)/d)+1 = v_p(n/d)$, this rewrites as $(\varphi_p(\alpha_d))^{(n/p)/d} \equiv a_d^{n/d} \pmod{p^\nu_p(n/d)A}$. Multiplying this by $d$ results in $d(\varphi_p(\alpha_d))^{(n/p)/d} \equiv da_d^{n/d} \pmod{dp^\nu_p(n/d)A}$. Since $dp^\nu_p(n/d)$ is divisible by $p^\nu_p(n)$, this yields (11.346). This completes the proof of (11.346), and thus also that of the implication $D \implies C$.

Proof of the implication $C \implies D$: Assume that Assertion $C$ holds. Thus, for every positive integer $n$ and every prime factor $p$ of $n$, we have

\[ \varphi_p(b_{n/p}) \equiv b_n \pmod{p^\nu_p(n)A}. \] (11.347)

We now need to prove that Assertion $D$ holds as well. In other words, we need to show that there exists a family $(a_n)_{n \geq 1} \in A^{\{1,2,3,\ldots\}}$ of elements of $A$ such that every positive integer $n$ satisfies $b_n = \sum_{d|n} da_d^{n/d}$. In other words (renaming $n$ as $m$), we need to show that there exists a family $(\alpha_m)_{m \geq 1} \in A^{\{1,2,3,\ldots\}}$ of elements of $A$ such that every positive integer $m$ satisfies $b_m = \sum_{d|m} da_d^{m/d}$.

We construct this family $(\alpha_m)_{m \geq 1}$ recursively. So we fix some $n \geq 1$, and assume that $\alpha_m$ is already constructed for every positive integer $m < n$ in such a way that

\[ \text{the equality } b_m = \sum_{d|m} da_d^{m/d} \text{ is satisfied for every positive integer } m < n. \] (11.348)

We now need to construct an $\alpha_n \in A$ such that $b_m = \sum_{d|m} da_d^{m/d}$ is satisfied for $m = n$. In other words, we need to construct an $\alpha_n \in A$ satisfying $b_n = \sum_{d|n} da_d^{n/d}$.

Let us first choose $\alpha_n$ arbitrarily (with the intention to tweak it later). Let $p$ be any prime factor of $n$. Then, applying (11.348) to $m = n/p$, we obtain $b_{n/p} = \sum_{d|n/p} da_d^{n/d}$. This allows us to prove that $\varphi_p(b_{n/p}) \equiv \sum_{d|n/p} da_d^{n/d} \pmod{p^\nu_p(n)A}$ holds (just as in the proof of the implication $D \implies C$). Compared with (11.347), this yields $b_n \equiv \sum_{d|n} da_d^{n/d} \pmod{p^\nu_p(n)A}$. That is, $b_n - \sum_{d|n} da_d^{n/d} \in p^\nu_p(n)A$.

Now, let us forget that we fixed $p$. We thus have shown that (with our arbitrary choice of $\alpha_n$) we have $b_n - \sum_{d|n} da_d^{n/d} \in p^\nu_p(n)A$ for every prime factor $p$ of $n$. As a consequence,

\[ b_n - \sum_{d|n} da_d^{n/d} \in \bigcap_{p \text{ is a prime factor of } n} p^\nu_p(n)A = \prod_{p \text{ is a prime factor of } n} \left( p^\nu_p(n)A \right) \]  

(by (11.343))

\[ = \left\{ \prod_{p \text{ is a prime factor of } n} p^\nu_p(n) \right\} A = nA. \]

In other words, there exists a $\gamma \in A$ such that $b_n - \sum_{d|n} da_d^{n/d} = n\gamma$. Consider this $\gamma$. Now, if we replace $\alpha_n$ by $\alpha_n + \gamma$, then the sum $\sum_{d|n} da_d^{n/d}$ increases by $n\gamma = b_n - \sum_{d|n} da_d^{n/d}$, and therefore becomes precisely $b_n$. Hence, by replacing $\alpha_n$ by $\alpha_n + \gamma$, we achieve that $b_n = \sum_{d|n} da_d^{n/d}$ holds. Thus, we have found the $\alpha_n$ we were searching for, and the recursive construction of the family $(\alpha_m)_{m \geq 1}$ has proceeded by one more step. The proof of the implication $C \implies D$ is thus complete.

Proof of the implication $E \implies C$: The proof of the implication $E \implies C$ proceeds exactly as our above proof of the implication $D \implies C$, with the following changes:

- Every appearance of $\alpha_i$, for some $i \geq 1$ must be replaced by the corresponding $\beta_i$.
- Every time an element of $A$ was taken to the $k$-th power (for some $k \in \{1,2,3,\ldots\}$) in our proof of the implication $D \implies C$, it needs now to be subjected to the ring endomorphism $\varphi_k$ instead. So, for example, $\alpha_d^{n/d}$ is replaced by $\varphi_d^{-1}(\beta_d)$ everywhere (remember that $\alpha_d$ becomes $\beta_d$). Note that this concerns only elements of $A$. We don’t replace the power $p^\nu_p(n)$ by anything.
The equality
\[ \varphi_p \left( \sum_{d|n/p} d \alpha_{\varphi(d)}(n/p)/d \right) = \sum_{d|n/p} d \left( \varphi_p \left( \alpha_d \right) \right)(n/p)/d \]
is replaced by
\[ \varphi_p \left( \sum_{d|n/p} d \varphi(n/p)/d (\beta_d) \right) = \sum_{d|n/p} d \varphi(n/p)/d (\varphi_p(\beta_d)), \]
whose proof uses the \( \mathbb{Z} \)-linearity of \( \varphi_p \) and the fact that \( \varphi_p \circ \varphi(n/p)/d = \varphi(n/d) = \varphi(n/p)/d \circ \varphi_p \).

The proof of (11.346) needs to be replaced by a proof of the congruence
\[
d\varphi(n/p)/d (\varphi_p(\beta_d)) \equiv d\varphi_{n/d}(\beta_d) \mod p^{v_{p}(n)}A \quad \text{for every } d \mid n/p.
\]

Fortunately, the latter congruence is obvious, since \( \varphi(n/p)/d \circ \varphi_p = \varphi_{n/d} \).

**Proof of the implication \( C \implies E \):** The proof of the implication \( C \implies E \) can be obtained from the proof of the implication \( D \implies C \) using the same changes that were made to transform the proof of the implication \( D \implies C \) into a proof of the implication \( E \implies C \). The only new “idea” is to use the fact that \( \varphi_1 = \text{id} \) (in showing that if we replace \( \beta_n \) by \( \beta_n + \gamma \), then the sum \( \sum_{d|n} d\varphi_{n/d}(\beta_d) \) increases by \( n\gamma \).

**Proof of the implication \( E \implies F \):** Assume that Assertion \( E \) holds. That is, there exists a family \((\beta_n)_{n \geq 1} \in A^{\{1,2,3,\ldots\}}\) of elements of \( A \) such that every positive integer \( n \) satisfies
\[
b_n = \sum_{d|n} d\varphi_{n/d}(\beta_d).
\]

Consider this family \((\beta_n)_{n \geq 1}\). We need to prove that Assertion \( F \) holds, i.e., that every positive integer \( n \) satisfies
\[
\sum_{d|n} \mu(d) \varphi_d(b_{n/d}) \in nA.
\]

We fix a positive integer \( n \). Then, every positive divisor \( d \) of \( n \) satisfies \( b_{n/d} = \sum_{e|n/d} e \varphi_{(n/d)/e}(\beta_e) \) (by (11.350), applied to \( n/d \) instead of \( n \), and with the summation index \( d \) renamed as \( e \)). Hence,
\[
\sum_{d|n} \mu(d) \varphi_d \left( \sum_{e|n/d} e \varphi_{(n/d)/e}(\beta_e) \right) = \sum_{d|n} \mu(d) \varphi_d \left( \sum_{e|n/d} e \varphi_{(n/d)/e}(\beta_e) \right) \quad \text{(since } \varphi_d \text{ is linear)}
\]
\[
= \sum_{d|n} \mu(d) \sum_{e|n/d} e \varphi_d \left( \varphi_{(n/d)/e}(\beta_e) \right)
\]
\[
= \sum_{d|n} \sum_{e|n/d} \mu(d) e \varphi_d \left( \varphi_{(n/d)/e}(\beta_e) \right)
\]
\[
= \sum_{e|n} \sum_{d|n/e} \mu(d) e \varphi_{n/e}(\beta_e) = \sum_{e|n} \sum_{d|n/e} \mu(d) e \varphi_{n/e}(\beta_e)
\]
\[
= \sum_{e|n} \delta_{n/e,1} e \varphi_{n/e}(\beta_e) = n \varphi_{n/e}(\beta_n) = n\beta_n \in nA.
\]

Thus, Assertion \( F \) holds, so that we have proven the implication \( E \implies F \).

**Proof of the implication \( F \implies E \):** Assume that Assertion \( F \) holds. That is, every positive integer \( n \) satisfies
\[
\sum_{d|n} \mu(d) \varphi_d(b_{n/d}) \in nA.
\]
Now we need to prove that Assertion $\mathcal{E}$ holds, i.e., that there exists a family $(\beta_n)_{n \geq 1} \in A^{\{1,2,3,\ldots\}}$ of elements of $A$ such that every positive integer $n$ satisfies

$$b_n = \sum_{d|n} d\varphi_{n/d} (\beta_d).$$

We shall construct such a family $(\beta_n)_{n \geq 1}$ recursively. That is, we fix some $N \in \{1,2,3,\ldots\}$, and we assume that we already have constructed a $\beta_n \in A$ for every positive integer $n < N$ in such a way that (11.352) is satisfied for every positive integer $n < N$. We now need to find a $\beta_N \in A$ such that (11.352) is satisfied for $n = N$ as well.

From (11.351) (applied to $n = N$), we have $\sum_{d|N} \mu (d) \varphi_d (b_{N/d}) \in nA$. Thus, there exists a $t \in A$ such that $\sum_{d|N} \mu (d) \varphi_d (b_{N/d}) = Nt$. Set Consider this $t$. Set $\beta_N = t$. We have

$$Nt = \sum_{d|N} \mu (d) \varphi_d (b_{N/d}) = \sum_{e|N} \mu (e) \varphi_e (b_{N/e}) = \sum_{e|N; e > 1} \mu (e) \varphi_e \left( \sum_{d|N/e} d\varphi_{(N/e)/d} (\beta_d) \right) + b_N = \sum_{e|N; e > 1} \mu (e) \varphi_e \left( \sum_{d|N/e} d\varphi_{(N/e)/d} (\beta_d) \right) + b_N (\text{since } \varphi_e \text{ is linear})$$

$$= \sum_{e|N; e > 1} \mu (e) \sum_{d|N/e} d\varphi_{e/d} (\beta_d) + b_N = \sum_{e|N; e > 1} \mu (e) d\varphi_{N/d} (\beta_d) + b_N = \sum_{e|N; e > 1} \mu (e) d\varphi_{N/d} (\beta_d) + b_N = \sum_{d|N; d < N} \sum_{e|N/d; e > 1} \mu (e) d\varphi_{N/d} (\beta_d) + b_N$$

$$= \sum_{d|N; d < N} \sum_{e|N/d; e > 1} \mu (e) d\varphi_{N/d} (\beta_d) + b_N = \sum_{d|N; d < N} \sum_{e|N/d; e > 1} \mu (e) d\varphi_{N/d} (\beta_d) + b_N = \sum_{d|N; d < N} \sum_{e|N/d; e > 1} \mu (e) d\varphi_{N/d} (\beta_d) + b_N (\text{by (11.337), applied to } N/d \text{ instead of } n)$$

$$= \sum_{d|N; d < N} \delta_{N/d,1} - \frac{1}{1} d\varphi_{N/d} (\beta_d) + b_N = \sum_{d|N; d < N} \delta_{N/d,1} - \frac{1}{1} d\varphi_{N/d} (\beta_d) + b_N (\text{since } N/d > 1)$$

(11.353)

$$= - \sum_{d|N; d < N} d\varphi_{N/d} (\beta_d) + b_N.$$
Now,
\[
\sum_{d|N} d \varphi_{N/d} (\beta_d) = \sum_{d|N; \ n/N \neq \varphi_1 = \text{id}} \varphi_{N/N} (\beta_N) = \sum_{d|N; \ n/N \neq \varphi_1 = \text{id}} d \varphi_{N/d} (\beta_d) + \sum_{d<N} d \varphi_{N/d} (\beta_d) + \sum_{d<N} d \varphi_{N/d} (\beta_d) + b_N = \sum_{d|N; \ n/N \neq \varphi_1 = \text{id}} d \varphi_{N/d} (\beta_d) + b_N.
\]

Thus, (11.352) is satisfied for \( n = N \). We have thus completed a step of our recursive construction of the family \((\beta_n)_{n \geq 1}\); this family therefore exists, and the implication \( F \implies E \) is proven.

**Proof of the implication** \( F \implies G \): Assume that Assertion \( F \) holds. That is, every positive integer \( n \) satisfies

\[
(11.354) \quad \sum_{d|n} \mu (d) \varphi_d (b_{n/d}) \in nA.
\]

On the other hand, every positive integer \( e \) satisfies

\[
(11.355) \quad \sum_{d|e} \mu (d) \frac{e}{d} = \phi (e)
\]

(by (11.337), applied to \( n = e \)). Now, every positive integer \( n \) satisfies

\[
\sum_{d|n} \phi (d) \varphi_d (b_{n/d}) = \sum_{e|n} \phi (e) \varphi_e (b_{n/e}) = \sum_{e|n} \mu (e) \frac{e}{d} \varphi_e (b_{n/e}) = \sum_{d|n} \mu (d) \frac{e}{d} \varphi_e (b_{n/d/e}) = \sum_{d|n} \mu (d) \frac{e}{d} \varphi_e (b_{(n/e)/ (e/d)})
\]

(here, we have substituted \( ed \) for \( e \) in the second sum)

\[
= \sum_{d|n} \sum_{e|n/e} \mu (d) e \varphi_e (b_{(n/e)/d}) = \sum_{e|n} \mu (e) \varphi_e (b_{(n/e)/d}) = \sum_{e|n} e \varphi_e (b_{(n/e)/d}) (\text{since } \varphi_e \text{ is } \mathbb{Z}-\text{linear})
\]

\[
= \sum_{e|n} e \varphi_e (b_{(n/e)/d}) (\text{since } \varphi_n \text{ in } \mathbb{Z}-\text{linear})
\]

Thus, Assertion \( G \) holds. We have thus proven the implication \( F \implies G \).
Proof of the implication $\mathcal{G} \implies \mathcal{F}$: Assume that Assertion $\mathcal{G}$ holds. That is, every positive integer $n$ satisfies
\begin{equation}
\sum_{d|n} \phi(d) \varphi_{d/d} (b_{n/d}) \in nA. \tag{11.356}
\end{equation}

On the other hand, every positive integer $e$ satisfies $\sum_{d|e} d\mu(d) \phi\left(\frac{e}{d}\right) = \mu(e)$ (by (11.340), applied to $n = e$). In other words, every positive integer $e$ satisfies
\begin{equation}
\mu(e) = \sum_{d|e} d\mu(d) \phi\left(\frac{e}{d}\right) = \sum_{d|e} e \frac{\mu(e)}{d} \phi(d) \tag{11.357}
\end{equation}
(here, we have substituted $e/d$ for $d$ in the sum). Now, every positive integer $n$ satisfies
\begin{align*}
\sum_{d|n} \mu(d) \varphi_{d} (b_{n/d}) & = \sum_{e|n} \mu(e) \varphi_{e} (b_{n/e}) \\
& = \sum_{e|n} \sum_{d|e} e \frac{\mu(e)}{d} \phi(d) \varphi_{e} \left(\frac{b_{n/e}}{\phi(d)}\right)
\end{align*}
(by (11.357))
\begin{align*}
& = \sum_{e|n} \sum_{d|e} \frac{e}{d} \phi(d) \varphi_{e} \left(\frac{b_{n/e}}{\phi(d)}\right) \\
& = \sum_{e|n} \sum_{d|e} e \mu(e) \phi(d) \varphi_{ed} \left(\frac{b_{n/e}}{\phi(d)}\right)
\end{align*}
(here, we have substituted $ed$ for $e$ in the second sum)
\begin{align*}
& = \sum_{e|n} \sum_{d|e} e \mu(e) \varphi_{e} \phi(d) \varphi_{d} (b_{n/e}) \left(\varphi_{e} \circ \varphi_{d} \right) (b_{n/e}) \\
& = \sum_{e|n} \sum_{d|e} e \mu(e) \varphi_{e} \left(\sum_{d|n/e} \varphi_{d} (b_{n/e}) \right)
\end{align*}
(since $\varphi_{e}$ is $\mathbb{Z}$-linear)
\begin{align*}
& \in \sum_{e|n} e \mu(e) \varphi_{e} \left(\left(\frac{n}{e}\right) A\right) \\
& \subset \sum_{e|n} e \mu(e) \left(\frac{n}{e}\right) A = \sum_{e|n} \eta_{e} \\
& \subset nA. \tag{by (11.356), applied to $n/e$ instead of $n$}
\end{align*}

Thus, Assertion $\mathcal{F}$ holds. We have thus proven the implication $\mathcal{G} \implies \mathcal{F}$.

Proof of the equivalence $\mathcal{G} \iff \mathcal{H}$: For every positive integer $n$, we have
\begin{align*}
\sum_{i=1}^{n} \varphi_{\gcd(i,n)} \left(\frac{b_{\gcd(i,n)}}{b_{\gcd(i,n)}}\right) & = \sum_{d|\gcd(i,n)} \sum_{\gcd(i,n)=d} \varphi_{n/d} (b_{d}) \\
& = e^\left(\frac{n}{d}\right) \varphi_{n/d} (b_{d}) \tag{by (11.356)}
\end{align*}
(here, we have substituted $n/d$ for $d$ in the sum). This makes it clear that Assertions $\mathcal{G}$ and $\mathcal{H}$ are equivalent.
The implications and equivalences that we have proven, combined, yield the equivalence of all seven assertions $C$, $D$, $E$, $F$, $G$, $H$ and $J$. This solves Exercise 2.82.

11.57. Solution to Exercise 2.84. Solution to Exercise 2.84. We know that the seven assertions $C$, $D$, $E$, $F$, $G$, $H$ and $J$ are equivalent; hence, for each of our families, it suffices to prove one of these assertions. We choose the assertion $C$, as it is the easiest to prove.

- Proof of Assertion $C$ for the family $(b_n)_{n \geq 1} = (q^n)_{n \geq 1}$, where $q$ is a given integer: Let $q$ be an integer. We need to prove Assertion $C$ for the family $(b_n)_{n \geq 1} = (q^n)_{n \geq 1}$. This means proving that for every positive integer $n$ and every prime factor $p$ of $n$, we have
\begin{equation}
\varphi_p \left( q^{n/p} \right) \equiv q^n \mod p^{v_p(n)} \mathbb{Z}.
\end{equation}

So let us prove this. Let $n$ be a positive integer, and let $p$ be a prime factor of $n$. Fermat’s little theorem yields $q^n \equiv q \mod p \mathbb{Z}$. Hence, $(11.342)$ (applied to $A = \mathbb{Z}$, $a = q^n$, $b = q$ and $N = n/p$) yields $(q^n)^{n/p} \equiv q^{n/p} \mod p^{v_p(n/p)+1} \mathbb{Z}$. Since $(q^n)^{n/p} = q^n$ and $v_p(n/p) + 1 = v_p(n)$, this rewrites as $q^n \equiv q^{n/p} \mod p^{v_p(n)} \mathbb{Z}$. Now, $\varphi_p = \text{id}$, so that $\varphi_p \left( q^{n/p} \right) \equiv q^n \mod p^{v_p(n)} \mathbb{Z}$. Thus, $(11.358)$ is proven, and we are done with the family $(b_n)_{n \geq 1} = (q^n)_{n \geq 1}$.

- Proof of Assertion $C$ for the family $(b_n)_{n \geq 1} = (q)_{n \geq 1}$, where $q$ is a given integer: Let $q$ be an integer. We need to prove Assertion $C$ for the family $(b_n)_{n \geq 1} = (q)_{n \geq 1}$. This means proving that for every positive integer $n$ and every prime factor $p$ of $n$, we have
$$\varphi_p \left( q \right) \equiv q \mod p^{v_p(n)} \mathbb{Z}.$$ But this is obvious, since $\varphi_p = \text{id}$. Thus, the family $(b_n)_{n \geq 1} = (q)_{n \geq 1}$ satisfies Assertion $C$.

- Proof of Assertion $C$ for the family $(b_n)_{n \geq 1} = \left( \begin{array}{c} qn \\ rn \end{array} \right)_{n \geq 1}$, where $r \in \mathbb{Q}$ and $q \in \mathbb{Z}$ are given: Let $r \in \mathbb{Q}$ and $q \in \mathbb{Z}$. We need to prove Assertion $C$ for the family $(b_n)_{n \geq 1} = \left( \begin{array}{c} qn \\ rn \end{array} \right)_{n \geq 1}$. This means proving that for every positive integer $n$ and every prime factor $p$ of $n$, we have
\begin{equation}
\varphi_p \left( \left( \begin{array}{c} qn/p \\ rn/p \end{array} \right) \right) \equiv \left( \begin{array}{c} qn \\ rn \end{array} \right) \mod p^{v_p(n)} \mathbb{Z}.
\end{equation}

So let us prove this. Let $n$ be a positive integer, and $p$ be a prime factor of $n$. We need to prove $(11.359)$. In other words, we need to prove that
\begin{equation}
\left( \begin{array}{c} qn/p \\ rn/p \end{array} \right) \equiv \left( \begin{array}{c} qn \\ rn \end{array} \right) \mod p^{v_p(n)} \mathbb{Z}
\end{equation}
(since $\varphi_p = \text{id}$). We WLOG assume that $rn \in \mathbb{Z}$ (since otherwise, both sides of this congruence $(11.360)$ are 0).

It is well-known that $(1 + X)^p \equiv 1 + X^p \mod p \mathbb{Z}[X]$ in the polynomial ring $\mathbb{Z}[X]$. Hence, $((1 + X)^p)^{n/p} \equiv (1 + X)^{n/p} \mod p^{v_p(n/p)+1} \mathbb{Z}[X]$ (by $(11.342)$, applied to $\mathbb{Z}[X]$, $(1 + X)^p$, $1 + X^p$ and $n/p$ instead of $A$, $a$, $b$ and $N$). Since $((1 + X)^p)^{n/p} = (1 + X)^n$ and $v_p(n/p) + 1 = v_p(n)$, this rewrites as
$$(1 + X)^n \equiv (1 + X)^{n/p} \mod p^{v_p(n)} \mathbb{Z}[X].$$
Hence,
$$(1 + X^n)^q \equiv (1 + X^{n/p})^q \mod p^{v_p(n)} \mathbb{Z}[X].$$
(since $\mathbb{Z}[X]$ is a subring of $\mathbb{Z}[[X]]$). We can take both sides of this congruence to the $q$-th power\footnote{This works even if $q$ is negative, since $(1 + X^n)^q$ and $p^{v_p(n)} \mathbb{Z}[X]$ are invertible in $\mathbb{Z}[[X]] / (p^{v_p(n)} \mathbb{Z}[[X]])$.}, and thus obtain
$$((((1 + X)^n)^q)^q)^q \equiv (1 + X^{n/p})^{q^q} \mod p^{v_p(n)} \mathbb{Z}[[X]].$$
In other words,
$$((1 + X)^n)^{q^q} \equiv (1 + X)^{n/p} \mod p^{v_p(n)} \mathbb{Z}. $$
Comparing the coefficients before $X^{rn}$ on both sides of this congruence, we obtain
\[
\binom{qn}{rn} \equiv \binom{qn/p}{rn/p} \mod p^{v_p(\ell)} Z.
\]
This proves (11.360) and thus (11.359). Hence, we are done with the family $(b_n)_{n \geq 1} = \binom{qn}{rn}_{n \geq 1}$.

[Remark: There is an alternative, combinatorial approach to proving (11.360) when $q$ is nonnegative. This approach proceeds by counting the $(rn)$-element subsets of the set $\mathbb{Z}/(qn)$. On the one hand, the number of such subsets is clearly $\binom{qn}{rn}$. On the other hand, these subsets fall into two classes:

- the subsets which are invariant under the permutation
  \[
  \mathbb{Z}/(qn) \rightarrow \mathbb{Z}/(qn),
  \]
  \[
  i \mapsto i + qn/p
  \]
  of $\mathbb{Z}/(qn)$;

- the subsets which are not invariant under this permutation.

It is easy to see that the number of all subsets in the first class is $\binom{qn/p}{rn/p}$ (indeed, the intersection of such a subset with $\{0, 1, \ldots, qn/p - 1\} \subset \mathbb{Z}/(qn)$ must have $rn/p$ elements, and uniquely determines the whole subset by “replication”), whereas the number of all subsets in the second class is divisible by $p^{v_p(n)}$ (because the permutation $\mathbb{Z}/(qn) \rightarrow \mathbb{Z}/(qn)$, $i \mapsto i + 1$ acts on these subsets, and thus splits them into orbits, each of which has size divisible by $p^{v_p(n)}$). Hence, the number of all subsets in both classes together is $\equiv \binom{qn/p}{rn/p} \mod p^{v_p(n)} Z$. Comparing these two answers, we obtain
\[
\binom{qn}{rn} \equiv \binom{qn/p}{rn/p} \mod p^{v_p(n)} Z.
\]

The downside of this nice approach is that it requires a modification in the case when $q$ is negative. Here, the upper negation formula \((-a \choose k) = (-1)^k \binom{a + k - 1}{k}\) needs to be used, along with the “stars-and-bars” formula \(\binom{a + k - 1}{k}\) for the number of $k$-element multisets whose elements belong to $\{1, 2, \ldots, a\}$. We leave the details to the reader.]

• Proof of Assertion C for the family $(b_n)_{n \geq 1} = \binom{qn - 1}{rn - 1}_{n \geq 1}$, where $r \in \mathbb{Z}$ and $q \in \mathbb{Z}$ are given:

Let $r \in \mathbb{Z}$ and $q \in \mathbb{Z}$. We need to prove Assertion C for the family $(b_n)_{n \geq 1} = \binom{qn - 1}{rn - 1}_{n \geq 1}$. This means proving that for every positive integer $n$ and every prime factor $p$ of $n$, we have
\[
\varphi_p \binom{qn/p - 1}{rn/p - 1} \equiv \binom{qn - 1}{rn - 1} \mod p^{v_p(\ell)} Z.
\]

So let us prove this. Let $n$ be a positive integer, and $p$ be a prime factor of $n$. We need to prove (11.361). In other words, we need to prove that
\[
\binom{qn/p - 1}{rn/p - 1} \equiv \binom{qn - 1}{rn - 1} \mod p^{v_p(n)} Z
\]
(since $\varphi_p = \text{id}$). By the recurrence of the binomial coefficients, we have $\binom{qn/p - 1}{rn/p - 1} = \binom{qn/p}{rn/p} - \binom{qn - 1}{rn - 1} = \binom{qn}{rn} - \binom{qn - 1}{rn}$). Hence, we can obtain the desired congruence (11.362) by subtracting the congruence
\[
\binom{qn/p - 1}{rn/p} \equiv \binom{qn - 1}{rn} \mod p^{v_p(n)} Z
\]
from the congruence \((11.360)\). It therefore remains to prove the congruence \((11.363)\) (since \((11.360)\) has already been proven).

We recall the (easy-to-check) formula \(\left(\frac{a-1}{k}\right) = (-1)^k \left(\frac{k-a}{k}\right)\) for every \(a \in \mathbb{Z}\) and \(k \in \mathbb{N}\). This formula allows us to rewrite both sides of \((11.363)\), and thus \((11.363)\) becomes

\[
(11.364) \quad (-1)^{rn/p} \left(\frac{rn/p - qn/p}{rn/p}\right) \equiv (-1)^{rn} \left(\frac{rn - qn}{rn}\right) \mod p^{v_p(n)}Z.
\]

So it remains to prove \((11.364)\). We have \((-1)^{rn/p} \equiv (-1)^{rn} \mod p^{v_p(n)}Z\). Thus, we can cancel the \((-1)^{rn/p}\) on the left hand side of \((11.364)\) against the \((-1)^{rn}\) on the right hand side, and therefore \((11.364)\) takes the equivalent form

\[
\left(\frac{rn/p - qn/p}{rn/p}\right) \equiv \left(\frac{rn - qn}{rn}\right) \mod p^{v_p(n)}Z.
\]

This further rewrites as

\[
\left(\frac{(r - q)n/p}{rn/p}\right) \equiv \left(\frac{(r - q)n}{rn}\right) \mod p^{v_p(n)}Z.
\]

But this follows from \((11.360)\), applied to \(r - q\) instead of \(q\). Hence, \((11.364)\) is proven, and consequently Assertion \(C\) holds for the family \((b_n)_{n \geq 1} = \left(\left(\frac{qn - 1}{rn - 1}\right)\right)_{n \geq 1}\).

Exercise 2.84 is thus solved.

---

11.58. **Solution to Exercise 2.85.** Solution to Exercise 2.85. (a) This is obvious since \(f_n\) is an evaluation homomorphism (in an appropriate sense).\(^{419}\)

(b) Let \(n\) and \(m\) be positive integers. Then,

\[ (f_n \circ f_m)(a) = f_n \left( f_m(a) \right) = f_n(a(x_1^m, x_2^m, x_3^m, \ldots)) = (a(x_1^m, x_2^m, x_3^m, \ldots)) (x_1^n, x_2^n, x_3^n, \ldots) \]

for all \(a \in \Lambda\). Thus, \(f_n \circ f_m = f_{nm}\), so that Exercise 2.85(b) is solved.

(c) For every \(a \in \Lambda\), we have \(f_1(a) = a(x_1, x_2, x_3, \ldots) = a\). Thus, \(f_1 = \text{id}\), so that Exercise 2.85(c) is solved.

(d) There are several ways to solve Exercise 2.85(d):

- One can check \(\Delta \circ f_n = (f_n \otimes f_n) \circ \Delta\) rather directly on the basis \((m_\lambda)_{\lambda \in \text{Par}}\) of \(\Lambda\), using \((2.3)\) and the definition of \(f_n\).
- One can check \(\Delta \circ f_n = (f_n \otimes f_n) \circ \Delta\) very easily on the elements \(p_1, p_2, p_3, \ldots\) of \(\Lambda\), and then “pretend” that these elements generate \(\Lambda\) (in fact, they do generate \(\Lambda\) when \(k = \mathbb{Q}\), so this solves Exercise 2.85(d) in the case of \(k = \mathbb{Q}\)); but this easily entails that Exercise 2.85(d) holds in the case of \(k = \mathbb{Z}\)\(^{420}\), and this, in turn, can be used to solve Exercise 2.85(d) for arbitrary \(k\)\(^{421}\).

---

\(^{418}\)Proof. The only situation in which this is not obvious is when one of the integers \(rn/p\) and \(rn\) is even and the other is odd. This means, of course, that \(rn/p\) is odd and \(rn\) is even (since \(rn/p \mid rn\)). Hence, in this situation, we have \(p = 2\) and \(v_p(n) \leq 1\), and therefore \(p^{v_p(n)} = 2\), so that \((-1)^{rn/p} \equiv (-1)^{rn} \mod p^{v_p(n)}Z\) follows immediately from the fact that any two integer powers of \(-1\) are congruent to each other modulo \(2Z\).

\(^{419}\)Or, in a more down-to-earth fashion, this is obvious because (for example) multiplying two power series and then replacing all variables in the product by their \(n\)-th powers gives the same result as first replacing all variables by their \(n\)-th powers and then multiplying the resulting two power series.

\(^{420}\)Getting from \(k = \mathbb{Q}\) to \(k = \mathbb{Z}\) requires a standard functoriality argument. See Step 2 of the solution of Exercise 2.86(g) further below for an example of such an argument.

\(^{421}\)Getting from \(k = \mathbb{Z}\) to arbitrary \(k\) requires a standard functoriality argument. See Step 3 of the solution of Exercise 2.86(g) further below for an example of such an argument.
• One can derive Exercise 2.85(d) from Exercise 2.86(e) using the selfduality of $\Lambda$ and Exercise 2.86(f).

This is somewhat complicated by the fact that $v_n$ and $f_n$ (generally) are not graded maps$^{\text{422}}$, but with the correct arguments, one can completely avoid any use of gradedness.$^{\text{423}}$

Probably the following solution of Exercise 2.85(d) is the shortest:

Fix $n \in \{1, 2, 3, \ldots \}$. We need to prove that $f_n : \Lambda \to \Lambda$ is a Hopf algebra homomorphism. We already know that $f_n$ is a $k$-algebra homomorphism. Therefore, if we can show that $f_n$ is a $k$-coalgebra homomorphism, then it will immediately follow that $f_n$ is a $k$-bialgebra homomorphism and thus a Hopf algebra homomorphism.

---

$^{\text{422}}$They nevertheless have a property close to being graded. Namely, let us say that a $k$-linear map $\varphi$ between two graded $k$-modules $V = \bigoplus_{n \geq 0} V_n$ and $W = \bigoplus_{n \geq 0} W_n$ scales the degree by $q$ (where $q$ is some fixed rational number) if it has the property that $\varphi(V_n) \subset W_{qn}$ for all $n \in \mathbb{N}$ (where $W_{qn}$ is understood to be 0 when $qn$ is not an integer). Then, $v_n$ scales the degree by $1/n$, whereas $f_n$ scales the degree by $n$. When a $k$-linear map $\varphi$ between two graded $k$-modules $V$ and $W$ scales the degree by a nonzero rational number $q$, the adjoint map $\varphi^* : W^* \to V^*$ restricts to a $k$-linear map $W^\circ \to V^\circ$, which scales the degree by $1/q$; thus a large part of the theory of adjoint maps which is usually formulated for graded maps can be carried over to the case of maps scaling the degree by $q$.

$^{\text{423}}$Here is a sketch of how these correct arguments look like.

The Hall inner product $(\cdot, \cdot)_{\Lambda}$ on $\Lambda$ gives rise to a symmetric bilinear form $(\cdot, \cdot)_{\Lambda \otimes \Lambda}$ on $\Lambda \otimes \Lambda$ (according to Definition 3.3(b)).

It is easy to see that every three elements $a, b$ and $c$ of $\Lambda$ satisfy

$$ (a, m(b \otimes c))_{\Lambda} = (\Delta(a), b \otimes c)_{\Lambda \otimes \Lambda}. $$

[Proof of (11.365): Let $a, b$ and $c$ be three elements of $\Lambda$. We need to prove the equality (11.365). Since this equality is $k$-linear in each of $a, b$ and $c$, we can WLOG assume that each of $a, b$ and $c$ belongs to the basis $\{(s)_{\lambda, \mu} \}_{\lambda, \mu \in \operatorname{Par}}$ of the $k$-module $\Lambda$. Assume this. Then, there exist three elements $\alpha, \beta$ and $\gamma$ of $\operatorname{Par}$ such that $a = s_\alpha, b = s_{\beta}$ and $c = s_{\gamma}$. Consider these $\alpha, \beta$ and $\gamma$. Now, let us use the notations of Corollary 2.35. Applying Corollary 2.35 to $\lambda = \alpha, \mu = \beta$ and $\nu = \gamma$ yields $c^\alpha_{\beta, \gamma} = c^\beta_{\gamma, \alpha}$. But the definition of the map $m$ yields

$$ m(b \otimes c) = \frac{b \otimes c}{s_{\beta} s_{\gamma}} = \sum_{\lambda \in \operatorname{Par}} c^\lambda_{\beta, \gamma} s_{\lambda} $$

(by the first equality in (2.26), applied to $\mu = \beta$ and $\nu = \gamma$). Hence,

$$ (a, m(b \otimes c))_{\Lambda} = \sum_{\lambda \in \operatorname{Par}} c^\lambda_{\beta, \gamma} s_{\lambda} = (\Delta(a), b \otimes c)_{\Lambda \otimes \Lambda}. $$

On the other hand, applying the map $\Delta$ to both sides of the equality $a = s_{\alpha}$, we obtain

$$ \Delta a = \delta_{s_{\alpha}} = \sum_{\mu, \nu \in \operatorname{Par}} c^\mu_{\nu} s_{\mu} \otimes s_{\nu} $$

(by the second equality in (2.26), applied to $\lambda = \alpha$). Hence,

$$ \begin{aligned}
(a, m(b \otimes c))_{\Lambda} & = \sum_{\mu, \nu \in \operatorname{Par}} c^\mu_{\nu} s_{\mu} \otimes s_{\nu} \otimes s_{\gamma} \\
& = \sum_{\mu, \nu \in \operatorname{Par}} c^\mu_{\nu} \delta_{s_{\mu}} \delta_{s_{\nu}} \delta_{s_{\gamma}} \end{aligned} $$

Compared with $(a, m(b \otimes c))_{\Lambda} = c^\alpha_{\beta, \gamma}$, this yields $(a, m(b \otimes c))_{\Lambda} = (\Delta(a), b \otimes c)_{\Lambda \otimes \Lambda}$. Thus, (11.365) is proven.]

Now, we can show that every $a \in \Lambda$ and $B \in \Lambda \otimes \Lambda$ satisfy

$$ (a, m(B))_{\Lambda} = (\Delta(a), B)_{\Lambda \otimes \Lambda}. $$

[Proof of (11.366): Let $a \in \Lambda$ and $B \in \Lambda \otimes \Lambda$. We need to prove the equality (11.366). Since this equality is $k$-linear in $B$, we can WLOG assume that $B$ is a pure tensor (because the pure tensors span the $k$-module $\Lambda \otimes \Lambda$). Assume this. Then, $B = b \otimes c$ for two elements $b$ and $c$ of $\Lambda$. Consider these $b$ and $c$. Then,

$$ (a, m(b \otimes c))_{\Lambda} = (\Delta(a), b \otimes c)_{\Lambda \otimes \Lambda} = (\Delta(a), B)_{\Lambda \otimes \Lambda} $$

This proves (11.366).]
(due to Proposition 1.46(c)). Therefore, it remains to show that $f_n$ is a k-coalgebra homomorphism. To do so, we need to check that $\epsilon \circ f_n = \epsilon$ and $\Delta \circ f_n = (f_n \otimes f_n) \circ \Delta$. We shall only prove $\Delta \circ f_n = (f_n \otimes f_n) \circ \Delta$, while the easy proof of $\epsilon \circ f_n = \epsilon$ is left to the reader.

Let us first notice that $\Lambda$ is a bialgebra, and therefore $\Delta$ is a $k$-algebra homomorphism (by the axioms of a bialgebra). Also, $f_n$ is a $k$-algebra homomorphism. Thus, $\Delta \circ f_n$ and $(f_n \otimes f_n) \circ \Delta$ are $k$-algebra homomorphisms.

Exercise 2.86(f) yields that the maps $f_n : \Lambda \to \Lambda$ and $v_n : \Lambda \to \Lambda$ are adjoint with respect to the Hall inner product on $\Lambda$. Thus,

$$
(f_n a, b)_\Lambda = (a, v_n b)_\Lambda \quad \text{for every } a, b \in \Lambda.
$$

From this, it is easy to see that

$$(f_n \otimes f_n) A, B)_{\Lambda \otimes \Lambda} = (A, (v_n \otimes v_n) B)_{\Lambda \otimes \Lambda} \quad \text{for every } A, B \in \Lambda \otimes \Lambda.
$$

[Proof of (11.368):] Let $A \in \Lambda \otimes \Lambda$ and $B \in \Lambda \otimes \Lambda$. We have to prove the equality (11.368). Since this equality is $k$-linear in each of $A$ and $B$, we can WLOG assume that $A$ and $B$ are pure tensors (since pure tensors span $\Lambda \otimes \Lambda$ as a $k$-module). Assume this. Then, we can write $A$ and $B$ in the forms $A = a_1 \otimes a_2$ and $B = b_1 \otimes b_2$ for some $a_1, a_2 \in \Lambda$ and $b_1, b_2 \in \Lambda$. Consider these $a_1, a_2$ and $b_1, b_2$. Now,

$$(f_n \otimes f_n) A, B)_{\Lambda \otimes \Lambda} = (a_1, \Lambda) \otimes (a_2, \Lambda) \otimes b_1 \otimes b_2)_{\Lambda \otimes \Lambda}
$$

Compared with

$$(A, (v_n \otimes v_n) B)_{\Lambda \otimes \Lambda} = (a_1, (v_n \otimes v_n) (b_1 \otimes b_2))_{\Lambda \otimes \Lambda}
$$

this yields $((f_n \otimes f_n) A, B)_{\Lambda \otimes \Lambda} = (A, (v_n \otimes v_n) B)_{\Lambda \otimes \Lambda}$. This proves (11.368).]

Let us now show that $(f_n \otimes f_n) \circ \Delta = \Delta \circ f_n$ (this is one of the axioms that need to be checked in order to show that $f_n$ is a Hopf algebra homomorphism).

Indeed, $v_n$ is a $k$-algebra homomorphism. Hence, $m \circ (v_n \otimes v_n) = v_n \circ m$.

Now, let $a \in \Lambda$. Every $B \in \Lambda \otimes \Lambda$ satisfies

$$
((f_n \otimes f_n) (\Delta a), B)_{\Lambda \otimes \Lambda} = (\Delta a, (v_n \otimes v_n) B)_{\Lambda \otimes \Lambda}
$$

(by (11.368), applied to $A = \Delta a$)

$$
= (a, m ((v_n \otimes v_n) B))_{\Lambda \otimes \Lambda}
$$

because $(a, m ((v_n \otimes v_n) B))_{\Lambda \otimes \Lambda} = (\Delta a, (v_n \otimes v_n) B)_{\Lambda \otimes \Lambda}$

(by (11.368), applied to $(v_n \otimes v_n) B$ instead of $B$)

$$
= (a, (m \circ (v_n \otimes v_n)) (B))_{\Lambda \otimes \Lambda}
$$

(by (11.367), applied to $mB$ instead of $B$)

$$
= (a, m (B))_{\Lambda \otimes \Lambda}
$$

(by (11.367), applied to $mB$ instead of $B$)

$$
= (\Delta (f_n a), B)_{\Lambda \otimes \Lambda}
$$

(by (11.366), applied to $f_n a$ instead of $a$)

$$
= ((\Delta \circ f_n) (a), B)_{\Lambda \otimes \Lambda}
$$

(by (11.367), applied to $(\Delta \circ f_n) (a)$ instead of $a$)

(11.369)
For every \( m \in \mathbb{N} \), let us define an element \( \tilde{h}_m \in k[[x]] \) by \( \tilde{h}_m = \sum_{i_1 \leq i_2 \leq \cdots \leq i_m} x_{i_1}^{n_{i_1}} x_{i_2}^{n_{i_2}} \cdots x_{i_m}^{n_{i_m}} \). Then, it is easy to see that
\[
(11.370) \quad f_n(h_m) = \tilde{h}_m \quad \text{for every } m \in \mathbb{N}.
\]

As a consequence, \( \tilde{h}_m = f_n(h_m) \in \Lambda \) for every \( m \in \mathbb{N} \). (Of course, this was obvious anyway.

Fix a positive integer \( r \). Proposition 2.24(iii) (applied to \( r \) instead of \( n \)) states that \( \Delta h_r = \sum_{i+j=r} h_i \otimes h_j \).

The proof of this proposition (which proceeded by observing that \( h_r(x,y) = \sum_{i+j=r} h_i(x) h_j(y) \) in \( k[[x,y]] \))
can be easily modified to obtain a proof of the equality \( \Delta h_r = \sum_{i+j=r} h_i \otimes h_j \). Thus, we have
\[
(11.371) \quad \Delta h_r = \sum_{i+j=r} h_i \otimes h_j.
\]

Now,
\[
(\Delta \circ f_n)(h_r) = \Delta \left( f_n(h_r) \right) = \sum_{i+j=r} h_i \otimes h_j \quad \text{(by (11.371))}.
\]

Compared with
\[
((f_n \otimes f_n) \circ \Delta)(h_r) = (f_n \otimes f_n) \left( \sum_{i+j=r} h_i \otimes h_j \right) = \sum_{i+j=r} f_n(h_i) \otimes f_n(h_j) = \sum_{i+j=r} \tilde{h}_i \otimes \tilde{h}_j,
\]

this yields \( (\Delta \circ f_n)(h_r) = ((f_n \otimes f_n) \circ \Delta)(h_r) \) for every positive integer \( r \).

Now, let us forget that we fixed \( r \). We thus have proven that
\[
(11.372) \quad (\Delta \circ f_n)(h_r) = ((f_n \otimes f_n) \circ \Delta)(h_r) \quad \text{for every positive integer } r.
\]

Now, recall that the family \( (h_r)_{r \geq 1} \) generates the \( k \)-algebra \( \Lambda \) (according to Proposition 2.28). In other words, \( (h_r)_{r \geq 1} \) is a generating set of the \( k \)-algebra \( \Lambda \). The two \( k \)-algebra homomorphisms \( \Delta \circ f_n \) and \( (f_n \otimes f_n) \circ \Delta \) are equal to each other on this generating set (according to (11.372)), and therefore must be identical (because if two \( k \)-algebra homomorphisms from the same domain are equal to each other on a generating set of their domain, then these two homomorphisms must be identical). In other words, \( \Delta \circ f_n = (f_n \otimes f_n) \circ \Delta \). This completes the solution of Exercise 2.85(d).

(c) Let \( m \in \mathbb{N} \). From (2.14), we have
\[
\prod_{i=1}^{\infty} (1 - x_i t_i^{-1}) = \sum_{n \geq 0} h_n(x) t^n \text{ in the ring } \Lambda[[t]].
\]

Substituting \( x_i^2 \) for \( x_i \) and \( t^2 \) for \( t \) in this equality, we obtain
\[
(11.373) \quad \prod_{i=1}^{\infty} (1 - x_i^2 t_i^2)^{-1} = \sum_{n \geq 0} h_n \left( x_1^2, x_2^2, x_3^2, \ldots \right) (t^2)^n.
\]

Note, the bilinear form \( (\cdot, \cdot)_{\Lambda \otimes \Lambda} \) is nondegenerate (in fact, \( (s \mu \otimes s \nu)_{(\mu,\nu)\in Par \times Par} \) is an orthonormal basis with respect to this bilinear form). Hence, if \( U \) and \( V \) are two elements of \( \Lambda \otimes \Lambda \) such that every \( B \in \Lambda \otimes \Lambda \) satisfies \( (U,B)_{\Lambda \otimes \Lambda} = (V,B)_{\Lambda \otimes \Lambda} \), then \( U = V \). Applying this to \( U = ((f_n \otimes f_n) \circ \Delta) \alpha \) and \( V = (\Delta \circ f_n) \alpha \), we obtain \( ((f_n \otimes f_n) \circ \Delta) \alpha = (\Delta \circ f_n) \alpha \) (because of (11.369)).

Since we have proven this for every \( \alpha \in \Lambda \), this yields that \( (f_n \otimes f_n) \circ \Delta = \Delta \circ f_n \).

Thus, \( (f_n \otimes f_n) \circ \Delta = \Delta \circ f_n \) is proven.

\[\text{Proof of (11.370): Let } m \in \mathbb{N}. \text{ Then, (2.6) (applied to } m \text{ instead of } n) \text{ yields } h_m = \sum_{i_1 \leq i_2 \leq \cdots \leq i_m} x_{i_1} x_{i_2} \cdots x_{i_m}. \text{ Thus,}
\]
\[
h_m(x_1^n, x_2^n, x_3^n, \ldots) = \sum_{i_1 \leq i_2 \leq \cdots \leq i_m} x_{i_1} x_{i_2} \cdots x_{i_m} (x_1^n, x_2^n, x_3^n, \ldots) = \sum_{i_1 \leq i_2 \leq \cdots \leq i_m} x_{i_1} x_{i_2} \cdots x_{i_m}.
\]

But the definition of \( f_n \) yields \( f_n(h_m) = f_n(x_1^n, x_2^n, x_3^n, \ldots) = \sum_{i_1 \leq i_2 \leq \cdots \leq i_m} x_{i_1} x_{i_2} \cdots x_{i_m} \). This proves (11.370).
But

\[ (11.374) \quad \prod_{i=1}^{\infty} (1 - x_i t)^{-1} = \sum_{n \geq 0} h_n(x) t^n = \sum_{n \geq 0} h_n t^n. \]

Substituting \(-t\) for \(t\) in this equality, we obtain \(\prod_{i=1}^{\infty} (1 - x_i (-t))^{-1} = \sum_{n \geq 0} h_n (-t)^n = \sum_{n \geq 0} (-1)^n h_n t^n\).

Thus,

\[ (11.375) \quad \sum_{n \geq 0} (-1)^n h_n t^n = \prod_{i=1}^{\infty} \left(1 - x_i (-t) \right)^{-1} = \prod_{i=1}^{\infty} (1 + x_i t)^{-1}. \]

Now,

\[ \sum_{n \geq 0} f_2(h_n) t^{2n} = \sum_{n \geq 0} h_n \left(x_1^2, x_2^2, x_3^2, \ldots \right)^n \quad \text{(by the definition of } f_2) \]

\[ = \prod_{i=1}^{\infty} \left((1 - x_i t)(1 + x_i t)^{-1} \right)^{-1} = \prod_{i=1}^{\infty} \left(1 - x_i t \right)^{-1} \prod_{i=1}^{\infty} \left(1 + x_i t \right)^{-1} \]

\[ = \prod_{i=1}^{\infty} \left(1 - x_i t \right)^{-1} = \prod_{i=1}^{\infty} \left(1 + x_i t \right)^{-1} \]

\[ = \sum_{n \geq 0} \sum_{i=0}^{n} (-1)^i h_i t^n \]

Comparing coefficients before \(t^{2n}\) on both sides of this equality, we obtain

\[ f_2(h_m) = \sum_{i=0}^{2m} h_i \cdot (-1)^{2m-i} = \sum_{i=0}^{2m} h_i \cdot (-1)^i t^{2m-i} \]

Exercise 2.85(e) is thus solved.

(f) Let \(p\) be a prime number, and let \(a \in \Lambda\). We need to prove that \(f_p(a) \equiv a^p \mod p\Lambda\).

Indeed, let us first check that \(f_p(a) \equiv a^p \mod p\mathbb{Z}[x]\). Since \(f_p(a) = (x_1^p, x_2^p, x_3^p, \ldots)\), this is equivalent to showing that \(a(x_1^p, x_2^p, x_3^p, \ldots) = a^p \mod p\mathbb{Z}[x]\). This, in turn, is equivalent to proving that \(\pi(x_1^p, x_2^p, x_3^p, \ldots) = \pi^p\), where \(\pi\) denotes the projection of the power series \(a \in \mathbb{Z}[x]\) onto the ring \((\mathbb{Z}/p\mathbb{Z})[x]\) (by reducing every coefficient modulo \(p\)). So let us prove \(\pi(x_1^p, x_2^p, x_3^p, \ldots) = \pi^p\) now.

Write the power series \(\pi\) in the form \(\pi = \sum_{\beta} \kappa_{\beta} x^\beta\), where the sum ranges over all weak compositions \(\beta\), and where \(\kappa_{\beta}\) is an element of \(\mathbb{Z}/p\mathbb{Z}\) for every weak composition \(\beta\). Taking both sides of this equality to the \(p\)-th power, we obtain

\[ (11.376) \quad \pi^p = \left( \sum_{\beta} \kappa_{\beta} x^\beta \right)^p. \]

In the commutative ring \((\mathbb{Z}/p\mathbb{Z})[[x]]\), we have \(p \cdot 1_{(\mathbb{Z}/p\mathbb{Z})[[x]]} = 0\). Thus, taking the \(p\)-th power is a ring endomorphism of \((\mathbb{Z}/p\mathbb{Z})[[x]]\). This ring endomorphism is moreover continuous (with respect to the usual topology on \((\mathbb{Z}/p\mathbb{Z})[[x]]\)) and \((\mathbb{Z}/p\mathbb{Z})\)-linear (by virtue of being a ring endomorphism). Thus, this endomorphism respects infinite \((\mathbb{Z}/p\mathbb{Z})\)-linear combinations; hence,

\[ (11.377) \quad \left( \sum_{\beta} \kappa_{\beta} x^\beta \right)^p = \sum_{\beta} \kappa_{\beta} (x^\beta)^p. \]
But $\pi = \sum_\beta \kappa_\beta x^\beta$, so that $\pi(x_1^p, x_2^p, x_3^p, \ldots) = \sum_\beta \kappa_\beta (x^\beta)^p$ (because replacing all variables $x_1, x_2, x_3, \ldots$ by their $p$-th powers transforms every monomial $x^\beta$ into $(x^\beta)^p$). Thus,

$$\pi(x_1^p, x_2^p, x_3^p, \ldots) = \sum_\beta \kappa_\beta (x^\beta)^p = \left( \sum_\beta \kappa_\beta x^\beta \right)^p$$

(by (11.377))

$$= \pi^p$$

(by (11.376)).

We thus have proven that $\pi(x_1^p, x_2^p, x_3^p, \ldots) = \pi^p$. As explained above, this yields $f_p(a) = a^p \mod pA[[X]]$. In other words, the power series $f_p(a) - a^p = (a^p - a^p) \mod pA$ follows that it lies in $A$. So we have $f_p(a) - a^p \in A$, thus $f_p(a) - a^p \in pA$ and hence $f_p(a) \equiv a^p \mod pA$. This solves Exercise 2.85(f).

(g) Set $k = Z$. Thus, the sign $\otimes$ will mean $\otimes_Z$ in the remainder of this solution. Also, $\Lambda = \Lambda_Z$. We define the notation $\psi_p(n)$ as in Exercise 2.82.

Let us introduce a notion from commutative algebra:

In the following, a special $\Psi$-ring will mean a pair $\left( A, (\varphi_n)_{n \in \{1,2,3,\ldots\}} \right)$, where $A$ is a commutative ring and $(\varphi_n)_{n \in \{1,2,3,\ldots\}}$ is a family of ring endomorphisms $\varphi_n : A \to A$ of $A$ satisfying the following properties:

- We have $\varphi_n \circ \varphi_m = \varphi_{nm}$ for any two positive integers $n$ and $m$.
- We have $\varphi_1 = \operatorname{id}$.
- We have $\varphi_p(a) \equiv a^p \mod pA$ for every $a \in A$ and every prime number $p$.

The tensor product of two special $\Psi$-rings $\left( A, (\varphi_n)_{n \in \{1,2,3,\ldots\}} \right)$ and $\left( B, (\psi_n)_{n \in \{1,2,3,\ldots\}} \right)$ is defined to be the pair $\left( A \otimes B, (\varphi_n \otimes \psi_m)_{n \in \{1,2,3,\ldots\}} \right)$. This pair $\left( A \otimes B, (\varphi_n \otimes \psi_m)_{n \in \{1,2,3,\ldots\}} \right)$ is a special $\Psi$-ring again.

Here are three examples of special $\Psi$-rings which we will need:

---

**Proof.** We need to prove the following five statements:

1. **Statement 1:** The ring $A \otimes B$ is a commutative ring.
2. **Statement 2:** The family $(\varphi_n \otimes \psi_n)_{n \in \{1,2,3,\ldots\}}$ is a family of ring endomorphisms $\varphi_n \otimes \psi_n : A \otimes B \to A \otimes B$ of $A \otimes B$.
3. **Statement 3:** We have $(\varphi_n \otimes \psi_n) \circ (\varphi_m \otimes \psi_m) = \varphi_{nm} \otimes \psi_{nm}$ for any two positive integers $n$ and $m$.
4. **Statement 4:** We have $\varphi_1 = \psi_1 = \operatorname{id}$.
5. **Statement 5:** We have $(\varphi_p \otimes \psi_p)(a) \equiv a^p \mod p(A \otimes B)$ for every $a \in A \otimes B$ and every prime number $p$.

**Proof of Statement 1:** This is obvious.

**Proof of Statement 2:** The family $(\varphi_n \otimes \psi_n)_{n \in \{1,2,3,\ldots\}}$ is a family of ring endomorphisms of $A$ (since $(A, (\varphi_n)_{n \in \{1,2,3,\ldots\}})$ is a special $\Psi$-ring). Thus, $\varphi_n$ is a ring endomorphism of $A$ for every positive integer $n$. Similarly, $\psi_n$ is a ring endomorphism of $B$ for every positive integer $n$. Thus, $\varphi_n \otimes \psi_n$ is a ring endomorphism of $A \otimes B$ for every positive integer $n$. In other words, Statement 2 holds.

**Proof of Statement 3:** Let $n$ and $m$ be positive integers. Then, $\varphi_n \circ \varphi_m = \varphi_{nm}$ (since $(A, (\varphi_n)_{n \in \{1,2,3,\ldots\}})$ is a special $\Psi$-ring) and $\psi_n \circ \psi_m = \psi_{nm}$ (similarly). Now, $(\varphi_n \otimes \psi_n) \circ (\varphi_m \otimes \psi_m) = (\varphi_n \circ \varphi_m) \otimes (\psi_n \circ \psi_m) = \varphi_{nm} \otimes \psi_{nm}$, and thus $\varphi_n \otimes \varphi_m = \varphi_{nm}$.

**Proof of Statement 4:** Since $(A, (\varphi_n)_{n \in \{1,2,3,\ldots\}})$ is a special $\Psi$-ring, we have $\varphi_1 = \operatorname{id}$. Similarly, $\psi_1 = \operatorname{id}$. Thus, $\varphi_1 \otimes \psi_1 = \operatorname{id} \otimes \operatorname{id} = \operatorname{id}$. This proves Statement 4.

**Proof of Statement 5:** Fix a prime number $p$. For every commutative ring $R$, we introduce three pieces of notation:

- We let $R$ denote the commutative ring $R/pR$.
- We let $\pi_R$ denote the canonical projection $R \to R/pR$.
- We let $\operatorname{pow}_R$ denote the map $R \to R$ which sends every $r \in R$ to $r^p$. This is not a linear map in general, but when $p \cdot 1_R = 0$, the map $\operatorname{pow}_R$ is a ring endomorphism of $R$.

Now, $(A, (\varphi_n)_{n \in \{1,2,3,\ldots\}})$ is a special $\Psi$-ring. Hence, $\varphi_p(a) \equiv a^p \mod pA$ for every $a \in A$. In other words, $\pi_A(\varphi_p(a)) = \pi_A\left( a^p \right) = \pi_A(\operatorname{pow}_A a) = \pi_A(\operatorname{pow}_A a) = \pi_A(\operatorname{pow}_A a) = \pi_A(\operatorname{pow}_A a)$.

Thus, $\pi_A \circ \varphi_p = \pi_A \circ \operatorname{pow}_A = \operatorname{pow}_A \circ \pi_A$.
• The pair \((\mathbb{Z}, (\text{id})_{n \in \{1,2,3,\ldots\}})\) is a special \(\Psi\)-ring. (The proof of this relies on Fermat’s little theorem.)
• The pair \((\Lambda, (\mathbf{f}_n)_{n \in \{1,2,3,\ldots\}})\) is a special \(\Psi\)-ring. (This follows from parts (a), (b), (c) and (f) of Exercise 2.85.)
• The pair \((\Lambda \otimes \mathbb{Z}, (\mathbf{f}_n \otimes \mathbf{f}_n)_{n \in \{1,2,3,\ldots\}})\) is a special \(\Psi\)-ring. (Indeed, this pair is the tensor product of the special \(\Psi\)-ring \((\Lambda, (\mathbf{f}_n)_{n \in \{1,2,3,\ldots\}})\) with itself.)

Now, we will establish a pattern which we will follow in our new solutions to parts (b), (c), (d), (e) and (f) of Exercise 2.80. Namely, let \((A, (\varphi_n)_{n \in \{1,2,3,\ldots\}})\) be a special \(\Psi\)-ring such that the \(\mathbb{Z}\)-module \(A\) is free. Then, \(A\) canonically injects into \(A \otimes \mathbb{Q}\). Identify \(A\) with a subring of \(A \otimes \mathbb{Q}\) using this injection. Also, consider \(\Lambda = \Lambda_2\) as a subring of \(\Lambda_\mathbb{Q}\) as in Exercise 2.80. Let \(f : \Lambda_\mathbb{Q} \to A \otimes \mathbb{Q}\) be a \(\mathbb{Q}\)-algebra homomorphism such that every \(n \in \{1,2,3,\ldots\}\) satisfies \(f(p_n) \in A\). We can then ask for a criterion for \(f(\Lambda) \subset A\). Using Exercise 2.82, we can obtain such an answer:

\[
\text{(11.378) } \begin{cases} 
\text{We have } f(\Lambda) \subset A \text{ if every positive integer } n \text{ and every prime factor } p \text{ of } n \text{ satisfy } \varphi_p \left( f \left( \frac{p_n}{p} \right) \right) \equiv f(p_n) \mod p^{\varphi(p)(n)} A \\
\end{cases}
\]

(426) (The “if” here can be extended to “if and only if”, but we do not need the “only if”.) Let us now give alternative solutions to parts (b), (c), (d), (e) and (f) of Exercise 2.80:

Similarly, \(\pi_B \circ \varphi_p = \pi_B \circ \text{pow}_{A,B} = \text{pow}_{\pi_A \circ \pi_B} \).

For every commutative ring \(R\), the ring \(\mathcal{R}\) satisfies \(p \cdot 1_\mathcal{R} = 0\), and thus \(\text{pow}_{\mathcal{R}, \mathcal{R}}\) is a ring endomorphism of \(\mathcal{R}\). Hence, \(\text{pow}_{\mathcal{R}, \mathcal{R}}\) makes sense. Moreover, we have \(\text{pow}_{\mathcal{R}, \mathcal{R}} \otimes \text{pow}_{\mathcal{R}, \mathcal{R}} = \text{pow}_{\mathcal{R} \otimes \mathcal{R}}\) (because both \(\text{pow}_{\mathcal{R}, \mathcal{R}} \otimes \text{pow}_{\mathcal{R}, \mathcal{R}}\) and \(\text{pow}_{\mathcal{R} \otimes \mathcal{R}}\) are ring endomorphisms of \(\mathcal{R} \otimes \mathcal{R}\) sending pure tensors \(\alpha \otimes \beta \mapsto \alpha^p \otimes \beta^p = (\alpha \otimes \beta)^p\), and this characterizes them uniquely).

There is a canonical isomorphism \(\mathcal{A} \otimes \mathcal{B} \to \mathcal{A} \otimes \mathcal{B}\) (because tensoring is right-exact). We identify \(\mathcal{A} \otimes \mathcal{B}\) with \(\mathcal{A} \otimes \mathcal{B}\) along this isomorphism. Then, \(\pi_{A \otimes B} = \pi_A \otimes \pi_B\). Thus,

\[
\text{(11.378) } \begin{cases} 
\text{We have } f(\Lambda) \subset A \text{ if every positive integer } n \text{ and every prime factor } p \text{ of } n \text{ satisfy } \varphi_p \left( f \left( \frac{p_n}{p} \right) \right) \equiv f(p_n) \mod p^{\varphi(p)(n)} A \\
\end{cases}
\]

(426) (The “if” here can be extended to “if and only if”, but we do not need the “only if”).
Thus, it remains to prove (11.379).

We need to prove that \( \Delta_x (p_n) = p_n \otimes p_n \in \Lambda \otimes \Lambda \).

This will follow from (11.378) (applied to \( \Lambda \otimes \Lambda \)) once we have showed that every positive integer \( n \) and every prime factor \( p \) of \( n \) satisfy

\[
(f_p \otimes f_p) (\Delta_x (p_{n/p})) \equiv \Delta_x (p_n) \mod p^{v_p(n)} (\Lambda \otimes \Lambda).
\]

Thus, it remains to prove (11.379).

Let \( n \) be a positive integer, and let \( p \) be a prime factor of \( n \). The definition of \( f_p (p_{n/p}) \) yields

\[
(f_p \otimes f_p) (p_{n/p}) = p_{n/p} (x_1^{n/p}, x_2^{n/p}, x_3^{n/p}, \ldots) = \left( x_1^{n/p} + x_2^{n/p} + x_3^{n/p} + \cdots \right)^p = p_n.
\]

Now,

\[
(f_p \otimes f_p) (p_{n/p}) = (f_p \otimes f_p) (p_{n/p} \otimes p_{n/p}) = f_p (p_{n/p} \otimes p_{n/p}) = f_p (p_{n/p} \otimes p_{n/p}) \equiv p_n \equiv p_n
\]

Hence, (11.379) holds. Thus, Exercise 2.80(b) is solved again.

Alternative solution to part (c) of Exercise 2.80: We identify \( \mathbb{Z} \otimes \mathbb{Q} \) with \( \mathbb{Q} \). Every \( n \in \{1, 2, 3, \ldots \} \) satisfies \( e_r (p_n) = r \in \mathbb{Z} \).

We need to prove that \( e_r (\Lambda) \subset \mathbb{Z} \). This will follow from (11.378) (applied to \( \Lambda \otimes \mathbb{Q} \)) once we have showed that every positive integer \( n \) and every prime factor \( p \) of \( n \) satisfy

\[
id (e_r (p_{n/p})) \equiv e_r (p_n) \mod p^{v_p(n) \mathbb{Z}}.
\]

But this congruence follows immediately from

\[
id (e_r (p_{n/p})) = e_r (p_{n/p}) = r (by \text{ the definition of } e_r)
\]

Thus, Exercise 2.80(c) is solved again.

Alternative solution to part (d) of Exercise 2.80: The \( \mathbb{Z} \)-module \( \Lambda \) is free, and thus canonically injects into \( \Lambda \otimes \mathbb{Q} \). We use this identification to regard \( \Lambda \) as a subring of \( \Lambda \otimes \mathbb{Q} \). Notice that every \( n \in \{1, 2, 3, \ldots \} \) satisfies \( i_r (p_n) = r p_n \in \Lambda \).

We need to prove that \( i_r (\Lambda) \subset \Lambda \). This will follow from (11.378) (applied to \( \Lambda \otimes \mathbb{Q} \)) once we have showed that every positive integer \( n \) and every prime factor \( p \) of \( n \) satisfy

\[
f_p (i_r (p_{n/p})) \equiv i_r (p_n) \mod p^{v_p(n)} \mathbb{Q}.
\]

But this congruence follows from

\[
f_p (i_r (p_{n/p})) \equiv f_p (r p_{n/p}) = r f_p (p_{n/p}) = r p_n = i_r (p_n) (by \text{ the definition of } i_r).
\]

This solves Exercise 2.80(d) again.

Alternative solution to part (e) of Exercise 2.80: The \( \mathbb{Z} \)-module \( \Lambda \) is free, and thus canonically injects into \( \Lambda \otimes \mathbb{Q} \). We use this identification to regard \( \Lambda \) as a subring of \( \Lambda \otimes \mathbb{Q} \). We also identify \( \Lambda \otimes \mathbb{Q} \) with \( \Lambda \otimes \mathbb{Q} \).
Applying this to the list \((11.382)\) we have showed that every positive integer \(n\) will satisfy
\[ f_p(p_{n/p}) \equiv (sq(p_n)) \mod p^{v_p(n)} \Lambda. \]
But this congruence follows from
\[
f_p \left(\frac{sq(p_{n/p})}{p_{n/p}}\right) = f_p \left(\frac{p_{n/p}^2}{p_{n/p}}\right) = f_p(p_{n/p})^2 \quad \text{(since } f_p \text{ is a ring homomorphism)}
\]
This solves Exercise 2.80(e) again.

Alternative solution to part (f) of Exercise 2.80: The \(\mathbb{Z}\)-module \(\Lambda \otimes \Lambda\) is free, and thus canonically injects into \((\Lambda \otimes \Lambda) \otimes \mathbb{Q}\). We use this identification to regard \(\Lambda \otimes \Lambda\) as a subring of \((\Lambda \otimes \Lambda) \otimes \mathbb{Q}\). We also identify \(\Lambda_{\mathbb{Q} \otimes \Lambda_{\mathbb{Q}}}\) with \((\Lambda \otimes \Lambda) \otimes \mathbb{Q}\). Notice that every \(n \in \{1, 2, 3, \ldots\}\) satisfies \(\Delta_r(p_{n/p}) = \sum_{i=1}^{n-1} \binom{n}{i} p_i \otimes p_{n-i} + q \otimes p_n + p_n \otimes r = \Lambda \otimes \Lambda\).

We need to prove that \(\Delta_r(p_{n/p}) \equiv (\Delta_r(p_n)) \mod p^{v_p(n)} (\Lambda \otimes \Lambda)\).

This will follow from (11.381) applied to \(p_{n/p} = \frac{p_n}{p}\).

Thus, it remains to prove (11.381).

Let \(n\) be a positive integer, and let \(p\) be a prime factor of \(n\). For the sake of brevity, we denote \(r\) by \(p_0\). Then,
\[(11.382) \quad \Delta_r(p_{n/p}) = \sum_{i=0}^{n/p} \binom{n/p}{i} p_i \otimes p_{n/p-i}.
\]

Applying this to \(n/p\) instead of \(n\), we obtain
\[(11.383) \quad \Delta_r(p_{n/p}) = \sum_{i=0}^{n/p} \binom{n/p}{i} p_i \otimes p_{n/p-i}.
\]

**Proof.** We need to check that \(Sq(ab) = (Sq(a))(Sq(b))\) for any \(a \in \Lambda_{\mathbb{Q}}\) and \(b \in \Lambda_{\mathbb{Q}}\). Since \(Sq\) is \(\mathbb{Q}\)-linear, this only needs to be checked on a basis of the \(\mathbb{Q}\)-module \(\Lambda_{\mathbb{Q}}\). For this we use the basis \((p_\lambda)_{\lambda \in \Lambda_{\mathbb{Q}}}\) of \(\Lambda_{\mathbb{Q}}\). Checking the identity \(Sq(ab) = (Sq(a))(Sq(b))\) on this basis amounts to proving that \(Sq(p_\lambda p_\mu) = (Sq(p_\lambda))(Sq(p_\mu))\) for any two partitions \(\lambda\) and \(\mu\). So let \(\lambda\) and \(\mu\) be two partitions. It is clear that there exists a partition \(\nu\) such that \(p_\lambda p_\mu = p_\nu\) (indeed, this \(\nu\) is the partition obtained by sorting the list \((\lambda_1, \lambda_2, \ldots, \lambda_\ell(\mu), \mu_1, \mu_2, \ldots, \mu_\ell(\mu))\) in decreasing order). Consider this \(\nu\). We have
\[
Sq(p_\lambda p_\mu) = Sq(p_\nu) = p_\nu^2 \quad \text{(by the definition of } Sq)\]
\[
= (p_\lambda p_\mu)^2 \quad \text{(since } p_\nu = p_\lambda p_\mu)\]
\[
= \left(\frac{p_\lambda}{p_\mu}\right)^2 \quad \text{(by the definition of } Sq)\]
\[
= (Sq(p_\lambda))(Sq(p_\mu)) \quad \text{(by the definition of } Sq).
\]
which is what we wanted to prove. Thus we have checked that \(Sq(ab) = (Sq(a))(Sq(b))\) for any \(a \in \Lambda_{\mathbb{Q}}\) and \(b \in \Lambda_{\mathbb{Q}}\). Hence, \(Sq\) is a \(\mathbb{Q}\)-algebra homomorphism (since \(Sq(1) = 1\)), qed.
Now, we will need two elementary congruences for binomial coefficients:

- For any $i \in \mathbb{N}$ satisfying $p \mid i$, we have

$$\binom{n/p}{i/p} \equiv \binom{n}{i} \mod p^{v_p(n)}.$$  \hspace{1cm} (This follows from (11.360), applied to $q = 1$ and $r = i/n$.)

- For any $i \in \mathbb{N}$ satisfying $p \nmid i$, we have

$$0 \equiv \binom{n}{i} \mod p^{v_p(n)}.$$  \hspace{1cm} (This follows from (11.360), applied to $q = 1$ and $r = i/n$, keeping in mind that $\binom{a}{b} = 0$ if $b \notin \mathbb{N}$.)

Now, applying the map $f_p \otimes f_p$ to both sides of the equality (11.383), we obtain

$$(f_p \otimes f_p)(\Delta_r(p_{n/p})) = (f_p \otimes f_p)\left(\sum_{i=0}^{n/p} \binom{n/p}{i} p_i \otimes p_{n/p-i}\right)$$

$$= \sum_{i=0}^{n/p} \binom{n/p}{i} f_p(p_i) \otimes f_p(p_{n/p-i})$$

$$= \sum_{i \in \{0, 1, \ldots, n/p\}} \binom{n/p}{i} p_{pi} \otimes p_{n-pi} = \sum_{i \in \{0, 1, \ldots, n\}; p \nmid i} \binom{n/p}{i} p_i \otimes p_{n-i}.$$  \hspace{1cm} (here, we have substituted $i/p$ for $i$ in the sum).

Comparing this with

$$\Delta_r(p_n) = \sum_{i=0}^{n} \binom{n}{i} p_i \otimes p_{n-i} \quad \text{(by (11.382))}$$

$$= \sum_{i \in \{0, 1, \ldots, n\}; p \mid i} \binom{n}{i} p_i \otimes p_{n-i} + \sum_{i \in \{0, 1, \ldots, n\}; p \nmid i} \binom{n}{i} p_i \otimes p_{n-i}$$

$$\equiv \left(\binom{n}{i/p \mod p^{v_p(n)}} \otimes \Lambda\right) \equiv 0 \mod p^{v_p(n)}(\Lambda \otimes \Lambda) \quad \text{(by (11.384))}$$

$$= \sum_{i \in \{0, 1, \ldots, n\}; p \mid i} \binom{n}{i/p} p_i \otimes p_{n-i} + \sum_{i \in \{0, 1, \ldots, n\}; p \nmid i} \binom{n}{i} p_i \otimes p_{n-i} = 0$$

$$= \sum_{i \in \{0, 1, \ldots, n\}; p \mid i} \binom{n}{i/p} p_i \otimes p_{n-i} \mod p^{v_p(n)}(\Lambda \otimes \Lambda),$$

we obtain

$$(f_p \otimes f_p)(\Delta_r(p_{n/p})) \equiv \Delta_r(p_n) \mod p^{v_p(n)}(\Lambda \otimes \Lambda).$$

Hence, (11.381) holds. Thus, Exercise 2.80(f) is solved again.
11.59. **Solution to Exercise 2.86.** Solution to Exercise 2.86. Let us first notice that every positive integer \( n \) satisfies

\[
(11.386) \quad v_n (h_m) = \begin{cases} h_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} \quad \text{for every } m \in \mathbb{N}
\]

(b) Let \( n \) be a positive integer. Recall that a ring of formal power series \( R[[t]] \) over a commutative ring \( R \) has a canonical topology which makes it into a topological \( R \)-algebra. Thus, in particular, \( \Lambda[[t]] \) becomes a topological \( \Lambda \)-algebra. We shall be considering this topology when we speak of continuity.

The \( k \)-algebra homomorphism \( v_n : \Lambda \to \Lambda \) induces a continuous \( k \)-algebra homomorphism \( v_n [[t]] : \Lambda[[t]] \to \Lambda[[t]] \) given by

\[
(v_n [[t]]) \left( \sum_{i \in \mathbb{N}} a_i t^i \right) = \sum_{i \in \mathbb{N}} v_n (a_i) t^i \quad \text{for all } (a_i)_{i \in \mathbb{N}} \in \Lambda[[t]].
\]

Consider the power series \( H(t) = 1 + h_1 (x) t + h_2 (x) t^2 + \cdots = \sum_{i \geq 0} h_i (x) t^i = \sum_{i \geq 0} h_i t^i \)

\]

and similarly \( E(t) = \sum_{i \geq 0} e_i t^i \).

Substituting \( t^n \) for \( t \) in the equality \( H(t) = \sum_{i \geq 0} h_i t^i \), we obtain \( H(t^n) = \sum_{i \geq 0} h_i (t^n)^i \). Substituting \(-t\) for \( t \) in the equality \( E(t) = \sum_{i \geq 0} e_i t^i \), we obtain \( E(-t) = \sum_{i \geq 0} e_i (-t)^i = \sum_{i \geq 0} (-1)^i e_i t^i \). Substituting \(-t^n\) for \( t \) in the equality \( E(t) = \sum_{i \geq 0} e_i t^i \), we obtain \( E(-t^n) = \sum_{i \geq 0} e_i (-t^n)^i \).

---

428 Proof of (11.386): Let \( n \) be a positive integer. Let \( m \in \mathbb{N} \). We need to prove that (11.386) holds.

If \( m \neq 0 \), then \( n \) is a positive integer. Hence, if \( m \neq 0 \), then (11.386) follows immediately from the definition of \( v_n (h_m) \).

Thus, for the rest of this proof of (11.386), we can WLOG assume that we don’t have \( m \neq 0 \). Assume this.

We don’t have \( m \neq 0 \). Thus, we have \( m = 0 \), so that \( h_m = h_0 = 1 \) and therefore \( v_n \left( \frac{h_m}{1} \right) = v_n (1) = 1 \) (since \( v_n \) is a \( k \)-algebra homomorphism).

Also,

\[
\begin{cases} h_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} = h_{m/n} \quad \text{(since } n \mid 0 = m \text{). Now,}
\]

\[
v_n (h_m) = 1 = h_0 = h_{m/n} \quad (\text{since } h_{m/n} = h_0 \quad (\text{since } \frac{m}{n} \text{ is integer}) \text{ and } n = 0/n = 0)
\]

\[
= \begin{cases} h_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases}.
\]

This proves (11.386).
Applying the map \( v_n([t]) \) to both sides of the equality \( H(t) = \sum_{i \geq 0} h_i t^i \), we obtain

\[
(v_n([t]))(H(t)) = (v_n([t])) \left( \sum_{i \geq 0} h_i t^i \right) = \sum_{i \geq 0} v_n(h_i) t^i = \left\{ \begin{array}{ll}
h_i/n, & \text{if } n \mid i; \\
0, & \text{if } n \nmid i
\end{array} \right. \quad \text{(by the definition of } v_n([t]))
\]

\[
= \sum_{i \geq 0} \left\{ \begin{array}{ll}
h_i/n, & \text{if } n \mid i; \\
0, & \text{if } n \nmid i
\end{array} \right. t^i = \sum_{i \geq 0} \left\{ \begin{array}{ll}
h_i/n, & \text{if } n \mid i; \\
0, & \text{if } n \nmid i
\end{array} \right. t^i + \sum_{i \geq 0} \left\{ \begin{array}{ll}
h_i/n, & \text{if } n \mid i; \\
0, & \text{if } n \nmid i
\end{array} \right. t^i
\]

\[
= \sum_{i \geq 0} h_i t^i = \sum_{i \geq 0} \frac{h_i}{n} t^{ni} = \sum_{i \geq 0} h_i t^{ni} = \sum_{i \geq 0} t^{ni} (\text{here, we substituted } ni \text{ for } i \text{ in the sum})
\]

\[
= \sum_{i \geq 0} h_i (t^n)^i = H(t^n).
\]

But (2.16) yields \( 1 = E(-t)H(t) \), so that \( E(-t) = \frac{1}{H(t)} \). Applying the map \( v_n([t]) \) to both sides of this equality, we obtain

\[
(v_n([t]))(E(-t)) = \left( \frac{1}{H(t)} \right) = \frac{1}{(v_n([t]))(H(t))} \quad \text{(since } v_n([t]) \text{ is a } k\text{-algebra homomorphism)}
\]

\[
= \frac{1}{H(t^n)} \quad \text{(by (11.387))}
\]

\[
= E(-t^n) = \left( \begin{array}{c}
\text{because } E(-t^n) = \frac{1}{H(t^n)} \quad \text{(this follows by substituting } t^n \\
\text{for } t \text{ in the equality } E(-t) = \frac{1}{H(t)}
\end{array} \right)
\]

\[
(11.388) \quad = \sum_{i \geq 0} e_i (-t^n)^i = \sum_{i \geq 0} (-1)^i e_it^{ni}.
\]

Now,

\[
\sum_{i \geq 0} \left\{ \begin{array}{ll}
(-1)^{i/n} e_i/n, & \text{if } n \mid i; \\
0, & \text{if } n \nmid i
\end{array} \right. t^i = \sum_{i \geq 0} \left\{ \begin{array}{ll}
(-1)^{i/n} e_i/n, & \text{if } n \mid i; \\
0, & \text{if } n \nmid i
\end{array} \right. t^i + \sum_{i \geq 0} \left\{ \begin{array}{ll}
(-1)^{i/n} e_i/n, & \text{if } n \mid i; \\
0, & \text{if } n \nmid i
\end{array} \right. t^i
\]

\[
= \sum_{i \geq 0} (-1)^{i/n} e_i/n t^i + \sum_{i \geq 0} 0 t^i = \sum_{i \geq 0} (-1)^{i/n} e_i/n t^i
\]

\[
= \sum_{i \geq 0} (-1)^i e_i t^{ni} \quad (\text{here, we substituted } ni \text{ for } i \text{ in the sum})
\]

\[
(11.389) \quad = (v_n([t]))(E(-t)) \quad \text{(by (11.388))}.
\]
But applying the map \( v_n [[t]] \) to both sides of the equality \( E (- t) = \sum_{i \geq 0} (-1)^i e_i t^i \), we obtain

\[
(v_n [[t]]) (E (- t)) = (v_n [[t]]) \left( \sum_{i \geq 0} (-1)^i e_i t^i \right) = \sum_{i \geq 0} v_n \left( (-1)^i e_i \right) t^i \tag{by the definition of \( v_n [[t]] \)}
\]

\[
= \sum_{i \geq 0} (-1)^i v_n (e_i) t^i.
\]

Hence,

\[
\sum_{i \geq 0} (-1)^i v_n (e_i) t^i = (v_n [[t]]) (E (- t)) = \sum_{i \geq 0} \left\{ \begin{array}{ll}
(-1)^{i/n} e_{i/n}, & \text{if } n \mid i; \\
0, & \text{if } n \nmid i
\end{array} \right. t^i
\]

(by (11.389)).

Now, let \( m \) be a positive integer. Comparing coefficients before \( t^m \) in the equality (11.390), we obtain

\[
(-1)^m v_n (e_m) = \left\{ \begin{array}{ll}
(-1)^{m/n} e_{m/n}, & \text{if } n \mid m; \\
0, & \text{if } n \nmid m
\end{array} \right.
\]

Dividing this by \((-1)^m\), we obtain

\[
v_n (e_m) = \left\{ \begin{array}{ll}
\frac{1}{(-1)^m} (-1)^{m/n} e_{m/n}, & \text{if } n \mid m; \\
0, & \text{if } n \nmid m
\end{array} \right.
\]

(because if \( n \mid m \), then \( \frac{1}{(-1)^m} (-1)^{m/n} = (-1)^{m/m} = (-1)^{m/m} \)). This solves Exercise 2.48(b).

(a) Let \( n \) be a positive integer. We define the continuous \( k \)-algebra homomorphism \( v_n [[t]] : \Lambda [[t]] \rightarrow [[t]] \) as in the solution of Exercise 2.86(a).

Consider the power series \( H (t) \in \Lambda [[t]] \) defined in the proof of Proposition 2.28. We have \( H (t) = \sum_{i \geq 0} h_i t^i \) (this can be proven just as in the solution of Exercise 2.86(a)), so that \( H' (t) = \sum_{i \geq 1} i h_i t^{i-1} \) (by the definition of the derivative of a power series). We can also see that the equality (11.387) holds (this can be proven just as in the solution of Exercise 2.86(a)). The power series \( H (t) \) is invertible (since its constant term is \( h_0 = 1 \)).

Exercise 2.48 yields \( \sum_{m \geq 0} p_{m+1} t^m = \frac{H' (t)}{H (t)} \), so that

\[
\frac{H' (t)}{H (t)} = \sum_{m \geq 0} p_{m+1} t^m = \sum_{i \geq 1} p_i t^{i-1} \tag{here, we substituted \( i \) for \( m + 1 \) in the sum}.
\]

Multiplying this equality with \( t \), we obtain

\[
t \cdot \frac{H' (t)}{H (t)} = t \cdot \sum_{i \geq 1} p_i t^{i-1} = \sum_{i \geq 1} p_i t^{i-1} = \sum_{i \geq 1} p_i t^i.
\]

Multiplying this equality with \( H (t) \), we obtain

\[
t \cdot H' (t) = H (t) \cdot \sum_{i \geq 1} p_i t^i.
\]

Hence,

\[
H (t) \cdot \sum_{i \geq 1} p_i t^i = t \cdot \frac{H' (t)}{H (t)} = t \cdot \sum_{i \geq 1} i h_i t^{i-1} = \sum_{i \geq 1} i h_i t^{i-1} = \sum_{i \geq 1} i h_i t^i.
\]

Substituting \( t^n \) for \( t \) in this equality, we obtain

\[
H (t^n) \cdot \sum_{i \geq 1} p_i (t^n)^i = \sum_{i \geq 0} i h_i (t^n)^i.
\]
Applying the map $v_n[[t]]$ to both sides of the equality (11.391), we obtain

$$
(v_n[[t]]) \left( H(t) \cdot \sum_{i \geq 1} p_i t^i \right)
$$

$$
= (v_n[[t]]) \left( \sum_{i \geq 1} ih_i t^i \right) = \sum_{i \geq 1} i \left( v_n[[t]](h_i) \right) \left( v_n[[t]](t^i) \right)
$$

(by the definition of $v_n[[t]]$) (by the definition of $v_n[[t]]$)

$$
= \sum_{i \geq 1} i \left( v_n(h_i) \right) t^i
$$

(since $v_n[[t]]$ is a continuous $k$-algebra homomorphism)

$$
= \sum_{i \geq 1} i \left( \frac{h_i}{n} \right) t^i
$$

(by the definition of $v_n$)

$$
= \sum_{i \geq 1} i \left( \frac{h_i}{n} \right) t^i = \sum_{i \geq 1; n \mid i} i \left( \frac{h_i}{n} \right) t^i + \sum_{i \geq 1; n \nmid i} i \left( \frac{h_i}{n} \right) t^i
$$

$$
= \sum_{i \geq 1; n \mid i} i \frac{h_i}{n} t^i = \sum_{i \geq 1; n \mid i} i \frac{h_i}{n} t^i + 0 t^i = \sum_{i \geq 1; n \mid i} i \frac{h_i}{n} t^i
$$

(since $\frac{h_i}{n}$ is a continuous $k$-algebra homomorphism)

$$
= \sum_{i \geq 1; n \mid i} i \frac{h_i}{n} t^i
$$

(here, we substituted $ni$ for $i$ in the sum)

$$
= \sum_{i \geq 1; n \mid i} i \frac{h_i}{n} t^i
$$

(by (11.392))

$$
= \sum_{i \geq 1; n \mid i} i \frac{h_i}{n} t^i = \sum_{i \geq 1; n \mid i} i h_i (t^n)^i = n \cdot H(t^n) \cdot \sum_{i \geq 1} p_i (t^n)^i = H(t^n) \cdot n \cdot \sum_{i \geq 1} p_i (t^n)^i.
$$

Compared with

$$
(v_n[[t]]) \left( H(t) \cdot \sum_{i \geq 1} p_i t^i \right)
$$

$$
= (v_n[[t]]) \left( H(t) \cdot \sum_{i \geq 1} p_i t^i \right)
$$

(since $v_n[[t]]$ is a $k$-algebra homomorphism)

$$
= H(t^n) \cdot \left( v_n[[t]] \right) \left( \sum_{i \geq 1} p_i t^i \right)
$$

(by (11.387))

$$
= H(t^n) \cdot \left( v_n[[t]] \right) \left( \sum_{i \geq 1} p_i t^i \right)
$$

this becomes

$$
H(t^n) \cdot n \cdot \sum_{i \geq 1} p_i (t^n)^i = H(t^n) \cdot \left( v_n[[t]] \right) \left( \sum_{i \geq 1} p_i t^i \right).
$$
We can divide both sides of this equality by $H (t^n)$ (since $H (t^n)$ is invertible\textsuperscript{429}). As a result, we obtain

$$n \cdot \sum_{i \geq 1} p_i (t^n)^i = (v_n [[t]]) \left( \sum_{i \geq 1} p_i t^i \right) = \sum_{i \geq 1} \left( v_n [t] \right) (p_i) \left( t^i \right)$$

(by the definition of $v_n [t]$) (by the definition of $v_n [t]$)

(since $v_n [[t]]$ is a continuous $k$-algebra homomorphism)

(11.393) \[ = \sum_{i \geq 1} v_n (p_i) t^i. \]

On the other hand,

$$\sum_{i \geq 1} \begin{cases} np_i/n, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases} t^i$$

$$= \sum_{i \geq 1; n \mid i} \begin{cases} np_i/n, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases} t^i + \sum_{i \geq 1; n \nmid i} \begin{cases} np_i/n, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases} t^i$$

$$= \sum_{i \geq 1; n \mid i} np_i/n t^i + \sum_{i \geq 1; n \nmid i} 0 t^i = \sum_{i \geq 1} np_i/n t^i = \sum_{i \geq 1} np_i t^{ni} (t^n)^i$$

(here, we have substituted $ni$ for $i$ in the sum)

(11.394) \[ = \sum_{i \geq 1} np_i (t^n)^i = n \cdot \sum_{i \geq 1} p_i (t^n)^i = \sum_{i \geq 1} v_n (p_i) t^i \quad \text{(by (11.393))}. \]

Now, let $m$ be a positive integer. Comparing coefficients before $t^m$ in the equality (11.394), we obtain

$$\begin{cases} np_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} = v_n (p_m).$$

This solves Exercise 2.48(a).

(c) Fix two positive integers $n$ and $m$. We need to prove that $v_n \circ v_m = v_{nm}$. But both $v_n \circ v_m$ and $v_{nm}$ are $k$-algebra homomorphisms (since $v_n$, $v_m$ and $v_{nm}$ are $k$-algebra homomorphisms).

Let us now show that

(11.395) \[ (v_n \circ v_m) (h_r) = v_{nm} (h_r) \quad \text{for every positive integer } r. \]

\textbf{Proof of (11.395):} Let $r$ be a positive integer. If $m \nmid r$, then (11.395) holds\textsuperscript{430}. Thus, for the rest of this proof, we can WLOG assume that we don’t have $m \nmid r$. Assume this.

\textsuperscript{429}Proof.} We know that the power series $H (t)$ is invertible. That is, there exists a power series $A (t) \in \Lambda [[t]]$ such that $A (t) \cdot H (t) = 1$. Consider this $A$. Substituting $t^n$ for $t$ in the equality $A (t) \cdot H (t) = 1$, we obtain $A (t^n) \cdot H (t^n) = 1$. Hence, the power series $H (t^n)$ is invertible (and its inverse is $A (t^n)$), qed.

\textsuperscript{430}Proof.} Assume that $m \mid r$. Then, $nm \mid r$ (because otherwise, we would have $nm \nmid r$, and therefore $m \mid nm \nmid r$, which would contradict $m \mid r$).

By the definition of $v_m$, we have $v_m (h_r) = \begin{cases} h_{r/m}, & \text{if } m \mid r; \\ 0, & \text{if } m \nmid r \end{cases} = 0$ (since $m \mid r$). By the definition of $v_{nm}$, we have

$$v_{nm} (h_r) = \begin{cases} h_{r/(nm)}, & \text{if } nm \mid r; \\ 0, & \text{if } nm \nmid r \end{cases} = 0$$

(since $nm \mid r$). Now, $(v_n \circ v_m) (h_r) = v_n \left( v_m (h_r) \right) = v_n (0) = 0$ (since $v_n$ is $k$-linear). Compared with $v_{nm} (h_r) = 0$, this yields $(v_n \circ v_m) (h_r) = v_{nm} (h_r)$. Hence, (11.395) holds is proven under the assumption that $m \mid r$. Qed.
We have \( m \mid r \) (since we don’t have \( m \upharpoonright r \)). Thus, \( r/m \) is a positive integer. By the definition of \( v_m \), we have \( v_m(h_r) = \begin{cases} h_{r/m}, & \text{if } m \mid r; \\ 0, & \text{if } m \upharpoonright r \end{cases} = h_{r/m} \) (since \( m \mid r \)). Now,

\[
(v_n \circ v_m)(h_r) = v_n \left( \frac{h_{r/m}}{h_{r/m}} \right) = v_n(h_{r/m}) = \begin{cases} h_{(r/m)/n}, & \text{if } n \mid r/m; \\ 0, & \text{if } n \upharpoonright r/m \end{cases} = \begin{cases} h_{(r/m)/n}, & \text{if } nm \mid r; \\ 0, & \text{if } nm \upharpoonright r \end{cases}
\]

since the condition \( n \mid r/m \) is equivalent to the condition \( nm \mid r \),

and since the condition \( n \mid r/m \) is equivalent to the condition \( nm \upharpoonright r \)

\[
= \begin{cases} h_{r/(nm)}, & \text{if } nm \mid r; \\ 0, & \text{if } nm \upharpoonright r \end{cases}
\]

(since \( (r/m)/n = r/(nm) \)).

Compared with

\[
v_{nm}(h_r) = \begin{cases} h_{r/(nm)}, & \text{if } nm \mid r; \\ 0, & \text{if } nm \upharpoonright r \end{cases}
\]

(by the definition of \( v_{nm} \)),

this yields \((v_n \circ v_m)(h_r) = v_{nm}(h_r)\). This proves (11.395).

Now, recall that the family \((h_r)_{r \geq 1}\) generates the \( k \)-algebra \( \Lambda \) (according to Proposition 2.28). In other words, \((h_r)_{r \geq 1}\) is a generating set of the \( k \)-algebra \( \Lambda \). The two \( k \)-algebra homomorphisms \( v_n \circ v_m \) and \( v_{nm} \) are equal to each other on this generating set (according to (11.395)), and therefore must be identical (because if two \( k \)-algebra homomorphisms from the same domain are equal to each other on a generating set of their domain, then these two homomorphisms must be identical). In other words, \( v_n \circ v_m = v_{nm} \). This solves Exercise 2.86(c).

(d) For every positive integer \( r \), we have

\[
v_1(h_r) = \begin{cases} h_{r/1}, & \text{if } 1 \mid r; \\ 0, & \text{if } 1 \upharpoonright r \end{cases} = \begin{cases} h_{r/1}, & \text{if } 1 \mid r; \\ 0, & \text{if } 1 \upharpoonright r \end{cases}
\]

(by the definition of \( v_1 \))

\[
= h_{r/1} \quad \text{(since } 1 \mid r \text{)}
\]

(11.396)

Now, recall that the family \((h_r)_{r \geq 1}\) generates the \( k \)-algebra \( \Lambda \) (according to Proposition 2.28). In other words, \((h_r)_{r \geq 1}\) is a generating set of the \( k \)-algebra \( \Lambda \). The two \( k \)-algebra homomorphisms \( v_1 \) and \( \text{id} \) are equal to each other on this generating set (according to (11.396)), and therefore must be identical (because if two \( k \)-algebra homomorphisms from the same domain are equal to each other on a generating set of their domain, then these two homomorphisms must be identical). In other words, \( v_1 = \text{id} \). This solves Exercise 2.86(d).

(e) Fix a positive integer \( n \). We need to show that \( v_n : \Lambda \to \Lambda \) is a Hopf algebra homomorphism.

We know that \( v_n : \Lambda \to \Lambda \) is a \( k \)-algebra homomorphism. Thus, \( v_n \otimes v_n : \Lambda \otimes \Lambda \to \Lambda \otimes \Lambda \) is a \( k \)-algebra homomorphism. Also, \( \Delta : \Lambda \to \Lambda \otimes \Lambda \) is a \( k \)-algebra homomorphism (due to the axioms of a bialgebra, since \( \Lambda \) is a bialgebra). Hence, \( \Delta \circ v_n \) and \( (v_n \otimes v_n) \circ \Delta \) are \( k \)-algebra homomorphisms.

For every \( q \in \mathbb{N} \), we have

\[
\Delta(h_q) = \sum_{i \in \{0, 1, \ldots, q\}} h_i \otimes h_{q-i}.
\]

(11.397)

Furthermore, for every \( q \in \mathbb{N} \), we have

\[
v_n(h_{nq}) = h_q
\]

(11.398)

\footnote{Proof of (11.397): Let \( q \in \mathbb{N} \). Then, Proposition 2.24(iii) (applied to \( q \) instead of \( n \)) yields

\[
\Delta(h_q) = \sum_{i+j=q} h_i \otimes h_j = \sum_{i+j=q} h_i \otimes h_{q-i}
\]

(here, we have substituted \((i, q-i)\) for \((i, j)\) in the sum), qed.}
Now, we shall prove that

\[(11.399) \quad (\Delta \circ v_n)(h_r) = ((v_n \otimes v_n) \circ \Delta)(h_r) \quad \text{for every positive integer } r.\]

\textbf{Proof of (11.399):} Let } r \text{ be a positive integer. We need to prove (11.399).

Let us first notice that

\[(11.400) \quad (v_n \otimes v_n)(h_i \otimes h_{r-i}) = 0 \quad \text{for every } i \in \{0, 1, \ldots, r\} \text{ which does not satisfy } (n \mid i \text{ and } n \mid r-i).\]

In particular,

\[(11.401) \quad (v_n \otimes v_n)(h_i \otimes h_{r-i}) = 0 \quad \text{for every } i \in \{0, 1, \ldots, r\} \text{ satisfying } n \nmid i\]

Let us now assume that \(n \nmid r\). Then, (11.386) (applied to \(m = r\)) yields \(v_n(h_r) = \begin{cases} h_{r/n}, & \text{if } n \mid r; \\ 0, & \text{if } n \nmid r \end{cases}\) = 0 (since \(n \nmid r\)), so that \((\Delta \circ v_n)(h_r) = \Delta \left(\frac{v_n(h_r)}{\Delta(0)}\right) = \Delta(0) = 0\). Also,

\[(11.402) \quad (v_n \otimes v_n)(h_i \otimes h_{r-i}) = 0 \quad \text{for every } i \in \{0, 1, \ldots, r\}\]

\textbf{Proof of (11.398):} Let } q \in \mathbb{N}. \text{ Applying (11.386) to } nq \text{ instead of } m, \text{ we obtain}

\[v_n(h_{nq}) = \begin{cases} h_{(nq)/n}, & \text{if } n \mid nq; \\ 0, & \text{if } n \nmid nq \end{cases}\]

\(= h_q, \) (since \(n \nmid nq\))

\text{qed.}

\textbf{Proof of (11.400):} Let } i \in \{0, 1, \ldots, r\} \text{ be such that we don’t have } (n \mid i \text{ and } n \mid r-i). \text{ If } n \nmid i, \text{ then}

\[(v_n \otimes v_n)(h_i \otimes h_{r-i}) = v_n(h_i) \otimes v_n(h_{r-i}) = \begin{cases} h_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{cases} = 0 \text{ (by (11.386), applied to } m=i)\]

Hence, if \(n \nmid i, \) then (11.400) is proven. Thus, for the rest of this proof of (11.400), we can WLOG assume that we don’t have \(n \mid i. \) Assume this.

We have \(n \nmid i \) (since we don’t have \(n \mid i). \) Hence, we don’t have \(n \mid r-i \) (because otherwise, we would have \((n \mid i \text{ and } n \mid r-i), \) which would contradict the fact that we don’t have \((n \mid i \text{ and } n \mid r-i)). \) In other words, we have \(n \nmid r-i. \) Now,

\[(v_n \otimes v_n)(h_i \otimes h_{r-i}) = v_n(h_i) \otimes v_n(h_{r-i}) = \begin{cases} h_{(r-i)/n}, & \text{if } n \mid r-i; \\ 0, & \text{if } n \nmid r-i \end{cases} = 0 \text{ (since } n\nmid r-i)\]

This proves (11.400).

\textbf{Proof of (11.401):} Let } i \in \{0, 1, \ldots, r\} \text{ be such that } n \nmid i. \text{ Then, we don’t have } n \mid i. \text{ Hence, we don’t have } (n \mid i \text{ and } n \mid r-i). \text{ Thus, (11.401) follows from (11.400), qed.}
Now, \[
((v_n \otimes v_n) \circ \Delta)(h_r) = (v_n \otimes v_n) \left( \frac{\Delta(h_r)}{\sum_{i \in \{0, \ldots, r\}} h_i \otimes h_{r-i}} \right) = (v_n \otimes v_n) \left( \sum_{i \in \{0, \ldots, r\}} h_i \otimes h_{r-i} \right)
\]

\[
= \sum_{i \in \{0, \ldots, r\}} (v_n \otimes v_n) (h_i \otimes h_{r-i}) \quad \text{(since the map } v_n \otimes v_n \text{ is } k\text{-linear})
\]

\[
= \sum_{i \in \{0, \ldots, r\}} 0 = 0.
\]

Compared to \((\Delta \circ v_n)(h_r) = 0\), this yields \((\Delta \circ v_n)(h_r) = ((v_n \otimes v_n) \circ \Delta)(h_r)\). In other words, \((11.399)\) holds.

Now, let us forget that we assumed that \(n \nmid r\). We thus have proven that \((11.399)\) holds under the assumption that \(n \nmid r\). Hence, for the rest of the proof of \((11.399)\), we can WLOG assume that we don’t have \(n \nmid r\). Assume this.

We have \(n \mid r\) (since we don’t have \(n \nmid r\)). Hence, \(r/n\) is a positive integer. Denote this positive integer \(r/n\) by \(s\). Then, \(s = r/n\), so that \(r = ns\). The equality \((11.386)\) (applied to \(m = r\)) yields \(v_n(h_r) = \left\{ \begin{array}{ll} h_{r/n}, & \text{if } n \mid r; \\ 0, & \text{if } n \nmid r \end{array} \right\} = h_{r/n} \) (since \(n \mid r\)), and we have

\[
(11.403) \quad (\Delta \circ v_n)(h_r) = \Delta \left( \frac{v_n(h_r)}{=} h_{r/n} = h_s \right. \left. \quad \text{(since } r/n = s) \right) = \Delta(h_s) = \sum_{i \in \{0, \ldots, s\}} h_i \otimes h_{s-i}
\]

(by \((11.397)\), applied to \(q = s\)).

---

\textit{Proof of }\((11.402)\): Let \(i \in \{0, 1, \ldots, r\}\).

Assume (for the sake of contradiction) that \((n \mid i \text{ and } n \mid r - i)\). Then, \(i \equiv 0 \mod n\) (since \(n \mid i\)) and \(r \equiv i \mod n\) (since \(n \mid r - i\)). Hence, \(r \equiv i \equiv 0 \mod n\), so that \(n \mid r\). This contradicts \(n \nmid r\). This contradiction shows that our assumption (that \((n \mid i \text{ and } n \mid r - i)\)) was wrong. Hence, we do not have \((n \mid i \text{ and } n \mid r - i)\). Thus, \((11.402)\) follows from \((11.400)\), qed.
On the other hand, 

\[(v_n \otimes v_n) \circ (h_r) = (v_n \otimes v_n) \left( \sum_{i \in \{0, \ldots, r\}} h_i \otimes h_{r-i} \right) = (v_n \otimes v_n) \left( \sum_{i \in \{0, \ldots, r\}} h_i \otimes h_{r-i} \right) \]

\[
= \sum_{i \in \{0, \ldots, r\}} (v_n \otimes v_n) (h_i \otimes h_{r-i}) \quad \text{(since the map } v_n \otimes v_n \text{ is } k\text{-linear)}
\]

\[
= \sum_{i \in \{0, \ldots, r\}; n | i} (v_n \otimes v_n) (h_i \otimes h_{r-i}) + \sum_{i \in \{0, \ldots, r\}; n \nmid i} (v_n \otimes v_n) (h_i \otimes h_{r-i}) \quad \text{(by } (11.401))
\]

\[
= \sum_{i \in \{0, \ldots, r\}; n | i} v_n (h_i) \otimes v_n (h_{r-i}) + \sum_{i \in \{0, \ldots, r\}; n \nmid i} 0 = \sum_{i \in \{0, \ldots, r\}; n | i} v_n (h_i) \otimes v_n (h_{r-i})
\]

\[
= \sum_{i \in \{0, \ldots, r\}/n} v_n (h_{ni}) \otimes v_n (h_{r-ni} \otimes (h_{r-ni})^{-1}) \quad \text{(since } r-ni=n(s-i))
\]

\[
= \sum_{i \in \{0, \ldots, s\}} v_n (h_{ni}) \otimes v_n (h_{n(s-i)}) = \sum_{i \in \{0, \ldots, s\}} h_i \otimes h_{s-i}.
\]

Compared with (11.403), this yields \((\Delta \circ v_n) (h_r) = ((v_n \otimes v_n) \circ \Delta) (h_r). Thus, (11.399) is proven.

Now, recall that the family \((h_r)_{r \geq 1}\) generates the \(k\text{-algebra } \Lambda \) (according to Proposition 2.28). In other words, \((h_r)_{r \geq 1}\) is a generating set of the \(k\text{-algebra } \Lambda \). The two \(k\text{-algebra homomorphisms } \Delta \circ v_n \) and \((v_n \otimes v_n) \circ \Delta \) are equal to each other on this generating set (according to (11.399)), and therefore must be identical (because if two \(k\text{-algebra homomorphisms from the same domain are equal to each other on a generating set of their domain, then these two homomorphisms must be identical}). In other words, \(\Delta \circ v_n = (v_n \otimes v_n) \circ \Delta \).

One can similarly check that \(e \circ v_n = e\). We can now conclude that the map \(v_n\) is a \(k\text{-coalgebra homomorphism (since it is } k\text{-linear and satisfies } \Delta \circ v_n = (v_n \otimes v_n) \circ \Delta \text{ and } e \circ v_n = e\), therefore a \(k\text{-bialgebra homomorphism (since it also is a } k\text{-algebra homomorphism), and therefore a Hopf algebra homomorphism (by Proposition 1.46(c))). This solves Exercise 2.86(e).}

(f) Recall first that \((h_\lambda)_{\lambda \in \Par} \) and \((m_\lambda)_{\lambda \in \Par} \) are mutually dual bases with respect to the Hall inner product on \(\Lambda \) (according to Corollary 2.45(a)). Thus,

\[(11.404) \quad (m_\lambda, h_\mu) = \delta_{\lambda, \mu} \quad \text{for any two partitions } \lambda \text{ and } \mu.\]

Let us introduce a notation: For every weak composition \(\lambda \) and every positive integer \(s\), let \(\lambda \{s\} \) denote the weak composition \((s\lambda_1, s\lambda_2, s\lambda_3, \ldots)\), where \(\lambda \) is written in the form \((\lambda_1, \lambda_2, \lambda_3, \ldots)\). Notice that if \(\lambda\) is a partition and \(s\) is a positive integer, then \(\lambda \{s\} \) is a partition as well, and satisfies

\[(11.405) \quad \ell \lambda \{s\} = \ell \lambda \cdot s.\]

We have

\[(11.406) \quad f_s m_\lambda = m_{\lambda \{s\}} \quad \text{for every partition } \lambda\]
We need to prove that the maps $f_n : \Lambda \to \Lambda$ and $v_n : \Lambda \to \Lambda$ are adjoint with respect to the Hall inner product on $\Lambda$. In other words, we need to prove that

$$\langle f_n a, b \rangle = \langle a, v_n b \rangle$$

for any $a \in \Lambda$ and $b \in \Lambda$.

**Proof of (11.413):** Fix $a \in \Lambda$ and $b \in \Lambda$.

**Proof of (11.406):** Let $\lambda$ be a partition. Write $\lambda$ in the form $\lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots)$. By the definition of $m_\lambda$, we have

$$m_\lambda = \sum_{\alpha \in \Theta(\infty) \lambda} x^\alpha.$$

But the definition of $f_n$ yields

$$f_n m_\lambda = \sum_{\alpha \in \Theta(\infty) \lambda} (x_1^n, x_2^n, \ldots) = \sum_{\alpha \in \Theta(\infty) \lambda} x^\alpha (x_1^n, x_2^n, \ldots).$$

Thus, (11.408) is proven.

Now, (11.407) becomes

$$f_n m_\lambda = \sum_{\alpha \in \Theta(\infty) \lambda} (x_1^n, x_2^n, \ldots) = \sum_{\alpha \in \Theta(\infty) \lambda} x^\alpha (n).$$

On the other hand, for every weak composition $\alpha$ and every permutation $\sigma \in \Theta(\infty)$, we have

$$\sigma(\alpha(n)) = (\sigma\alpha)(n).$$

**Proof of (11.410):** Let $\alpha$ be a weak composition. Let $\sigma \in \Theta(\infty)$.

Theorem (11.408) yields $x^\alpha = (x_1^{\alpha_1}, x_2^{\alpha_2}, \ldots)$ (by the definition of $x^\alpha$). Then

$$x^\alpha(n) = x_1^{\alpha_1(n)} x_2^{\alpha_2(n)} \cdots.$$

This proves (11.410).

Now, (11.411) and (11.412) are also proved.

**Proof of (11.411):** Let $\alpha \in \Theta(\infty)$.

Theorem (11.408) yields $x^\alpha = (x_1^{\alpha_1}, x_2^{\alpha_2}, \ldots)$ (by the definition of $x^\alpha$). Then

$$\sigma(\alpha(n)) = (\sigma\alpha)(n).$$

This proves (11.411).

Now, (11.412) is also proved.

**Proof of (11.412):** Let $\sigma$ be an element of $\Theta(\infty)$. Then, every element $\sigma \in \Theta(\infty)$ has the form $\alpha \{n\}$ for some $\alpha \in \Theta(\infty)$.
Recall that \((m_{\lambda})_{\lambda \in \text{Par}}\) is a basis of the \(k\)-module \(\Lambda\). Hence, in proving (11.413), we can WLOG assume that \(a\) is an element of this basis \((m_{\lambda})_{\lambda \in \text{Par}}\) (because the equality (11.413) is \(k\)-linear in \(a\)). Assume this. Thus, \(a\) is an element of the basis \((m_{\lambda})_{\lambda \in \text{Par}}\). In other words, there exists a \(\mu \in \text{Par}\) such that \(a = m_{\mu}\). Consider this \(\mu\). We have \(f_{n}a = f_{n}m_{\mu} = m_{\mu(n)}\) (by (11.406), applied to \(\lambda = \mu\)).

Recall that \((h_{\lambda})_{\lambda \in \text{Par}}\) is a basis of the \(k\)-module \(\Lambda\). Hence, in proving (11.413), we can WLOG assume that \(b\) is an element of this basis \((h_{\lambda})_{\lambda \in \text{Par}}\) (because the equality (11.413) is \(k\)-linear in \(b\)). Assume this. Thus, \(b\) is an element of the basis \((h_{\lambda})_{\lambda \in \text{Par}}\). In other words, there exists a \(\nu \in \text{Par}\) such that \(b = h_{\nu}\). Consider this \(\nu\). We have

\[
(11.414) \quad \left( f_{n}a, \frac{b}{m_{\mu(n)}} = h_{\nu} \right) = (m_{\mu(n)}, h_{\nu}) = \delta_{\mu(n), \nu}
\]

(by (11.404), applied to \(\mu \{n\}\) and \(\nu\) instead of \(\lambda\) and \(\mu\)).

Let us write the partition \(\nu\) in the form \((\nu_{1}, \nu_{2}, \nu_{3}, \ldots)\). Then, \(\nu = (\nu_{1}, \nu_{2}, \ldots, \nu_{\ell(\nu)})\), so that \(h_{\nu} = h_{\nu_{1}}h_{\nu_{2}} \cdots h_{\nu_{\ell(\nu)}}\) (by the definition of \(h_{\nu}\)).

Let us first assume that

\[
(11.415) \quad \text{(there exists an} \ i \ \text{in} \ \{1, 2, 3, \ldots \} \ \text{such that} \ n \nmid \nu_{i}).
\]

Then, \(v_{n}(h_{\nu}) = 0\). Also, \(\{n\} \neq \nu\), so that \(\delta_{\mu(n), \nu} = 0\). Thus, (11.414) becomes \((f_{n}a, b) = \delta_{\mu(n), \nu} = 0\). Compared with \((a, v_{n}b) = (a, v_{n}(h_{\nu})) = (a, 0) = 0\) (since the Hall inner product is \(k\)-bilinear), this yields \((f_{n}a, b) = (a, v_{n}b)\). Thus, (11.413) holds.

Now, let us forget that we assumed that (11.415) holds. We thus have proven (11.413) under the assumption that (11.415) holds. Hence, for the rest of our proof of (11.413), we can WLOG assume that (11.415) does not hold. Assume this.

\(\Phi_{(n)} : S_{(n)} \rightarrow S_{(n + 1)}\) is bijective. Compared with

\[
m_{\lambda(n)} = \sum_{\alpha \in S_{(n)}(\lambda(n))} x^{\alpha} \quad \text{(by the definition of} \ m_{\lambda(n)})
\]

this yields \(f_{n}m_{\lambda} = m_{\lambda(n)}\). Thus, (11.406) is proven.

437\textbf{Proof.} There exists an \(i \in \{1, 2, 3, \ldots \}\) such that \(n \nmid \nu_{i}\). Consider this \(i\). We have \(n \nmid \nu_{i}\), thus \(\nu_{i} \neq 0\), and therefore \(\nu_{i}\) is a positive integer. Hence, \(i \leq \ell(\nu)\), so that \(i \in \{1, 2, \ldots, \ell(\nu)\}\). Also, the definition of \(v_{n}\) yields

\[
v_{n}(h_{\nu}) = \begin{cases} h_{\nu_{i}/n} & \text{if} \ n \nmid \nu_{i}; \\ 0 & \text{if} \ n \nmid \nu_{i} \end{cases} = 0 \quad \text{(since} \ n \nmid \nu_{i})
\]

Applying the map \(v_{n}\) to both sides of this equality, we obtain

\[
v_{n}(h_{\nu}) = v_{n} \left( h_{\nu_{1}} h_{\nu_{2}} \cdots h_{\nu_{\ell(\nu)}} \right) = v_{n} \left( h_{\nu_{1}} h_{\nu_{2}} \cdots h_{\nu_{\ell(\nu)}} \right) \quad \text{(since} \ v_{n} \text{is a} \ k\text{-algebra homomorphism)}
\]

\[
= 0,
\]

\(\text{qed.}

438\textbf{Proof.} Assume the contrary. Thus, \(\mu \{n\} = \nu\). Let us write the partition \(\mu\) in the form \((\mu_{1}, \mu_{2}, \mu_{3}, \ldots)\). Then, \(\mu \{n\} = (n\mu_{1}, n\mu_{2}, n\mu_{3}, \ldots)\) (by the definition of \(\mu \{n\}\)). Hence, \((\nu_{1}, \nu_{2}, \nu_{3}, \ldots) = (\nu_{1}, \nu_{2}, \nu_{3}, \ldots) = (n\mu_{1}, n\mu_{2}, n\mu_{3}, \ldots)\). Thus, every positive integer \(j\) satisfies \(\nu_{j} = n\mu_{j}\).

But (11.415) yields that there exists an \(i \in \{1, 2, 3, \ldots \}\) such that \(n \nmid \nu_{i}\). Consider this \(i\). Now, recall that every positive integer \(j\) satisfies \(\nu_{j} = n\mu_{j}\). Applied to \(j = i\), this yields \(\nu_{i} = n\mu_{i}\), so that \(n \nmid \nu_{i}\). This contradicts \(n \nmid \nu_{i}\). This contradiction proves that our assumption was wrong. \(\text{qed.}\)
We have assumed that (11.415) does not hold. In other words, there exists no \( i \in \{1, 2, 3, \ldots \} \) such that \( n \nmid \nu_i \).

\[(11.416)\]

\[
every \ i \in \{1, 2, 3, \ldots \} \ satisfies \ n \mid \nu_i.\]

Thus, \( \nu_i/n \) is a nonnegative integer for every \( i \in \{1, 2, 3, \ldots \} \). These nonnegative integers satisfy \( \nu_1/n \geq \nu_2/n \geq \nu_3/n \geq \cdots \) (since \( \nu_1 \geq \nu_2 \geq \nu_3 \geq \cdots \) (since \( (\nu_1, \nu_2, \nu_3, \ldots) = \nu \) is a partition)) and \( (\nu_1/n = 0 \ for \ all \ sufficiently \ high \ i) \) (since \( (\nu_1, \nu_2, \nu_3, \ldots) = \nu \) is a partition). Hence, \( (\nu_1/n, \nu_2/n, \nu_3/n, \ldots) \) is a partition. Let us denote this partition by \( \kappa \). We have \( \kappa \{n\} = \nu \)

and thus \( \ell(\kappa) = \ell(\nu) \)

440.

We have \( \nu_n(\nu) = h_\kappa \) 441 and thus

\[(11.417)\]

\[
\left( \begin{array}{cc}
q & b \\
= m_\mu & = h_\nu
\end{array} \right) = \left( \begin{array}{cc}
m_\mu, \nu_n(\nu) \\
= h_\mu & = h_\kappa
\end{array} \right) = (m_\mu, h_\kappa) = \delta_{\mu, \kappa}
\]

(by (11.404), applied to \( \mu \) and \( \kappa \) instead of \( \lambda \) and \( \mu \)).

Now, let us write the partition \( \mu \) in the form \((\mu_1, \mu_2, \mu_3, \ldots)\). Then, \( \mu \{n\} = (n\mu_1, n\mu_2, n\mu_3, \ldots) \) (by the definition of \( \mu \{n\} \)).

Now, we have the following equivalence of assertions:

\[
(\mu \{n\} = \nu) \iff ((n\mu_1, n\mu_2, n\mu_3, \ldots) = (\nu_1, \nu_2, \nu_3, \ldots)) \iff \text{(every } i \in \{1, 2, 3, \ldots \} \text{ satisfies } n\mu_i = \nu_i) \iff \text{(every } i \in \{1, 2, 3, \ldots \} \text{ satisfies } \mu_i = \nu_i) \iff ((\mu_1, \mu_2, \mu_3, \ldots) = (\nu_1/n, \nu_2/n, \nu_3/n, \ldots)) \iff (\mu = \kappa) \text{ (since } (\mu_1, \mu_2, \mu_3, \ldots) = \mu \text{ and } (\nu_1/n, \nu_2/n, \nu_3/n, \ldots) = \kappa)\]

439 \text{Proof.} We have \( \kappa = (\nu_1/n, \nu_2/n, \nu_3/n, \ldots) \). Hence, the definition of \( \kappa \{n\} \) yields \( \kappa \{n\} = (n(\nu_1/n), n(\nu_2/n), n(\nu_3/n), \ldots) = (\nu_1, \nu_2, \nu_3, \ldots) = \nu \), qed.

440 \text{Proof.} Applying (11.405) to \( \kappa \) and \( n \) instead of \( \lambda \) and \( s \), we obtain \( \ell(\kappa \{n\}) = \ell(\kappa) \). Since \( \kappa \{n\} = \nu \), this rewrites as \( \ell(\nu) = \ell(\nu) \), qed.

441 \text{Proof.} We have \( \kappa = (\nu_1/n, \nu_2/n, \nu_3/n, \ldots) \) and therefore \( \kappa = (\nu_1/n, \nu_2/n, \ldots, \nu_{\ell(\nu)/n})/n \). Since \( \ell(\kappa) = \ell(\nu) \), this rewrites as \( \kappa = (\nu_1/n, \nu_2/n, \ldots, \nu_{\ell(\nu)/n}/n) \). Hence, the definition of \( h_\kappa \) yields \( h_\kappa = h_{\nu_1/n}h_{\nu_2/n} \cdots h_{\nu_{\ell(\nu)/n}/n} = \prod_{i=1}^{\ell(\nu)} h_{\nu_i}/n \).

On the other hand, \( \nu = (\nu_1, \nu_2, \ldots, \nu_{\ell(\nu)}) \), so that the definition of \( h_\nu \) yields \( h_\nu = h_{\nu_1}h_{\nu_2} \cdots h_{\nu_{\ell(\nu)}} = \prod_{i=1}^{\ell(\nu)} h_{\nu_i} \). Applying the map \( \nu_n \) to both sides of this equality, we obtain

\[
\nu_n(\nu) = \nu_n \left( \prod_{i=1}^{\ell(\nu)} h_{\nu_i} \right) = \prod_{i=1}^{\ell(\nu)} \nu_n(\nu_{\nu_i}) = \left\{ \begin{array}{ll}
\nu_{\nu_1}/n, & \text{if } n \mid \nu_i; \\
0, & \text{if } n \nmid \nu_i
\end{array} \right.
\]

(by \( (11.436), \text{applied to } m=n \))

\[
= \prod_{i=1}^{\ell(\nu)} \left\{ \begin{array}{ll}
h_{\nu_1}/n, & \text{if } n \mid \nu_i; \\
0, & \text{if } n \nmid \nu_i
\end{array} \right.
\]

\[
= \prod_{i=1}^{\ell(\nu)} h_{\nu_i}/n = h_\kappa
\]

\[
= \prod_{i=1}^{\ell(\nu)} h_{\nu_i}/n, \text{qed.}
\]

(since \( h_\kappa = \prod_{i=1}^{\ell(\nu)} h_{\nu_i}/n \)).
But \((11.414)\) becomes
\[
(f_n(a, b) = \delta_{\mu(n), \nu} = \begin{cases} 1, & \text{if } \mu \{n\} = \nu; \\ 0, & \text{otherwise} \end{cases} \quad \text{(by the definition of } \delta_{\mu(n), \nu})
\]
\[
= \begin{cases} 1, & \text{if } \mu = \kappa; \\ 0, & \text{otherwise} \end{cases} \quad \text{(since } \mu \{n\} = \nu \text{ is equivalent to } \mu = \kappa \text{ (by } \ref{11.418})).
\]
\[
= \delta_{\mu, \kappa} \quad \text{(since } \delta_{\mu, \kappa} = \begin{cases} 1, & \text{if } \mu = \kappa; \\ 0, & \text{otherwise} \end{cases} \text{ (by the definition of } \delta_{\mu, \kappa}))
\]
\[
= (a, v_n, b) \quad \text{(by } \ref{11.417}).
\]
Thus, \((11.413)\) is proven. As we know, this completes the solution of Exercise 2.86(f).

\text{(g) Our solution to Exercise 2.86(g) shall proceed in three steps:}

\begin{itemize}
\item \textit{Step 1:} proving that Exercise 2.86(g) holds if \(k = \mathbb{Q}\).
\item \textit{Step 2:} proving that Exercise 2.86(g) holds if \(k = \mathbb{Z}\).
\item \textit{Step 3:} proving that Exercise 2.86(g) holds in the general case.
\end{itemize}

Let us now get to the details of these three steps:

\textit{Step 1:} We shall prove that Exercise 2.86(g) holds if \(k = \mathbb{Q}\).

Indeed, assume that \(k = \mathbb{Q}\). We know that \(f_n\) is a \(k\)-algebra homomorphism (due to Exercise 2.85(a)), and that \(v_n\) is a \(k\)-algebra homomorphism. Hence, \(v_n \circ f_n\) is a \(k\)-algebra homomorphism. Also, \(id_A \ast id_A \ast \cdots \ast id_A\) is a \(k\)-algebra homomorphism (due to Exercise 1.63(b), applied to \(H = \Lambda, A = \Lambda, k = n\) and \(f_i = id_A\)). In other words, \(id_A^n\) is a \(k\)-algebra homomorphism.

Let \(r\) be a positive integer. We shall now show that \((v_n \circ f_n)(p_r) = id_A^n(p_r)\).

Indeed, it is easy to see that
\[
(11.419) \quad id_A^a(p_r) = ar \quad \text{for every } a \in \mathbb{N}
\]

\textit{Proof of }\((11.419)\): We shall prove \((11.419)\) by induction over \(a\):

\textit{Induction base:} We have \(id_A^0(p_r) = e(p_r) = 0\) (since \(r\) is positive), so that \(id_A^0(p_r) = 0 = 0p_r\). In other words, \((11.419)\)
holds for \(a = 0\). This completes the induction base.

\textit{Induction step:} Let \(A\) be a positive integer. Assume that \((11.419)\) holds for \(a = A - 1\). We now need to show that \((11.419)\) holds for \(a = A\).

We know that \((11.419)\) holds for \(a = A - 1\). In other words, \(id_A^{(A-1)}(p_r) = (A - 1) p_r\). Now, recall that \(\Delta p_r = 1 \otimes p_r + p_r \otimes 1\) (by Proposition 2.24(i), applied to \(r\) instead of \(n\)). On the other hand, \(id_A^{(A-1)}\) is a \(k\)-algebra homomorphism (in fact, this is shown in the same way as we have shown that \(id_A^n\) is a \(k\)-algebra homomorphism), so that we have \(id_A^{(A-1)}(1) = 1\). Now, \(id_A^A = id_A^{(A-1)} \ast id_A = m \circ (id_A^{(A-1)} \otimes id_A) \circ \Delta\) (by the definition of convolution). Applying both sides of this equality to \(p_r\), we obtain
\[
id_A^A(p_r) = m \circ (id_A^{(A-1)} \otimes id_A) \circ \Delta)(p_r) = m \left( \left( id_A^{(A-1)} \otimes id_A \right) \circ \Delta(p_r) \right)
\]
\[
= m \left( \left( id_A^{(A-1)} \otimes id_A \right) (1 \otimes p_r + p_r \otimes 1) \right)
\]
\[
= m \left( id_A^{(A-1)}(1) \otimes id_A(p_r) + id_A^{(A-1)}(p_r) \otimes id_A(1) \right)
\]
\[
= id_A^{(A-1)}(1) \cdot id_A(p_r) + id_A^{(A-1)}(p_r) \cdot id_A(1) = p_r + (A - 1) p_r = Ap_r.
\]
Thus, \((11.419)\) holds for \(a = A\). This completes the induction step, so that \((11.419)\) is proven.
Hence,

\[
(v_n \circ f_n)(p_r) = v_n \left( f_n(p_r) \right) = v_n(p_{rn}) = \begin{cases} np_{rn}/n, & \text{if } n \mid rn; \\ 0, & \text{if } n \nmid rn \end{cases}
\]

(by Exercise 2.86(a), applied to \( m = rn \))

\[= np_{rn}/n \quad \text{(since } n \mid rn)\]

\[= np_r = id^*_{\Lambda}(p_r)\]

(since \( id^*_{\Lambda}(p_r) = np_r \)).

Now, let us forget that we fixed \( r \). We thus have proven that

\[(v_n \circ f_n)(p_r) = id^*_{\Lambda}(p_r) \quad \text{for every positive integer } r.\]

We have assumed that \( k = \mathbb{Q} \). Hence, the elements \( p_1, p_2, p_3, \ldots \) of \( \Lambda \) generate the \( k \)-algebra \( \Lambda \) (due to Proposition 2.28). But \( v_n \circ f_n \) and \( id^*_{\Lambda} \) are two \( k \)-algebra homomorphisms from \( \Lambda \). These two homomorphisms \( v_n \circ f_n \) and \( id^*_{\Lambda} \) are equal to each other on each of the elements \( p_1, p_2, p_3, \ldots \) (due to (11.421)), and therefore are identical (because if two \( k \)-algebra homomorphisms with one and the same domain are equal to each other on a generating set of the domain, then these homomorphisms must be identical). In other words, \( v_n \circ f_n = id^*_{\Lambda} \). Thus, Exercise 2.86(g) is solved under the assumption that \( k = \mathbb{Q} \). Our Step 1 is complete.

**Step 2:** We shall prove that Exercise 2.86(g) holds if \( k = \mathbb{Z} \).

Indeed, let us first consider the general case, without any assumptions on \( k \).

We shall now formalize the intuitive concept that the constructions relevant to Exercise 2.86(g) (the Hopf algebra \( \Lambda \), its elements \( m_\lambda, p_\lambda, e_\lambda \) etc., and the maps \( v_n, f_n \) and \( id^*_{\Lambda} \)) are “functorial” with respect to the base ring \( k \). This is chiefly a matter of introducing notations:

- We denote the \( k \)-algebra \( \Lambda = \Lambda_k \) by \( \Lambda^{[k]} \). (This is just a minor change of notation that serves to make it more similar to the notations \( f_n^{[k]}, v_n^{[k]} \) and \( m_\lambda^{[k]} \) further below.)
- If \( m \) and \( n \) are two commutative rings, and \( \varphi : m \to n \) is a ring homomorphism, then \( \varphi \) canonically induces a ring homomorphism \( \Lambda^{[m]} \to \Lambda^{[n]} \). We denote this latter homomorphism by \( \Lambda^{[\varphi]} \).

Notice that while \( \Lambda^{[k]} \) denotes a ring, \( \Lambda^{[\varphi]} \) denotes a map. This might look confusing, but it makes sense from the viewpoint of category theory: There is a functor from the category of commutative rings to itself, which sends every commutative ring \( k \) to \( \Lambda^{[k]} \), and every morphism \( \varphi : m \to n \) of commutative rings to \( \Lambda^{[\varphi]} \). Unsurprisingly, this functor is denoted by \( \Lambda \).

- A ring homomorphism \( \varphi : m \to n \) also induces a ring homomorphism \( \Lambda^{[m]} \otimes_m \Lambda^{[m]} \to \Lambda^{[n]} \otimes_n \Lambda^{[n]} \).

This makes \( \Lambda \otimes \Lambda \) into a functor of \( k \).

- For every partition \( \lambda \), we shall denote the monomial symmetric function \( m_\lambda \in \Lambda \) by \( m^{[k]}_\lambda \). This notation makes the dependency of \( m_\lambda \) on the base ring \( k \) more explicit, and thus allows us to talk about these elements defined over several different base rings at the same time (without the danger of confusing them). Of course, the element \( m_\lambda \) does not “really” depend on \( k \), in the sense that it is defined in the same way for every \( k \). This entails that for any two commutative rings \( m \) and \( n \) and any ring homomorphism \( \varphi : m \to n \), we have

\[(m^{[k]}_\lambda) = \Lambda^{[\varphi]} \left( m^{[m]}_\lambda \right)\]

for every partition \( \lambda \).

\[\text{443 In fact, this ring homomorphism } \Lambda^{[m]} \to \Lambda^{[n]} \text{ can be defined by taking the ring homomorphism } \varphi \left( [x] \right) : m \left( [x] \right) \to n \left( [x] \right) \text{ canonically induced by } \varphi \text{ and restricting it to the subring } \Lambda^{[m]} \text{ of } m \left( [x] \right).\]

\[\text{444 There are several ways to define this homomorphism. One of them is to define it as the } m \text{-module homomorphism } \Lambda^{[m]} \otimes_m \Lambda^{[m]} \to \Lambda^{[n]} \otimes_n \Lambda^{[n]} \text{ canonically induced by the } m \text{-bilinear map}
\]

\[
\Lambda^{[m]} \times \Lambda^{[m]} \to \Lambda^{[n]} \otimes_n \Lambda^{[n]},
\]

\[(a, b) \mapsto \left( \Lambda^{[\varphi]}(a) \right) \otimes_n \left( \Lambda^{[\varphi]}(b) \right) .
\]

\(\text{ (Here, } \Lambda^{[m]} \otimes_n \Lambda^{[n]} \text{ is an } m \text{-module because the ring homomorphism } \varphi : m \to n \text{ makes } n \text{ into an } m \text{-algebra.)}\)
• We shall denote the homomorphism \( f_n : \Lambda \to \Lambda \) by \( f_n^{[k]} \). This notation makes the dependency of \( f_n \) on the base ring \( k \) more explicit.

The definition of \( f_n \) was functorial in \( k \). Thus, for any two commutative rings \( m \) and \( n \) and any ring homomorphism \( \varphi : m \to n \), we have

\[
f_n^{[m]} \circ \Lambda^{[\varphi]} = \Lambda^{[\varphi]} \circ f_n^{[m]}
\]

as maps from \( \Lambda^{[m]} \) to \( \Lambda^{[n]} \).

• Similarly, we shall denote the homomorphism \( v_n : \Lambda \to \Lambda \) by \( v_n^{[k]} \). This homomorphism is again defined in a way that is functorial in \( k \). Thus, for any two commutative rings \( m \) and \( n \) and any ring homomorphism \( \varphi : m \to n \), we have

\[
v_n^{[m]} \circ \Lambda^{[\varphi]} = \Lambda^{[\varphi]} \circ v_n^{[m]}
\]

as maps from \( \Lambda^{[m]} \) to \( \Lambda^{[n]} \).

• Notice that the Hopf algebra structure on \( \Lambda \) is functorial in \( k \).

As a consequence, for any two commutative rings \( m \) and \( n \) and any ring homomorphism \( \varphi : m \to n \), we have

\[
\text{id}_{\Lambda_n} \circ \Lambda^{[\varphi]} = \Lambda^{[\varphi]} \circ \text{id}_{\Lambda_m}
\]

as maps from \( \Lambda^{[m]} \) to \( \Lambda^{[n]} \).

• For every commutative ring \( k \), there exists a canonical (and unique) ring homomorphism \( \mathbb{Z} \to k \). Denote this homomorphism by \( \rho_k \).

In Step 1, we have shown that Exercise 2.86(g) holds if \( k = \mathbb{Q} \). In other words, we have \( v_n \circ f_n = \text{id}_{\Lambda_k}^n \) if \( k = \mathbb{Q} \). In other words,

\[
v_n^{[\mathbb{Q}]} \circ f_n^{[\mathbb{Q}]} = \text{id}_{\Lambda_{\mathbb{Q}}}^n.
\]

Recall that \( (m_{\lambda}^{[k]})_{\lambda \in \text{Par}} = (m_{\lambda})_{\lambda \in \text{Par}} \) is a basis of the \( k \)-module \( \Lambda_k \) for every commutative ring \( k \).

Applying this to \( k = \mathbb{Z} \), we conclude that \( (m_{\lambda}^{[\mathbb{Z}]})_{\lambda \in \text{Par}} \) is a basis of the \( \mathbb{Z} \)-module \( \Lambda_{\mathbb{Z}} \).

The ring homomorphism \( \rho_{\mathbb{Q}} : \mathbb{Z} \to \mathbb{Q} \) is just the canonical inclusion of \( \mathbb{Z} \) into \( \mathbb{Q} \), and thus is injective. Hence, the map \( \Lambda^{[\rho_{\mathbb{Q}}]} \) is injective (because \( \Lambda^{[\varphi]} \) is injective for every injective ring homomorphism \( \varphi \)).

Now, let us recall that our goal (in Step 2) is to show that Exercise 2.86(g) holds if \( k = \mathbb{Z} \). In other words, our goal is to show that

\[
v_n^{[\mathbb{Z}]} \circ f_n^{[\mathbb{Z}]} = \text{id}_{\Lambda_{\mathbb{Z}}}^n.
\]

This is an equality between \( \mathbb{Z} \)-module homomorphisms (indeed, it is clear that both \( v_n^{[\mathbb{Z}]} \circ f_n^{[\mathbb{Z}]} \) and \( \text{id}_{\Lambda_{\mathbb{Z}}}^n \) are \( \mathbb{Z} \)-module homomorphisms). Hence, in order to prove it, it is sufficient to verify it on the basis \( (m_{\lambda}^{[\mathbb{Z}]})_{\lambda \in \text{Par}} \) of the \( \mathbb{Z} \)-module \( \Lambda_{\mathbb{Z}} \). In other words, it is sufficient to prove that

\[
(v_n^{[\mathbb{Z}]} \circ f_n^{[\mathbb{Z}]}) (m_{\lambda}^{[\mathbb{Z}]}) = \text{id}_{\Lambda_{\mathbb{Z}}}^n (m_{\lambda}^{[\mathbb{Z}]}) \quad \text{for every } \lambda \in \text{Par}.
\]

So let us prove (11.427) now:

\[\text{(11.427)}\]

\[v_n^{[\mathbb{Z}]} \circ f_n^{[\mathbb{Z}]}, m_{\lambda}^{[\mathbb{Z}]} = \text{id}_{\Lambda_{\mathbb{Z}}}^n (m_{\lambda}^{[\mathbb{Z}]}) \quad \text{for every } \lambda \in \text{Par}.\]

\[445\]To make sense of this statement, we need to recall that \( \Lambda \otimes \Lambda \) has been made into a functor of \( k \).
Proof of (11.427): Let $\lambda \in \text{Par}$. We have

$$
\Lambda^{[\rho_0]} \left( \left( v_n^{[z]} \circ f_n^{[z]} \right) \left( m_\lambda^{[z]} \right) \right)
= \left( \lambda^{[\rho_0]} \circ v_n^{[z]} \circ f_n^{[z]} \right) \left( m_\lambda^{[z]} \right)
= \left( m_\lambda^{[z]} \right)
= \left( \Lambda^{[\rho_0]} \circ f_n^{[z]} \right) \left( m_\lambda^{[z]} \right),

$$

(by of $v_n^{[z]} \circ \Lambda^{[\rho_0]} \circ v_n^{[z]}$ (by (11.424), applied to $m=\mathbb{Z}$, $n=\mathbb{Q}$ and $\varphi=\rho_0$))

$$
= \left( \Lambda^{[\rho_0]} \circ \text{id}^{*n} \right) \left( m_\lambda^{[z]} \right)
= \left( \Lambda^{[\rho_0]} \circ \text{id}^{*n} \right) \left( m_\lambda^{[z]} \right),

$$

(by (11.425), applied to $m=\mathbb{Z}$, $n=\mathbb{Q}$ and $\varphi=\rho_0$)

$$
= \left( \Lambda^{[\rho_0]} \circ \text{id}^{*n} \right) \left( m_\lambda^{[z]} \right).

$$

Since the map $\Lambda^{[\rho_0]}$ is injective, this yields $\left( v_n^{[z]} \circ f_n^{[z]} \right) \left( m_\lambda^{[z]} \right) = \text{id}^{*n} \left( m_\lambda^{[z]} \right)$. Thus, (11.427) is proven.

As we said, proving (11.427) is sufficient to showing that $v_n^{[z]} \circ f_n^{[z]} = \text{id}^{*n}$. Hence, $v_n^{[z]} \circ f_n^{[z]} = \text{id}^{*n}$ is shown (since (11.427) is proven). In other words, Exercise 2.86(g) holds if $k = \mathbb{Z}$. This completes Step 2.

Step 3: We shall now prove that Exercise 2.86(g) holds in the general case.

Let us use all the notations that we introduced in Step 2.

In Step 2, we have shown that Exercise 2.86(g) holds if $k = \mathbb{Z}$. In other words, we have $v_n \circ f_n = \text{id}^{*n}$ if $k = \mathbb{Z}$. In other words,

$$
\left( \Lambda^{[\rho_0]} \circ \text{id}^{*n} \right) \left( m_\lambda^{[z]} \right).

$$

Recall that $(m_\lambda)_{\lambda \in \text{Par}}$ is a basis of the $k$-module $\Lambda_k$.

Now, let us recall that our goal (in Step 3) is to show that Exercise 2.86(g) holds. In other words, our goal is to show that

$$
v_n \circ f_n = \text{id}^{*n}.

$$

This is an equality between $k$-module homomorphisms (indeed, it is clear that both $v_n \circ f_n$ and $\text{id}^{*n}$ are $k$-module homomorphisms). Hence, in order to prove it, it is sufficient to verify it on the basis $(m_\lambda)_{\lambda \in \text{Par}}$ of the $k$-module $\Lambda_k$. In other words, it is sufficient to prove that

$$
\left( v_n \circ f_n \right) \left( m_\lambda \right) = \text{id}^{*n} \left( m_\lambda \right) \quad \text{for every } \lambda \in \text{Par}.

$$

So let us prove (11.429) now:

Proof of (11.429): Let $\lambda \in \text{Par}$. Then,

$$
m_\lambda = n_{\lambda}^{[k]} = \Lambda^{[\rho_0]} \left( m_\lambda^{[z]} \right) \quad \text{(by (11.422) (applied to } m=\mathbb{Z}, n=k \text{ and } \varphi=\rho_k).}

$$
We shall prove that Exercise 2.86(h) holds if (This can be proven just as in (11.420)). In other words, Exercise 2.86(g) holds in the general case. This completes Step 3.

Indeed, we first notice that

Now, let us first assume that $m$ \mid k.

Then, Exercise 2.86(a) (applied to $m$ and $r$ instead of $n$ and $m$) yields

\[ \mathbf{v}_m (p_r) = \begin{cases} mp_{r/m}, & \text{if } m \mid r; \\ 0, & \text{if } m \nmid r \end{cases} = 0 \quad (\text{since } m \nmid r), \]

Thus, (11.429) is proven.

As we said, proving (11.429) is sufficient to showing that $\mathbf{v}_n \circ \mathbf{f}_n = \mathbf{id}^* n$. Hence, $\mathbf{v}_n \circ \mathbf{f}_n = \mathbf{id}^* n$ is shown (since (11.429) is proven). In other words, Exercise 2.86(g) holds in the general case. This completes Step 3 and, with it, the solution of Exercise 2.86(g).

(h) The structure of our solution of Exercise 2.86(h) is similar to that of our solution of Exercise 2.86(g) above. It proceeds in three steps:

- **Step 1**: proving that Exercise 2.86(h) holds if $k = \mathbb{Q}$.
- **Step 2**: proving that Exercise 2.86(h) holds if $k = \mathbb{Z}$.
- **Step 3**: proving that Exercise 2.86(h) holds in the general case.

The details of Steps 2 and 3 are very similar to the details of the corresponding steps in the solution of Exercise 2.86(g), and so we will not dwell on these details. However, we need to give the details of Step 1:

**Step 1**: We shall prove that Exercise 2.86(h) holds if $k = \mathbb{Q}$.

Indeed, assume that $k = \mathbb{Q}$. We know that $\mathbf{f}_n$ is a $k$-algebra homomorphism (due to Exercise 2.85(a)), and that $\mathbf{v}_m$ is a $k$-algebra homomorphism. Hence, $\mathbf{f}_n \circ \mathbf{v}_m$ and $\mathbf{v}_m \circ \mathbf{f}_n$ are $k$-algebra homomorphisms.

Let $r$ be a positive integer. We shall show that $(\mathbf{f}_n \circ \mathbf{v}_m) (p_r) = (\mathbf{v}_m \circ \mathbf{f}_n) (p_r)$.

Indeed, we first notice that

\begin{equation}
\mathbf{f}_n (p_r) = p_{rn}.
\end{equation}

(This can be proven just as in (11.420).)

Now, let us first assume that $m \nmid r$. Then, Exercise 2.86(a) (applied to $m$ and $r$ instead of $n$ and $m$) yields

\[ \mathbf{v}_m (p_r) = \begin{cases} mp_{r/m}, & \text{if } m \mid r; \\ 0, & \text{if } m \nmid r \end{cases} = 0 \quad (\text{since } m \nmid r), \]
so that \((f_n \circ v_m)(p_r) = f_n \left( v_m \left( p_r \right) \right) = f_n (0) = 0\) (since the map \(f_n\) is \(k\)-linear). On the other hand, if we
had \(m \mid rn\), then we would have \(m \mid r\) (since \(m\) is coprime to \(n\)), which would contradict \(m \nmid r\). Hence, we cannot have \(m \mid rn\). Thus, we have \(m \nmid rn\). Now,
\[
(v_m \circ f_n)(p_r) = v_m \left( f_n \left( p_r \right) \right) = v_m \left( \frac{mp_{rn/m}}{mp_{f_r/m}} \right) = v_m \left( \frac{mp_{rn/m}}{mpr/m} \right) = 0,
\]
(by Exercise 2.86(a), applied to \(m\) and \(rn\) instead of \(n\) and \(m\))
\[
= 0 \quad (since \ m \nmid rn).
\]

Compared with \((f_n \circ v_m)(p_r) = 0\), this yields \((f_n \circ v_m)(p_r) = (v_m \circ f_n)(p_r)\).

Now, let us forget our assumption that \(m \nmid r\). Hence, we have proven that \((f_n \circ v_m)(p_r) = (v_m \circ f_n)(p_r)\) under the assumption that \(m \nmid r\). As a consequence, for the rest of the proof of \((f_n \circ v_m)(p_r) = (v_m \circ f_n)(p_r)\), we can WLOG assume that we don’t have \(m \nmid r\). Assume this.

We have \(m \nmid r\) (since we don’t have \(m \nmid r\)). Now, Exercise 2.86(a) (applied to \(m\) and \(rn\) instead of \(n\) and \(m\)) yields
\[
v_m(p_r) = \begin{cases} \frac{mp_{r/m}}{m}, & \text{if } m \mid r; \\ 0, & \text{if } m \nmid r \end{cases}
= \frac{mp_{r/m}}{m} \quad (since \ m \nmid r),
\]
so that
\[
(f_n \circ v_m)(p_r) = f_n \left( v_m \left( p_r \right) \right) = f_n \left( \frac{mp_{r/m}}{m} \right) = m \frac{mp_{r/m}}{m} = mp_{r/m} \quad (since \ the \ map \ f_n \ is \ k-linear)
\]
\[
= mp_{(r/m)n} = mp_{rn/m} \quad (since \ (r/m)n = rn/m).
\]

Compared with
\[
(v_m \circ f_n)(p_r) = v_m \left( f_n \left( p_r \right) \right) = v_m \left( \frac{mp_{rn/m}}{mp_{f_r/m}} \right) = v_m \left( \frac{mp_{rn/m}}{mp_{f_r/m}} \right) = 0,
\]
(by Exercise 2.86(a), applied to \(m\) and \(rn\) instead of \(n\) and \(m\))
\[
= mp_{rn/m} \quad (since \ m \mid rn \ (since \ m \mid r \ | \ rn)),
\]
this yields \((f_n \circ v_m)(p_r) = (v_m \circ f_n)(p_r)\). Hence, \((f_n \circ v_m)(p_r) = (v_m \circ f_n)(p_r)\) is proven.

Now, let us forget that we fixed \(r\). We thus have proven that
\[
(f_n \circ v_m)(p_r) = (v_m \circ f_n)(p_r) \quad for \ every \ positive \ integer \ r.
\]

We have assumed that \(k = \mathbb{Q}\). Hence, the elements \(p_1, p_2, p_3, \ldots\) of \(\Lambda\) generate the \(k\)-algebra \(\Lambda\) (due to Proposition 2.28). But \(f_n \circ v_m\) and \(v_m \circ f_n\) are two \(k\)-algebra homomorphisms from \(\Lambda\). These two homomorphisms \(f_n \circ v_m\) and \(v_m \circ f_n\) are equal to each other on each of the elements \(p_1, p_2, p_3, \ldots\) (due to (11.432)), and therefore are identical (because if two \(k\)-algebra homomorphisms with one and the same domain are equal to each other on a generating set of the domain, then these homomorphisms must be identical). In other words, \(f_n \circ v_m = v_m \circ f_n\). Thus, Exercise 2.86(h) is solved under the assumption that \(k = \mathbb{Q}\). Our Step 1 is complete.

As already mentioned, Steps 2 and Steps 3 are very similar to the corresponding steps in our above solution of Exercise 2.86(g). Thus, we forego showing these steps. Exercise 2.86(h) is thus solved.

(i) Our solution of Exercise 2.86(i) will be somewhat similar to that of Exercise 2.86(g), but simpler. We will only need two steps:

- **Step 1**: proving that Exercise 2.86(i) holds if \(k = \mathbb{Z}\).
- **Step 2**: proving that Exercise 2.86(i) holds in the general case.
We shall prove that Exercise 2.86(i) holds if \( k = \mathbb{Z} \).

Indeed, assume that \( k = \mathbb{Z} \). Every positive integer \( r \) satisfies

\[
pr = \sum_{d|r} dw_d^{r/d}
\]

(as per Exercise 2.79(e), applied to \( r \) instead of \( n \)).

Now, for every positive integer \( m \), we define an element \( \tilde{w}_m \) of \( \Lambda \) by

\[
\tilde{w}_m = \begin{cases} \frac{w_m}{m}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases}
\]

It is now easy to see that every positive integer \( r \) satisfies

\[
v_n (pr) = \sum_{d|r} dw_d^{r/d}.
\]

We shall now prove that every positive integer \( m \) satisfies

\[
v_n (w_m) = \tilde{w}_m.
\]

**Proof of (11.437):** We proceed by strong induction over \( m \):

**Induction step:** Fix a positive integer \( M \). Assume that (11.437) holds for every positive integer \( m < M \). We now must show that (11.437) holds for \( m = M \).

The \( k \)-module \( \Lambda \) is free, and thus torsionfree. Hence,

\[
every element \ a \ of \ \Lambda \ satisfying \ \text{Ma} = 0 \ satisfies \ a = 0
\]

(since \( k = \mathbb{Z} \)).

We have assumed that (11.437) holds for every positive integer \( m < M \). In other words,

\[
v_n (w_m) = \tilde{w}_m \quad \text{for every positive integer } m < M.
\]

**Proof of (11.435):** Let \( r \) be a positive integer. Let us first assume that \( n \mid r \). Then, every positive divisor \( d \) of \( r \) satisfies

\[
\tilde{w}_d = \begin{cases} \frac{w_d}{d}, & \text{if } n \mid d; \\ 0, & \text{if } n \nmid d \end{cases} \quad \text{(by the definition of } \tilde{w}_d \text{)}
\]

\[
= 0 \quad \text{(since } n \mid d \text{ (because otherwise, we would have } n \mid d \text{ and thus } n \mid d \mid r, \text{ which would contradict } n \nmid r \text{).}
\]

But Exercise 2.86(a) (applied to \( r \) instead of \( m \)) yields

\[
v_n (pr) = \begin{cases} \frac{np}{r}, & \text{if } n \mid r; \\ 0, & \text{if } n \nmid r \end{cases} \quad \text{(since } n \nmid r \text{).}
\]

Compared with

\[
\sum_{d|r} dw_d^{r/d} = \sum_{d|r} \left( \frac{\tilde{w}_d}{d} \right)^{r/d} = \sum_{d|r} d \left( \frac{\tilde{w}_d}{d} \right)^{r/d} = \sum_{d|r} \frac{r/d}{0} = 0,
\]

this yields \( v_n (pr) = \sum_{d|r} dw_d^{r/d} \). Thus, (11.435) holds.

Let us now forget that we assumed that \( n \mid r \). We thus have proven (11.435) under the assumption that \( n \nmid r \). Hence, for the rest of the proof of (11.435), we can WLOG assume that \( m \nmid r \). Assume this.

We have \( n \mid r \) (since we don’t have \( n \nmid r \)). Hence, \( r/n \) is a positive integer. Exercise 2.86(a) (applied to \( r \) instead of \( m \)) yields

\[
v_n (pr) = \begin{cases} \frac{np}{r}, & \text{if } n \mid r; \\ 0, & \text{if } n \nmid r \end{cases} \quad \text{for every positive integer } n \mid r.
\]

holding for every positive integer \( m < M \). In other words,
Applying (11.433) to \( r = M \), we obtain \( p_M = \sum_{d|M} d w^M_d \). Applying the map \( v_n \) to both sides of this equality, we obtain

\[
v_n(p_M) = v_n \left( \sum_{d|M} d w^M_d \right) = \sum_{d|M} d \left( v_n(w_d) \right)^{M/d} \quad \text{(since \( v_n \) is a \( k \)-algebra homomorphism)}
\]

\[
= M \left( v_n(w_M) \right)^{M/M} + \sum_{d|M; \ d \neq M} d \left( v_n(w_d) \right)^{M/d}
\]

\( = v_n(w_M) \quad \text{(by (11.439), applied to } m=d) \)

\[
\text{(here, we have split off the addend for } d = M \text{ from the sum)}
\]

\[
v_n(w_M) + \sum_{d|M; \ d < M} d \left( v_n(w_d) \right)^{M/d} = M v_n(w_M) + \sum_{d|M; \ d < M} d \tilde{w}^{M/d}_d.
\]

Compared with

\[
\sum_{d|r} dw^r/d_d = \sum_{d|r} \left\{ \begin{array}{ll}
w_{d/n}, & \text{if } n \mid d; \\
0, & \text{if } n \nmid d
\end{array} \right\} = \sum_{d|r} d \left( \begin{array}{ll}
w_{d/n}, & \text{if } n \mid d; \\
0, & \text{if } n \nmid d
\end{array} \right)^{r/d}
\]

\[
= \sum_{d|r; \ n \mid d} d w_{d/n}^{r/d} + \sum_{d|r; \ n \nmid d} d \tilde{w}_{d/n}^{r/d} = \sum_{d|r; \ n \mid d} d w_{d/n}^{r/d} + \sum_{d|r; \ n \nmid d} d \tilde{w}_{d/n}^{r/d} = \sum_{d|r; \ n \mid d} d^{r/d} + \sum_{d|r; \ n \nmid d} d^{r/d} = \sum_{d|r} d^{r/d}
\]

\[
= \sum_{d|r} nd \left( \begin{array}{ll}
w_{r/(nd)}^{r/(nd)}, & \text{if } n \mid d; \\
0, & \text{if } n \nmid d
\end{array} \right) + \sum_{d|r; \ n \mid d} d = \sum_{d|r; \ n \mid d} d^{r/(nd)} + \sum_{d|r; \ n \nmid d} d = \sum_{d|r} d^{r/(nd)}
\]

\[
\text{(here, we have substituted } nd \text{ for } d \text{ in the sum)}
\]

\[
\text{this yields } v_n(p_r) = \sum_{d|r} d \tilde{w}^{r/d}_d. \quad \text{Thus, (11.435) is proven.}
\]
Compared with
\[
\nu_n (p_M) = \sum_{d|M} dw_{d}^{M/d}
\]
(by (11.435), applied to \( r = M \))
\[
= M \overline{w}_M - \sum_{d|M; d \neq M} dw_{d}^{M/d}
\]
(here, we have split off the addend for \( d = M \) from the sum)
\[
= M \overline{w}_M + \sum_{d|M; d < M} dw_{d}^{M/d},
\]
this yields
\[
M \nu_n (w_M) + \sum_{d|M; d < M} dw_{d}^{M/d} = M \overline{w}_M + \sum_{d|M; d < M} dw_{d}^{M/d}.
\]

Subtracting \( \sum_{d < M} dw_{d}^{M/d} \) from both sides of this equality, we obtain
\[
M \nu_n (w_M) = M \overline{w}_M.
\]
Thus, (11.438) (applied to \( a = \nu_n (w_M) - \overline{w}_M \)) yields \( \nu_n (w_M) = \overline{w}_M \). In other words, (11.437) holds for \( m = M \). This completes the induction step. Thus, (11.437) is proven.

Now, every positive integer \( m \) satisfies
\[
\nu_n (w_m) = \overline{w}_m
\]
(by (11.437))
\[
= \begin{cases} w_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases}
\]
(by the definition of \( \overline{w}_m \)).

Thus, Exercise 2.86(i) is solved in the case when \( k = \mathbb{Z} \). In other words, Step 1 is finished.

**Step 2:** We shall now show that Exercise 2.86(i) holds in the general case.

Indeed, let us use all the notations that we introduced in Step 2 of the solution of Exercise 2.86(g). Let us furthermore introduce one more notation:

- For every positive integer \( m \), we shall denote the element \( w_m \) of \( \Lambda \) by \( w_m^{[k]} \). This notation makes the dependency of \( w_m \) on the base ring \( k \) more explicit. Of course, the element \( w_m \) does not “really” depend on \( k \), in the sense that it is defined in the same way for every \( k \). This entails that for any two commutative rings \( m \) and \( n \) and any ring homomorphism \( \varphi : m \rightarrow n \), we have
\[
w_m^{[n]} = \Lambda^{[\varphi]} \left( w_m^{[m]} \right)
\]
for every positive integer \( m \).

Now, fix a positive integer \( m \). In Step 1, we have shown that Exercise 2.86(i) holds if \( k = \mathbb{Z} \). In other words, we have
\[
\nu_n (w_m) = \begin{cases} w_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} \quad \text{if } k = \mathbb{Z}.
\]
In other words,
\[
\nu_n^{[2]} (w_m^{[2]}) = \begin{cases} w_{m/n}^{[2]}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases}.
\]
Now,
\[
w_m = w_m^{[k]} = \Lambda^{[\varphi]} \left( w_m^{[m]} \right)
\]
(by (11.440) (applied to \(m = \mathbb{Z}, n = k\) and \(\varphi = \rho_k\)). Thus,

\[
\begin{align*}
\nu_n = \nu_k & = \lambda [\rho_k] \left( \nu_k \right) \\
\nu_n & = \lambda [\rho_k] \left( \nu_n^{[2]} \right) = \lambda [\rho_k] \left( \nu_n^{[2]} \right) = \left( \nu_n^{[2]} \right) \quad \text{(by (11.441))}
\end{align*}
\]

\[
= \left( \lambda [\rho_k] \circ \nu_n^{[2]} \right) \left( \nu_n^{[2]} \right) = \lambda [\rho_k] \left( \nu_n^{[2]} \right) = \left\{ \begin{array}{ll}
\nu_n^{[2]} & \text{if } n \mid m; \\
0 & \text{if } n \nmid m
\end{array} \right.
\]

(11.443)

On the other hand, we have

\[
\lambda [\rho_k] \left( \nu_n^{[2]} \right) = \nu_n^{[2]} \quad \text{in the case when } n \mid m
\]

(11.444)

Now, (11.443) becomes

\[
\nu_n \left( \nu_m \right) = \left\{ \begin{array}{ll}
\lambda [\rho_k] \left( \nu_m^{[2]} \right) & \text{if } n \mid m; \\
0 & \text{if } n \nmid m
\end{array} \right.
\]

(11.445)

Thus, Exercise 2.86(i) holds. This completes Step 2, and thus Exercise 2.86(i) is solved.

11.60. **Solution to Exercise 2.87.**

**Solution to Exercise 2.87.** (b) Let us first notice that

\[
X_{n,d,s} = \sum_{w = \left( x_1, x_2, \ldots, x_n \right) \in \{1,2,3,\ldots\}^n} x_{i_1} x_{i_2} \cdots x_{i_n} \quad \text{for all } d \in \mathbb{N} \text{ and } s \in \mathbb{N}.
\]

(11.445)

(This is just the definition of \(X_{n,d,s}\), written as a formula.)

Let us first show that if \(n\) is a positive integer, then any \(d \in \mathbb{N}\) and \(s \in \mathbb{N}\) satisfy

\[
(d + 1) X_{n,d+1,s} + (s + 1) X_{n,d,s+1} + (n - 1 - d - s) X_{n,d,s}
\]

(11.446)

Proof of (11.444): Assume that \(n \mid m\). Then, \(m/n\) is a positive integer. Hence, \(w_m/n = \lambda [\rho_k] \left( w_m^{[2]} \right) \) (in fact, this follows from the same argument that was used to prove (11.442), but with \(m\) replaced by \(m/n\)). In other words, \(\lambda [\rho_k] \left( w_m^{[2]} \right) = w_{m/n}\). This proves (11.444).
Proof of (11.446): Let \( n \) be a positive integer. Let \( d \in \mathbb{N} \) and \( s \in \mathbb{N} \). Let \( i \in \{1, 2, \ldots, n-1\} \) be arbitrary. We make some more definitions:

- Define a power series \( \mathcal{D}_i \in \mathbb{k}[x] \) as the sum of the monomials \( x_{w_1} x_{w_2} \cdots x_{w_n} \) over all \( n \)-tuples \( w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n \) satisfying \(|\text{Des}(w)| = d + 1\), \(|\text{Stag}(w)| = s\) and \( w_i > w_{i+1} \).
- Define a power series \( \mathcal{S}_i \in \mathbb{k}[x] \) as the sum of the monomials \( x_{w_1} x_{w_2} \cdots x_{w_n} \) over all \( n \)-tuples \( w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n \) satisfying \(|\text{Des}(w)| = d\), \(|\text{Stag}(w)| = s + 1\) and \( w_i = w_{i+1} \).
- Define a power series \( \mathcal{A}_i \in \mathbb{k}[x] \) as the sum of the monomials \( x_{w_1} x_{w_2} \cdots x_{w_n} \) over all \( n \)-tuples \( w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n \) satisfying \(|\text{Des}(w)| = d\), \(|\text{Stag}(w)| = s\) and \( w_i < w_{i+1} \).

These three power series \( \mathcal{D}_i, \mathcal{S}_i \) and \( \mathcal{A}_i \) are not necessarily symmetric (but will nevertheless come useful). Let us notice that the definitions of \( \mathcal{D}_i, \mathcal{S}_i \) and \( \mathcal{A}_i \) can be rewritten as follows:

- The power series \( \mathcal{D}_i \in \mathbb{k}[x] \) is the sum of the monomials \( x_{w_1} x_{w_2} \cdots x_{w_n} \) over all \( n \)-tuples \( w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n \) satisfying \(|\text{Des}(w)| = d\), \(|\text{Stag}(w)| = i\) and \( w_i > w_{i+1} \).
- The power series \( \mathcal{S}_i \in \mathbb{k}[x] \) is the sum of the monomials \( x_{w_1} x_{w_2} \cdots x_{w_n} \) over all \( n \)-tuples \( w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n \) satisfying \(|\text{Des}(w)| = d\), \(|\text{Stag}(w)| = i\) and \( w_i = w_{i+1} \).
- The power series \( \mathcal{A}_i \in \mathbb{k}[x] \) is the sum of the monomials \( x_{w_1} x_{w_2} \cdots x_{w_n} \) over all \( n \)-tuples \( w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n \) satisfying \(|\text{Des}(w)| = d\), \(|\text{Stag}(w)| = i\) and \( w_i < w_{i+1} \).

These reformulations make it obvious that the sum \( \mathcal{D}_i + \mathcal{S}_i + \mathcal{A}_i \) is precisely the sum of the monomials \( x_{w_1} x_{w_2} \cdots x_{w_n} \) over all \( n \)-tuples \( w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n \) satisfying \(|\text{Des}(w)| = d\) and \(|\text{Stag}(w)| = i\). In other words,

\[
\mathcal{D}_i + \mathcal{S}_i + \mathcal{A}_i = \sum_{w=(w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n; \ (|\text{Des}(w)| \setminus \{i\}) = d; \ (|\text{Stag}(w)| \setminus \{i\}) = s} x_{w_1} x_{w_2} \cdots x_{w_n}.
\]

But the sum on the right hand side of (11.447) can be rewritten in terms of more familiar sums. In fact, the \( n \)-tuples \( w \in \{1, 2, 3, \ldots\}^n \) are in bijection with the pairs \( (u, v) \) consisting of an \( i \)-tuple \( u \in \{1, 2, 3, \ldots\}^i \) and an \((n-i)\)-tuple \( v \in \{1, 2, 3, \ldots\}^{n-i} \). This bijection sends an \( n \)-tuple \((w_1, w_2, \ldots, w_n)\) to the pair \((u_1, u_2, \ldots, u_i), (w_{i+1}, w_{i+2}, \ldots, w_n)\), and has the property that \(|\text{Des}(w)| \setminus \{i\}| = |\text{Des}(u)| + |\text{Des}(v)|\) and \(|\text{Stag}(w)| \setminus \{i\}| = |\text{Stag}(u)| + |\text{Stag}(v)|\). Hence,

\[
\mathcal{D}_i + \mathcal{S}_i + \mathcal{A}_i = \sum_{w=(u_1, u_2, \ldots, u_i) \in \{1,2,3,\ldots\}^i; \ v=(v_1, v_2, \ldots, v_{n-i}) \in \{1,2,3,\ldots\}^{n-i}; \ (|\text{Des}(u)| + |\text{Des}(v)|) = d; \ (|\text{Stag}(u)| + |\text{Stag}(v)|) = s-t} x_{u_1} x_{u_2} \cdots x_{u_i} x_{v_1} x_{v_2} \cdots x_{v_{n-i}}.
\]

\[
= \sum_{e=0}^{d} \sum_{t=0}^{s} \sum_{u=(u_1, u_2, \ldots, u_i) \in \{1,2,3,\ldots\}^i; \ v=(v_1, v_2, \ldots, v_{n-i}) \in \{1,2,3,\ldots\}^{n-i}; \ (|\text{Des}(u)|) = e; \ (|\text{Stag}(u)|) = e} x_{u_1} x_{u_2} \cdots x_{u_i} x_{v_1} x_{v_2} \cdots x_{v_{n-i}}
\]

\[
= \sum_{e=0}^{d} \sum_{t=0}^{s} \left( \sum_{u=(u_1, u_2, \ldots, u_i) \in \{1,2,3,\ldots\}^i; \ (|\text{Des}(u)|) = e; \ (|\text{Stag}(u)|) = t} x_{u_1} x_{u_2} \cdots x_{u_i} \right) \left( \sum_{v=(v_1, v_2, \ldots, v_{n-i}) \in \{1,2,3,\ldots\}^{n-i}; \ (|\text{Des}(v)|) = d-e; \ (|\text{Stag}(v)|) = s-t} x_{v_1} x_{v_2} \cdots x_{v_{n-i}} \right).
\]
But for every $e \in \{0, 1, \ldots, d\}$ and $t \in \{0, 1, \ldots, s\}$, the equality (11.445) (applied to $i$, $e$ and $t$ instead of $n$, $d$ and $s$) yields

\[(11.449)\quad X_{i,e,t} = \sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n; \ |\text{Des}(w)|=e; \ |\text{Stag}(w)|=t} x_{w_1} x_{w_2} \cdots x_{w_n} = \sum_{u=(u_1,u_2,\ldots,u_t)\in\{1,2,3,\ldots\}^n; \ |\text{Des}(u)|=e; \ |\text{Stag}(u)|=t} x_{u_1} x_{u_2} \cdots x_{u_t}.\]

Also, for every $e \in \{0, 1, \ldots, d\}$ and $t \in \{0, 1, \ldots, s\}$, the equality (11.445) (applied to $n-i$, $d-e$ and $s-t$ instead of $n$, $d$ and $s$) yields

\[(11.450)\quad X_{n-i,d-e,s-t} = \sum_{w=(w_1,w_2,\ldots,w_{n-i})\in\{1,2,3,\ldots\}^{n-i}; \ |\text{Des}(w)|=d-e; \ |\text{Stag}(w)|=s-t} x_{w_1} x_{w_2} \cdots x_{w_{n-i}} = \sum_{v=(v_1,v_2,\ldots,v_{n-i})\in\{1,2,3,\ldots\}^{n-i}; \ |\text{Des}(v)|=d-e; \ |\text{Stag}(v)|=s-t} x_{v_1} x_{v_2} \cdots x_{v_{n-i}}.\]

Now, (11.448) becomes

\[
\sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n; \ |(\text{Des}(w))\setminus\{i\}|=d; \ |\text{Stag}(w))\setminus\{i\}|=s} x_{w_1} x_{w_2} \cdots x_{w_n}
= \sum_{e=0}^{d} \sum_{t=0}^{s} \left( \sum_{u=(u_1,u_2,\ldots,u_t)\in\{1,2,3,\ldots\}^n; \ |\text{Des}(u)|=e; \ |\text{Stag}(u)|=t} x_{u_1} x_{u_2} \cdots x_{u_t} \right) \times \left( \sum_{v=(v_1,v_2,\ldots,v_{n-i})\in\{1,2,3,\ldots\}^{n-i}; \ |\text{Des}(v)|=d-e; \ |\text{Stag}(v)|=s-t} x_{v_1} x_{v_2} \cdots x_{v_{n-i}} \right)
= X_{i,e,t} \quad \text{(by (11.449))}
= X_{n-i,d-e,s-t} \quad \text{(by (11.450))}
= \sum_{e=0}^{d} \sum_{t=0}^{s} X_{i,e,t} X_{n-i,d-e,s-t}.
\]

Hence, (11.447) becomes

\[(11.451)\quad D_i + S_i + A_i = \sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n; \ |(\text{Des}(w))\setminus\{i\}|=d; \ |\text{Stag}(w))\setminus\{i\}|=s} x_{w_1} x_{w_2} \cdots x_{w_n} = \sum_{e=0}^{d} \sum_{t=0}^{s} X_{i,e,t} X_{n-i,d-e,s-t}.
\]

Now, let us forget that we fixed $i$. We thus have defined $D_i$, $S_i$ and $A_i$ and proven the equality (11.451) for all $i \in \{1, 2, \ldots, n-1\}$.

Now, let us take a closer look at the sums $\sum_{i=1}^{n-1} D_i$, $\sum_{i=1}^{n-1} S_i$ and $\sum_{i=1}^{n-1} A_i$:

- The definition of $D_i$ can be rewritten as follows:

\[
D_i = \sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,3,\ldots\}^n; \ |\text{Des}(w)|=d+1; \ |\text{Stag}(w)|=s; \ w_i > w_{i+1} \quad \text{for every } i \in \{1, 2, \ldots, n-1\}.
\]
Summing up these equations over all \( i \in \{1, 2, \ldots, n - 1\} \), we obtain

\[
\sum_{i=1}^{n-1} D_i = \sum_{i=1}^{n-1} \sum_{\substack{w=(w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n; \\
|\text{Des}(w)|=d+1; |\text{Stag}(w)|=s; w_i > w_{i+1}} x_{w_1} x_{w_2} \cdots x_{w_n}
\]

\[
= \sum_{\substack{w=(w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n; \\
|\text{Des}(w)|=d+1; |\text{Stag}(w)|=s}} x_{w_1} x_{w_2} \cdots x_{w_n}
\]

\[
= \sum_{\substack{w=(w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n; \\
|\text{Des}(w)|=d+1; |\text{Stag}(w)|=s}} |\{i \in \{1,2,\ldots,n-1\} \mid w_i > w_{i+1}\}| x_{w_1} x_{w_2} \cdots x_{w_n}
\]

\[
= \sum_{\substack{w=(w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n; \\
|\text{Des}(w)|=d+1; |\text{Stag}(w)|=s}} (d+1) x_{w_1} x_{w_2} \cdots x_{w_n}
\]

\[
= (d+1) X_{n,d+1,s} \tag{11.452}
\]

- Similarly, we can see that

\[
\sum_{i=1}^{n-1} S_i = (s+1) X_{n,d,s+1} \tag{11.453}
\]

- Similarly, we can see that

\[
\sum_{i=1}^{n-1} A_i = (n-1-d-s) X_{n,d,s} \tag{11.454}
\]

Now, summing the equality (11.451) over all \( i \in \{1, 2, \ldots, n - 1\} \) yields

\[
\sum_{i=1}^{n-1} (D_i + S_i + A_i) = \sum_{i=1}^{n-1} \sum_{d=0}^{d} \sum_{t=0}^{s} X_{i,e,t} X_{n-i,d-e,s-t}
\]

Hence,

\[
\sum_{i=1}^{n-1} \sum_{d=0}^{d} \sum_{t=0}^{s} X_{i,e,t} X_{n-i,d-e,s-t}
\]

\[
= \sum_{i=1}^{n-1} (D_i + S_i + A_i) = \sum_{i=1}^{n-1} D_i + \sum_{i=1}^{n-1} S_i + \sum_{i=1}^{n-1} A_i
\]

\[
= (d+1) X_{n,d+1,s} + (s+1) X_{n,d,s+1} + (n-1-d-s) X_{n,d,s} \tag{11.456}
\]

This proves (11.446).

In solving Exercise 2.87(b), we WLOG assume that \( k = \mathbb{Z} \) (because there is a canonical ring homomorphism \( \mathbb{Z}[x] \rightarrow k [x] \) for any commutative ring \( k \), and because the power series \( X_{n,d,s} \) defined over \( k \) is clearly the image of the power series \( X_{n,d,s} \) defined over \( \mathbb{Z} \) under this homomorphism, whereas the image of the ring \( A_{\mathbb{Z}} \) under this homomorphism is a subring of \( A_k \)).

\footnote{In proving this, we have to observe that \(|\{i \in \{1,2,\ldots,n-1\} \mid w_i < w_{i+1}\}| = n - 1 - |\text{Des}(w)| - |\text{Stag}(w)| \) for every \( w = (w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n \). This is clear from realizing that the sets \( \{i \in \{1,2,\ldots,n-1\} \mid w_i < w_{i+1}\} \), Des \( w \) and Stag \( w \) are pairwise disjoint and their union is \( \{1,2,\ldots,n-1\} \).}
Now, we are going to solve Exercise 2.87(b) by strong induction over \( n + d \). So (for the induction step) we need to show that \( X_{n,d,s} \in \Lambda \), and we can assume (as the induction hypothesis) that \( X_{n',d',s'} \in \Lambda \) is already known to hold for any nonnegative integers \( n', d', s' \) satisfying \( n' + d' < n + d \).

We must be in one of the following two cases:

**Case 1:** We have \( d > 0 \).

**Case 2:** We have \( d = 0 \).

Let us first consider Case 1. In this case, we have \( d > 0 \). Hence, \( d - 1 \in \mathbb{N} \). We also WLOG assume that \( n \) is positive (otherwise, \( X_{n,d,s} \) is a constant and thus lies in \( \Lambda \) for sure). Applying (11.446) to \( d - 1 \) instead of \( d \), we obtain

\[
dX_{n,d,s} + (s + 1) X_{n,d-1,s+1} + (n - d - s) X_{n,d-1,s} = \sum_{i=1}^{n-1} \sum_{c=0}^{d-1} \sum_{t=0}^{s} X_{i,c,t} X_{n-i,d-1-c,s-t}.
\]

Thus,

(11.455) \[
dX_{n,d,s} = \sum_{i=1}^{n-1} \sum_{c=0}^{d-1} \sum_{t=0}^{s} X_{i,c,t} X_{n-i,d-1-c,s-t} - (s + 1) X_{n,d-1,s+1} + (n - d - s) X_{n,d-1,s}.
\]

Each of the power series \( X_{i,c,t}, X_{n-i,d-1-c,s-t}, X_{n,d-1,s+1} \) and \( X_{n,d-1,s} \) on the right hand side of this equality is already known to lie in \( \Lambda \) (by the induction hypothesis). Hence, the right hand side of (11.455) lies in \( \Lambda \) (since \( \Lambda \) is a ring), and thus (11.455) shows that \( dX_{n,d,s} \in \Lambda \). Since \( d > 0 \), this yields \( X_{n,d,s} \in \Lambda \) (because if a power series \( Q \in \mathbb{Z}[x] \) satisfies \( dQ \in \Lambda \), then \( Q \) must belong to \( \Lambda \) itself\(^{451}\)). Hence, the induction step is complete in Case 1.

Let us now consider Case 2. In this case, we have \( d = 0 \). Hence, \( X_{n,d,s} = X_{n,0,s} \).

We shall now show that

(11.456) \[
X_{n,0,s} = \sum_{\lambda \in \text{Par}_n; \ell(\lambda) = n - s} m_\lambda.
\]

**Proof of (11.456):** Summing up the equality (2.1) over all \( \lambda \in \text{Par}_n \) satisfying \( \ell(\lambda) = n - s \), we obtain

(11.457) \[
\sum_{\lambda \in \text{Par}_n; \ell(\lambda) = n - s} m_\lambda = \sum_{\lambda \in \text{Par}_n; \ell(\lambda) = n - s} \sum_{\alpha \in \Theta(\infty) \lambda} x^\alpha.
\]

On the other hand, the equality (11.445) (applied to 0 instead of \( d \)) yields

(11.458) \[
X_{n,0,s} = \sum_{w=(w_1,w_2,\ldots,w_n) \in \{1,2,3,\ldots\}^n; |\text{Des}(w)|=0; |\text{Stag}(w)|=s} x_{w_1} x_{w_2} \cdots x_{w_n}.
\]

The condition \( |\text{Des}(w)| = 0 \) on an \( n \)-tuple \( w = (w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n \) is equivalent to \( w_1 \leq w_2 \leq \ldots \leq w_n \), and therefore (11.458) rewrites as

(11.459) \[
X_{n,0,s} = \sum_{w=(w_1,w_2,\ldots,w_n) \in \{1,2,3,\ldots\}^n; w_1 \leq w_2 \leq \ldots \leq w_n; |\text{Stag}(w)|=s} x_{w_1} x_{w_2} \cdots x_{w_n}.
\]

For an \( n \)-tuple \( w = (w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n \) satisfying \( w_1 \leq w_2 \leq \ldots \leq w_n \), the condition \( |\text{Stag}(w)| = s \) is equivalent to the condition that \( |\{w_1, w_2, \ldots, w_n\}| = n - s \). Hence, (11.459) rewrites as

(11.460) \[
X_{n,0,s} = \sum_{w=(w_1,w_2,\ldots,w_n) \in \{1,2,3,\ldots\}^n; w_1 \leq w_2 \leq \ldots \leq w_n; |\{w_1,w_2,\ldots,w_n\}|=n-s} x_{w_1} x_{w_2} \cdots x_{w_n}.
\]

For any fixed weak composition \( \alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots) \), the monomial \( x^\alpha \) occurs at most once on the right hand side of (11.460) (because there is at most one way to write \( x^\alpha \) in the form \( x_{w_1} x_{w_2} \cdots x_{w_n} \) for a \( w = (w_1, w_2, \ldots, w_n) \in \{1,2,3,\ldots\}^n \) satisfying \( w_1 \leq w_2 \leq \ldots \leq w_n \)). We can easily tell whether it occurs or not by looking at the size \( |\alpha| \) of \( \alpha \) and the number of positive integers \( i \) satisfying \( \alpha_i \neq 0 \): Namely, the monomial \( x^\alpha \) for a fixed weak composition \( \alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots) \) occurs on the right hand side of (11.460) if and only if the following two properties:

\(^{451}\)It is here that we are using our assumption that \( k = \mathbb{Z} \).
We have \( |a| = n \).

There are precisely \( n - s \) positive integers \( i \) satisfying \( \alpha_i \neq 0 \).

Thus, the monomials satisfying these two properties occur exactly once on the right hand side of (11.460), while all other monomials don’t occur there at all. But the same conclusion can be reached for the right hand side of (11.457). Hence, for any fixed weak composition \( \alpha = (\alpha_1, \alpha_2, \alpha_3, ...) \), the monomial \( x^\alpha \) occurs on the right hand side of (11.460) precisely as often as it occurs on the right hand side of (11.457). Hence, the right hand side of (11.460) equals the right hand side of (11.457). Therefore, the left hand side of (11.460) also equals the left hand side of (11.457). In other words, \( X_{n,0,s} = \sum_{\lambda \in \text{Par}_n; \ell(\lambda) = n-s} m_\lambda \). This proves (11.456).

Now, (11.456) yields \( X_{n,0,s} = \sum_{\lambda \in \text{Par}_n; \ell(\lambda) = n-s} m_\lambda \in \Lambda \), so that \( X_{n,d,s} = X_{n,0,s} \in \Lambda \). This completes the induction step in Case 2.

Thus, the induction step is complete in both Cases 1 and 2. This finally completes the induction step. Exercise 2.87(b) is thus solved.

(a) First solution of Exercise 2.87(a): Recall the power series \( X_{n,d,s} \) defined in Exercise 2.87(b) for any \( d \in \mathbb{N} \) and \( s \in \mathbb{N} \). Exercise 2.87(b) yields that \( X_{n,k,0} \in \Lambda \). But an \( n \)-tuple \( w \) is Smirnov if and only if its stagnation set \( \text{Stag}(w) \) is empty, i.e., if and only if \( |\text{Stag}(w)| = 0 \). Hence, \( X_{n,k,0} \) is the sum of the monomials \( x_{w_1}x_{w_2} \cdots x_{w_n} \) over all Smirnov \( n \)-tuples \( w \in \{1, 2, 3, \ldots\}^n \) satisfying \( |\text{Des}(w)| = k \). This shows that \( X_{n,k,0} = X_{n,k} \). Thus, \( X_{n,k} \in \Lambda \), so that Exercise 2.87(a) is solved.

Second solution of Exercise 2.87(a): Here is an alternative solution to Exercise 2.87(a), which avoids using part (b). It is an application of [151, proof of Theorem 4.5] to our special setting.

We proceed similarly to the proof of Proposition 2.11: It suffices to show that for every positive integer \( p \), the power series \( X_{n,k} \) is invariant under swapping the variables \( x_p \) and \( x_{p+1} \). So let us fix a positive integer \( p \).

Let \( S_{n,k} \) denote the set of all Smirnov \( n \)-tuples \( w \in \{1, 2, 3, \ldots\}^n \) satisfying \( |\text{Des}(w)| = k \). Then, the definition of \( X_{n,k} \) rewrites as follows:

\[
X_{n,k} = \sum_{w=(w_1,w_2,\ldots,w_n) \in S_{n,k}} x_{w_1}x_{w_2}\cdots x_{w_n}.
\]

For every \( w \in S_{n,k} \), let \( x_w \) denote the monomial \( x_{w_1}x_{w_2} \cdots x_{w_n} \), where \( w \) is written in the form \( w = (w_1, w_2, \ldots, w_n) \). Then, (11.461) rewrites as

\[
X_{n,k} = \sum_{w \in S_{n,k}} x_w.
\]

We need to show that this power series \( X_{n,k} \) is invariant under swapping the variables \( x_p \) and \( x_{p+1} \). In order to do so, it is clearly enough to provide an involution \( J : S_{n,k} \to S_{n,k} \) which has the property that for every \( w \in S_{n,k} \), the monomial \( x_{J(w)} \) is obtained from the monomial \( x_w \) by swapping the variables \( x_p \) and \( x_{p+1} \).

To define the involution \( J \), we introduce some notations. If \( w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n \) is any \( n \)-tuple, then a nonempty interval \( I \) of \{1, 2, 3, \ldots, n\} will be called a \((p, p+1)\)-interval of \( w \) if every \( i \in I \) satisfies \( w_i \in \{p, p+1\} \). A \((p, p+1)\)-run of \( w \) will mean a \((p, p+1)\)-interval of \( w \) maximal with respect to inclusion. For instance, if \( n = 9 \), \( w = (3, 1, 2, 5, 4, 2, 3, 2, 4) \) and \( p = 2 \), then the \((p, p+1)\)-intervals of \( w \) are \{1\}, \{3\}, \{6\}, \{7\}, \{8\}, \{6, 7\}, \{7, 8\} and \{6, 7, 8\} (since the letters of \( w \) belonging to \{p, p+1\} are the 1-st, 3-rd, the 6-th, the 7-th and the 8-th letter), while only \{1\}, \{3\} and \{6, 7, 8\} are \((p, p+1)\)-runs of \( w \).

An interval of \{1, 2, 3, \ldots, n\} is said to be even if it has even size, and odd if it has odd size.

It is easy to see that if \( w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n \) is any Smirnov \( n \)-tuple, and \( \{a+1, a+2, \ldots, b\} \) is any \((p, p+1)\)-interval of \( w \), then the \((b - a)\)-tuple \((w_{a+1}, w_{a+2}, \ldots, w_b)\) alternates between \( p \)'s and \((p+1)\)'s, that is, has one of the forms

\[
(p, p+1, p, p+1, p, \ldots, p+1)
\]

and

\[
(p+1, p+1, p, p+1, p, \ldots, p+1)
\]

(if \( b - a \) is even) or one of the forms

\[
(p, p+1, p, p+1, p, \ldots, p)
\]

and

\[
(p+1, p+1, p, p+1, p, \ldots, p+1)
\]

(if \( b - a \) is odd).
Now, we define a map $J : S_{n,k} \to S_{n,k}$ as follows: Let $w = (w_1, w_2, \ldots, w_n) \in S_{n,k}$. We know that $w$ is Smirnov and satisfies $|\text{Des}(w)| = k$. It is clear that the $(p, p + 1)$-runs of $w$ are pairwise disjoint and even separated from each other by at least 1 (this means that if $\alpha$ and $\beta$ are elements of two distinct $(p, p + 1)$-runs of $w$, then $|\alpha - \beta| > 1$). For every $i \in \{1, 2, \ldots, n\}$, define a positive integer $w'_i$ as follows:

- If $i$ belongs to an odd $(p, p + 1)$-run of $w$, then set $w'_i = \begin{cases} p + 1, & \text{if } w_i = p; \\ p, & \text{if } w_i = p + 1. \end{cases}$
- Otherwise, set $w'_i = w_i$.

Thus, we have defined an $n$-tuple $(w'_1, w'_2, \ldots, w'_n)$ of positive integers. Denote this $n$-tuple $(w'_1, w'_2, \ldots, w'_n)$ by $w'$. The reader can easily check that this new $n$-tuple $w'$ is again Smirnov and satisfies $|\text{Des}(w')| = k$.

In other words, $w' \in S_{n,k}$. Now, set $J(w) = w'$. We have thus defined a map $J : S_{n,k} \to S_{n,k}$. It is easy to see that, for every $w \in S_{n,k}$, the $(p, p + 1)$-runs of $J(w)$ are exactly the $(p, p + 1)$-runs of $w$, and applying the map $J$ to $J(w)$ precisely reverts the changes made by the map $J$ to $w$. In other words, $J \circ J = \text{id}$, so that the map $J$ is an involution. Finally, it is straightforward to see that for every $w \in S_{n,k}$, we have the following facts:

1. The number of entries equal to $p + 1$ in $J(w)$ equals the number of entries equal to $p$ in $w$.
2. The number of entries equal to $p$ in $J(w)$ equals the number of entries equal to $p + 1$ in $w$.
3. For every $j \in \{1, 2, 3, \ldots\} \setminus \{p, p + 1\}$, the number of entries equal to $j$ in $J(w)$ equals the number of entries equal to $j$ in $w$.

These three statements, combined, show that for every $w \in S_{n,k}$, the monomial $x_{J(w)}$ is obtained from the monomial $x_w$ by swapping the variables $x_p$ and $x_{p+1}$. This concludes our solution of Exercise 2.87(a).

**Remark:** We could have given an alternative solution to Exercise 2.87(b) that would still rely on (11.446), but proceed by induction on $n + s$ (rather than $n + d$) and handle the case $s = 0$ separately (rather than the case $d = 0$ as we did). In the case $s = 0$, the assertion of Exercise 2.87(b) follows from Exercise 2.87(a). (But this only works combined with a solution to Exercise 2.87(b) that does not rely on Exercise 2.87(b).)

(c) For every $d \in \mathbb{N}$, $s \in \mathbb{N}$ and $i \in \{1, 2, \ldots, n - 1\}$, we define the power series $D_i, S_i$ and $A_i$ as in the solution to Exercise 2.87(b) above.

Let us first show that if $n$ is a positive integer, then any $d \in \mathbb{N}$ and $s \in \mathbb{N}$ satisfy

$$(d + 1) U_{n,d+1,s} + (s + 1) U_{n,d,s+1} + (n - 1 - d - s) U_{n,d,s} = (d + 1) X_{n,d+1,s}. \quad (11.463)$$

**Proof of (11.463):** Let $n$ be a positive integer. Let $d \in \mathbb{N}$ and $s \in \mathbb{N}$. Let $i \in \{1, 2, \ldots, n - 1\}$ be arbitrary.

Before we make any new definitions, let us recall a statement that we have shown during our proof of (11.446) (back when we were solving Exercise 2.87(b)): The power series $D_{i} \in \mathbb{k}[x]$ is the sum of the monomials $x_{w_{1}}x_{w_{2}} \cdots x_{w_{n}}$ over all $n$-tuples $w = (w_{1}, w_{2}, \ldots, w_{n}) \in \{1, 2, 3, \ldots\}^{n}$ satisfying $|\text{Des}(w)| \setminus \{i\} = d$, $|\text{Stag}(w)| \setminus \{i\} = s$ and $w_{i} > w_{i+1}$. Written as a formula, this yields

$$D_{i} = \sum_{\substack{w=(w_{1}, w_{2}, \ldots, w_{n}) \in \{1, 2, 3, \ldots\}^{n}; \\ |\text{Des}(w)| \setminus \{i\} = d; \\ |\text{Stag}(w)| \setminus \{i\} = s; \\ w_{i} > w_{i+1}}} x_{w_{1}}x_{w_{2}} \cdots x_{w_{n}}. \quad (11.464)$$

Applying this to $n - i$ instead of $i$, we obtain

$$D_{n-i} = \sum_{\substack{w=(w_{1}, w_{2}, \ldots, w_{n}) \in \{1, 2, 3, \ldots\}^{n}; \\ |\text{Des}(w)| \setminus \{n-i\} = d; \\ |\text{Stag}(w)| \setminus \{n-i\} = s; \\ w_{n-i} > w_{n-i+1}}} x_{w_{1}}x_{w_{2}} \cdots x_{w_{n}}. \quad (11.464)$$

Now, let us make some more definitions:

- Define a power series $D'_{i} \in \mathbb{k}[x]$ as the sum of the monomials $x_{w_{1}}x_{w_{2}} \cdots x_{w_{n}}$ over all $n$-tuples $w = (w_{1}, w_{2}, \ldots, w_{n}) \in \{1, 2, 3, \ldots\}^{n}$ satisfying $|\text{Des}(w)| = d + 1$, $|\text{Stag}(w)| = s$, $w_{1} < w_{n}$ and $w_{i} > w_{i+1}$.

45. Informally speaking, $(w'_1, w'_2, \ldots, w'_n)$ is simply obtained by changing those $p$'s and $p+1$'s in $w$ whose positions belong to odd $(p, p+1)$-runs of $w$ into $p+1$'s and $p$'s, respectively, while leaving all other entries of $w$ intact. For instance, in our above example of $n = 9$, $w = (3, 1, 2, 5, 4, 2, 3, 2, 4)$ and $p = 2$, we would have $(w'_1, w'_2, \ldots, w'_n) = (2, 1, 3, 5, 4, 2, 3, 2, 4)$. (All letters 2 and 3 have been changed here because all $(p, p+1)$-runs of this $w$ were odd.) For another example, if $n = 5$, $w = (2, 3, 1, 5, 2)$ and $p = 2$, then $(w'_1, w'_2, \ldots, w'_n) = (2, 3, 1, 5, 3)$ (the first two letters are unchanged since the $(p, p+1)$-run $(1, 2)$ is even).

46. The idea is that the map $J$ switches the number of $p$'s with the number of $(p+1)$'s in any odd $(p, p+1)$-run, while the numbers in an even $(p, p+1)$-run are already equal to begin with.
• Define a power series \( S'_i \in k[[x]] \) as the sum of the monomials \( x_{w_1}x_{w_2} \cdots x_{w_n} \) over all \( n \)-tuples \( w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n \) satisfying \( |\text{Des}(w)| = d, \ |\text{Stag}(w)| = s + 1, \ w_1 < w_n \) and \( w_i = w_{i+1} \).

• Define a power series \( A'_i \in k[[x]] \) as the sum of the monomials \( x_{w_1}x_{w_2} \cdots x_{w_n} \) over all \( n \)-tuples \( w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n \) satisfying \( |\text{Des}(w)| = d, \ |\text{Stag}(w)| = s, \ w_1 < w_n \) and \( w_i < w_{i+1} \).

These three power series \( D'_i, S'_i \) and \( A'_i \) were defined in obvious analogy to the power series \( D_i, S_i \) and \( A_i \) defined in our solution to Exercise 2.87(b) above. In the same way as we have proved (11.447) back there, we can see that

\[
(11.465) \quad D'_i + S'_i + A'_i = \sum_{\substack{w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n; \\
|\text{Des}(w)\} \{i\}| = d; \ |\text{Stag}(w)\} \{i\}| = s;}} x_{w_1}x_{w_2} \cdots x_{w_n}.
\]

But the sum on the right hand side of (11.465) can be rewritten in terms of more familiar sums. In fact, the \( n \)-tuples \( w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots\}^n \) are in bijection with the pairs \((u, v)\) consisting of an \( i \)-tuple \( u = (u_1, u_2, \ldots, u_i) \in \{1, 2, 3, \ldots \}^i \) and an \((n - i)\)-tuple \( v = (v_1, v_2, \ldots, v_{n-i}) \in \{1, 2, 3, \ldots \}^{n-i} \). This bijection sends an \( n \)-tuple \((w_1, w_2, \ldots, w_n)\) to the pair ((\(w_1, w_2, \ldots, w_i\)), (\(w_{i+1}, w_{i+2}, \ldots, w_n\)), and has the property that \(|\text{Des}(w)\} \{i\}| = |\text{Des}(u)| + |\text{Des}(v)| \) \[454\], \( |\text{Stag}(w)\} \{i\}| = |\text{Stag}(u)| + |\text{Stag}(v)| \) \[455\], \( x_{w_1}x_{w_2} \cdots x_{w_n} = x_{u_1}x_{u_2} \cdots x_{u_i}x_{v_1}x_{v_2} \cdots x_{v_{n-i}}, \) \( u_1 = w_1 \) and \( w_n = v_{n-i} \). Hence,

\[
(11.466) \quad \sum_{\substack{u = (u_1, u_2, \ldots, u_i) \in \{1, 2, 3, \ldots \}^i; \\
|\text{Des}(u)| + |\text{Des}(v)| = d; \ |\text{Stag}(u)| + |\text{Stag}(v)| = s; \ u_1 < v_{n-i}}} x_{u_1}x_{u_2} \cdots x_{u_i}x_{v_1}x_{v_2} \cdots x_{v_{n-i}} = \sum_{\substack{u = (u_1, u_2, \ldots, u_i) \in \{1, 2, 3, \ldots \}^i; \\
|\text{Des}(u)| + |\text{Des}(v)| = d; \ |\text{Stag}(u)| + |\text{Stag}(v)| = s; \ u_1 < v_{n-i}}} x_{u_1}x_{u_2} \cdots x_{u_i}x_{v_1}x_{v_2} \cdots x_{v_{n-i}}.
\]

On the other hand, the pairs \((v, u)\) consisting of an \((n - i)\)-tuple \( v = (v_1, v_2, \ldots, v_{n-i}) \in \{1, 2, 3, \ldots \}^{n-i} \) and an \( i \)-tuple \( u = (u_1, u_2, \ldots, u_i) \in \{1, 2, 3, \ldots \}^i \) are in bijection with the \( n \)-tuples \( w = (w_1, w_2, \ldots, w_n) \in \{1, 2, 3, \ldots \}^n \). This bijection sends a pair \((v, u)\) with \( v = (v_1, v_2, \ldots, v_{n-i}) \) and \( u = (u_1, u_2, \ldots, u_i) \) to the \( n \)-tuple \( (v_1, v_2, \ldots, v_{n-i}, u_1, u_2, \ldots, u_i) \in \{1, 2, 3, \ldots \}^n \), and has the property that \(|\text{Des}(w)\} \{n-i\}| = |\text{Des}(u)| + |\text{Des}(v)| \) \[456\], \( |\text{Stag}(w)\} \{n-i\}| = |\text{Stag}(u)| + |\text{Stag}(v)| \) \[457\], \( x_{u_1}x_{u_2} \cdots x_{u_i}x_{v_1}x_{v_2} \cdots x_{v_{n-i}} = x_{u_1}x_{u_2} \cdots x_{u_i}x_{v_1}x_{v_2} \cdots x_{v_{n-i}}, \) \( u_{n-i+1} = u_1 \) and \( w_{n-i} = v_{n-i} \). We can use this bijection to transform the sum on the right hand side of

\[454\] In fact, \(|\text{Des}(w)\} \{i\}| = |\text{Des}(u)| + |\text{Des}(v)| \) if \( i \) is the union of the set \(|\text{Des}(u)|\) with the set \(|\text{Des}(v)|\) shifted by \( i \) (that is, the set \(|p+i| \ p \in |\text{Des}(v)|\)).

\[455\] For similar reasons

\[456\] In fact, \(|\text{Des}(w)\} \{n-i\}| = |\text{Des}(u)| + |\text{Des}(v)| \) if \( n-i \) is the union of the set \(|\text{Des}(u)|\) with the set \(|\text{Des}(u)|\) shifted by \( n-i \) (that is, the set \(|p+(n-i)| \ p \in |\text{Des}(u)|\)).

\[457\] For similar reasons
and obtain
\[
\sum_{v=(v_1, v_2, \ldots, v_{n-i})\in\{1, 2, 3, \ldots\}^{n-i};
\quad u=(u_1, u_2, \ldots, u_i)\in\{1, 2, 3, \ldots\}^i; 
\quad |\text{Des}(u)|+|\text{Des}(v)|=d; 
\quad |\text{Stag}(u)|+|\text{Stag}(v)|=s; 
\quad u_1 < v_{n-i}} \cdots x_{n-1} x_1 x_2 \cdots x_{n+i} = \sum_{w=(w_1, w_2, \ldots, w_n)\in\{1, 2, 3, \ldots\}^n; 
\quad |\text{Des}(w)|\{n-i\}=d; 
\quad |\text{Stag}(w)|\{n-i\}=s; 
\quad w_{n-i+1} < w_n} x_{n+1} x_{n+2} \cdots x_{n+i} \quad \text{(11.467)}
\]

Now, (11.465) becomes
\[
D'_i + S'_i + A'_i = \sum_{v=(v_1, v_2, \ldots, v_{n-i})\in\{1, 2, 3, \ldots\}^{n-i}; 
\quad w=(w_1, w_2, \ldots, w_n)\in\{1, 2, 3, \ldots\}^n; 
\quad (\text{by (11.466))}} \sum_{u=(u_1, u_2, \ldots, u_i)\in\{1, 2, 3, \ldots\}^i; 
\quad |\text{Des}(u)|+|\text{Des}(v)|=d; 
\quad |\text{Stag}(u)|+|\text{Stag}(v)|=s; 
\quad u_1 < v_{n-i}} \cdots x_{n-1} x_1 x_2 \cdots x_{n+i}
\]

Now, let us forget that we fixed \(i\). We thus have defined \(D'_i, S'_i\) and \(A'_i\) and proven the equality (11.468) for all \(i \in \{1, 2, \ldots, n-1\}\).

Now, let us take a closer look at the sums \(\sum_{i=1}^{n-1} D'_i, \sum_{i=1}^{n-1} S'_i\) and \(\sum_{i=1}^{n-1} A'_i\):

- Similarly to how we proved (11.452), we can show that

\[
\sum_{i=1}^{n-1} D'_i = (d+1) U_{n,d+1,s} \quad \text{(11.469)}
\]

- Similarly to how we proved (11.453), we can see that

\[
\sum_{i=1}^{n-1} S'_i = (s+1) U_{n,d,s+1} \quad \text{(11.470)}
\]

- Similarly to how we proved (11.454), we can see that

\[
\sum_{i=1}^{n-1} A'_i = (n-1 - d - s) U_{n,d,s} \quad \text{(11.471)}
\]

Now, summing the equality (11.468) over all \(i \in \{1, 2, \ldots, n-1\}\) yields
\[
\sum_{i=1}^{n-1} (D'_i + S'_i + A'_i) = \sum_{i=1}^{n-1} D' - \sum_{i=1}^{n-1} A' = (d+1) X_{n,d+1,s} \quad \text{(by (11.452))}
\]

Hence,
\[
(d+1) X_{n,d+1,s} = \sum_{i=1}^{n-1} (D'_i + S'_i + A'_i) = \sum_{i=1}^{n-1} D'_i + \sum_{i=1}^{n-1} S'_i + \sum_{i=1}^{n-1} A'_i \quad \text{(by (11.469))} = (d+1) U_{n,d+1,s} \]
\[
= (s+1) U_{n,d,s+1} + (n-1 - d - s) U_{n,d,s} \quad \text{(by (11.471))}
\]

This proves (11.463).

In solving Exercise 2.87(c), we WLOG assume that \(k = \mathbb{Z}\) (for the same reason why we could assume \(k = \mathbb{Z}\) in solving Exercise 2.87(b)).
Now, we are going to prove $U_{n,d,s} \in \Lambda$ by strong induction over $n + d$. So (for the induction step) we need to show that $U_{n',d',s'} \in \Lambda$, and we can assume (as the induction hypothesis) that $U_{n',d',s'} \in \Lambda$ is already known to hold for any nonnegative integers $n'$, $d'$ and $s'$ satisfying $n' + d' < n + d$.

We must be in one of the following two cases:

**Case 1:** We have $d > 0$.

**Case 2:** We have $d = 0$.

In Case 1, we can proceed in the same way as in the corresponding case of the solution of Exercise 2.87(b) (with the only difference that we now have to use $X_{n,d,s} \in \Lambda$, but this follows from the already solved Exercise 2.87(b)).

Let us now consider Case 2. In this case, we have $d = 0$. We now distinguish between two subcases:

**Subcase 2.1:** We have $s = n - 1$.

**Subcase 2.2:** We have $s \neq n - 1$.

Let us consider Subcase 2.1 first. In this subcase, we have $s = n - 1$. We will show that $U_{n,d,s} = 0$ in this subcase.

Indeed, let $w = (w_1, w_2, \ldots , w_n) \in \{1, 2, 3, \ldots \}^n$ be such that $|\text{Des}(w)| = d$, $|\text{Stag}(w)| = s$ and $w_1 < w_n$. Then, Stag($w$) is a subset of $\{1, 2, \ldots , n - 1\}$ whose cardinality is $|\text{Stag}(w)| = s = n - 1 = |\{1, 2, \ldots , n - 1\}|$. Obviously, the only such subset is $\{1, 2, \ldots , n - 1\}$ itself, and so Stag($w$) must be $\{1, 2, \ldots , n - 1\}$. Hence, every $j \in \{1, 2, \ldots , n - 1\}$ satisfies $j \in \{1, 2, \ldots , n - 1\} = \text{Stag}(w) = \{i \in \{1, 2, \ldots , n - 1\} : w_i = w_{i+1}\}$ and thus $w_j = w_{j+1}$.

Now, forget that we fixed $w$. We thus have obtained a contradiction for every $w = (w_1, w_2, \ldots , w_n) \in \{1, 2, 3, \ldots \}^n$ satisfying $|\text{Des}(w)| = d$ and $w_1 < w_n$. Therefore, there exists no such $w$. Hence, the sum on the right hand side of (2.49) is empty and thus equals 0. Thus, (2.49) rewrites as $U_{n,d,s} = 0 \in \Lambda$.

Hence, the induction step is complete in Subcase 2.1.

Let us now consider Subcase 2.2. In this case, $s \neq n - 1$. We will show that $U_{n,d,s} = X_{n,d,s}$ in this subcase.

Indeed, let $w = (w_1, w_2, \ldots , w_n) \in \{1, 2, 3, \ldots \}^n$ be such that $|\text{Des}(w)| = d$ and $|\text{Stag}(w)| = s$. We will show that $w_1 < w_n$.

We have $|\text{Des}(w)| = d = 0$, so that the set Des($w$) is empty. In other words, every $j \in \{1, 2, \ldots , n - 1\}$ satisfies $w_j \leq w_{j+1}$. Hence, we have the chain of inequalities $w_1 \leq w_2 \leq \cdots \leq w_n$. At least one inequality in this chain must be strict (because otherwise, we would have $w_1 = w_2 = \cdots = w_n$, thus $w_j = w_{j+1}$ for every $j \in \{1, 2, \ldots , n - 1\}$, thus $j \in \text{Stag}(w)$ for every $j \in \{1, 2, \ldots , n - 1\}$, which would lead to Stag($w$) = $\{1, 2, \ldots , n - 1\}$ and thus $|\text{Stag}(w)| = |\{1, 2, \ldots , n - 1\}| = n - 1$, in contradiction to Stag($w$) = $s \neq n - 1$, and thus we have $w_1 < w_n$.

Now, let us forget that we fixed $w$. We thus have proven that every $w = (w_1, w_2, \ldots , w_n) \in \{1, 2, 3, \ldots \}^n$ satisfying $|\text{Des}(w)| = d$ and $|\text{Stag}(w)| = s$ automatically satisfies $w_1 < w_n$. Hence, the condition $w_1 < w_n$ under the summation sign $\sum_{\substack{w = (w_1, w_2, \ldots , w_n) \in \{1, 2, 3, \ldots \}^n; \\
|\text{Des}(w)| = d; |\text{Stag}(w)| = s}} w_1 < w_n$ is redundant (i.e., can be removed without changing the range of the summation). Thus,

$$
\sum_{\substack{w = (w_1, w_2, \ldots , w_n) \in \{1, 2, 3, \ldots \}^n; \\
|\text{Des}(w)| = d; |\text{Stag}(w)| = s}} x_{w_1} x_{w_2} \cdots x_{w_n} = \sum_{\substack{w = (w_1, w_2, \ldots , w_n) \in \{1, 2, 3, \ldots \}^n; \\
|\text{Des}(w)| = d; |\text{Stag}(w)| = s}} x_{w_1} x_{w_2} \cdots x_{w_n} = X_{n,d,s}
$$

(by (11.445)). Now, (2.49) becomes

$$
U_{n,d,s} = \sum_{\substack{w = (w_1, w_2, \ldots , w_n) \in \{1, 2, 3, \ldots \}^n; \\
|\text{Des}(w)| = d; |\text{Stag}(w)| = s}} x_{w_1} x_{w_2} \cdots x_{w_n} = X_{n,d,s} \in \Lambda
$$

(by Exercise 2.87(b)). Hence, the induction step is complete in Subcase 2.2.

The induction step is thus complete in Case 1 and in each of the two Subcases 2.1 and 2.2. These are all cases, and so the induction step is finally complete. We have thus proven that

(11.472) $U_{n,d,s} \in \Lambda$ for all positive integers $n$, all $d \in \mathbb{N}$ and all $s \in \mathbb{N}$.

In order to complete the solution of Exercise 2.87(e), we still need to prove that $V_{n,d,s}$ and $W_{n,d,s}$ belong to $\Lambda$ for all positive integers $n$, all $d \in \mathbb{N}$ and all $s \in \mathbb{N}$.
Let $n$ be a positive integer, let $d \in \mathbb{N}$ and let $s \in \mathbb{N}$. The equality (2.51) becomes

\[
W_{n,d,s} = \sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,\ldots,n\}^n; \atop \text{Des}(w)=d; \text{Stag}(w)=s; \atop w_1 \geq w_n} x_{w_1} x_{w_2} \cdots x_{w_n} = \sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,\ldots,n\}^n; \atop \text{Des}((w_n,w_{n-1},\ldots,w_1))=d; \text{Stag}((w_n,w_{n-1},\ldots,w_1))=s; \atop w_1 \geq w_n} x_{w_n} x_{w_{n-1}} \cdots x_{w_1}
\]

(here, we substituted $(w_n,w_{n-1},\ldots,w_1)$ for $w=(w_1,w_2,\ldots,w_n)$ in the sum)

\[
= \sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,\ldots,n\}^n; \atop \text{Des}((w_n,w_{n-1},\ldots,w_1))=d; \text{Stag}((w_n,w_{n-1},\ldots,w_1))=s; \atop w_1 > w_n} x_{w_1} x_{w_2} \cdots x_{w_n}
\]

\[
= \sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,\ldots,n\}^n; \atop \text{Des}((w_n,w_{n-1},\ldots,w_1))=d; \text{Stag}((w_n,w_{n-1},\ldots,w_1))=s; \atop w_n > w_1} x_{w_1} x_{w_2} \cdots x_{w_n}
\]

\[
= \sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,\ldots,n\}^n; \atop \text{Des}((w_n,w_{n-1},\ldots,w_1))=d; \text{Stag}((w_n,w_{n-1},\ldots,w_1))=s; \atop w_n > w_1} \left( \begin{array}{c}
\text{since every } w=(w_1,w_2,\ldots,w_n) \in \{1,2,\ldots,n\}^n \text{ satisfies} \\
\text{Des}((w_n,w_{n-1},\ldots,w_1)) = n-1 - \text{Des}(w) - \text{Stag}(w) \\
\text{and } \text{Stag}((w_n,w_{n-1},\ldots,w_1)) = \text{Stag}(w) \\
\text{is the union of its three disjoint subsets} \\
\{i \in \{1,2,\ldots,n-1\} \mid w_i < w_{i+1}\}, \text{ Des } w \text{ and Stag } w
\end{array} \right)
\]

\[
= \sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,\ldots,n\}^n; \atop \text{Des}(w)=n-1-d-s; \text{Stag}(w)=s; \atop w_1 < w_n} x_{w_1} x_{w_2} \cdots x_{w_n}
\]

\[
\left( \begin{array}{c}
\text{because for a } w=(w_1,w_2,\ldots,w_n) \in \{1,2,\ldots,n\}^n \text{ satisfying} \\
\text{Stag}(w)=s, \text{ the assertions } n-1 - \text{Des}(w) - \text{Stag}(w) = d \\
\text{and } w_n > w_1 \text{ are equivalent to} \ \text{Des}(w)=n-1-d-s \text{ and } w_n < w_1
\end{array} \right)
\]

\[
= U_{n,n-1-d-s,s} \quad \text{(by (2.49), applied to } n-1-d-s \text{ instead of } d)
\]

Finally, (11.445) becomes

\[
X_{n,d,s} = \sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,\ldots,n\}^n; \atop \text{Des}(w)=d; \text{Stag}(w)=s} x_{w_1} x_{w_2} \cdots x_{w_n}
\]

\[
= \sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,\ldots,n\}^n; \atop \text{Des}(w)=d; \text{Stag}(w)=s; \atop w_1 < w_n} x_{w_1} x_{w_2} \cdots x_{w_n}
\]

\[
\left( \begin{array}{c}
\text{by (2.49)}
\end{array} \right)
\]

\[
= U_{n,d,s}
\]

\[
= \sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,\ldots,n\}^n; \atop \text{Des}(w)=d; \text{Stag}(w)=s; \atop w_1 = w_n} x_{w_1} x_{w_2} \cdots x_{w_n}
\]

\[
\left( \begin{array}{c}
\text{by (2.50)}
\end{array} \right)
\]

\[
= V_{n,d,s}
\]

\[
\left( \begin{array}{c}
\text{by (2.51)}
\end{array} \right)
\]

\[
\left( \begin{array}{c}
\text{by Exercise 2.87(b)}
\end{array} \right)
\]

so that $V_{n,d,s} = \sum_{w=(w_1,w_2,\ldots,w_n)\in\{1,2,\ldots,n\}^n; \atop \text{Des}(w)=d; \text{Stag}(w)=s; \atop w_1 < w_n} x_{w_1} x_{w_2} \cdots x_{w_n}$ and $W_{n,d,s}$, where $V_{n,d,s} \in \Lambda - \Lambda \subseteq \Lambda$. 


We have thus shown that $V_{n,d,s} \in \Lambda$ and $W_{n,d,s} \in \Lambda$. Combined with $U_{n,d,s} \in \Lambda$ (this follows from (11.472)), this completes the solution of Exercise 2.87(c).

11.61. Solution to Exercise 2.89. Solution to Exercise 2.89. We start out with two definitions:

- If $m$ and $q$ are integers satisfying $0 \leq q \leq m$, and if $U = (u_{i,j})_{i,j=1,2,...,m} \in k^{m \times m}$ is an $m \times m$-matrix, then NWsm$_qU$ will mean the matrix $(u_{i,j})_{i,j=1,2,...,m,q} \in k^{q \times q}$. This is the submatrix of $U$ obtained by removing all rows other than the first $q$ rows and then removing all columns other than the first $q$ columns.\(^{458}\)

- A square matrix $(u_{i,j})_{i,j=1,2,...,m} \in k^{m \times m}$ is said to be nearly lower-triangular if we have

$$u_{i,j} = 0 \text{ for every } (i,j) \in \{1,2,...,m\}^2 \text{ satisfying } j > i + 1.$$

(Thus, informally, a square matrix is nearly lower-triangular if and only if all its entries above the superdiagonal are 0, where the superdiagonal is the set of all cells which lie just one step north of a cell on the diagonal.)

We will now show a lemma:

**Lemma 11.9.** Let $m \in \mathbb{N}$. Let $U = (u_{i,j})_{i,j=1,2,...,m} \in k^{m \times m}$ be a nearly lower-triangular $m \times m$-matrix. Then,

$$\det U = \sum_{r=1}^{m} (-1)^{m-r} u_{m,r} \det (\text{NWsm}_{r-1} U) \cdot \prod_{k=r}^{m-1} u_{k,k+1}.$$

*Proof of Lemma 11.9.* Fix some $r \in \{1,2,...,m\}$. We define the following four matrices:

- the $(r - 1) \times (r - 1)$-matrix $P = (u_{i,j})_{i,j=1,2,...,r-1}$, which is obtained from the matrix $U$ by removing all rows other than the first $r - 1$ rows and then removing all columns other than the first $r - 1$ columns;

- the $(r - 1) \times (m - r)$-matrix $Q = (u_{i,r+j})_{i=1,2,...,r-1; j=1,2,...,m-r}$, which is obtained from the matrix $U$ by removing all rows other than the first $r - 1$ rows and then removing all columns other than the last $m - r$ columns;

- the $(m - r) \times (r - 1)$-matrix $R = (u_{r-1+i,j})_{i=1,2,...,m-r; j=1,2,...,r-1}$, which is obtained from the matrix $U$ by removing all rows other than the $r$-th, the $(r+1)$-st, etc., the $(m-1)$-st row, and then removing all columns other than the first $r - 1$ columns;

- the $(m - r) \times (m - r)$-matrix $S = (u_{r-1+i,r+j})_{i,j=1,2,...,m-r}$, which is obtained from the matrix $U$ by removing all rows other than the $r$-th, the $(r+1)$-st, etc., the $(m-1)$-st row, and then removing all columns other than the last $m - r$ columns.

Now, the matrix $U$ can be written as a block matrix as follows: $U =$ \( \begin{pmatrix} P & v & Q \\ R & w & S \\ x & y & z \end{pmatrix} \), where \( \begin{pmatrix} v \\ w \\ y \end{pmatrix} \) is the $r$-th column of $U$, and where \( \begin{pmatrix} x & y & z \end{pmatrix} \) is the $m$-th row of $U$.\(^{459}\) Hence,

\[
(11.473)
\]

But the matrix $Q = (u_{r+i,j})_{i,j=1,2,...,m-r}$ is the zero matrix (since $U$ is nearly lower-triangular), and the matrices $P$ and $S$ are square matrices. Hence, \( \begin{pmatrix} P & Q \\ R & S \end{pmatrix} \) is a block lower-triangular matrix with

\(^{458}\) The notation NWsm$_qU$ stands short for “$q$-th northwest submatrix of $U$”. It is the kind of submatrices whose determinants usually figure in the Sylvester criterion for the positive definiteness of a matrix.

\(^{459}\) Despite being labelled with lowercase letters, $v$, $w$, $x$, $y$ and $z$ are still blocks, although each has (at least) one of its dimensions equal to 1 (and $y$ is a $1 \times 1$-block).
diagonal blocks $P$ and $S$. Thus, its determinant is

$$\det \begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \det P \cdot \det S$$

(since the determinant of a block lower-triangular matrix is known to equal the product of the determinants of its diagonal blocks).

But the matrix $S = (u_{r-1+i,j})_{i,j=1,2,\ldots,m-r}$ is lower-triangular (since $U$ is nearly lower-triangular). Since it is well-known that the determinant of a lower-triangular matrix equals the product of its diagonal entries, we can therefore compute the determinant $\det S$ of $S$ as follows:

$$\det S = \prod_{i=1}^{m-r} u_{r-1+i,i} = \prod_{k=r}^{m-1} u_{k,k+1} \quad \text{(here, we have substituted $k$ for $r-1+i$ in the sum).}$$

Moreover, $P = NWs_{m-1}U$ (since comparing the definitions of $P$ and of $NWs_{m-1}U$ shows that these two matrices are the same). Now, applying the map $\det$ to both sides of $(11.473)$, we obtain

$$\det (\text{the matrix obtained from } U \text{ by removing the } m\text{-th row and the } r\text{-th column}) = \det \begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \det P \cdot \det S \quad \text{(by } (11.474)\text{)}$$

$$\det \left( NWs_{m-1}U \right) \cdot \prod_{k=r}^{m-1} u_{k,k+1}. \quad \text{(by } (11.475)\text{)}$$

Now, forget that we fixed $r$. We can compute the determinant $\det U$ by Laplace expansion along the $m$-th row, thus obtaining

$$\det U = \sum_{r=1}^{m} \quad \begin{equation} \tag{11.476} \end{equation}$$

This proves Lemma 11.9. \qed

Before we solve the actual exercise, we record one further identity that we will be using twice. Namely, we claim that

$$me_m = \sum_{i=1}^{m} (-1)^{i-1} e_{m-i}p_i \quad \text{for every } m \in \mathbb{N}. \quad \text{(11.477)}$$

**Proof of (11.477):** Recall the power series $H(t)$ defined in (2.14), and the power series $E(t)$ defined in (2.15). From (2.16), we know that $E(-t)H(t) = 1$. Differentiating both sides of this equation with respect to $t$, we obtain $(E(-t)H(t))' = 1' = 0$, so that

$$0 = (E(-t)H(t))' = (E(-t))'H(t) + E(-t)H'(t) \quad \text{(by the Leibniz rule)}$$

$$= -E'(-t)H(t) + E(-t)H'(t),$$

and thus $E'(-t)H(t) = E(-t)H'(t)$. Hence, $\frac{E'(-t)}{E(-t)} = \frac{H'(t)}{H(t)}$. 

But Exercise 2.48 yields \( \sum_{m \geq 0} p_{m+1} t^m = \frac{H'(t)}{H(t)} \) Compared with \( \frac{E'(t)}{E(t)} = \frac{H'(t)}{H(t)} \), this yields \( \frac{E'(t)}{E(t)} = \sum_{m \geq 0} p_{m+1} t^m \). Substituting \(-t\) for \(t\) in this equality, we obtain
\[
\frac{E'(t)}{E(t)} = \sum_{m \geq 0} p_{m+1} (-t)^m = \sum_{m \geq 0} p_{m+1} (-1)^m t^m.
\]
Thus,
\[
\frac{E'(t)}{E(t)} = \left( \sum_{m \geq 0} p_{m+1} (-1)^m t^m \right) \left( \sum_{m \geq 0} e_m t^m \right) = \sum_{m \geq 0} \left( \sum_{i=1}^{m-1} p_{i+1} (-1)^i e_{m-i} \right) t^m \quad \text{(by the definition of the product of two power series)}
\]
\[
= \sum_{m \geq 1} \left( \sum_{i=0}^{m-1} p_{i+1} (-1)^i e_{m-1-i} \right) t^{m-1} \quad \text{(here, we substituted \(m-1\) for \(m\) in the first sum)}
\]
\[
= \sum_{m \geq 0} \left( \sum_{i=1}^{m} p_i (-1)^i e_{m-i} \right) t^{m-1} = \sum_{m \geq 0} \left( \sum_{i=1}^{m} (-1)^{i-1} e_{m-i} p_i \right) t^{m-1}.
\]
Compared with
\[
\frac{E'(t)}{E(t)} = \left( \sum_{m \geq 0} e_m t^m \right) \left( \sum_{m \geq 0} e_m t^m \right) = \sum_{m \geq 0} e_m \cdot m t^{m-1} = \sum_{m \geq 0} me_m t^{m-1},
\]
this yields
\[
\sum_{m \geq 0} me_m t^{m-1} = \sum_{m \geq 0} \left( \sum_{i=1}^{m} (-1)^{i-1} e_{m-i} p_i \right) t^{m-1}.
\]
Multiplying both sides of this equality by \(t\), we obtain
\[
\sum_{m \geq 0} me_m t^m = \sum_{m \geq 0} \left( \sum_{i=1}^{m} (-1)^{i-1} e_{m-i} p_i \right) t^m.
\]
Comparing coefficients in this equality of power series, we conclude that every \(m \in \mathbb{N}\) satisfies
\[
me_m = \sum_{i=1}^{m} (-1)^{i-1} e_{m-i} p_i.
\]
This proves (11.477).

Now, let us solve the exercise.
(a) We will solve Exercise 2.89(a) by strong induction over \( n \). Thus, we assume (as the induction hypothesis) that

\[
\det (A_k) = k!e_k \quad \text{for all } k \in \mathbb{N} \text{ satisfying } k < n.
\]

We now need to prove that \( \det (A_n) = n!e_n \).

The matrix \( A_n = (a_{i,j})_{i,j=1,2,...,n} \) is nearly lower-triangular\(^{460}\). Hence, Lemma 11.9 (applied to \( m=n \), \( U=A_n \) and \( u_{i,j} = a_{i,j} \)) yields

\[
\det (A_n) = \sum_{r=1}^{n} (-1)^{n-r} p_{n-r+1} \det (A_{r-1}) \prod_{k=r}^{n-1} k = \sum_{r=1}^{n} (-1)^{n-r} (r-1)! e_{r-1} \prod_{k=r}^{n-1} k \cdot \prod_{k=r}^{n-1} k p_{n-r+1} e_{r-1}
\]

\[
= (n-1)! \sum_{r=1}^{n} (-1)^{n-r} p_{n-r+1} e_{r-1} = (n-1)! \sum_{r=1}^{n} (-1)^{r-1} p_i e_{n-i}
\]

(here, we have substituted \( n - i + 1 \) for \( r \) in the sum)

\[
= (n-1)! n! e_n = n!e_n.
\]

This completes the induction step, and so Exercise 2.89(a) is solved.

(b) We will solve Exercise 2.89(b) by strong induction over \( n \). Thus, we assume (as the induction hypothesis) that

\[
\det (B_k) = p_k \quad \text{for all positive integers } k \text{ satisfying } k < n.
\]

We now need to prove that \( \det (B_n) = p_n \).
The matrix $B_n = (b_{i,j})_{i,j=1,\ldots,n}$ is nearly lower-triangular.\footnote{because for every $(i,j) \in \{1,2,\ldots,n\}^2$ satisfying $j > i + 1$, we have $b_{i,j} = \begin{cases} i e_i, & \text{if } j = 1; \\ e_{i-j+1}, & \text{if } j > 1 \quad \text{(since } j > 1 \text{ because } j > i + 1 \geq 1) \end{cases}$}
Hence, Lemma 11.9 (applied to $m = n$, $U = B_n$ and $u_{i,j} = b_{i,j}$) yields
\[
\det(B_n) = \sum_{r=1}^{n} (-1)^{n-r} b_{n,r} \det \left( NW_{\text{sm}_{r-1}}(B_n) \right) \cdot \prod_{k=r}^{n-1} \frac{b_{k,k+1}}{e_{k-(k+1)+1}} = e_n = 1 \quad \text{(this is easy to see by the definitions of } B_n \text{ and } B_{r-1})
\]
\[
= \sum_{r=1}^{n} (-1)^{n-r} b_{n,r} \det(B_{r-1}) \cdot \prod_{k=r}^{n-1} e_{k-(k+1)+1} = e_{n-r+1} = 1 \quad \text{(by the definition of } b_{k,k+1}, \text{ since } k+1 > k \geq r \geq 1)
\]
\[
= \sum_{r=1}^{n} (-1)^{n-r} b_{n,r} \det(B_{r-1}) \cdot \prod_{k=r}^{n-1} 1 = \sum_{r=1}^{n} (-1)^{n-r} b_{n,r} \det(B_{r-1})
\]
\[
= (-1)^{n-1} b_{n,1} \underbrace{\det(B_{1-1})}_{1=1} + \sum_{r=2}^{n} (-1)^{n-r} e_{n-r+1} p_{r-1}
\]
\[
= (-1)^{n-1} \sum_{i=1}^{n} (-1)^{i-1} e_{n-i} p_i + \sum_{r=2}^{n} (-1)^{n-r} e_{n-r+1} p_{r-1}
\]
\[
= \sum_{i=1}^{n} (-1)^{n-1} \underbrace{(-1)^{i-1} e_{n-i} p_i}_{i=1} + \sum_{i=1}^{n-1} (-1)^{n-i} e_{n-i} p_i
\]
\[
= \sum_{i=1}^{n} (-1)^{n-i} e_{n-i} p_i - \sum_{i=1}^{n} (-1)^{n-i} e_{n-i} p_i
\]
\[
= \sum_{i=1}^{n-1} (-1)^{n-i} e_{n-i} p_i + (-1)^{n-n} e_{n-n} p_n - \sum_{i=1}^{n-1} (-1)^{n-i} e_{n-i} p_i = (-1)^{n-n} e_{n-n} p_n = p_n.
\]
This completes the induction step, and so Exercise 2.89(b) is solved.

Remark: The above solution follows closely the solution of Exercise 9.3 in the first author’s “λ-rings: Definitions and basic properties” (http://sites.google.com/site/darijgrinberg/lambda, version 0.0.17), which is more or less a restatement of this exercise (since the elements of $\Lambda_\mathbb{Z}$ are in a 1-to-1 correspondence with unary functorial operations defined on every $\lambda$-ring).

Whenever $\ell \in \mathbb{N}$ and two partitions $\lambda$ and $\mu$ of length $\ell$ have the property that the transpose of the matrix $(h_{\lambda_i,\mu_j-i+j})_{i,j=1,2,\ldots,\ell}$ is nearly lower-triangular, we can use Lemma 11.9 to obtain a recursive formula for the determinant of this matrix, which is $s_{\lambda/\mu}$ according to (2.22). This does not seem to be of much use, however.
11.62. **Solution to Exercise 2.90.** Solution to Exercise 2.90. For any integers \( a \) and \( b \), we define an element \( s(a, b) \) of \( \Lambda \) by

\[
 s(a, b) = \begin{cases} 
 s_{(a+1,1^b)}, & \text{if } a \geq 0 \text{ and } b \geq 0; \\
 (-1)^b \delta_{a+b,-1}, & \text{if } a < 0 \text{ and } b \geq 0; \\
 0, & \text{if } b < 0 
\end{cases}
\]

(We are introducing this \( s(a, b) \) chiefly for reasons of convenience: it will allow us to unify parts (b) and (c) of the exercise.)

Using the definition of \( s(a, b) \) and straightforward case analysis, we can see that:

\[(11.480) \quad s(a, b) = s_{(a+1,1^b)} \quad \text{for any } a \in \mathbb{N} \text{ and } b \in \mathbb{N}; \)

\[(11.481) \quad s(a, b) = (-1)^b \delta_{a+b,-1} \quad \text{for every negative integer } a \text{ and every } b \in \mathbb{Z}; \)

\[(11.482) \quad s(a, 0) = h_{a+1} \quad \text{for every } a \in \mathbb{Z}; \)

\[(11.483) \quad s(a, b) = 0 \quad \text{for any } a \in \mathbb{Z} \text{ and any negative } b \in \mathbb{Z}. \]

Let us now show that

\[(11.484) \quad e_n h_m = s(m, n - 1) + s(m - 1, n) \quad \text{for every } n \in \mathbb{Z} \text{ and every } m \in \mathbb{Z}. \]

*Proof of (11.484):* Let \( n \in \mathbb{Z} \) and \( m \in \mathbb{Z} \). It is easy to prove (11.484) in the case \( n < 0 \) (since both sides of (11.484) vanish in this case) and in the case \( n = 0 \) (here, it follows from (11.482)). We can therefore WLOG assume that \( n > 0 \). Assume this.

It is easy to prove (11.484) in the case \( m < 0 \) (in which case both sides of (11.484) vanish) and in the case \( m = 0 \) (in which case both sides of (11.484) equal \( e_n \)). Thus, we can WLOG assume that \( m > 0 \). Assume this. Since \( m > 0 \), we have \( h_m = s(m) \), so that

\[(11.485) \quad e_n h_m = e_n s(m) = s(m) e_n = \sum_{\lambda^+, \lambda^+/(m) \text{ is a vertical } n\text{-strip}} s_{\lambda^+} \quad \text{(by (2.37), applied to } \lambda = (m)). \]

It is easy to see that there are exactly two partitions \( \lambda^+ \) for which \( \lambda^+/(m) \) is a vertical \( n\text{-strip} \), namely the partitions \( (m, 1^n) \) and \( (m+1, 1^{n-1}) \). Hence, the sum \( \sum_{\lambda^+, \lambda^+/(m) \text{ is a vertical } n\text{-strip}} s_{\lambda^+} \) has exactly two addends: that for \( \lambda^+ = (m, 1^n) \) and that for \( \lambda^+ = (m+1, 1^{n-1}) \). Thus, this sum simplifies to \( s_{(m,1^n)} + s_{(m+1,1^{n-1})} \). Hence, (11.485) rewrites as

\[(11.486) \quad e_n h_m = s_{(m,1^n)} + s_{(m+1,1^{n-1})}. \]

But (11.480) (applied to \( a = m - 1 \) and \( b = n \)) yields \( s_{(m-1,n)} = s_{(m,1^n)} \). Also, (11.480) (applied to \( a = m \) and \( b = n - 1 \)) yields \( s_{(m,n-1)} = s_{(m+1,1^{n-1})} \). Thus,

\[ e_n h_m = \underbrace{s_{(m,1^n)}}_{=s_{(m-1,n)}} + \underbrace{s_{(m+1,1^{n-1})}}_{=s_{(m,n-1)}} = s_{(m-1,n)} + s_{(m,n-1)} = s_{(m,n-1)} + s_{(m-1,n)}. \]

This proves (11.484).

Let us now prove a statement which encompasses both parts (b) and (c) of Exercise 2.90. Namely, we are going to prove that

\[(11.487) \quad \sum_{i=0}^{b} (-1)^i h_{a+i+1} e_{b-i} = s(a, b) \quad \text{for any } a \in \mathbb{Z} \text{ and } b \in \mathbb{Z}. \]
Proof of (11.487): Let $a \in \mathbb{Z}$ and $b \in \mathbb{Z}$. We WLOG assume that $b \in \mathbb{N}$ (since otherwise, the left hand side of (11.487) is an empty sum, and the right hand side is 0 by definition). For every $i \in \{0, 1, \ldots, b\}$, we have

$$h_{a+i+1}c_{b-i} = e_{b-i}h_{a+i+1} = s \left( \frac{a+i+1, b-i-1}{a+i+1}, \frac{a+i+1, b-i}{a+i+1} \right) + s \left( \frac{(a+i+1)-1, b-i}{a+i+1} \right)$$

(by (11.484), applied to $n = b-i$ and $m = a+i+1$)

(11.488)

$$= s \left( a+(i+1), b-(i+1) \right) + s \left( a+i, b-i \right).$$

Now,

$$\sum_{i=0}^{b} (-1)^i \left[ h_{a+i+1}c_{b-i} \right] = s(a+(i+1), b-(i+1)) + s(a+i, b-i)$$

(by (11.488))

$$= \sum_{i=0}^{b} (-1)^i \left[ s(a+(i+1), b-(i+1)) + s(a+i, b-i) \right]$$

$$= \sum_{i=0}^{b} \left[ (-1)^i s(a+(i+1), b-(i+1)) - (-1)^{i-1} s(a+i, b-i) \right]$$

$$= (-1)^b s(a+b+1, b-b+1) - (-1)^0 s(a+0, b-0)$$

(by the telescope principle)

$$= (-1)^b s(a+1, 0) - (-1) s(a, b) = (-1) s(a, b) = s(a, b).$$

(11.489)

This proves (11.487).

Here is a modified version of (11.487) which will be used in our solution of Exercise 2.90(d) further below: For any $a \in \mathbb{Z}$, $b \in \mathbb{Z}$ and $c \in \mathbb{Z}$ satisfying $c \geq b$, we have

$$\sum_{k=0}^{c} (-1)^k h_{a+k+1}c_{b-k} = s(a, b).$$

(11.489)

462

Now, the first three parts of Exercise 2.90 are as good as solved: The claim of Exercise 2.90(a) has already been proven in (11.486), and the claims of Exercise 2.90(b) and Exercise 2.90(c) follow from (11.487).

Before we come to the solution of Exercise 2.90(d), we simplify our life by introducing another definition. Namely, let us define a $\mathbb{k}$-module endomorphism $\widetilde{\id}$ of $\Lambda$ by $\widetilde{\id} = \id - uc$. Then, it is easy to see that $\widetilde{\id}(s_\lambda) = s_\lambda$ for every nonempty partition $\lambda$, whereas $\widetilde{\id}(c) = 0$ for every $c \in \mathbb{k}$. Now, straightforward

462 Proof of (11.487): Let $a \in \mathbb{Z}$, $b \in \mathbb{Z}$ and $c \in \mathbb{Z}$ be such that $c \geq b$. We WLOG assume that $b \geq 0$ (since otherwise, both sides of (11.489) vanish). Then,

$$\sum_{k=0}^{c} (-1)^k h_{a+k+1}c_{b-k} = \sum_{k=0}^{b} (-1)^k h_{a+k+1}c_{b-k} + \sum_{k=b+1}^{c} (-1)^k h_{a+k+1}c_{b-k} \quad (\text{since } b-k < 0)$$

$$= \sum_{k=0}^{b} (-1)^k h_{a+k+1}c_{b-k} + \sum_{k=b+1}^{c} (-1)^k h_{a+k+1}c_{b-k} \quad (\text{since } k > b)$$

$$= \sum_{k=0}^{b} (-1)^k h_{a+k+1}c_{b-k} = \sum_{i=0}^{b} (-1)^i h_{a+i+1}c_{b-i} = s(a, b) \quad (\text{by (11.487)}).$$

This proves (11.489).
computation shows that

\[(11.490) \quad \text{id} (s(a, b)) = s(a+1, 1^b) \quad \text{for all } a \in \mathbb{N} \text{ and } b \in \mathbb{N};\]

\[(11.491) \quad \text{id} (s(a, b)) = 0 \quad \text{for every } (a, b) \in \mathbb{Z}^2 \text{ satisfying } (a, b) \notin \mathbb{N}^2.\]

But every \(x \in \Lambda\) satisfies

\[(11.492) \quad \Delta x = 1 \otimes x - x \otimes 1 = (\text{id} \otimes \text{id})(\Delta x) - \epsilon(x) \cdot 1 \otimes 1 \quad \text{in } \Lambda \otimes \Lambda.\]

(This is an identity which holds not only in \(\Lambda\) but in every \(k\)-bialgebra, and which is proven by applying the bialgebra axioms.)

(d) Recall that \(\Lambda\) is a \(k\)-bialgebra. Hence, \(\Delta : \Lambda \rightarrow \Lambda \otimes \Lambda\) is a \(k\)-algebra homomorphism.

We can rewrite Proposition 2.24(ii) as follows: Every \(n \in \mathbb{N}\) satisfies

\[(11.493) \quad \Delta e_n = \sum_{i=0}^{n} e_i \otimes e_{n-i}.\]

Thus, every \(n \in \mathbb{N}\) satisfies

\[(11.494) \quad \Delta e_n = \sum_{i \in \mathbb{N}} e_i \otimes e_{n-i}.\]

Similarly, every \(n \in \mathbb{N}\) satisfies

\[(11.495) \quad \Delta h_n = \sum_{i \in \mathbb{N}} h_i \otimes h_{n-i}.\]

Let \(a \in \mathbb{N}\) and \(b \in \mathbb{N}\). Exercise 2.90(b) yields \(\sum_{i=0}^{b} (-1)^i h_{a+i+1} e_{b-i} = s(a+1, 1^b)\), so that

\[s(a+1, 1^b) = \sum_{i=0}^{b} (-1)^i h_{a+i+1} e_{b-i} = \sum_{k=0}^{b} (-1)^k h_{a+k+1} e_{b-k}\]

\[\sum_{i \in \mathbb{N}} e_i \otimes e_{n-i} = \sum_{i=0}^{n} e_i \otimes e_{n-i} + \sum_{i=n+1}^{n} e_i \otimes e_{n-i} = \sum_{i=0}^{n} e_i \otimes e_{n-i} + \sum_{i=n+1}^{n} e_i \otimes e_{n-i} = \sum_{i=0}^{n} e_i \otimes e_{n-i}. \quad \text{(since } n-i<0\text{)}\]

\[\sum_{i \in \mathbb{N}} e_i \otimes e_{n-i} = \sum_{i=0}^{n} e_i \otimes e_{n-i} + \sum_{i=n+1}^{n} e_i \otimes e_{n-i} = \sum_{i=0}^{n} e_i \otimes e_{n-i} + \sum_{i=n+1}^{n} e_i \otimes e_{n-i} = \sum_{i=0}^{n} e_i \otimes e_{n-i}. \quad \text{(since } n-i<0\text{)}\]
(here, we have renamed the summation index $i$ as $k$). Applying the map $\Delta$ to both sides of this equality, we obtain

\[
\Delta s_{(a+1,b)} = \Delta \left( \sum_{k=0}^{b} (-1)^k h_{a+k+1} e_{b-k} \right) = \sum_{k=0}^{b} (-1)^k \Delta \left( h_{a+k+1} \right) \cdot \Delta \left( e_{b-k} \right)
\]

(since $\Delta$ is a $k$-algebra homomorphism)

\[
= \sum_{k=0}^{b} (-1)^k \left( \sum_{i \in N} h_i \otimes h_{a+k+1-i} \right) \left( \sum_{j \in N} e_j \otimes e_{b-k-j} \right)
\]

\[
= \sum_{i \in N} \sum_{j \in N} h_i e_j \otimes \left( \sum_{k=0}^{b} (-1)^k h_{a-i+k+1} e_{b-j-k} \right)
\]

\[
= \sum_{(i,j) \in \mathbb{N}^2} (s(i,j-1) + s(i-1,j)) \otimes s(a-i,b-j)
\]

(11.496)

Now, applying (11.492) to $x = s_{(a+1,b)}$, we obtain

\[
\Delta s_{(a+1,b)} - s_{(a+1,b)} \otimes 1 = (\overline{id} \otimes \overline{id}) \left( \sum_{(i,j) \in \mathbb{N}^2} (s(i,j-1)+s(i-1,j)) \otimes s(a-i,b-j) \right) - 0 \otimes 1
\]

(11.497)

We can now observe that

\[
\sum_{(i,j) \in \mathbb{N}^2} \overline{id} (s(i+1,1) \otimes s(a-i,b-j)) = \sum_{(c,d,e,f) \in \mathbb{N}^4} s_{(c+1,1)^4} \otimes s_{(c+1,1)^4}.
\]
Similarly,

\[(11.500)\quad \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} \text{id} (s(i, j - 1)) \otimes \text{id} (s(a - i, b - j)) = \sum_{(c,d,e,f) \in \mathbb{N}^4; \atop c+e=a-1; \atop d+f=b} s(c+1,1^d) \otimes s(e+1,1^f).\]

Now, \(11.497\) becomes

\[
\Delta s(a+1,1^b) - 1 \otimes s(a+1,1^b) - s(a+1,1^b) \otimes 1
\]

\[
= \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} \text{id} (s(i, j - 1)) \otimes \text{id} (s(a - i, b - j)) + \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} \text{id} (s(i, j - 1)) \otimes \text{id} (s(a - i, b - j))
\]

\[
= \sum_{(c,d,e,f) \in \mathbb{N}^4; \atop c+e=a-1; \atop d+f=b} s(c+1,1^d) \otimes s(e+1,1^f) + \sum_{(c,d,e,f) \in \mathbb{N}^4; \atop c+e=a-1; \atop d+f=b} s(c+1,1^d) \otimes s(e+1,1^f).
\]

Adding \(1 \otimes s(a+1,1^b) + s(a+1,1^b) \otimes 1\) to both sides of this equality, we obtain the claim of Exercise 2.90(d).

Remark: In our above solution, we have first proved \(11.484\) and then used it to derive \(11.487\). There is also an alternative way to prove these two identities, which proceeds the other way round. The main idea is to obtain \(11.487\) by applying Exercise 2.77(b) to \(\lambda = (1^b)\) and \(m = a + 1\). (This works in the case of \(a \in \mathbb{N}\) and \(b \in \mathbb{N}\) only. The remaining cases, however, are easy to either check directly or reduce to (2.17).) Once this is done, \(11.484\) can be verified by rewriting both \(s(m, n-1)\) and \(s(m-1, n)\) using \(11.487\).

---

Proof of \(11.498\): Let \(\mathbb{N} = \{-1, 0, 1, 2, \ldots\} = \{-1\} \cup \mathbb{N}\). We have

\[(11.499)\quad \sum_{(c,d,e,f) \in \mathbb{N} \times \mathbb{N} \times \mathbb{Z} \times \mathbb{Z}; \atop c+e=a-1; \atop d+f=b} \text{id} (s(c,d)) \otimes \text{id} (s(e,f)) = \sum_{(c,d,e,f) \in \mathbb{N}^4; \atop c+e=a-1; \atop d+f=b} \text{id} (s(c,d)) \otimes \text{id} (s(e,f)).\]

(In fact, the sum on the left hand side of \(11.499\) differs from that on the right hand side of \(11.499\) only by the presence of addends of the form \(\text{id} (s(c,d)) \otimes \text{id} (s(e,f))\) for certain quadruples \((c,d,e,f) \in \mathbb{N} \times \mathbb{N} \times \mathbb{Z} \times \mathbb{Z}\) which don’t belong to \(\mathbb{N}^4\). But all such addends are 0 (as can be easily seen using \(11.491\)), and so the two sums have the same value, and \(11.499\) is proven.)

The map

\[\mathbb{N}^2 \to \{(c,d,e,f) \in \mathbb{N} \times \mathbb{N} \times \mathbb{Z} \times \mathbb{Z}; \text{ } c+e=a-1; \text{ } d+f=b\}, \quad (i,j) \mapsto (i-1,j,a-i,b-j),\]

is a bijection. Hence, we can substitute \((i-1,j,a-i,b-j)\) for \((c,d,e,f)\) in the sum \(\sum_{(c,d,e,f) \in \mathbb{N} \times \mathbb{N} \times \mathbb{Z} \times \mathbb{Z}; \atop c+e=a-1; \atop d+f=b} \text{id} (s(c,d)) \otimes \text{id} (s(e,f)),\) and thus obtain

\[
\sum_{(c,d,e,f) \in \mathbb{N} \times \mathbb{N} \times \mathbb{Z} \times \mathbb{Z}; \atop c+e=a-1; \atop d+f=b} \text{id} (s(c,d)) \otimes \text{id} (s(e,f)) = \sum_{(i,j) \in \mathbb{N}^2} \text{id} (s(i-1,j)) \otimes \text{id} (s(a-i,b-j)).
\]

Comparing this with \(11.499\), we obtain

\[
\sum_{(i,j) \in \mathbb{N}^2} \text{id} (s(i-1,j)) \otimes \text{id} (s(a-i,b-j)) = \sum_{(c,d,e,f) \in \mathbb{N}^4; \atop c+e=a-1; \atop d+f=b} \text{id} (s(c,d)) \otimes \text{id} (s(e,f)) = \sum_{(c,d,e,f) \in \mathbb{N}^4; \atop c+e=a-1; \atop d+f=b} s(c+1,1^d) \otimes s(e+1,1^f).
\]

This proves \(11.498\).
11.63. **Solution to Exercise 2.91.** Consider the partition \((m^k) = \left(\frac{m, \ldots, m}{k \text{ times}}\right)\).

(a) The fact that \(\lambda^\nu\) and \(\mu^\nu\) are partitions is easy to check. It remains to show that \(s_{\lambda/\mu} = s_{\mu^\nu/\lambda^\nu}\).

It is easy to show that if \(\mu \subseteq \lambda\) does not hold, then \(\lambda^\nu \subseteq \mu^\nu\) does not hold either (because if some positive integer \(i\) fails to satisfy \(\mu_i \leq \lambda_i\), then this \(i\) belongs to \(\{1, 2, \ldots, k\}\) and fails to satisfy \(m - \lambda_i \leq m - \mu_i\) as well). Hence, if \(\mu \subseteq \lambda\) does not hold, then both \(s_{\lambda/\mu}\) and \(s_{\mu^\nu/\lambda^\nu}\) are 0, and therefore the equality \(s_{\lambda/\mu} = s_{\mu^\nu/\lambda^\nu}\) is obvious. Hence, for the rest of the proof of \(s_{\lambda/\mu} = s_{\mu^\nu/\lambda^\nu}\), we WLOG assume that \(\mu \subseteq \lambda\) does hold. Then, it is easy to see that \(\lambda^\nu \subseteq \mu^\nu\) holds, too.

Now, let \(Z\) denote the 180° rotation around the center of the Ferrers diagram of \((m^k)\). Let \(Y(\rho)\) denote the Ferrers diagram of \(\rho\) whenever \(\rho\) is a partition or a skew partition. It is straightforward to see that \(Z(Y((m^k)) \setminus Y(\mu)) = Y(\mu^\nu)\) and \(Z(Y((m^k)) \setminus Y(\lambda)) = Y(\lambda^\nu)\). Now,

\[
Y(\mu^\nu/\lambda^\nu) = \frac{Y(\mu)}{Y(\lambda)} = \frac{Y((m^k)) \setminus Y(\mu)}{Y((m^k)) \setminus Y(\lambda)} = Z(Y((m^k)) \setminus Y(\lambda)) = Z(Y((m^k)) \setminus Y(\mu)).
\]

Hence, the skew Ferrers diagram \(\mu^\nu/\lambda^\nu\) can be obtained from the skew Ferrers diagram \(\lambda/\mu\) by a 180° rotation (namely, by the 180° rotation \(Z\)). Thus, Exercise 2.22(b) (applied to \(\lambda' = \mu^\nu\) and \(\mu' = \lambda^\nu\)) yields \(s_{\lambda/\mu} = s_{\mu^\nu/\lambda^\nu}\). This completes the solution of Exercise 2.91(a).

(b) According to Remark 2.37, we have \(s_{\lambda/\mu} = \sum_{\nu} c_{\mu,\nu}^\lambda s_{\nu}\), where the sum ranges over all partitions \(\nu\). In other words, we have

\[(11.501)\]

Exercise 2.91(b) yields that \(\lambda^\nu\) and \(\mu^\nu\) are partitions, and that \(s_{\lambda/\mu} = s_{\mu^\nu/\lambda^\nu}\). Applying (11.501) to \(\mu^\nu\) and \(\lambda^\nu\) instead of \(\lambda\) and \(\mu\), we obtain

\[s_{\mu^\nu/\lambda^\nu} = \sum_{\nu} c_{\lambda^\nu,\nu}^\mu s_{\nu}\].

Now, (11.501) yields

\[
\sum_{\nu \in \text{Par}} c_{\mu,\nu}^\lambda s_{\nu} = s_{\lambda/\mu} = s_{\mu^\nu/\lambda^\nu} = \sum_{\nu \in \text{Par}} c_{\lambda^\nu,\nu}^\mu s_{\nu}.
\]

Comparing coefficients before \(s_{\nu}\) in this equality, we conclude that \(c_{\mu,\nu}^\lambda = c_{\lambda^\nu,\nu}^\mu\) for every \(\nu \in \text{Par}\) (since \((s_{\nu})_{\nu \in \text{Par}}\) is a basis of the \(k\)-module \(\Lambda\)). This solves Exercise 2.91(b).

(c) Exercise 2.91(a) (applied to \(\nu\) instead of \(\mu\)) yields that \(\lambda^\nu\) and \(\nu^\nu\) are partitions and that \(s_{\lambda/\nu} = s_{\nu^\nu/\lambda^\nu}\).

From (2.27), we obtain

\[c_{\mu,\nu}^\lambda = c_{\nu^\nu,\nu}^\lambda = c_{\lambda^\nu,\nu}^\mu = c_{\mu^\nu,\mu}^\nu\]

(by Exercise 2.91(b) (applied to \(\nu\) and \(\mu\) instead of \(\mu\) and \(\nu\)))

(by (2.27) (applied to \(\nu^\nu\), \(\lambda^\nu\) and \(\mu\) instead of \(\lambda\), \(\mu\) and \(\nu\)).

On the other hand, Exercise 2.91(b) yields \(c_{\mu,\nu}^\lambda = c_{\nu^\nu,\nu}^\lambda = c_{\lambda^\nu,\nu}^\mu\) (by (2.27) (applied to \(\mu^\nu\), \(\lambda^\nu\) and \(\nu\) instead of \(\lambda\), \(\mu\) and \(\nu\)). Thus, \(c_{\mu,\nu}^\lambda = c_{\nu^\nu,\nu}^\lambda = c_{\lambda^\nu,\nu}^\mu = c_{\mu^\nu,\mu}^\nu = c_{\lambda^\nu,\nu}^\mu = c_{\mu^\nu,\mu}^\nu\). This solves Exercise 2.91(c).

(d) **First solution to Exercise 2.91(d):** Notice that \(\lambda^\nu = (m - \lambda_k, m - \lambda_{k-1}, \ldots, m - \lambda_1)\). Thus, \(\ell(\lambda^\nu) \leq k\).
Let \( n = k \). We shall use the notations of Section 2.6; in particular, we set \( x = (x_1, x_2, \ldots, x_n) \) and \( \rho = (n - 1, n - 2, \ldots, 2, 1, 0) \). Clearly, \( x = (x_1, x_2, \ldots, x_n) = (x_1, x_2, \ldots, x_k) \) (since \( n = k \)). Notice that \( \ell(\lambda) \leq k = n \). Hence, we can regard \( \lambda \) as an element of \( \mathbb{N}^n \); therefore, \( \lambda + \rho \) is a well-defined element of \( \mathbb{N}^n \), and the alternant \( a_{\lambda + \rho} \) is well-defined. Corollary 2.54 yields that \( s_\lambda(x) = \frac{a_{\lambda + \rho}}{a_\rho} \), so that

\[
(11.502) \quad a_\rho \cdot s_\lambda(x) = a_{\lambda + \rho}.
\]

Applying this equality to \( \lambda^\nu \) instead of \( \lambda \), we obtain

\[
(11.503) \quad a_\rho \cdot s_{\lambda^\nu}(x) = a_{\lambda^\nu + \rho}
\]

(since \( \lambda^\nu \) is also a partition satisfying \( \ell(\lambda^\nu) \leq k = n \)).

Substituting the variables \( x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1} \) for \( x_1, x_2, \ldots, x_n \) in the equality (11.502), we obtain

\[
(11.504) \quad a_\rho(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}) \cdot s_\lambda(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}) = a_{\lambda + \rho}(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}).
\]

Let \( w_\circ: \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\} \) be the map which sends every \( i \in \{1, 2, \ldots, n\} \) to \( n + 1 - i \). Then, \( w_\circ \) is a permutation of \( \{1, 2, \ldots, n\} \), thus an element of \( \mathfrak{S}_n \). For every \( i \in \{1, 2, \ldots, n\} \), we have

\[
(11.505) \quad (w \circ w_\circ)(i) = w(n + 1 - i).
\]

The map \( \mathfrak{S}_n \to \mathfrak{S}_n, \ w \mapsto w \circ w_\circ \) is a bijection (since \( \mathfrak{S}_n \) is a group, and since \( w_\circ \in \mathfrak{S}_n \)).

But every \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}^n \) satisfies

\[
(11.506) \quad a_\alpha = \sum_{w \in \mathfrak{S}_n} \text{sgn}(w) \frac{w(x^n)}{=\prod_{i=1}^{n} x_{w(i)}^{n_i}} \quad \text{(by the definition of} \ a_\alpha) \quad \text{and} \quad w(x^n))
\]

But \( \lambda = (\lambda_1, \lambda_2, \lambda_3, \ldots) \), thus \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \) (since \( \ell(\lambda) \leq k = n \)). Hence,

\[
= (\lambda_1, \lambda_2, \ldots, \lambda_n) + (n - 1, n - 2, \ldots, n - n) = (\lambda_1 + n - 1, \lambda_2 + n - 2, \ldots, \lambda_n + n - n).
\]

Thus, (11.506) (applied to \( \lambda + \rho \) and \( (\lambda_1 + n - 1, \lambda_2 + n - 2, \ldots, \lambda_n + n - n) \) instead of \( \alpha \) and \((\alpha_1, \alpha_2, \ldots, \alpha_n)) \) yields

\[
(11.507) \quad a_{\lambda + \rho} = \sum_{w \in \mathfrak{S}_n} \text{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{\lambda_1+n-i}.
\]
Substituting $x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}$ for $x_1, x_2, \ldots, x_n$ on both sides of this equality, we obtain

$$a_{\lambda+\rho} \left( x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1} \right)$$

$$= \sum_{w \in \mathfrak{S}_n} \text{sgn} \left( w \right) \prod_{i=1}^{n} \frac{\lambda_i + n - i}{x_{w(i)}} = \sum_{w \in \mathfrak{S}_n} \text{sgn} \left( w \right) \prod_{i=1}^{n} x_{w(i)}^{-\left( \lambda_i + n - i \right)}$$

$$= \sum_{w \in \mathfrak{S}_n} \frac{1}{\text{sgn} \left( w \right) \text{sgn}(w_0)} \prod_{i=1}^{n} x_{w(n+1-i)}^{-\left( \lambda_{n+1-i} + n - (n+1-i) \right)}$$

$$= \sum_{w \in \mathfrak{S}_n} \frac{1}{\text{sgn} \left( w \circ w_0 \right) \text{sgn}(w_0)} \prod_{i=1}^{n} x_{w(w_0(i))}^{-\left( \lambda_{n+1-i} + n - (n+1-i) \right)}$$

$$= \sum_{w \in \mathfrak{S}_n} \frac{1}{\text{sgn} \left( w \circ w_0 \right) \text{sgn}(w_0)} \prod_{i=1}^{n} x_{w_0(i)}^{-\left( \lambda_{n+1-i} + n - (n+1-i) \right)}$$

$$= \frac{1}{\text{sgn} \left( w_0 \right)} \sum_{w \in \mathfrak{S}_n} \text{sgn} \left( w \right) \prod_{i=1}^{n} x_{w(i)}^{-\left( \lambda_{n+1-i} + n - (n+1-i) \right)}.$$

Multiplying both sides of this equality by $\text{sgn} \left( w_0 \right)$, we obtain

$$(11.508) \quad \text{sgn} \left( w_0 \right) \cdot a_{\lambda+\rho} \left( x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1} \right) = \sum_{w \in \mathfrak{S}_n} \text{sgn} \left( w \right) \prod_{i=1}^{n} x_{w(i)}^{-\left( \lambda_{n+1-i} + n - (n+1-i) \right)}.$$

But $\lambda^\vee = (m - \lambda_k, m - \lambda_{k-1}, \ldots, m - \lambda_1) = (m - \lambda_n, m - \lambda_{n-1}, \ldots, m - \lambda_1)$ (since $k = n$), so that

$$\lambda^\vee = (m - \lambda_n, m - \lambda_{n-1}, \ldots, m - \lambda_1)$$

$$= (m - \lambda_{n+1-i}, m - \lambda_{n+2-i}, \ldots, m - \lambda_{n+1-i} - n)$$

$$= (m - \lambda_{n+1-i} + n - 2, \ldots, m - \lambda_{n+1-i} - n + (n - 1), n, \ldots, n)$$

$$= (m - \lambda_{n+1-i} + n - 2, \ldots, m - \lambda_{n+1-i} - n + (n - 1), n, \ldots, n - n).$$

Hence, $(11.506)$ (applied to $\lambda^\vee + \rho$ and
to $\lambda$) instead of $\alpha$ and
$(\alpha_1, \alpha_2, \ldots, \alpha_n)$ yields

$$a_{\lambda + \rho} = \sum_{w \in S_n} \text{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{\left( m - \lambda_{n+1-i} + (n-i) \right)}$$

$$= \sum_{w \in S_n} \text{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{-\left( \lambda_{n+1-i} + i - 1 \right)} x_{w(i)}^{n+m-1}$$

$$= \sum_{w \in S_n} \text{sgn}(w) \left( \prod_{i=1}^{n} x_{w(i)}^{-\lambda_{n+1-i} + i - 1} \right) \left( \prod_{i=1}^{n} x_{w(i)}^{n+m-1} \right)$$

$$= \left( \prod_{i=1}^{n} x_{i}^{n+m-1} \right) \sum_{w \in S_n} \text{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{\left( \lambda_{n+1-i} + i - 1 \right)}$$

$$= \left( \prod_{i=1}^{n} x_{i}^{n+m-1} \right) \text{sgn}(w) \cdot a_{\lambda + \rho} \left( x_{1}^{-1}, x_{2}^{-1}, \ldots, x_{n}^{-1} \right) \cdot \left( \prod_{i=1}^{n} x_{w(i)}^{\left( \lambda_{n+1-i} + i - 1 \right)} \right)$$

\[(11.509)\]

Finally, $\rho = (n-1, n-2, \ldots, 2, 1, 0) = (n-1, n-2, \ldots, n-n)$. Thus, $(11.506)$ (applied to $\rho$ and $(n-1, n-2, \ldots, n-n)$ instead of $\alpha$ and $(\alpha_1, \alpha_2, \ldots, \alpha_n)$) yields

\[(11.510)\]

$$a_{\rho} = \sum_{w \in S_n} \text{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{n-i}.$$
Substituting \( x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1} \) for \( x_1, x_2, \ldots, x_n \) on both sides of this equality, we obtain

\[
\begin{align*}
&\quad a_\rho \left( x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1} \right) \\
= & \sum_{w \in \mathfrak{S}_n} \text{sgn}(w) \prod_{i=1}^{n} \left( x_{w(i)}^{-1} \right)^{n-i} = \sum_{w \in \mathfrak{S}_n} \text{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{-n} \\
= & \sum_{w \in \mathfrak{S}_n} \frac{1}{\text{sgn}(w)} \cdot \sum_{w \in \mathfrak{S}_n} \text{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{-(n+1-i)} \tag{since \( w(w \circ w_i) = (w \circ w_i)(i) \)} \\
= & \frac{1}{\text{sgn}(w)} \sum_{w \in \mathfrak{S}_n} \text{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{n-i+1} \tag{since \( w \circ w_i \) is a bijection} \\
= & \frac{1}{\text{sgn}(w)} \sum_{w \in \mathfrak{S}_n} \text{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{n+1} \tag{since \( w \) is a permutation of \( \{1, 2, \ldots, n\} \)} \\
= & \frac{1}{\text{sgn}(w)} \sum_{w \in \mathfrak{S}_n} \text{sgn}(w) \prod_{i=1}^{n} x_{w(i)}^{1-n} \tag{by (11.510)} \\
= & \frac{1}{\text{sgn}(w)} \left( \prod_{i=1}^{n} x_i \right)^{1-n} a_\rho. \\
\end{align*}
\]

Multiplying both sides of this equality by \( \text{sgn}(w) \), we obtain

\[
(11.511) \quad \text{sgn}(w) \cdot a_\rho \left( x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1} \right) = \left( \prod_{i=1}^{n} x_i \right)^{1-n} a_\rho.
\]
Now, \((11.503)\) becomes

\[
a_\rho \cdot s_{\lambda^\vee}(x) = a_{\lambda^\vee + \rho} = \left( \prod_{i=1}^{n} x_i \right)^{n+m-1} \cdot \sgn(w_\rho) \cdot a_{\lambda^\vee + \rho}(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}) = a_\rho(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}) s_\lambda(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1})
\]

(by \((11.509)\))

\[
= \left( \prod_{i=1}^{n} x_i \right)^{n+m-1} \cdot \sgn(w_\rho) \cdot a_\rho(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}) \cdot s_\lambda(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}) = (\prod_{i=1}^{n} x_i)^{1-n} a_\rho \cdot s_\lambda(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1})
\]

(by \((11.511)\))

\[
= \left( \prod_{i=1}^{n} x_i \right)^{n+m-1} \cdot \left( \prod_{i=1}^{n} x_i \right)^{1-n} a_\rho \cdot s_\lambda(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}) = \left( \prod_{i=1}^{n} x_i \right)^{m} a_\rho \cdot s_\lambda(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1})
\]

\[
= a_\rho \cdot (x_1 x_2 \cdots x_n)^m \cdot s_\lambda(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}).
\]

\((11.512)\)

But\(^{465}\)

\(a_\rho\) is a non-zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}]\).

\(^{466}\)Hence, we can cancel the factor \(a_\rho\) from the equality \((11.512)\). As a result, we obtain \(s_{\lambda^\vee}(x) = (x_1 x_2 \cdots x_n)^m \cdot s_\lambda(x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1})\). Compared with \(s_{\lambda^\vee}(x) = s_{\lambda^\vee}(x_1, x_2, \ldots, x_k)\) (since \(x = (x_1, x_2, \ldots, x_k)\)),

In the following, a non-zero-divisor in a commutative ring \(B\) means an element \(b \in B\) such that every element \(c \in B\) satisfying \(bc = 0\) must satisfy \(c = 0\).

\(^{466}\)Proof. We know (from a footnote in Corollary 2.54) that \(a_\rho\) is not a zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n]\). In other words, \(a_\rho\) is a non-zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n]\).

Now, let \(b\) be any element of the ring \(k[x_1, x_2, \ldots, x_n]\) such that \(b\) is a non-zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n]\). We shall show that \(b\) is a non-zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}]\) as well.

Indeed, let \(c \in k[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}]\) be such that \(bc = 0\). It is known that every element of \(k[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}]\) has the form \(p = \frac{m}{x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n}}\) for some \(p \in k[x_1, x_2, \ldots, x_n]\) and some \((q_1, q_2, \ldots, q_n) \in \mathbb{N}^n\).

So let us write \(c\) in this form, and consider the corresponding \(p\) and \((g_1, g_2, \ldots, g_n)\). We thus have \(c = \frac{p}{x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n}} \cdot \frac{1}{x_1^{g_1} x_2^{g_2} \cdots x_n^{g_n}} = 0\). Thus, \(p = 0\) (since \(b\) is a non-zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n]\) and since \(x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n} = 0\)).

\(bc = 0\). This rewrites as \(b \cdot \frac{p}{x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n}} = 0\) (since \(c = \frac{p}{x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n}}\)). Multiplying both sides of this equality by \(x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n}\), we obtain \(bp = 0\). This equality holds in the ring \(k[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}]\), and thus also in the ring \(k[x_1, x_2, \ldots, x_n]\) (since \(k[x_1, x_2, \ldots, x_n]\) is a subring of \(k[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}]\)). Thus, \(p = 0\) (since \(b\) is a non-zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n]\)). Hence,

\[
c = \frac{p}{x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n}} \cdot \frac{1}{x_1^{g_1} x_2^{g_2} \cdots x_n^{g_n}} = 0 \quad \text{for} \quad p = 0\]

Now, let us forget that we fixed \(c\). We thus have proven that every element \(c \in k[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}]\) satisfying \(bc = 0\) must satisfy \(c = 0\). In other words, \(b\) is a non-zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}]\).

Now, let us forget that we fixed \(b\). We thus have proven that if \(b\) is any element of the ring \(k[x_1, x_2, \ldots, x_n]\) such that \(b\) is a non-zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n]\), then \(b\) is a non-zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}]\) as well. Applying this to \(b = a_\rho\), we conclude that \(a_\rho\) is a non-zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n, x_1^{-1}, x_2^{-1}, \ldots, x_n^{-1}]\) (since we know that \(a_\rho\) is a non-zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n]\)). This proves \((11.513)\).
this yields
\[ s_{\lambda^\vee} (x_1, x_2, \ldots, x_k) = (x_1 x_2 \cdots x_n)^m \cdot s_{\lambda} (x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}) = (x_1 x_2 \cdots x_k)^m \cdot s_{\lambda} (x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}) \]
(since \( n = k \)). This solves Exercise 2.91(d).

Second solution to Exercise 2.91(d) (sketched): Let us give a more combinatorial solution now. In the following, if \( \alpha \) is a partition, then an \((\alpha, k)\)-CST will mean a column-strict tableau \( T \) of shape \( \alpha \) such that all entries of \( T \) belong to \( \{1, 2, \ldots, k\} \).

We have
\[ s_{\lambda^\vee} (x_1, x_2, \ldots, x_k) = s_{\lambda^\vee / \emptyset} (x_1, x_2, \ldots, x_k) = \sum_{\substack{T \text{ is a column-strict} \\
\text{tableau of shape } \lambda^\vee / \emptyset; \\
\text{all entries of } T \text{ belong} \\
to \{1, 2, \ldots, k\}}} x^{\cont(T)} \]
(by Exercise 2.26(a), applied to \( k, \lambda^\vee \) and \( \emptyset \) instead of \( n, \lambda \) and \( \mu \)). This rewrites as
\[ (11.514) \quad s_{\lambda^\vee} (x_1, x_2, \ldots, x_k) = \sum_{\substack{T \text{ is a } (\lambda^\vee, k)\text{-CST}}} x^{\cont(T)} \]
(because the column-strict tableaux \( T \) of shape \( \lambda^\vee / \emptyset \) such that all entries of \( T \) belong to \( \{1, 2, \ldots, k\} \) are precisely the \((\lambda^\vee, k)\)-CSTs).

On the other hand,
\[ s_{\lambda} (x_1, x_2, \ldots, x_k) = s_{\lambda / \emptyset} (x_1, x_2, \ldots, x_k) = \sum_{\substack{T \text{ is a column-strict} \\
\text{tableau of shape } \lambda / \emptyset; \\
\text{all entries of } T \text{ belong} \\
to \{1, 2, \ldots, k\}}} x^{\cont(T)} \]
(by Exercise 2.26(a), applied to \( k \) and \( \emptyset \) instead of \( n \) and \( \mu \)). Substituting \( x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1} \) for \( x_1, x_2, \ldots, x_k \) on both sides of this equality, we obtain
\[ (11.515) \quad s_{\lambda} (x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}) = \sum_{\substack{T \text{ is a column-strict} \\
\text{tableau of shape } \lambda / \emptyset; \\
\text{all entries of } T \text{ belong} \\
to \{1, 2, \ldots, k\}}} (x^{-1})^{\cont(T)}, \]
where \((x^{-1})^{\cont(T)}\) is defined as \( \prod_{i \geq 1} (x_i^{-1})^{T^{-1}(i)} \) for any column-strict tableau \( T \). The equality (11.515) rewrites as
\[ s_{\lambda} (x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}) = \sum_{\substack{T \text{ is a } (\lambda, k)\text{-CST}}} (x^{-1})^{\cont(T)} \]
(since the column-strict tableaux \( T \) of shape \( \lambda / \emptyset \) such that all entries of \( T \) belong to \( \{1, 2, \ldots, k\} \) are precisely the \((\lambda, k)\)-CSTs). Hence,
\[ (x_1 x_2 \cdots x_k)^m \cdot s_{\lambda} (x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}) = (x_1 x_2 \cdots x_k)^m \cdot \sum_{\substack{T \text{ is a } (\lambda, k)\text{-CST}}} (x^{-1})^{\cont(T)} \]
\[ = \sum_{\substack{T \text{ is a } (\lambda, k)\text{-CST}}} (x_1 x_2 \cdots x_k)^m \cdot (x^{-1})^{\cont(T)}. \]
(11.516)

We need to prove that
\[ s_{\lambda^\vee} (x_1, x_2, \ldots, x_k) = (x_1 x_2 \cdots x_k)^m \cdot s_{\lambda} (x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}). \]
Due to (11.514) and (11.516), this rewrites as
\[ (11.517) \quad \sum_{\substack{T \text{ is a } (\lambda^\vee, k)\text{-CST}}} x^{\cont(T)} = \sum_{\substack{T \text{ is a } (\lambda, k)\text{-CST}}} (x_1 x_2 \cdots x_k)^m \cdot (x^{-1})^{\cont(T)}. \]
So it remains to prove (11.517).

In order to prove (11.517), it is clearly sufficient to construct a bijection
\[ \Omega : (\text{the set of all } (\lambda, k)\text{-CSTs}) \to (\text{the set of all } (\lambda^\vee, k)\text{-CSTs}) \]
with the property that every $((\lambda,k),\text{CST})$ $T$ satisfies

(11.518) \[ x^{\text{cont}(\Omega(T))} = (x_1 x_2 \cdots x_k)^m \cdot (x^{-1})^{\text{cont}(T)}. \]

Let us construct such a $\Omega$ now. We begin by doing some elementary combinatorics.

We define a relation $\leq_\#$ on the set of all subsets of $\{1, 2, \ldots, k\}$ as follows:

**Definition 11.10.** Let $k$ be a nonnegative integer. Let $[k] = \{1, 2, \ldots, k\}$. Let $I$ and $J$ be two subsets of $[k]$. We say that $I \leq_\# J$ if the following two properties hold:

- We have $|I| \geq |J|$.
- For every $r \in \{1, 2, \ldots, |J|\}$, the $r$-th smallest element of $I$ is $\leq$ to the $r$-th smallest element of $J$.

We notice that this relation $\leq_\#$ is the less-or-equal relation of a partial order (as follows easily from the definition); but we will not have any use for this fact. Instead, we need a symmetry property:

**Proposition 11.11.** Let $k$ be a nonnegative integer. Let $[k] = \{1, 2, \ldots, k\}$. Let $I$ and $J$ be two subsets of $[k]$.

(a) We have $I \leq_\# J$ if and only if $[k] \setminus J \leq_\# [k] \setminus I$.

(b) For every $\ell \in [k]$ and $S \subseteq [k]$, let $\alpha_S(\ell)$ denote the number $|\{s \in S \mid s \leq \ell\}|$. Then,

(11.519) \[ I \leq_\# J \text{ holds if and only if every } \ell \in [k] \text{ satisfies } \alpha_I(\ell) \geq \alpha_J(\ell). \]

**Proof of Proposition 11.11.** (b) Proof of (11.519): $\implies$: Assume that $I \leq_\# J$. In other words, the following two properties hold:

Property $\alpha$: We have $|I| \geq |J|$.

Property $\beta$: For every $r \in \{1, 2, \ldots, |J|\}$, the $r$-th smallest element of $I$ is $\leq$ to the $r$-th smallest element of $J$.

Now, let $\ell \in [k]$. Then, we need to show that $\alpha_I(\ell) \geq \alpha_J(\ell)$. Since this is obvious if $\alpha_J(\ell) = 0$ (because $\alpha_I(\ell) \geq 0$), we can WLOG assume that $\alpha_J(\ell) \neq 0$. Assume this. Thus, $\alpha_J(\ell) \geq 1$. Also, \[ \alpha_J(\ell) = \left| \{s \in J \mid s \leq \ell\} \right| \leq |J| \leq |I| \text{ (since } |I| \geq |J|). \]

Hence, both the $\alpha_I(\ell)$-th smallest element of $J$ and the $\alpha_J(\ell)$-th smallest element of $I$ are well-defined.

Since $\alpha_J(\ell) = |\{s \in J \mid s \leq \ell\}|$, we know that the elements of $J$ which are $\leq \ell$ are precisely the $\alpha_J(\ell)$ smallest elements of $J$. Thus,

\[ (\text{the } \alpha_J(\ell) \text{-th smallest element of } J) = (\text{the largest element of } J \text{ which is } \leq \ell). \]

But by Property $\beta$ (applied to $r = \alpha_J(\ell)$), we have

\[ (\text{the } \alpha_J(\ell) \text{-th smallest element of } I) \leq (\text{the } \alpha_J(\ell) \text{-th smallest element of } J) = (\text{the largest element of } J \text{ which is } \leq \ell). \]

Hence, there are at least $\alpha_J(\ell)$ elements of $I$ which are $\leq \ell$ (namely, the $\alpha_J(\ell)$ smallest ones). In other words, $|\{s \in I \mid s \leq \ell\}| \geq \alpha_J(\ell)$. Now, $\alpha_I(\ell) = |\{s \in I \mid s \leq \ell\}| \geq \alpha_J(\ell)$. We thus have proven the $\implies$ direction of (11.519).

$\Longleftarrow$: Assume that every $\ell \in [k]$ satisfies $\alpha_I(\ell) \geq \alpha_J(\ell)$. We need to prove that $I \leq_\# J$. In other words, we need to prove that the following two properties hold:

Property $\alpha$: We have $|I| \geq |J|$.

Property $\beta$: For every $\ell \in \{1, 2, \ldots, |J|\}$, the $r$-th smallest element of $I$ is $\leq$ to the $r$-th smallest element of $J$.

First of all, $\{s \in I \mid s \leq k\} = I$ (since every $s \in I$ satisfies $s \leq k$), and the definition of $\alpha_I(k)$ yields

\[ \alpha_I(k) = |\{s \in I \mid s \leq k\}| = |I|. \]

Similarly, $\alpha_J(k) = |J|$. Applying $\alpha_I(\ell) \geq \alpha_J(\ell)$ to $\ell = k$, we obtain $\alpha_I(k) \geq \alpha_J(k)$, so that $|I| = \alpha_I(k) \geq \alpha_J(k) = |J|$, and thus Property $\alpha$ is proven.

Now, let $r \in \{1, 2, \ldots, |J|\}$. The $r$-th smallest element of $I$ and the $r$-th smallest element of $J$ are then well-defined (because of $r \leq |J| \leq |I|$). Let $\ell$ be the $r$-th smallest element of $J$. Then, $\{s \in J \mid s \leq \ell\}$ is the set
consisting of the \( r \) smallest elements of \( J \), so that \(|\{s \in J \mid s \leq \ell\}| = r \). Now, \( \alpha_J(\ell) = |\{s \in J \mid s \leq \ell\}| = r \).

But \( \alpha_I(\ell) = |\{s \in I \mid s \leq \ell\}| \), so that

\[
|\{s \in I \mid s \leq \ell\}| = \alpha_I(\ell) \geq \alpha_J(\ell) = r.
\]

In other words, there exist at least \( r \) elements of \( I \) which are \( \leq \ell \). Hence, the \( r \)-th smallest element of \( I \) must be \( \leq \ell \). Since \( \ell \) is the \( r \)-th smallest element of \( J \), this rewrites as follows: The \( r \)-th smallest element of \( I \) is \( \leq \) to the \( r \)-th smallest element of \( J \). Thus, Property \( \beta \) holds. Now we know that both Properties \( \alpha \) and \( \beta \) hold. Hence, \( I \leq \# J \) holds (which, as we know, is equivalent to the conjunction of said properties).

This proves the \( \iff \) direction of (11.519). Thus, (11.519) is proven. In other words, Proposition 11.11(b) is proven.

(a) For every \( \ell \in [k] \) and \( S \subset [k] \), let \( \alpha_S(\ell) \) denote the number \(|\{s \in S \mid s \leq \ell\}| \). Thus, every \( \ell \in [k] \) satisfies

\[
\alpha_I(\ell) + \alpha_{[k] \setminus I}(\ell) = |\{s \in I \mid s \leq \ell\}| + |\{s \in [k] \setminus I \mid s \leq \ell\}|
\]

\[
= \begin{cases} 
\{ s \in I \cup ([k] \setminus I) \mid s \leq \ell \} & \text{if } I \text{ and } [k] \setminus I \text{ are disjoint} \\
|\{s \in [k] \mid s \leq \ell\}| & \text{else}
\end{cases}
\]

so that \( \alpha_{[k] \setminus I}(\ell) = \ell - \alpha_I(\ell) \). Similarly, every \( \ell \in [k] \) satisfies \( \alpha_{[k] \setminus J}(\ell) = \ell - \alpha_J(\ell) \).

Applying (11.519) to \([k] \setminus J\) and \([k] \setminus I\) in lieu of \( I \) and \( J \), we obtain that

(11.520) \quad \[ k \setminus J \leq \# [k] \setminus I \quad \text{holds if and only if every } \ell \in [k] \text{ satisfies } \alpha_{[k] \setminus J}(\ell) \geq \alpha_{[k] \setminus I}(\ell). \]

Now, we have the following equivalence of assertions:

\[
\begin{align*}
\iff & \quad \text{every } \ell \in [k] \text{ satisfies } \alpha_{[k] \setminus J}(\ell) \geq \alpha_{[k] \setminus I}(\ell) \quad \text{(by (11.520))} \\
\iff & \quad \text{every } \ell \in [k] \text{ satisfies } \ell - \alpha_J(\ell) \geq \ell - \alpha_I(\ell) \\
\iff & \quad \text{every } \ell \in [k] \text{ satisfies } \alpha_I(\ell) \geq \alpha_J(\ell) \\
\iff & \quad (I \leq \# J) \quad \text{(by (11.519))}.
\end{align*}
\]

This proves Proposition 11.11(a).

Returning to the solution of Exercise 2.91(d), we now define some more notations.

If \( T \) is a column-strict tableau and \( i \) is an integer, then the \( i \)-th set column of \( T \) will mean the set of the entries in the \( i \)-th column of \( T \). Notice that the cardinality of the \( i \)-th set column of \( T \) is the length of the \( i \)-th column of the shape of \( T \) (since every column of a column-strict tableau has all its entries distinct), and that the \( i \)-th column of \( T \) can be uniquely reconstructed from the \( i \)-th set column of \( T \) (because the order of the entries in a column can only be increasing).

For every subset \( S \subset [k] \), we define \( x_S \) to be the monomial \( \prod_{s \in S} x_s \) in \( k[x_1, x_2, \ldots, x_k] \). If \((S_1, S_2, \ldots, S_m)\) is an \( m \)-tuple of subsets of \([k]\), then we set \( x_{(S_1, S_2, \ldots, S_m)} = \prod_{i=1}^{m} x_{S_i} \).

Let \( \lambda^i \) denote the conjugate of the partition \( \lambda \). Then, \( (\lambda^i)_i \) is the length of the \( i \)-th column of the Ferrers diagram of \( \lambda \) for every \( i \in \{1, 2, \ldots, m\} \).

Let \( A(\lambda) \) denote the set of all \( m \)-tuples \((I_1, I_2, \ldots, I_m)\) of subsets of \([k]\) satisfying \( I_1 \leq \# I_2 \leq \# \cdots \leq \# I_m \) and \(|I_i| = (\lambda^i)_i \) for every \( i \in \{1, 2, \ldots, m\} \). Note that we denote it by \( A(\lambda) \) to stress its dependency on \( \lambda \).

Let \( (\lambda^\vee)^i \) denote the conjugate of the partition \( \lambda^\vee \). It is easy to see that

(11.521) \quad \( (\lambda^\vee)^i = k - (\lambda^i)_{m+1-i} \quad \text{for every } i \in \{1, 2, \ldots, m\} \).

\[467\text{Proof of (11.521): Fix } i \in \{1, 2, \ldots, m\}. \text{ We recall that } \lambda^\vee = (m - \lambda_k, m - \lambda_{k-1}, \ldots, m - \lambda_1). \text{ Hence,}
\]

(11.522) \quad \( (\lambda^\vee)_j = m - \lambda_{k+1-j} \quad \text{for every } j \in \{1, 2, \ldots, k\} \).
It is easy to see (from the definition of $\lambda^\vee$) that $\lambda^\vee$ is a partition satisfying $\ell(\lambda^\vee) \leq k$, and all parts of $\lambda^\vee$ are $\leq m$. Hence, we can define a set $A(\lambda^\vee)$ in the same way as we defined the set $A(\lambda)$ (but with every $\lambda$ replaced by $\lambda^\vee$). Explicitly, $A(\lambda^\vee)$ is the set of all $m$-tuples $(I_1, I_2, \ldots, I_m)$ of subsets of $[k]$ satisfying $I_1 \leq_\# I_2 \leq_\# \cdots \leq_\# I_m$ and $\left| I_i \right| = (\lambda^\vee)^\ell_i$ for every $i \in \{1, 2, \ldots, m\}$.

Define a map

$$\varphi(\lambda): \text{the set of all } (\lambda, k)\text{-CSTs} \rightarrow A(\lambda)$$

Also, $\ell(\lambda^\vee) \leq k$ (since $\lambda^\vee = (m - \lambda_k, m - \lambda_{k-1}, \ldots, m - \lambda_1)$), so that every positive integer $j > k$ satisfies $(\lambda^\vee)^_j = 0$. Thus, every positive integer $j$ satisfying $(\lambda^\vee)^_j \geq i$ must belong to $\{1, 2, \ldots, k\}$ (since otherwise, this $j$ would satisfy $j > k$, and thus $(\lambda^\vee)^_j = 0$, which would contradict $(\lambda^\vee)^_j \geq i > 0$). Hence,

$$\left\{ j \mid (\lambda^\vee)^_j \geq i \right\} = \left\{ j \in \{1, 2, \ldots, k\} \mid (\lambda^\vee)^_j \geq i \right\} = \left\{ j \in \{1, 2, \ldots, k\} \mid m - \lambda_{k+1-j} \geq i \right\} \quad \text{(by (11.524))}$$

But applying (2.10) to $\lambda^\vee$ instead of $\lambda$, we obtain

$$\left( (\lambda^\vee)^{+}_i \right)_j = \left\{ j \mid (\lambda^\vee)^_j \geq i \right\} = \left\{ j \in \{1, 2, \ldots, k\} \setminus \{ j \in \{1, 2, \ldots, k\} \mid \lambda_{k+1-j} \geq m - i + 1 \} \right\} \quad \text{and}
\left( (\lambda^\vee)^{+}_i \right)_j = k - \left| \{ j \in \{1, 2, \ldots, k\} \mid \lambda_j \geq m - i + 1 \} \right|.$$

But $\ell(\lambda) \leq k$, so that every positive integer $j > k$ satisfies $\lambda_j = 0$. Thus, every positive integer $j$ satisfying $\lambda_j \geq m - i + 1$ must belong to $\{1, 2, \ldots, k\}$ (since otherwise, this $j$ would satisfy $j > k$, and thus $\lambda_j = 0$, which would contradict $\lambda_j \geq m - i + 1 > 0$). Hence,

$$\left\{ j \mid \lambda_j \geq m - i + 1 \right\} = \left\{ j \in \{1, 2, \ldots, k\} \mid \lambda_j \geq m - i + 1 \right\}.$$

Now, (2.10) (applied to $m - i + 1$ instead of $i$) yields

$$\left( (\lambda^\vee)^{+}_{m-i+1} \right)_m = \left\{ j \mid \lambda_j \geq m - i + 1 \right\} = \left| \{ j \in \{1, 2, \ldots, k\} \mid \lambda_j \geq m - i + 1 \} \right|.$$

Thus, (11.523) becomes

$$\left( (\lambda^\vee)^{+}_i \right)_j = k - \left| \{ j \in \{1, 2, \ldots, k\} \mid \lambda_j \geq m - i + 1 \} \right| = k - (\lambda^\vee)^{+}_i = k - (\lambda^\vee)^{+}_{m-i+1}.$$

This proves (11.521).
by sending every \((\lambda, k)\)-CST \(T\) to the \(m\)-tuple whose \(i\)-th entry is the \(i\)-th set column of \(T\). This map is well-defined \(^{468}\), injective \(^{469}\) and surjective \(^{470}\). Hence, \(\varphi(\lambda)\) is a bijection. This bijection \(\varphi(\lambda)\) furthermore satisfies

\[
(11.525) \quad x_{(\varphi(\lambda))(T)} = x^{\text{cont}(T)} \quad \text{for every } (\lambda, k)\text{-CST } T
\]

\(^{471}\). Substituting \(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\) for \(x_1, x_2, \ldots, x_k\) on both sides of this equality, we obtain

\[
(11.526) \quad (x_{(\varphi(\lambda))(T)})^{-1} = (x^{-1})^{\text{cont}(T)} \quad \text{for every } (\lambda, k)\text{-CST } T
\]

(in fact, \(x_{(\varphi(\lambda))(T)}\) is a monomial, so that substituting \(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\) for \(x_1, x_2, \ldots, x_k\) transforms it into its inverse \((x_{(\varphi(\lambda))(T)})^{-1}\).

Recall that \(\lambda^\vee\) is a partition satisfying \(\ell(\lambda^\vee) \leq k\), and all parts of \(\lambda^\vee\) are \(\leq m\). Hence, we can define a bijection

\[
\varphi(\lambda^\vee): (\text{the set of all } (\lambda^\vee, k)\text{-CSTs}) \rightarrow A(\lambda^\vee)
\]

in the same way as we defined the bijection \(\varphi(\lambda): (\text{the set of all } (\lambda, k)\text{-CSTs}) \rightarrow A(\lambda)\) (but with \(\lambda^\vee\) instead of \(\lambda\)). This bijection satisfies

\[
(11.527) \quad x_{(\varphi(\lambda^\vee))(T)} = x^{\text{cont}(T)} \quad \text{for every } (\lambda^\vee, k)\text{-CST } T
\]

(this can be shown in analogy to \((11.525)\)).

Finally, we define a map \(\psi(\lambda): A(\lambda) \rightarrow A(\lambda^\vee)\) as follows: For every \((I_1, I_2, \ldots, I_m) \in A(\lambda)\), let

\[
(\psi(\lambda))(I_1, I_2, \ldots, I_m) = ([k] \setminus I_m, [k] \setminus I_{m-1}, \ldots, [k] \setminus I_1).
\]

\(^{468}\)Indeed, if \(T\) is an \((\lambda, k)\)-CST, then the \(m\)-tuple whose \(i\)-th entry is the \(i\)-th set column of \(T\) belongs to \(A(\lambda)\) (because if we denote this \(m\)-tuple by \((I_1, I_2, \ldots, I_m)\), then

\[
|I_i| = |(\text{the } i\text{-th set column of } T)| = (\text{the length of the } i\text{-th column of } T) = (\text{the length of the } i\text{-th column of the Ferrers diagram of } \lambda)
\]

(since \(T\) is a \((\lambda, k)\)-CST, thus a column-strict tableau of shape \(\lambda\))

\[= \lambda^\vee_i,
\]

and the fact that the entries of \(T\) increase weakly along rows (because \(T\) is a column-strict tableau) translates precisely into the inequality chain \(I_1 \leq_{\#} I_2 \leq_{\#} \cdots \leq_{\#} I_m\).

\(^{469}\)This is because the \(i\)-th column of a \((\lambda, k)\)-CST \(T\) can be uniquely reconstructed from the \(i\)-th set column of \(T\) (indeed, the entries of the \(i\)-th column of \(T\) must be strictly increasing down the column, and therefore the knowledge of the set of these entries is sufficient to recover the \(i\)-th column). Here, we are using the fact that \(T\) has at most \(m\) columns (since every part of \(\lambda\) is \(\leq m\)).

\(^{470}\)Indeed, given any \((I_1, I_2, \ldots, I_m) \in A(\lambda)\). Then, \((I_1, I_2, \ldots, I_m)\) is an \(m\)-tuple of subsets of \([k]\) satisfying \(I_1 \leq_{\#} I_2 \leq_{\#} \cdots \leq_{\#} I_m\) and \(|I_i| = \lambda^\vee_i\) for every \(i \in \{1, 2, \ldots, m\}\). Now, we can fill in each column of the Ferrers diagram of \(\lambda\) with the entries of the corresponding set \(I_i\) in increasing order, and then the resulting filling is a column-strict tableau (indeed, its entries increase weakly along its rows, due to \(I_1 \leq_{\#} I_2 \leq_{\#} \cdots \leq_{\#} I_m\)), and more precisely a \((\lambda, k)\)-CST. Our \(m\)-tuple \((I_1, I_2, \ldots, I_m)\) is the image of this \((\lambda, k)\)-CST under \(\varphi(\lambda)\); therefore, \((I_1, I_2, \ldots, I_m)\) lies in the image of \(\varphi(\lambda)\). Hence, we have shown that every \((I_1, I_2, \ldots, I_m) \in A(\lambda)\) lies in the image of \(\varphi(\lambda)\). The map \(\varphi(\lambda)\) is thus surjective, q.d.

\(^{471}\)Indeed, it is easy to check that both \(x_{(\varphi(\lambda))(T)}\) and \(x^{\text{cont}(T)}\) are equal to the product \(\prod c_i\text{ is a cell of } T^x(\text{entry of } T \text{ in } c)\).
This map $\psi(\lambda)$ is well-defined and bijective. This bijective map $\psi(\lambda)$ has the property that

$$
(11.528) \quad x_{(\psi(\lambda))(S)} = (x_1 x_2 \cdots x_k)^m \cdot (x_S)^{-1} \quad \text{for every } S \in A(\lambda)
$$

Using the three bijective maps

$$
\varphi(\lambda) : \text{(the set of all } (\lambda, k)\text{-CSTs}) \to A(\lambda),
$$

$$
\psi(\lambda) : A(\lambda) \to A(\lambda'),
$$

$$
\varphi(\lambda') : \text{(the set of all } (\lambda', k)\text{-CSTs}) \to A(\lambda'),
$$

we can define a map

$$
\Omega : \text{(the set of all } (\lambda, k)\text{-CSTs}) \to \text{(the set of all } (\lambda', k)\text{-CSTs})
$$

by $\Omega = (\varphi(\lambda'))^{-1} \circ (\psi(\lambda)) \circ (\varphi(\lambda))$. Clearly, this $\Omega$ is a bijection. If we can show that every $(\lambda, k)$-CST $T$ satisfies (11.518), then (11.517) will be proven, and thus Exercise 2.91(d) will be solved again. Hence, all that remains to prove now is that every $(\lambda, k)$-CST $T$ satisfies (11.518).

---

**Proof.** Let $(I_1, I_2, \ldots, I_m) \in A(\lambda)$. Then, $(I_1, I_2, \ldots, I_m)$ is an $m$-tuple of subsets of $[k]$ satisfying $I_1 \subseteq I_2 \subseteq \cdots \subseteq I_m$ and $(|I_i| = (\lambda')^i)$ for every $i \in \{1, 2, \ldots, m\}$.

In order to prove the well-definedness of $\psi(\lambda)$, we need to show that $(|k| \setminus I_m) \setminus k \subseteq I_{m-1} \subseteq \cdots \subseteq \{1\}$ is an $m$-tuple of subsets of $[k]$ satisfying $|k| \setminus I_m \subseteq \cdots \subseteq |k| \setminus I_1 = \{1, 2, \ldots, m\}$.

First of all, it is clear that $(|k| \setminus I_m, |k| \setminus I_{m-1}, \ldots, |k| \setminus I_1)$ is an $m$-tuple of subsets of $[k]$. Furthermore, $|k| \setminus I_m \subseteq \cdots \subseteq |k| \setminus I_1$ follows from $I_1 \subseteq I_2 \subseteq \cdots \subseteq I_m$ (according to Proposition 11.11(a)). Thus, it remains to prove that $(|k| \setminus I_{m+1-i} = (\lambda')^i)$ for every $i \in \{1, 2, \ldots, m\}$.

So fix some $i \in \{1, 2, \ldots, m\}$. Then, $|I_{m+1-i}| = (\lambda')^i$ (this follows from the $|I_i| = (\lambda')^i$ formula, but applied to $m+1-i$ instead of $i$). Since $I_{m+1-i} \subseteq [k]$, we have

$$
|k| \setminus I_{m+1-i} = |k| - |I_{m+1-i}| = k - (\lambda')^i = (\lambda')^i,
$$

$qed.$

Indeed, we can define a map $\psi(\lambda')$ in the same way as we have defined $\psi(\lambda)$ (but with $\lambda'$ instead of $\lambda$). It is then easy to see that the maps $\psi(\lambda)$ and $\psi(\lambda')$ are mutually inverse, so that $\psi(\lambda)$ is bijective. qed.

**Proof of (11.528):** Let $S \in A(\lambda)$. Then, $S$ is an $m$-tuple of subsets of $[k]$. Write $S$ as $S = (I_1, I_2, \ldots, I_m)$. The definition of $x_S$ then yields

$$
x_S = \prod_{i=1}^m x_{I_i}.
$$

Hence,

$$
(11.529) \quad (x_1 x_2 \cdots x_k)^m \cdot (x_S)^{-1} = (x_1 x_2 \cdots x_k)^m \cdot \left(\prod_{i=1}^m x_{I_i}\right)^{-1} = \prod_{i=1}^m x_{\frac{x_1 x_2 \cdots x_k}{x_{I_i}}}.
$$

On the other hand, the definition of $\psi(\lambda)$ yields $\psi(\lambda)(S) = (|k| \setminus I_m, |k| \setminus I_{m-1}, \ldots, |k| \setminus I_1)$, and thus (by the definition of $x_{(\psi(\lambda))(S)}$) we have

$$
x_{(\psi(\lambda))(S)} = \prod_{i=1}^m x_{|k| \setminus I_{m+1-i}} = \prod_{i=1}^m x_{\frac{|k| \setminus I_i}{x_S}} \quad \text{(here, we substituted } m+1-i \text{ for } i \text{ in the product}).
$$

But we need to prove (11.528). In view of (11.529) and (11.530), this boils down to proving that $\prod_{i=1}^m x_{|k| \setminus I_i} = \prod_{i=1}^m \frac{x_1 x_2 \cdots x_k}{x_{I_i}}$. This will immediately follow if we can prove the identity $x_{|k| \setminus I_i} = \frac{x_1 x_2 \cdots x_k}{x_{I_i}}$ for every $i \in \{1, 2, \ldots, m\}$.

But the latter identity follows from the (obvious) fact that $x_{|k| \setminus S} = \frac{x_1 x_2 \cdots x_k}{x_S}$ for every $S \subseteq [k]$. Thus, (11.528) is proven.
Let $T$ be a $(\lambda, k)$-CST. We have $\Omega = (\varphi(\lambda^\vee))^{-1} \circ (\psi(\lambda)) \circ (\varphi(\lambda))$, thus $(\varphi(\lambda^\vee)) \circ \Omega = (\psi(\lambda)) \circ (\varphi(\lambda))$. Applying (11.527) to $\Omega(T)$ instead of $T$, we obtain $x_{(\varphi(\lambda^\vee))(\Omega(T))} = x^{\text{cont}(\Omega(T))}$, whence

$$x^{\text{cont}(\Omega(T))} = x_{(\varphi(\lambda^\vee))(\Omega(T))} = x_{(\psi(\lambda))(\varphi(\lambda))}(T)$$

since $(\varphi(\lambda^\vee))(\Omega(T)) = ((\varphi(\lambda^\vee)) \circ \Omega)(T) = ((\psi(\lambda)) \circ (\varphi(\lambda)))(T) = (\psi(\lambda))(\varphi(\lambda))(T)$

(by (11.528), applied to $S = (\varphi(\lambda))(T)$)

$$= (x_1x_2 \cdots x_k)^m \cdot (x_{(\varphi(\lambda))(T)})^{-1}$$

(by (11.526))

$$= (x_1x_2 \cdots x_k)^m \cdot (x^{-1})^{\text{cont}(T)}.$$

Thus, (11.518) holds. We thus have proven that every $(\lambda, k)$-CST $T$ satisfies (11.518). Exercise 2.91(d) is thus solved.

(e) First solution to Exercise 2.91(e): Obviously, the $k$-tuple $(r + \lambda_1, r + \lambda_2, \ldots, r + \lambda_k)$ is a partition. Let us denote this partition by $\lambda^{[r]}$. Then, $\lambda^{[r]} = (r + \lambda_1, r + \lambda_2, \ldots, r + \lambda_k)$ and $\ell(\lambda^{[r]}) \leq k$.

Let $n = k$. We shall use the notations of Section 2.6; in particular, we set $x = (x_1, x_2, \ldots, x_n)$ and $\rho = (n - 1, n - 2, \ldots, 2, 1, 0)$. Clearly, $x = (x_1, x_2, \ldots, x_n) = (x_1, x_2, \ldots, x_k)$ (since $n = k$). Notice that $\ell(\lambda) \leq k = n$. Hence, we can regard $\lambda$ as an element of $\mathbb{N}^n$; therefore, $\lambda + \rho$ is a well-defined element of $\mathbb{N}^n$, and the alternant $a_{\lambda + \rho}$ is well-defined. Corollary 2.54 yields that $s_{\lambda}(x) = \frac{a_{\lambda + \rho}}{a_{\rho}}$, so that

$$a_{\rho} \cdot s_{\lambda}(x) = a_{\lambda + \rho}.$$

Applying this equality to $\lambda^{[r]}$ instead of $\lambda$, we obtain

$$a_{\rho} \cdot s_{\lambda^{[r]}}(x) = a_{\lambda^{[r]} + \rho}$$

(since $\lambda^{[r]}$ is also a partition satisfying $\ell(\lambda^{[r]}) \leq k = n$).

We have (11.507), and every $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}^n$ satisfies (11.506). (This can be proven just as in the First solution to Exercise 2.91(d).) We have $\lambda^{[r]} = (r + \lambda_1, r + \lambda_2, \ldots, r + \lambda_k) = (r + \lambda_1, r + \lambda_2, \ldots, r + \lambda_n)$ (since $k = n$), so that

$$\frac{\lambda^{[r]}}{\rho} = (r + \lambda_1, r + \lambda_2, \ldots, r + \lambda_n) + (n - 1, n - 2, \ldots, n - n) = ((r + \lambda_1) + (n - 1), (r + \lambda_2) + (n - 2), \ldots, (r + \lambda_n) + (n - n)).$$
Hence, \((11.506)\) (applied to \(\lambda^{[r]} + \rho\) and 
\(((r + \lambda_1) + (n - 1), (r + \lambda_2) + (n - 2), \ldots, (r + \lambda_n) + (n - n))\) instead of \(\alpha\) and \((\alpha_1, \alpha_2, \ldots, \alpha_n)\) yields

\[
\begin{align*}
\tilde{a}_{\lambda^{[r]} + \rho} &= \sum_{w \in \mathcal{S}_n} \text{sgn}(w) \prod_{i=1}^{n} x_w^{(r+\lambda_i)+(n-i)} = \sum_{w \in \mathcal{S}_n} \text{sgn}(w) \left( \prod_{i=1}^{n} x_w^{r} \right) \left( \prod_{i=1}^{n} x_w^{\lambda_i+n-i} \right) \\
&= \sum_{w \in \mathcal{S}_n} \text{sgn}(w) \left( \prod_{i=1}^{n} x_w^{r} \right) \left( \prod_{i=1}^{n} x_w^{\lambda_i+n-i} \right) = \sum_{w \in \mathcal{S}_n} \text{sgn}(w) \left( \prod_{i=1}^{n} x_w^{r} \right) \left( \prod_{i=1}^{n} x_w^{\lambda_i+n-i} \right) \\
&= \sum_{w \in \mathcal{S}_n} \text{sgn}(w) \left( \prod_{i=1}^{n} x_w^{r} \right) \left( \prod_{i=1}^{n} x_w^{\lambda_i+n-i} \right) = \sum_{w \in \mathcal{S}_n} \text{sgn}(w) \left( \prod_{i=1}^{n} x_w^{r} \right) \left( \prod_{i=1}^{n} x_w^{\lambda_i+n-i} \right) \\
&= (x_1 x_2 \cdots x_n)^r \cdot a_{\lambda^{[r]} + \rho} = (x_1 x_2 \cdots x_n)^r \cdot \alpha_{\rho} \cdot s_\lambda (\mathbf{x}) = a_{\rho} \cdot (x_1 x_2 \cdots x_n)^r \cdot s_\lambda (\mathbf{x}).
\end{align*}
\]

Hence, \(a_{\rho} \cdot (x_1 x_2 \cdots x_n)^r \cdot s_\lambda (\mathbf{x}) = a_{\lambda^{[r]} + \rho} = a_{\rho} \cdot s_\lambda^{[r]} (\mathbf{x}) \) by \((11.531)\).}

But we know (from a footnote in Corollary 2.54) that \(a_{\rho}\) is not a zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n]\).

In other words, \(a_{\rho}\) is a non-zero-divisor in the ring \(k[x_1, x_2, \ldots, x_n]\). Hence, we can cancel the factor \(a_{\rho}\) from the equality \(a_{\rho} \cdot (x_1 x_2 \cdots x_n)^r \cdot s_\lambda (\mathbf{x}) = a_{\rho} \cdot s_\lambda^{[r]} (\mathbf{x})\). As a result, we obtain \((x_1 x_2 \cdots x_n)^r \cdot s_\lambda (\mathbf{x}) = s_{\lambda^{[r]}} (\mathbf{x}) = s_{\lambda^{[r]}} (x_1, x_2, \ldots, x_k)\) (since \(\mathbf{x} = (x_1, x_2, \ldots, x_k)\)), so that

\[
s_{\lambda^{[r]}} (x_1, x_2, \ldots, x_k) = \left( \frac{x_1 x_2 \cdots x_n}{\frac{x_1 x_2 \cdots x_k}{\text{since } n=k}} \right)^r \cdot \frac{s_\lambda (\mathbf{x})}{s_\lambda (x_1, x_2, \ldots, x_k)} = \left( \frac{x_1 x_2 \cdots x_n}{\frac{x_1 x_2 \cdots x_k}{\text{since } n=k}} \right)^r \cdot \frac{s_\lambda (\mathbf{x})}{s_\lambda (x_1, x_2, \ldots, x_k)}.
\]

Since \(\lambda^{[r]} = s_{(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k)}\), this rewrites as \(s_{(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k)} (x_1, x_2, \ldots, x_k) = (x_1 x_2 \cdots x_k)^r \cdot s_\lambda (x_1, x_2, \ldots, x_k)\). This solves Exercise 2.91(e).

**Second solution to Exercise 2.91(e) (sketched):** Obviously, the \(k\)-tuple \((r + \lambda_1, r + \lambda_2, \ldots, r + \lambda_k)\) is a partition. Let us denote this partition by \(\lambda^{[r]}\). Then, \(\lambda^{[r]} = (r + \lambda_1, r + \lambda_2, \ldots, r + \lambda_k)\) and \(\ell (\lambda^{[r]}) \leq k\).

We have \((r + \lambda_1, r + \lambda_2, \ldots, r + \lambda_k) = \lambda^{[r]}\), thus \(s_{(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k)} = s_{\lambda^{[r]}} = s_{\lambda^{[r]}/\emptyset}\), hence

\[
s_{(r+\lambda_1, r+\lambda_2, \ldots, r+\lambda_k)} (x_1, x_2, \ldots, x_k) = \sum_{T \text{ is a column-strict tableau of shape } \lambda^{[r]}/\emptyset; \text{ all entries of } T \text{ belong to } \{1, 2, \ldots, k\}} x^{\text{cont}(T)}
\]

(by Exercise 2.26(a), applied to \(k, \lambda^{[r]}\) and \(\emptyset\) instead of \(n, \lambda\) and \(\mu\))

\[(11.533) \quad = \sum_{T \text{ is a column-strict tableau of shape } \lambda^{[r]}; \text{ all entries of } T \text{ belong to } \{1, 2, \ldots, k\}} x^{\text{cont}(T)}
\]

(since column-strict tableaux of shape \(\lambda^{[r]}/\emptyset\) are the same as column-strict tableaux of shape \(\lambda^{[r]}\)).
But an argument analogous to the one we just used to prove (11.533) (but with \( \lambda \) in place of \( \lambda^{[r]} \)) shows that

\[
s_{\lambda}(x_1, x_2, \ldots, x_k) = \sum_{T \text{ is a column-strict tableau of shape } \lambda; \text{ all entries of } T \text{ belong to } \{1, 2, \ldots, k\}} x^{\text{cont}(T)}.
\]

Multiplied by \((x_1 x_2 \cdots x_k)^r\), this equality becomes

\[
(x_1 x_2 \cdots x_k)^r \cdot s_{\lambda}(x_1, x_2, \ldots, x_k) = (x_1 x_2 \cdots x_k)^r \cdot \sum_{T \text{ is a column-strict tableau of shape } \lambda; \text{ all entries of } T \text{ belong to } \{1, 2, \ldots, k\}} x^{\text{cont}(T)}.
\]

(11.534)

We need to prove that

\[
s_{(r+1, 1, r+2, \ldots, r+k)}(x_1, x_2, \ldots, x_k) = (x_1 x_2 \cdots x_k)^r \cdot s_{\lambda}(x_1, x_2, \ldots, x_k).
\]

Due to (11.533) and (11.534), this rewrites as

(11.535)

\[
\sum_{T \text{ is a column-strict tableau of shape } \lambda^{[r]}; \text{ all entries of } T \text{ belong to } \{1, 2, \ldots, k\}} x^{\text{cont}(T)} = \sum_{T \text{ is a column-strict tableau of shape } \lambda; \text{ all entries of } T \text{ belong to } \{1, 2, \ldots, k\}} (x_1 x_2 \cdots x_k)^r \cdot x^{\text{cont}(T)}.
\]

So it remains to prove (11.535).

In order to prove (11.535), it is clearly sufficient to construct a bijection

\[
\Omega: \left( \text{the set of all column-strict tableaux } T \text{ of shape } \lambda^{[r]} \text{ such that all entries of } T \text{ belong to } \{1, 2, \ldots, k\} \right) \rightarrow \left( \text{the set of all column-strict tableaux } T \text{ of shape } \lambda \text{ such that all entries of } T \text{ belong to } \{1, 2, \ldots, k\} \right)
\]

with the property that every column-strict tableau \( T \) of shape \( \lambda^{[r]} \) such that all entries of \( T \) belong to \( \{1, 2, \ldots, k\} \) satisfies

(11.536)

\[
(x_1 x_2 \cdots x_k)^r \cdot x^{\text{cont}(\Omega(T))} = x^{\text{cont}(T)}.
\]

Constructing such an \( \Omega \) is easy: The map \( \Omega \) just maps every column-strict tableau \( T \) of shape \( \lambda^{[r]} \) to the tableau of shape \( \lambda \) obtained by removing the first \( r \) entries of each row of \( T \), and moving all other entries to the left by \( r \) cells. To prove that \( \Omega \) is bijective, we need to construct a map inverse to \( \Omega \); this latter map sends every column-strict tableau \( T \) of shape \( \lambda \) to the tableau of shape \( \lambda^{[r]} \) obtained by moving all entries of \( T \) to the right by \( r \) cells, and filling the now-vacant first \( r \) columns as follows:

\[
\begin{array}{cccc}
1 & 1 & \cdots & 1 \\
2 & 2 & \cdots & 2 \\
\vdots & \vdots & \ddots & \vdots \\
k & k & \cdots & k \\
\end{array}
\]

Proving that this map is well-defined and really inverse to \( \Omega \) is left to the reader\(^{475}\). Anyway, we now know that \( \Omega \) is bijective, and it is easy to check that (11.536) holds. This completes the second solution of Exercise 2.91(e).

\(^{475}\)The main ingredient of this proof is the observation that if \( T \) is a column-strict tableau of shape \( \lambda^{[r]} \) such that all entries of \( T \) belong to \( \{1, 2, \ldots, k\} \), then the first \( r \) columns of \( T \) must look like this:

\[
\begin{array}{cccc}
1 & 1 & \cdots & 1 \\
2 & 2 & \cdots & 2 \\
\vdots & \vdots & \ddots & \vdots \\
k & k & \cdots & k \\
\end{array}
\]
11.64. Solution to Exercise 2.92. Solution to Exercise 2.92. (a) We can apply Exercise 2.91(a) to $\mu$, $\mu$, $\nu^\lambda$, and $\nu^\mu$ instead of $\lambda$, $\mu$, $\lambda^\nu$ and $\mu^\nu$. As a consequence, we conclude that $\nu^\lambda$ and $\nu^\mu$ are partitions, and that $s_{\mu/\nu} = s_{\nu^\lambda/\nu^\mu}$. Thus, $\nu^\lambda$ is a partition. The same argument (but with $m$, $\mu$ and $\nu^\lambda$ replaced by $n, \nu$ and $\nu^\lambda$) yields that $\nu^\lambda$ is a partition. This solves Exercise 2.92(a).

(b) Assume that not all parts of $\lambda$ are $\leq m + n$. Then, some part of $\lambda$ is $> m + n$; hence, the greatest part of $\lambda$ is $> m + n$. Thus, $\lambda_1 > m + n$ (since $\lambda_1$ is the greatest part of $\lambda$). As a consequence, $\lambda_1 - \mu_1 > \nu_1$.

Now, we shall show that

\[(11.537)\quad \left( \text{there exists no column-strict tableau } T \text{ of shape } \lambda/\mu \text{ with } \text{cont}(T) = \nu \right) \quad \text{having the property that each } \text{cont}(T |_{\text{cols} \geq j}) \text{ is a partition} \]

\[\text{Proof of (11.537): Assume the contrary. Thus, there exists a column-strict tableau } T \text{ of shape } \lambda/\mu \text{ with } \text{cont}(T) = \nu \text{ having the property that each } \text{cont}(T |_{\text{cols} \geq j}) \text{ is a partition. Consider this } T. \]

The column-strict tableau $T$ has shape $\lambda/\mu$, and thus its 1-st row has $\lambda_1 - \mu_1$ entries. In other words,

\[(\text{the number of entries in the 1-st row of } T) = \lambda_1 - \mu_1 > \nu_1 = (\text{cont}(T))_1 \quad (\text{since } \nu = \text{cont}(T)) \]

\[\quad = \vert T^{-1}(1) \vert \quad (\text{by the definition of } \text{cont}(T)) \]

\[\quad = (\text{the number of entries of } T \text{ equal to } 1). \]

Hence, the tableau $T$ has more entries in its 1-st row than it has entries equal to 1. Consequently, not every entry in the 1-st row is equal to 1. Thus, there exists an entry in the 1-st row of $T$ which is $> 1$. Let $k$ be this entry, and let $c$ be the cell it occupies. Then, $c$ is a cell in the 1-st row of $T$, and thus can be written in the form $(1, j)$ for some positive integer $j$. Consider this $j$. The cell $c$ lies in column $j$, and thus is a cell of the skew tableau $T |_{\text{cols} \geq j}$; its entry is $T |_{\text{cols} \geq j}(c) = T(c) = k$ (since we know that $k$ is the entry of $T$ occupying cell $c$). Hence, $c \in (T |_{\text{cols} \geq j})^{-1}(k)$, so that the set $(T |_{\text{cols} \geq j})^{-1}(k)$ is nonempty. Thus, $\left( (T |_{\text{cols} \geq j})^{-1}(k) \right) > 0$.

Recalling the definition of cont $(T |_{\text{cols} \geq j})$, we see that $(\text{cont}(T |_{\text{cols} \geq j}))_k = \left( (T |_{\text{cols} \geq j})^{-1}(k) \right) > 0$.

But we know that cont $(T |_{\text{cols} \geq j})$ is a partition. Thus, $(\text{cont}(T |_{\text{cols} \geq j}))_1 \geq (\text{cont}(T |_{\text{cols} \geq j}))_2 \geq (\text{cont}(T |_{\text{cols} \geq j}))_3 \geq \cdots$, so that $(\text{cont}(T |_{\text{cols} \geq j}))_1 \geq (\text{cont}(T |_{\text{cols} \geq j}))_k > 0$. But the definition of cont $(T |_{\text{cols} \geq j})$ yields

\[(\text{cont}(T |_{\text{cols} \geq j}))_1 = \left( (T |_{\text{cols} \geq j})^{-1}(1) \right) = (\text{the number of entries of } T \text{ equal to } 1). \]

Hence, $(\text{the number of entries of } T \text{ equal to } 1) = (\text{cont}(T |_{\text{cols} \geq j}))_1 > 0$. Hence, the skew tableau $T |_{\text{cols} \geq j}$ must have at least one entry equal to 1.

But each cell of the skew tableau $T |_{\text{cols} \geq j}$ lies in one of the columns $j, j + 1, j + 2, \ldots$ (by the definition of $T |_{\text{cols} \geq j}$) and in one of the rows $2, 3, \ldots$ (obviously). Hence, each cell of the skew tableau $T |_{\text{cols} \geq j}$ lies (weakly) southeast of the cell $(1, j) = c$. As a consequence, each entry of the skew tableau $T |_{\text{cols} \geq j}$ is greater or equal to the entry of $T |_{\text{cols} \geq j}$ in the cell $c$. 477. Since the entry of $T |_{\text{cols} \geq j}$ in the cell $c$

476Proof. All parts of $\mu$ are $\leq m$. Thus, $\mu_1 \leq m$ (since $\mu_1$ is either a part of $\mu$ and therefore $\leq m$, or is zero and therefore $\leq m$ as well). Hence, $m \geq \mu_1$. Similarly, $n \geq \nu_1$. Now, $\lambda_1 > m + \nu_1 \geq m_1 + \nu_1$, so that $\lambda_1 - \mu_1 > \nu_1$, qed.

477Indeed, $T$ is a column-strict tableau, and thus the entries of $T$ increase weakly left-to-right along rows, and increase strictly top-to-bottom in columns. Consequently, the entries of $T$ increase weakly as one moves to the southeast. The same holds for the entries of $T |_{\text{cols} \geq j}$ (since $T |_{\text{cols} \geq j}$ is a restriction of $T$), and thus each entry of $T |_{\text{cols} \geq j}$ is greater or equal to the entry of $T |_{\text{cols} \geq j}$ in the cell $c$ (because each cell of $T |_{\text{cols} \geq j}$ lies (weakly) southeast of the cell $c$). Qed.
is \((T_{\text{cols} \geq j}) (c) = k\), this yields that each entry of the skew tableau \(T_{\text{cols} \geq j}\) is greater or equal to \(k\), and thus greater than 1 (since \(k > 1\)). As a consequence, no entry of the skew tableau \(T_{\text{cols} \geq j}\) can be equal to 1. This contradicts the fact that the skew tableau \(T_{\text{cols} \geq j}\) must have at least one entry equal to 1. This contradiction proves that our assumption was wrong. Hence, (11.537) is proven.

Now, Corollary 2.58 yields that \(c^\lambda_{\mu, \nu}\) counts column-strict tableaux \(T\) of shape \(\lambda/\mu\) with \(\text{cont}(T) = \nu\) having the property that each \(\text{cont}(T_{\text{cols} \geq j})\) is a partition. Since there exists no such \(T\) (according to (11.537)), this yields that \(c^\lambda_{\mu, \nu} = 0\). This solves Exercise 2.92(b).

Remark: An alternative solution of Exercise 2.92(b) can be obtained easily from Exercise 2.93(c).

(c) Let us forget that \(\ell(\lambda) > k\), then

\[(11.538) \quad s_\lambda (x_1, x_2, \ldots, x_k) = 0.\]

We first notice that

\[(11.539) \quad s_\mu (x_1, x_2, \ldots, x_k) \cdot s_\nu (x_1, x_2, \ldots, x_k) = \sum_{\lambda \in \text{Par}; \ell(\lambda) \leq k} c^\lambda_{\mu, \nu} s_\lambda (x_1, x_2, \ldots, x_k).\]

Hence,

\[(11.540) \quad s_\mu (x_1, x_2, \ldots, x_k) \cdot s_\nu (x_1, x_2, \ldots, x_k) = \sum_{\lambda \in \text{Par}; \ell(\lambda) \leq k; \text{all parts of } \lambda \leq m+n} c^\lambda_{\mu, \nu} s_\lambda (x_1, x_2, \ldots, x_k).\]

\(^{478}\)Proof of (11.538): Let \(\lambda \in \text{Par}\) be such that \(\ell(\lambda) > k\). Then, the number of parts of \(\lambda\) is \(\ell(\lambda) > k\). Hence, Exercise 2.26(b) (applied to \(k\) instead of \(n\)) yields \(s_\lambda (x_1, x_2, \ldots, x_k) = 0\), qed.

\(^{479}\)Proof of (11.539): We have \(s_\mu s_\nu = \sum_\lambda c^\lambda_{\mu, \nu} s_\lambda\), where the sum ranges over all partitions \(\lambda\) (according to the definition of the coefficients \(c^\lambda_{\mu, \nu}\)). In other words, \(s_\mu s_\nu = \sum_{\lambda \in \text{Par}} c^\lambda_{\mu, \nu} s_\lambda\). Evaluating both sides of this equality at \((x_1, x_2, \ldots, x_k)\), we obtain

\[
\begin{align*}
    s_\mu (x_1, x_2, \ldots, x_k) & \cdot s_\nu (x_1, x_2, \ldots, x_k) \\
    &= \sum_{\lambda \in \text{Par}; \ell(\lambda) \leq k} c^\lambda_{\mu, \nu} s_\lambda (x_1, x_2, \ldots, x_k) \\
    &= \sum_{\lambda \in \text{Par}; \ell(\lambda) \leq k} c^\lambda_{\mu, \nu} s_\lambda (x_1, x_2, \ldots, x_k) + \sum_{\lambda \in \text{Par}; \ell(\lambda) > k} c^\lambda_{\mu, \nu} s_\lambda (x_1, x_2, \ldots, x_k) \\
    &= \sum_{\lambda \in \text{Par}; \ell(\lambda) \leq k} c^\lambda_{\mu, \nu} s_\lambda (x_1, x_2, \ldots, x_k) + \sum_{\lambda \in \text{Par}; \ell(\lambda) > k} c^\lambda_{\mu, \nu} 0 \\
    &= \sum_{\lambda \in \text{Par}; \ell(\lambda) \leq k} c^\lambda_{\mu, \nu} s_\lambda (x_1, x_2, \ldots, x_k).
\end{align*}
\]

qed.

\(^{480}\)Proof of (11.540): From (11.539), we obtain

\[
\begin{align*}
    s_\mu (x_1, x_2, \ldots, x_k) & \cdot s_\nu (x_1, x_2, \ldots, x_k) \\
    &= \sum_{\lambda \in \text{Par}; \ell(\lambda) \leq k} c^\lambda_{\mu, \nu} s_\lambda (x_1, x_2, \ldots, x_k) \\
    &= \sum_{\lambda \in \text{Par}; \ell(\lambda) \leq k; \text{all parts of } \lambda \leq m+n} c^\lambda_{\mu, \nu} s_\lambda (x_1, x_2, \ldots, x_k) + \sum_{\lambda \in \text{Par}; \ell(\lambda) \leq k; \text{not all parts of } \lambda \leq m+n} c^\lambda_{\mu, \nu} s_\lambda (x_1, x_2, \ldots, x_k) \\
    &= \sum_{\lambda \in \text{Par}; \ell(\lambda) \leq k; \text{all parts of } \lambda \leq m+n} c^\lambda_{\mu, \nu} s_\lambda (x_1, x_2, \ldots, x_k) + \sum_{\lambda \in \text{Par}; \ell(\lambda) \leq k; \text{not all parts of } \lambda \leq m+n} 0 s_\lambda (x_1, x_2, \ldots, x_k) \\
    &= \sum_{\lambda \in \text{Par}; \ell(\lambda) \leq k; \text{all parts of } \lambda \leq m+n} c^\lambda_{\mu, \nu} s_\lambda (x_1, x_2, \ldots, x_k).
\end{align*}
\]

This proves (11.540).
Define a set $\mathfrak{A}$ by
\begin{equation}
\mathfrak{A} = \{ \alpha \in \text{Par} \mid \ell(\alpha) \leq k; \text{ all parts of } \alpha \text{ are } \leq m+n \}.
\end{equation}
For every partition $\lambda$, let $\lambda^\nu(m+n)$ denote the $k$-tuple $(m + n - \lambda_k, m + n - \lambda_{k-1}, \ldots, m + n - \lambda_1)$. It is straightforward to see that for every $\lambda \in \mathfrak{A}$, we have
\begin{equation}
\lambda^\nu(m+n) \in \mathfrak{A}.
\end{equation}
Thus, the map
$$
\mathfrak{A} \to \mathfrak{A}, \quad \lambda \mapsto \lambda^\nu(m+n)
$$
is well-defined. It is easy to see that this map is an involution (i.e., every $\lambda \in \mathfrak{A}$ satisfies $(\lambda^\nu(m+n))^\nu(m+n) = \lambda$), thus a bijection.

Now, the summation sign $\sum_{\lambda \in \text{Par} : \ell(\lambda) \leq k; \text{ all parts of } \lambda \text{ are } \leq m+n}$ in (11.540) rewrites as $\sum_{\lambda \in \mathfrak{A}}$ (because of how we defined $\mathfrak{A}$). Hence, (11.540) rewrites as
\begin{equation}
\sum_{\lambda \in \mathfrak{A}} c_{\mu,\nu}^\lambda s_{\lambda}(x_1, x_2, \ldots, x_k) = \sum_{\lambda \in \mathfrak{A}} c_{\mu,\nu}^\lambda s_{\lambda}(x_1, x_2, \ldots, x_k).
\end{equation}

Hence, in the Laurent polynomial ring $k\left[x_1, \ldots, x_k, x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}\right]$, we have
\begin{equation}
s_{\mu}(x_1, x_2, \ldots, x_k) \cdot s_{\nu}(x_1, x_2, \ldots, x_k) = \sum_{\lambda \in \mathfrak{A}} c_{\mu,\nu}^\lambda s_{\lambda}(x_1, x_2, \ldots, x_k).
\end{equation}

481. Now, it is straightforward to see that $\mu^\nu(m)$ and $\nu^\nu(n)$ are partitions satisfying $\ell(\mu^\nu(m)) \leq k$ and $\ell(\nu^\nu(n)) \leq k$; also, all parts of the partition $\mu^\nu(m)$ are $\leq m$, and all parts of the partition $\nu^\nu(n)$ are $\leq n$. Moreover,
\begin{equation}
\mu = \left(m - \left(\mu^\nu(m)\right)_k, m - \left(\mu^\nu(m)\right)_{k-1}, \ldots, m - \left(\mu^\nu(m)\right)_1\right)
\end{equation}
and
\begin{equation}
\nu = \left(n - \left(\nu^\nu(n)\right)_k, n - \left(\nu^\nu(n)\right)_{k-1}, \ldots, n - \left(\nu^\nu(n)\right)_1\right)
\end{equation}
(this is still easy to verify). These observations lead us to the conclusion that we can apply (11.544) to $\mu^\nu(m)$, $\mu$, $\nu^\nu(n)$ and $\nu$ instead of $\mu$, $\mu^\nu(m)$, $\mu$ and $\nu^\nu(n)$. As a result, we obtain
\begin{equation}
s_{\mu^\nu(m)}(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}) \cdot s_{\nu^\nu(n)}(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}) = \sum_{\lambda \in \mathfrak{A}} c_{\mu^\nu(m),\nu^\nu(n)}^\lambda s_{\lambda}(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1})
\end{equation}

\begin{equation}
= \sum_{\lambda \in \mathfrak{A}} c_{\mu^\nu(m),\nu^\nu(n)}^\lambda s_{\lambda^\nu(m+n)}(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1})
\end{equation}
(here, we have substituted $\lambda^\nu(m+n)$ for the summation index $\lambda$ (since the map $\mathfrak{A} \to \mathfrak{A}$, $\lambda \mapsto \lambda^\nu(m+n)$ is a bijection)).

Now, we can apply Exercise 2.91(d) to $\mu^\nu(m)$, $\mu^\nu(m)$, $\mu$ and $\mu$ instead of $\lambda$, $\mu$, $\lambda^\nu$ and $\mu^\nu$ (because of (11.545)). As a result, we obtain
\begin{equation}
s_{\mu^\nu(m)}(x_1, x_2, \ldots, x_k) = (x_1 x_2 \cdots x_k)^m \cdot s_{\mu^\nu(m)}(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}).
\end{equation}

Also, we can apply Exercise 2.91(d) to $n$, $\nu^\nu(n)$, $\nu^\nu(n)$, $\nu$ and $\nu$ instead of $m$, $\lambda$, $\mu$, $\lambda^\nu$ and $\mu^\nu$ (because of (11.546)). As a result, we obtain
\begin{equation}
s_{\nu^\nu(n)}(x_1, x_2, \ldots, x_k) = (x_1 x_2 \cdots x_k)^n \cdot s_{\nu^\nu(n)}(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}).
\end{equation}

Furthermore, every $\lambda \in \mathfrak{A}$ satisfies
\begin{equation}
s_{\lambda}(x_1, x_2, \ldots, x_k) = (x_1 x_2 \cdots x_k)^{m+n} \cdot s_{\lambda^\nu(m+n)}(x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}).
\end{equation}

481. This follows by substituting the variables $x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}$ for $x_1, x_2, \ldots, x_k$ in the equality (11.543).
Now, (11.543) yields
\[
\sum_{\lambda \in \mathfrak{A}} c_{\mu,\nu}^{\lambda} s_\lambda (x_1, x_2, \ldots, x_k) = s_\mu (x_1, x_2, \ldots, x_k) \cdot s_\nu (x_1, x_2, \ldots, x_k)
\]
\[
= (x_1 x_2 \cdots x_k)^{m-n} s_\mu (x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}) = (x_1 x_2 \cdots x_k)^{n-n} s_\nu (x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1})
\]
(by (11.548))
\[
= (x_1 x_2 \cdots x_k)^{m+n} \cdot s_\mu^{(m+n)} (x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1}) \cdot (x_1 x_2 \cdots x_k)^{n} s_\nu^{(n)} (x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1})
\]
\[
= \sum_{\lambda \in \mathfrak{A}} c_{\mu,\nu}^{\lambda v(m+n)} s_\lambda^{(m+n)} (x_1^{-1}, x_2^{-1}, \ldots, x_k^{-1})
\]
(by (11.547))
\[
= \sum_{\lambda \in \mathfrak{A}} c_{\mu,\nu}^{\lambda v(m+n)} s_\lambda (x_1, x_2, \ldots, x_k)
\]
(by (11.550))
\[
(11.551)
\]
But Remark 2.27(d) (applied to \( N = k \)) yields that the set \( \{ s_\lambda (x_1, x_2, \ldots, x_k) \} \), as \( \lambda \) runs through all partitions having length \( \leq k \), is a basis of the \( k \)-module \( \Lambda (x_1, x_2, \ldots, x_k) \). In other words, the family \( \{ s_\lambda (x_1, x_2, \ldots, x_k) \}_{\lambda \in \text{Par}; \ell(\lambda) \leq k} \) is a basis of the \( k \)-module \( \Lambda (x_1, x_2, \ldots, x_k) \). In particular, the family \( \{ s_\lambda (x_1, x_2, \ldots, x_k) \}_{\lambda \in \text{Par}; \ell(\lambda) \leq k} \) is \( k \)-linearly independent.

But the set \( \mathfrak{A} \) is a subset of \( \{ \alpha \in \text{Par} \mid \ell(\alpha) \leq k \} \). Hence, the family \( \{ s_\lambda (x_1, x_2, \ldots, x_k) \}_{\lambda \in \mathfrak{A}} \) is a subfamily of the family \( \{ s_\lambda (x_1, x_2, \ldots, x_k) \}_{\lambda \in \text{Par}; \ell(\lambda) \leq k} \). Since the family \( \{ s_\lambda (x_1, x_2, \ldots, x_k) \}_{\lambda \in \mathfrak{A}} \) is \( k \)-linearly independent, we thus conclude that its subfamily \( \{ s_\lambda (x_1, x_2, \ldots, x_k) \}_{\lambda \in \mathfrak{A}} \) is \( k \)-linearly independent. As a consequence, if two \( k \)-linear combinations of the family \( \{ s_\lambda (x_1, x_2, \ldots, x_k) \}_{\lambda \in \mathfrak{A}} \) are equal, then their respective coefficients must be equal. Thus, from (11.552), we conclude that
\[
(11.552)
\]
Now, let \( \lambda \) be a partition such that \( \ell(\lambda) \leq k \). Assume that all parts of \( \lambda \) are \( \leq m + n \). Then, \( \lambda \) is an element of \( \text{Par} \) such that \( \ell(\lambda) \leq k \) and such that all parts of \( \lambda \) are \( \leq m + n \). In other words, \( \lambda \in \mathfrak{A} \) (by the definition of \( \mathfrak{A} \)). Hence, (11.552) yields \( c_{\mu,\nu}^{\lambda} = c_{\mu,\nu}^{\lambda v(m+n)} \). This solves Exercise 2.92(c).
11.65. **Solution to Exercise 2.93. Solution to Exercise 2.93.** (a) We notice that every partition \( \lambda \) and every positive integer \( i \) satisfy

\[
\left( \lambda^i \right)_i = \left| \{ j \in \{1, 2, 3, \ldots \} \mid \lambda_j \geq i \} \right| \quad \text{(by (2.10))}
\]

(by (11.553), applied to \( \lambda = \mu \sqcup \nu \))

(11.553)

(because \( \lambda_j \geq i \) can happen only when \( j \in \{1, 2, \ldots, \ell(\lambda)\} \) (since \( i \) is positive)).

Let \( \mu \) and \( \nu \) be two partitions. For every positive integer \( i \), we have

\[
\left( (\mu \sqcup \nu)^i \right)_i = \left| \left\{ j \in \{1, 2, \ldots, \ell(\mu \sqcup \nu) \} \mid (\mu \sqcup \nu)_j \geq i \right\} \right| \quad \text{(by (11.553), applied to } \lambda = \mu \sqcup \nu \text{)}
\]

(11.554)

This solves Exercise 2.93(a).

(b) Let \( \mu \) and \( \nu \) be two partitions. We are going to prove that \( c^{\mu+\nu}_{\mu, \nu} = 1 \).

\( \left( \mu \sqcup \nu \right)^i \)-LR-tableau will mean a column-strict tableau \( T \) of shape \( (\mu + \nu)/\mu \) with \( \text{cont} (T) = \nu \) having the property that each cont \( (T)_{\text{cols} \geq j} \) is a partition (where we are using the notations of Corollary 2.58). Corollary 2.58 shows that \( c^{\mu+\nu}_{\mu, \nu} \) is the number of \( (\mu + \nu, \mu, \nu) \)-LR-tableaux. Hence, in order to prove that \( c^{\mu+\nu}_{\mu, \nu} = 1 \), it will be enough to show that there exists one and only one \( (\mu + \nu, \mu, \nu) \)-LR-tableau.

First of all, let \( T_0 \) be the filling of the skew shape \( (\mu + \nu)/\mu \) which assigns to every cell in row \( i \) the number \( i \), for all \( i \in \{1, 2, 3, \ldots \} \). This \( T_0 \) is clearly a column-strict tableau, and satisfies \( \text{cont} (T_0) = \nu \) (because for every positive integer \( i \), the \( i \)-th row of the skew shape \( (\mu + \nu)/\mu \) has \( (\mu + \nu)_i = \nu_i \) cells). Moreover, for every \( j \in \{1, 2, 3, \ldots \} \), the weak composition \( \text{cont} (T_{0|\text{cols} \geq j}) \) is a partition of \( 483 \). Therefore, \( T_0 \) is a \( (\mu + \nu, \mu, \nu) \)-LR-tableau. It remains to prove that it is the only \( (\mu + \nu, \mu, \nu) \)-LR-tableau.

\[=\mu_i+\nu_i\]

\[\begin{array}{c}
\frac{\mu + \nu}{\nu} \quad \text{and } \quad \text{cont} (T_{0|\text{cols} \geq j}) \quad \text{is a partition.qed.}
\end{array}\]
So fix any \((\mu + \nu, \mu, \nu)\)-LR-tableau \(T\). Thus, \(T\) is a column-strict tableau of shape \((\mu + \nu) / \mu\) with \(\text{cont}(T) = \nu\) having the property that each \(\text{cont}(T_{|\text{cols} \geq j})\) is a partition.

We shall show that

\[
(11.555) \quad \text{for every } i \in \{1, 2, 3, \ldots\}, \text{ all entries in the } i\text{-th row of } T \text{ are } \leq i.
\]

**Proof of (11.555):** Assume the contrary. Then, there exists an \(i \in \{1, 2, 3, \ldots\}\) such that not all entries in the \(i\)-th row of \(T\) are \(\leq i\). Let \(p\) be the smallest such \(i\) (this is clearly well-defined). Hence, not all entries in the \(p\)-th row of \(T\) are \(\leq p\). But since \(p\) is minimal, we know that \((11.555)\) holds for every \(i < p\).

Since not all entries in the \(p\)-th row of \(T\) are \(\leq p\), this shows that at least one entry of the \(p\)-th row of \(T\) must be \(> p\). Since the entries of \(T\) weakly increase along rows, this yields that the rightmost entry of the \(p\)-th row of \(T\) is \(> p\). Let \(q\) be this entry, and let \(j\) be the column in which the rightmost cell of the \(p\)-th row of \(T\) lies. Thus, \(q\) is the entry in cell \((p, j)\) of \(T\), and we have \(\text{cont}(T_{|\text{cols} \geq j})_q \geq 1\) (since \(q\) appears in \(T_{|\text{cols} \geq j}\)).

We know that \(\text{cont}(T_{|\text{cols} \geq j})\) is a partition, so that

\[
(\text{cont}(T_{|\text{cols} \geq j}))_p \geq \text{cont}(T_{|\text{cols} \geq j})_q \quad \text{(since } p < q) \geq 1.
\]

In other words, the entry \(p\) appears somewhere in the tableau \(T_{|\text{cols} \geq j}\). Where can it appear? It cannot appear in any of the first \(p - 1\) rows, because all entries in these rows are \(< p\) (since \((11.555)\) holds for every \(i < p\)). Hence, it must appear in the \(p\)-th row or further down. In other words, this entry \(p\) appears in a cell \((u, v)\) of \(T\) with \(u \geq p\) and \(v \geq j\). As a consequence, this entry \(p\) is \(\geq\) to the entry in cell \((p, j)\) of \(T\) (since \(T\) is column-strict). Since the entry in cell \((p, j)\) of \(T\) is \(q\), this yields \(p \geq q\), which contradicts \(q > p\). This contradiction shows that our assumption was wrong, and \((11.555)\) is proven.

We furthermore claim that

\[
(11.556) \quad \text{for every } i \in \{1, 2, 3, \ldots\}, \text{ all entries in the } i\text{-th row of } T \text{ are } i.
\]

**Proof of (11.556):** Assume the contrary. Then, there exists an \(i \in \{1, 2, 3, \ldots\}\) such that not all entries in the \(i\)-th row of \(T\) are \(i\). Let \(p\) be the smallest such \(i\) (this is clearly well-defined). Hence, not all entries in the \(p\)-th row of \(T\) are \(\leq p\). But since \(p\) is minimal, we know that \((11.556)\) holds for every \(i < p\).

All entries in the \(p\)-th row of \(T\) are \(\leq p\) (by \((11.555)\)), but not all of them are \(p\). Hence, the \(p\)-th row of \(T\) has an entry \(< p\). Let \(k\) be this entry. We can apply \((11.556)\) to \(k = i\) (since \((11.556)\) holds for every \(i < p\)), and conclude that all entries in the \(k\)-th row of \(T\) are \(k\). Thus, in the tableau \(T\), the entry \(k\) appears \(\nu_k\) times in row \(k\) (since the length of the \(k\)-th row of \(T\) is \((\mu + \nu)_k = (\mu_k + \nu_k) - \mu_k = \nu_k\) and at least \(1\) time in row \(p\) (by the definition of \(k\)). In total, \(k\) must thus appear at least \(\nu_k + 1\) times in \(T\), which contradicts the fact that \(\text{cont}(T) = \nu\). This contradiction disproves our assumption, and thus \((11.556)\) is proven.

Now that \((11.556)\) is proven, we immediately conclude that \(T = T_0\). Now, forget that we fixed \(T\). We thus have shown that every \((\mu + \nu, \mu, \nu)\)-LR-tableau \(T\) equals \(T_0\). Hence, there exists one and only one \((\mu + \nu, \mu, \nu)\)-LR-tableau, namely \(T = T_0\) (because we have already seen that \(T_0\) is a \((\mu + \nu, \mu, \nu)\)-LR-tableau).

As we said above, this proves that

\[
(11.557) \quad e_{\mu,\nu}^{\mu + \nu} = 1.
\]

It remains to show that \(e_{\mu,\nu}^{\mu + \nu} = 1\). Exercise 2.69(c) (applied to \(\lambda = \mu \cup \nu\)) shows that

\[
e_{\mu,\nu}^{\mu + \nu} = e_{\mu^t,\nu^t}^{\mu^t + \nu^t} = e_{\mu^t,\nu^t}^{\mu^t + \nu^t} \quad \text{(since } (\mu \cup \nu)^t = \mu^t + \nu^t \text{ (by Exercise 2.93(a)))} \]

\[
= 1 \quad \text{(by } (11.557)\text{), applied to } \mu^t \text{ and } \nu^t \text{ instead of } \mu \text{ and } \nu\).
\]

This solves Exercise 2.93(b).

(c) Let \(k \in \mathbb{N}\) and \(n \in \mathbb{N}\) satisfy \(k \leq n\), and let \(\mu \in \text{Par}_k\), \(\nu \in \text{Par}_{n-k}\) and \(\lambda \in \text{Par}_n\) be such that \(e_{\mu,\nu}^\lambda \neq 0\). We need to prove that \(\mu + \nu \triangleright \lambda \triangleright \mu \cup \nu\).

We have \(|\lambda| = n\), \(|\mu| = k\) and \(|\nu| = n - k\), so that \(|\lambda| = n = k + (n - k) = |\mu| + |\nu|\). We will first show that

\[
(11.558) \quad \mu + \nu \triangleright \lambda.
\]
A \((\lambda, \mu, \nu)\)-LR-tableau will mean a column-strict tableau \(T\) of shape \(\lambda/\mu\) with \(\text{cont}(T) = \nu\) having the property that each \(\text{cont}(T)_{\text{coh} \geq j}\) is a partition (where we are using the notations of Corollary 2.58). Corollary 2.58 shows that \(c_{\lambda, \mu, \nu}^\lambda\) is the number of \((\lambda, \mu, \nu)\)-LR-tableaux. Since \(c_{\lambda, \mu, \nu}^\lambda \neq 0\), we thus see that there exists at least one \((\lambda, \mu, \nu)\)-LR-tableau. Let this \((\lambda, \mu, \nu)\)-LR-tableau be \(T\).

Just as in the solution of Exercise 2.93(b), we can prove that (11.555) holds. Let \(k\) be a positive integer. Applying (11.555) to all \(i \in \{1, 2, \ldots, k\}\), we see that all entries in the first \(k\) rows of \(T\) (meaning the 1-st row, the 2-nd row, etc., the \(k\)-th row) are \(\leq k\). Hence,

\[
\text{(the number of all entries in the first} \ k \text{ rows of} \ T) \leq \text{(the number of all entries} \leq k \text{ in} \ T) = \sum_{i=1}^{k} (\text{the number of all entries} \ i \text{ in} \ T) = \sum_{i=1}^{k} \nu_i. 
\]

Since

\[
\text{(the number of all entries in the first} \ k \text{ rows of} \ T) = \sum_{i=1}^{k} (\text{the number of all entries in the} \ i \text{-th row of} \ T) = \sum_{i=1}^{k} (\lambda_i - \mu_i) = \sum_{i=1}^{k} \lambda_i - \sum_{i=1}^{k} \mu_i,
\]

this rewrites as \(\sum_{i=1}^{k} \lambda_i - \sum_{i=1}^{k} \mu_i \leq \sum_{i=1}^{k} \nu_i\). Hence,

\[
\sum_{i=1}^{k} \lambda_i \leq \sum_{i=1}^{k} \mu_i + \sum_{i=1}^{k} \nu_i = \sum_{i=1}^{k} (\mu_i + \nu_i) = \sum_{i=1}^{(\mu + \nu)_i},
\]

so that \(\sum_{i=1}^{k} (\mu + \nu)_i \geq \sum_{i=1}^{k} \lambda_i\). In other words, \((\mu + \nu)_1 + (\mu + \nu)_2 + \cdots + (\mu + \nu)_k \geq \lambda_1 + \lambda_2 + \cdots + \lambda_k\).

Now, let us forget that we fixed \(k\). We have shown that \((\mu + \nu)_1 + (\mu + \nu)_2 + \cdots + (\mu + \nu)_k \geq \lambda_1 + \lambda_2 + \cdots + \lambda_k\) for every positive integer \(k\). Combined with \(|\lambda| = |\mu| + |\nu| = |\mu + \nu|\), this yields that \(\mu + \nu \triangleright \lambda\). This proves (11.558).

It now remains to prove that \(\lambda \triangleright \mu \sqcup \nu\). To do so, we notice that Exercise 2.69(c) yields \(c_{\mu, \nu}^\lambda = c_{\nu, \mu}^\lambda\), so that \(c_{\mu, \nu}^\lambda = c_{\nu, \mu}^\lambda \neq 0\). Hence, we can apply (11.558) to \(\lambda', \mu'\) and \(\nu'\) instead of \(\lambda\), \(\mu\) and \(\nu\). As a result, we obtain \(\mu' + \nu' \triangleright \lambda'\). Since \((\mu \sqcup \nu)' = \mu' + \nu'\) (by Exercise 2.93(a)), this rewrites as \((\mu \sqcup \nu) \triangleright \lambda'\). But Exercise 2.16 (applied to \(\mu \sqcup \nu\) instead of \(\mu\)) yields that \(\lambda \triangleright \mu \sqcup \nu\) if and only if \((\mu \sqcup \nu) \triangleright \lambda'\). Since we already know that \((\mu \sqcup \nu)' \triangleright \lambda'\), we can thus conclude that \(\lambda \triangleright \mu \sqcup \nu\). Combined with (11.558), this yields \(\mu + \nu \triangleright \lambda \sqcup \mu \sqcup \nu\). This solves Exercise 2.93(c).

(d) Let \(n \in \mathbb{N}\) and \(m \in \mathbb{N}\) and \(\alpha, \beta \in \text{Par}_n\) and \(\gamma, \delta \in \text{Par}_m\) be such that \(\alpha \triangleright \beta\) and \(\gamma \triangleright \delta\). It is completely straightforward to check that (11.559)

\[
\alpha + \gamma \triangleright \beta + \delta
\]

It remains to prove that \(\alpha \sqcup \gamma \triangleright \beta \sqcup \delta\).

Exercise 2.16 (applied to \(\lambda = \alpha\) and \(\mu = \beta\)) yields that \(\alpha \triangleright \beta\) if and only if \(\beta' \triangleright \alpha'\). Since \(\alpha \triangleright \beta\), we thus have \(\beta' \triangleright \alpha'\). Hence, we can apply (11.559) to \(\beta', \alpha', \delta'\) and \(\gamma'\) instead of \(\alpha, \beta, \gamma\) and \(\delta\). As a result, we obtain \(\beta' + \delta' \triangleright \alpha' + \gamma'\). Now, (11.554) (applied to \(\mu = \beta\) and \(\nu = \delta\)) yields \((\beta \sqcup \delta)' = \beta' + \delta'\). Similarly, \((\alpha \sqcup \nu)' = \alpha' + \nu'\). Thus, \((\beta \sqcup \delta)' = \beta' + \delta' \triangleright \alpha' + \gamma' = (\alpha \sqcup \nu)'\).

Finally, Exercise 2.16 (applied to \(n + m\), \(\alpha \sqcup \nu\) and \(\beta \sqcup \delta\) instead of \(n\), \(\lambda\) and \(\mu\)) yields that \(\alpha \sqcup \gamma \triangleright \beta \sqcup \delta\) if and only if \((\beta \sqcup \delta)' \triangleright (\alpha \sqcup \nu)'\). Since we have \((\beta \sqcup \delta)' \triangleright (\alpha \sqcup \nu)'\), we thus obtain \(\alpha \sqcup \gamma \triangleright \beta \sqcup \delta\). This completes the solution of Exercise 2.93(d).

(e) The Ferrers diagram of the partition \(\lambda = (m^k)\) is a rectangle. Let \(C\) denote the center of this rectangle. For every partition \(\mu\) satisfying \(\mu \subseteq \lambda\), let us define a partition \(\mu^c\) by \(\mu^c = (m - \mu_k, m - \mu_{k-1}, \ldots, m - \mu_1)\). The Ferrers diagram of this partition \(\mu^c\) is obtained from the skew Ferrers diagram of the skew partition \(\lambda/\mu\) by the 180° rotation around \(C\). In other words, the skew Ferrers diagram of \(\mu^c/\emptyset\) is obtained from

\[\mu^c = (m - \mu_k, m - \mu_{k-1}, \ldots, m - \mu_1)\].
the skew Ferrers diagram of the skew partition \( \lambda/\mu \) by the 180° rotation around \( C \). Hence, Exercise 2.22(b) (applied to \( \lambda' = \mu^c \) and \( \mu' = \emptyset \)) yields that

\[
s_{\lambda/\mu} = s_{\mu^c/\emptyset} = s_{\mu^c}.
\]

Now, Proposition 2.24(iv) yields

\[
\Delta s_\lambda = \sum_{\mu \subseteq \lambda} s_\mu \otimes s_{\lambda/\mu} = \sum_{\mu \subseteq \lambda} s_\mu \otimes s_{\mu^c}.
\]

(by (11.560))

Compared with

\[
\Delta s_\lambda = \sum_{\mu,\nu} c_{\mu,\nu}^\lambda s_\mu \otimes s_\nu = \sum_{\mu,\nu} c_{\mu,\nu}^\lambda s_\mu \otimes s_\nu,
\]

this yields \( \sum_{\mu,\nu} c_{\mu,\nu}^\lambda s_\mu \otimes s_\nu = \sum_{\mu \subseteq \lambda} s_\mu \otimes s_{\mu^c} \). Comparing the coefficients in front of \( s_\mu \otimes s_\nu \) on both sides of this equality, we obtain: Any two partitions \( \mu \) and \( \nu \) satisfy

\[
c_{\mu,\nu}^\lambda = \begin{cases} 1, & \text{if } \mu \subseteq \lambda \text{ and } \nu = \mu^c; \\ 0, & \text{otherwise} \end{cases} \in \{0,1\}.
\]

This solves Exercise 2.93(e).

(f) We know that \((s_\mu)_{\mu \in \text{Par}}\) is a basis of the \( k \)-module \( \Lambda \). Hence, \((s_\mu \otimes s_\nu)_{\mu,\nu \in \text{Par}}\) is a basis of the \( k \)-module \( \Lambda \otimes \Lambda \). We will refer to this basis as the Schur basis of \( \Lambda \otimes \Lambda \).

We have

\[
\Delta s_\lambda = \sum_{\mu,\nu} c_{\mu,\nu}^\lambda s_\mu \otimes s_\nu = \sum_{\mu,\nu} c_{\mu,\nu}^\lambda s_\mu \otimes s_\nu.
\]

Thus, for every \( \mu,\nu \in \text{Par} \), the coefficient of \( s_\mu \otimes s_\nu \) with respect to the Schur basis of \( \Lambda \otimes \Lambda \) is \( c_{\mu,\nu}^\lambda \).

But \( \lambda = (a + 1,1^b) \). Hence, Exercise 2.90(d) gives a formula for \( \Delta s_\lambda = \Delta s_{(a+1,1^b)} \) as a sum of pure tensors of the form \( s_\mu \otimes s_\nu \) with \( \mu,\nu \in \text{Par} \). Every such pure tensor occurs at most once in this formula (as can be easily verified). In other words, for every \( \mu,\nu \in \text{Par} \), the coefficient of \( s_\mu \otimes s_\nu \) with respect to the Schur basis of \( \Lambda \otimes \Lambda \) is either 0 or 1. Since the coefficient of \( s_\mu \otimes s_\nu \) with respect to the Schur basis of \( \Lambda \otimes \Lambda \) is \( c_{\mu,\nu}^\lambda \), this rewrites as follows: For every \( \mu,\nu \in \text{Par} \), the scalar \( c_{\mu,\nu}^\lambda \) is either 0 or 1. This solves Exercise 2.93(f).

(g) The following solution is inspired by [161, proof of Thm. 2.1(iv)].

Before we start with the solution, we recall two formulas. Firstly, any two partitions \( \mu \) and \( \nu \) satisfy

\[
s_\mu s_\nu = \sum_{\lambda \in \text{Par}} c_{\mu,\nu}^\lambda s_\lambda \quad \text{(this is just a restatement of the first equality of (2.26))}
\]

(11.561)

\[
= \sum_{\tau \in \text{Par}} c_{\mu,\nu}^\tau s_\tau \quad \text{(here, we renamed the summation index } \lambda \text{ as } \tau \).
\]

Secondly, any two partitions \( \lambda \) and \( \mu \) satisfy

\[
s_{\lambda/\mu} = \sum_{\nu \in \text{Par}} c_{\mu,\nu}^\lambda s_\nu \quad \text{(this is one of the identities in Remark 2.37)}
\]

(11.562)

\[
= \sum_{\tau \in \text{Par}} c_{\mu,\nu}^\lambda s_\tau \quad \text{(here, we renamed the summation index } \nu \text{ as } \tau \).
\]

Now, let \( \lambda \) be any partition, and let \( \mu \) and \( \nu \) be two rectangular partitions. We need to show that \( c_{\mu,\nu}^\lambda \in \{0,1\} \).

Assume the contrary. Thus, \( c_{\mu,\nu}^\lambda \notin \{0,1\} \), so that \( c_{\mu,\nu}^\lambda \neq 0 \).

We have \( c_{\mu,\nu}^\lambda = c_{\nu,\mu}^\lambda \). Hence, we can WLOG assume that \( \ell(\mu) \geq \ell(\nu) \) (since otherwise, we can just switch \( \mu \) and \( \nu \)). Assume this. Assume WLOG that \( \mu \neq \emptyset \) (since otherwise, \( \mu = \emptyset \) and thus \( \nu = \emptyset \) (because \( \ell(\mu) \geq \ell(\nu) \)), which makes the claim \( c_{\mu,\nu}^\lambda \in \{0,1\} \) a rather obvious fact).
The partition \( \mu \) is rectangular, i.e., has the form \( (m^k) = \left( m, m, \ldots, m \right) \) for some \( m \in \mathbb{N} \) and \( k \in \mathbb{N} \).

Consider these \( m \) and \( k \). Both \( m \) and \( k \) are positive (since \( (m^k) = \mu \neq \emptyset \)). Thus, \( k = \ell(\mu) \) (since \( \mu = (m^k) \)), so that \( k = \ell(\mu) \geq \ell(\nu) \), thus \( \ell(\nu) \leq k \).

Corollary 2.58 shows that \( c^\lambda_{\mu,\nu} \) counts column-strict tableaux \( T \) of shape \( \lambda/\mu \) with cont \( (T) = \nu \) having the property that each cont \( (T) \) is a partition (where we are using the notations of Corollary 2.58). Since \( c^\lambda_{\mu,\nu} \neq 0 \), we see that there exists at least one such tableau \( T \). Consider this \( T \). All entries of the tableau \( T \) are \( \leq \ell(\nu) \) (because cont \( (T) = \nu \)). This quickly yields that \( \mu_k \geq \lambda_{k+1} \). Of course, we also have \( \mu \subseteq \lambda \) (since \( T \) is a tableau of shape \( \lambda/\mu \)), and thus \( \lambda_i \geq \mu_i \) for every \( i \in \{1, 2, 3, \ldots \} \). In particular, every \( i \in \{1, 2, \ldots, k\} \) satisfies \( \lambda_i \geq \mu_i = m \) (since \( \mu = (m^k) \)).

Now, define \( F, F_{\text{rows} \leq k} \) and \( F_{\text{rows} > k} \) as in Exercise 2.23. Define four partitions \( \alpha, \beta, \gamma \) and \( \delta \) by

\[
\alpha = (\lambda_1 - m, \lambda_2 - m, \ldots, \lambda_k - m), \quad \beta = \emptyset, \quad \gamma = (\lambda_{k+1}, \lambda_{k+2}, \lambda_{k+3}, \ldots), \quad \delta = \emptyset
\]

(notice that \( \alpha \) is well-defined because every \( i \in \{1, 2, \ldots, k\} \) satisfies \( \lambda_i \geq m \)). It is now easy to see that the skew Ferrers diagram \( \alpha/\beta \) can be obtained from \( F_{\text{rows} \leq k} \) by parallel translation (namely, the translation by \( m \) steps to the west), and that the skew Ferrers diagram \( \gamma/\delta \) can be obtained from \( F_{\text{rows} > k} \) by parallel translation (namely, the translation by \( k \) steps to the north). Hence, Exercise 2.23 yields

\[
s_{\lambda/\mu} = s_{\alpha/\beta} s_{\gamma/\delta} = s_{\alpha/\emptyset} s_{\gamma/\emptyset} = s_{\alpha} s_{\gamma} \quad \text{(since } \beta = \emptyset \text{ and } \delta = \emptyset)\]

\[
= s_{\alpha} s_{\gamma} = \sum_{\tau \in \text{Par}} c^\alpha_{\alpha,\gamma} s_{\tau} \quad \text{(by (11.561))}.
\]

Compared with (11.562), this yields \( \sum_{\tau \in \text{Par}} c^\lambda_{\mu,\tau} s_{\tau} = \sum_{\tau \in \text{Par}} c^\alpha_{\alpha,\gamma} s_{\tau} \). Since \( (s_{\tau})_{\tau \in \text{Par}} \) is a \( k \)-basis of \( \Lambda \), we can compare coefficients before \( s_{\nu} \) in this equality, and thus obtain \( c^\lambda_{\mu,\tau} = c^\alpha_{\alpha,\gamma} \). But \( \nu \) is a rectangular partition, and thus has the form \( \left( \bar{m}^k \right) \) for some \( \bar{m} \in \mathbb{N} \) and \( \bar{k} \in \mathbb{N} \). Hence, \( c^\alpha_{\alpha,\gamma} \in \{0, 1\} \) (by Exercise 2.93(e), applied to \( \bar{m}, \bar{k}, \nu, \alpha \) and \( \gamma \) instead of \( m, k, \lambda, \mu \) and \( \nu \)). Thus, \( c^\lambda_{\mu,\tau} = c^\nu_{\alpha,\gamma} \in \{0, 1\} \), which contradicts \( c^\lambda_{\mu,\tau} \notin \{0, 1\} \). This contradiction proves that our assumption was wrong, and Exercise 2.93(g) is solved.

11.66. Solution to Exercise 2.94. Solution to Exercise 2.94. (a) We shall prove the implications \( \mathcal{A} \implies \mathcal{B} \) and \( \mathcal{B} \implies \mathcal{A} \).

Proof of the implication \( \mathcal{A} \implies \mathcal{B} \): Assume that Assertion \( \mathcal{A} \) holds. That is, there exist a partition \( \lambda \) and a column-strict tableau \( T \) of shape \( \lambda/\mu \) such that all \( (i, j) \in \{1, 2, 3, \ldots \}^2 \) satisfy (2.52). Consider this \( \lambda \) and this \( T \). Since \( T \) is column-strict, the entries of \( T \) increase weakly left-to-right along rows, and increase strictly top-to-bottom along columns.

\[486\text{Proof}\] Assume the contrary. Thus, \( \mu_k < \lambda_{k+1} \). Since \( \mu = (m^k) \), we have \( \mu_k = m \), so that \( m = \mu_k < \lambda_{k+1} \). Thus, \( m \leq \lambda_{k+1} - 1 \) (since \( m \) and \( \lambda_{k+1} \) are integers), so that \( m + 1 \leq \lambda_{k+1} \). Now, for every \( i \in \{1, 2, \ldots, k+1\} \), we have

\[
\mu_i \leq m \quad \text{(since } \mu = (m^k)\}
\]

\[
< m + 1 \leq \lambda_{k+1} \leq \lambda_i \quad \text{(since } k + 1 \geq i\}
\]

Hence, \( (i, m + 1) \) is a cell of the skew Ferrers diagram \( \lambda/\mu \) for every \( i \in \{1, 2, \ldots, k + 1\} \). Altogether, the \( (m + 1) \)-th column of the skew Ferrers diagram \( \lambda/\mu \) contains at least \( k + 1 \) different cells (namely, the cells \( (i, m + 1) \) for all \( i \in \{1, 2, \ldots, k + 1\} \)). In the tableau \( T \), these \( k + 1 \) different cells must be filled with \( k + 1 \) distinct values (because the entries of \( T \) are strictly decreasing top-to-bottom in columns). As a consequence, there must be at least \( k + 1 \) distinct values among the entries of \( T \); but this is impossible, because all entries of \( T \) are \( \leq \ell(\nu) \leq k \). This contradiction proves that our assumption was wrong, qed.
For every $u \in \mathbb{N}$ and $j \in \{1, 2, 3, \ldots\}$, we have

\[
\sum_{i=1}^{u} \left(\text{the number of all entries } \leq u \text{ in the } j\text{-th row of } T\right) = \sum_{i=1}^{u} b_{i,j} \quad \text{(by (11.562))}
\]

\[
(11.563) \quad \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) \leq b_{1,j} + b_{2,j} + \cdots + b_{u,j}.
\]

Let $(i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}$. We are going to prove that

\[
(11.564) \quad \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) \leq \mu_{j} + (b_{1,j} + b_{2,j} + \cdots + b_{i,j}).
\]

Indeed, assume the contrary. Then,

\[
(11.565) \quad \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) > \mu_{j} + (b_{1,j} + b_{2,j} + \cdots + b_{i,j}).
\]

Hence,

\[
\mu_{j+1} + \left(\text{the number of all entries } \leq i+1 \text{ in the } (j+1)\text{-th row of } T\right) = \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) \geq b_{1,j} + b_{2,j} + \cdots + b_{u,j} \quad \text{(since } \mu \text{ is a partition)}.
\]

Subtracting $\mu_{j+1}$ from both sides of this inequality, we obtain

\[
(\text{the number of all entries } \leq i+1 \text{ in the } (j+1)\text{-th row of } T) > 0.
\]

In other words, there exists at least one entry $\leq i+1$ in the $(j+1)$-th row of $T$. Let $c$ be the rightmost cell of the $(j+1)$-th row of $T$ which contains such an entry. That is, $c$ is the rightmost cell of the $(j+1)$-th row of $T$ which contains an entry $\leq i+1$.

The cell $c$ lies in the $(j+1)$-th row of $T$. Hence, we can write the cell $c$ in the form $c = (j+1, y)$ for some positive integer $y$. Consider this $y$.

The entries of $T$ increase weakly left-to-right along rows. Thus, the cells of the $(j+1)$-th row of $T$ which contain entries $\leq i+1$ form a contiguous segment of the $(j+1)$-th row of $T$. This segment begins in cell $(j+1, \mu_{j+1}+1)$ (since the entries of the $(j+1)$-th row of $T$ begin in cell $(j+1, \mu_{j+1}+1)$ (because $T$ has shape $\lambda/\mu$)), and ends in cell $(j+1, y)$ (since $(j+1, y) = c$ is the rightmost cell of the $(j+1)$-th row of $T$ which contains an entry $\leq i+1$). Hence, this segment contains precisely $y - \mu_{j+1}$ cells. In other words, there exist exactly $y - \mu_{j+1}$ cells of the $(j+1)$-th row of $T$ which contain entries $\leq i+1$. In other words, the number of all cells of the $(j+1)$-th row of $T$ which contain entries $\leq i+1$ is $y - \mu_{j+1}$. In other words,

\[
(\text{the number of all entries } \leq i+1 \text{ in the } (j+1)\text{-th row of } T) = y - \mu_{j+1}.
\]

Solving this for $y$, we obtain

\[
y = \mu_{j+1} + (\text{the number of all entries } \leq i+1 \text{ in the } (j+1)\text{-th row of } T),
\]

Since $c$ is a cell of the tableau $T$, it is clear that the cell $c$ lies inside the Ferrers diagram of $\lambda$, and therefore (due to $j \in \{1, 2, 3, \ldots\}$) the cell $(j, y)$ must also lie inside the Ferrers diagram of $\lambda$ (because the cell $(j, y)$ is the northern neighbor of the cell $(j+1, y) = c$).

We have

\[
y = \mu_{j+1} + (\text{the number of all entries } \leq i+1 \text{ in the } (j+1)\text{-th row of } T) = \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) > \mu_{j}.
\]

Therefore, the cell $(j, y)$ lies outside the Ferrers diagram of $\mu$. Since $(j, y)$ lies inside the Ferrers diagram of $\lambda$ but outside the Ferrers diagram of $\mu$, we see that $(j, y)$ is a cell of $T$.

But recall that the entries of $T$ increase strictly top-to-bottom along columns. Hence, the entry of $T$ in the cell $(j, y)$ must be strictly smaller than the entry of $T$ in the cell $c$ (since the cell $(j, y)$ is the northern neighbor of the cell $(j+1, y) = c$). Since the latter entry is $\leq i+1$ (by the definition of $c$), this shows that the entry of $T$ in the cell $(j, y)$ must be strictly smaller than $i+1$. Hence, this entry must be $\leq i$. As a
consequence, all cells in the \( j \)-th row of \( T \) which lie weakly to the left of the cell \((j, y)\) must also have entries \( \leq i \) (because the entries of \( T \) increase weakly left-to-right along rows). The number of such cells is \( y - \mu_j \) (because the entries in the \( j \)-th row of \( T \) begin in cell \((j, \mu_j + 1)\) (since \( T \) has shape \( \lambda/\mu) \)). Thus, there are at least \( y - \mu_j \) cells in the \( j \)-th row of \( T \) which have entries \( \leq i \); in other words, the number of all entries \( \leq i \) in the \( j \)-th row of \( T \) is at least \( y - \mu_j \). In other words,

\[
(\text{the number of all entries } \leq i \text{ in the } j\text{-th row of } T) \geq y - \mu_j.
\]

Since

\[
(\text{the number of all entries } \leq i \text{ in the } j\text{-th row of } T) = b_{1,j} + b_{2,j} + \cdots + b_{i,j}
\]

(by (11.563), applied to \( u = i \)), this rewrites as follows:

\[
b_{1,j} + b_{2,j} + \cdots + b_{i,j} \geq y - \mu_j.
\]

Now, (11.565) becomes

\[
\mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) > \mu_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j}) \geq y - \mu_j
\]

\[
y = \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}).
\]

This is absurd. This contradiction proves that our assumption was wrong. Hence, (11.564) holds.

Now, forget that we have fixed \((i, j)\). We thus have proven that the inequality (11.564) holds for all \((i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}\). In other words, Assertion \( \mathcal{B} \) holds. We thus have proven the implication \( \mathcal{A} \implies \mathcal{B} \).

**Proof of the implication \( \mathcal{B} \implies \mathcal{A} \):** Assume that Assertion \( \mathcal{B} \) holds. That is, the inequality (2.53) holds for all \((i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}\).

For every \( j \in \{1, 2, 3, \ldots\} \), the sum \( b_{1,j} + b_{2,j} + b_{3,j} + \cdots \) has only finitely many nonzero addends (since \( b_{1,j} = 0 \) for all but finitely many pairs \((i, j)\) ), and can be computed as the following limit with respect to the discrete topology:

\[
(11.566) \quad b_{1,j} + b_{2,j} + b_{3,j} + \cdots = \lim_{i \to \infty} (b_{1,j} + b_{2,j} + \cdots + b_{i,j}).
\]

For every \( j \in \{1, 2, 3, \ldots\} \), we have

\[
(11.567) \quad \mu_j + (b_{1,j} + b_{2,j} + b_{3,j} + \cdots) \geq \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + b_{3,j+1} + \cdots)
\]

\(^{487}\) Hence, we can define a partition \( \lambda \) by

\[
(11.568) \quad (\lambda_j = \mu_j + (b_{1,j} + b_{2,j} + b_{3,j} + \cdots) \quad \text{for every } j \in \{1, 2, 3, \ldots\})
\]

\(^{487}\) **Proof.** Let \( j \in \{1, 2, 3, \ldots\} \). Then,

\[
\mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + b_{3,j+1} + \cdots)
\]

\[
= \lim_{i \to \infty} (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i,j+1})
\]

(by (11.566), applied to \( j+1 \) instead of \( j \))

\[
= \mu_{j+1} + \lim_{i \to \infty} (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i,j+1})
\]

(\( \text{here, we substituted } i+1 \text{ for } i \text{ in the limit} \))

\[
= \mu_{j+1} + \lim_{i \to \infty} (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) = \lim_{i \to \infty} \left( \mu_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) \right)
\]

\[
\leq \lim_{i \to \infty} (\mu_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j})) = \mu_j + \lim_{i \to \infty} (b_{1,j} + b_{2,j} + \cdots + b_{i,j}) = \mu_j + (b_{1,j} + b_{2,j} + b_{3,j} + \cdots),
\]

and this proves (11.567).
Consider this partition $\lambda$. We have $\mu \subseteq \lambda$ (since every $j \in \{1, 2, 3, \ldots\}$ satisfies $\lambda_j = \mu_j + (b_{1,j} + b_{2,j} + b_{3,j} + \cdots) \geq 0$)

We also have

$\lambda_j - \mu_j = b_{1,j} + b_{2,j} + b_{3,j} + \cdots$ for every $j \in \{1, 2, 3, \ldots\}$

(because of (11.568)).

Now, we construct a filling $T$ of the Ferrers diagram of $\lambda/\mu$ with positive integers as follows: For every $j \in \{1, 2, 3, \ldots\}$, the $j$-th row of this Ferrers diagram of $\lambda/\mu$ has $\lambda_j - \mu_j = b_{1,j} + b_{2,j} + b_{3,j} + \cdots$ cells. We fill in the lefmost $b_{1,j}$ of these cells with 1’s, the lefmost $b_{2,j}$ of the remaining cells with 2’s, the lefmost $b_{3,j}$ of the still remaining cells with 3’s, and so on. Once this has been done for all positive integers $j$ (of course, for all sufficiently high $j$, the $j$-th row of the Ferrers diagram of $\lambda/\mu$ has no cells, and therefore nothing has to be filled), we are left with a filling of the Ferrers diagram of $\lambda/\mu$ with positive integers. Denote this filling by $T$. It is clear that the entries of $T$ increase strictly top-to-bottom in rows (by the construction of $T$). We shall soon show that the entries of $T$ increase strictly top-to-bottom in columns.

First, however, let us observe that (2.52) holds for all $(i,j) \in \{1, 2, 3, \ldots\}$ (by the construction of $T$). Hence, every $u \in \mathbb{N}$ and $j \in \{1, 2, 3, \ldots\}$ satisfy (11.563) (this is proven just as in our proof of the implication $A \Rightarrow B$).

Now, we are going to prove that the entries of $T$ increase strictly top-to-bottom in columns.

Indeed, assume the contrary. Then, there exists at least one column of $T$ in which the entries don’t increase strictly top-to-bottom. Let this be the $k$-th column. So the entries in the $k$-th column of $T$ don’t increase strictly top-to-bottom. As a consequence, there exist two cells $c$ and $d$ in the $k$-th column of $T$ such that $d$ is the northern neighbor of $c$, but the entry of $T$ in cell $d$ is not smaller than the entry of $T$ in cell $c$. Consider these cells $c$ and $d$.

Write the cell $c$ as $c = (x,y)$. Then, $d = (x - 1, y)$ (since $d$ is the northern neighbor of $c$), so that $x - 1 \geq 1$.

Let $p$ be the entry of $T$ in cell $c$. Then, the entry of $T$ in cell $d$ is not smaller than $p$ (since the entry of $T$ in cell $d$ is not smaller than the entry of $T$ in cell $c$). In other words, the entry of $T$ in cell $(x - 1, y)$ is not smaller than $p$ (since $d = (x - 1, y)$). In other words,

$$\text{(the entry of } T \text{ in cell } (x-1,y) \geq p).$$

But let $i = p - 1$. Then, $i + 1 = p$ is the entry of $T$ in cell $c = (x,y)$. Thus, the cell $(x,y)$ of $T$ has the entry $i + 1$. Hence, all columns of $T$ which lie weakly to the left of the cell $(x,y)$ must have entries $\leq i + 1$ (since the entries of $T$ increase weakly left-to-right in rows). The number of such cells is $y - \mu_x$ (since the entries in the $x$-th row of $T$ begin in cell $(x,\mu_x + 1)$ (since $T$ has shape $\lambda/\mu$)). Thus, there are at least $y - \mu_x$ cells in the $x$-th row of $T$ which have entries $\leq i + 1$; in other words, the number of all entries $\leq i + 1$ in the $x$-th row of $T$ is at least $y - \mu_x$. In other words,

$$(\text{the number of all entries } \leq i + 1 \text{ in the } x\text{-th row of } T) \geq y - \mu_x.$$

Since (the number of all entries $\leq i + 1$ in the $x$-th row of $T) = b_{1,x} + b_{2,x} + \cdots + b_{i+1,x}$ (by (11.563), applied to $u = i + 1$ and $j = x$), this rewrites as follows:

$$b_{1,x} + b_{2,x} + \cdots + b_{i+1,x} \geq y - \mu_x.$$

Now, recall that $x - 1 \geq 0$, so that $x - 1 \in \{1, 2, 3, \ldots\}$. Hence, (2.53) (applied to $j = x - 1$) yields

$$\mu_{(x-1)+1} + (b_{1,(x-1)+1} + b_{2,(x-1)+1} + \cdots + b_{i+1,(x-1)+1}) \leq \mu_{x-1} + (b_{1,x-1} + b_{2,x-1} + \cdots + b_{i,x-1}),$$

so that

$$\mu_{x-1} + (b_{1,x-1} + b_{2,x-1} + \cdots + b_{i,x-1}) \geq \mu_{(x-1)+1} + (b_{1,(x-1)+1} + b_{2,(x-1)+1} + \cdots + b_{i+1,(x-1)+1})$$

$$= \mu_x = b_{1,x} + b_{2,x} + \cdots + b_{i+1,x} \geq y - \mu_x$$

$$= \mu_x + y - \mu_x \geq y.$$

[488] Here, we are using the fact that $\mu_j + (b_{1,j} + b_{2,j} + b_{3,j} + \cdots) = 0$ for all sufficiently high positive integers $j$. The proof of this is easy: We have $\mu_j = 0$ for all sufficiently high positive integers $j$ (since $\mu$ is a partition), and we also have $b_{1,j} + b_{2,j} + b_{3,j} + \cdots = 0$ for all sufficiently high positive integers $j$ (since $b_{1,j} = 0$ for all but finitely many pairs $(i,j)$). Thus, if a positive integer $j$ is sufficiently high, we have both $\mu_j = 0$ and $b_{1,j} + b_{2,j} + b_{3,j} + \cdots = 0$, and therefore $\mu_j + (b_{1,j} + b_{2,j} + b_{3,j} + \cdots) = 0 + 0 = 0$, qed.
But every entry $\leq i$ in the $(x-1)$-st row of $T$ must lie in a cell strictly left of the cell $(x-1, y)$ \footnote{Proof. Assume the contrary. Then, there exists an entry $\leq i$ in the $(x-1)$-st row of $T$ which lies in a cell not strictly left of the cell $(x-1, y)$. Let $c$ be this entry. Since $c$ lies in a cell not strictly left of the cell $(x-1, y)$, the entry $c$ must lie in a cell weakly to the right of the cell $(x-1, y)$, and therefore this entry $c$ must be $\geq x$ to the entry of $T$ in cell $(x-1, y)$ (because the entries of $T$ increase weakly left-to-right in rows). Hence, $c \geq (\text{the entry of } T \text{ in cell } (x-1, y)) \geq p$ (by (11.569)) $\geq i$ (since $i = p - 1 < p$), which contradicts the fact that $c \leq i$ (by definition of $c$). This contradiction shows that our assumption was wrong, qed.}. Since the number of such cells \footnote{Here, “such cells” means cells of the $(x-1)$-st row of $T$ which lie strictly left of the cell $(x-1, y)$.} is $y - 1 - \mu_{x-1}$ (because the entries in the $(x-1)$-st row of $T$ begin in cell $(x-1, \mu_{x-1} + 1)$ (since $T$ has shape $\lambda/\mu$)), this yields that there are at most $y - 1 - \mu_{x-1}$ entries $\leq i$ in the $(x-1)$-st row of $T$. In other words, the number of all entries $\leq i$ in the $(x-1)$-st row of $T$ is at most $y - 1 - \mu_{x-1}$. In other words,

\[(\text{the number of all entries } \leq i \text{ in the } (x-1)\text{-st row of } T) \leq y - 1 - \mu_{x-1}.
\]

Since (the number of all entries $\leq i$ in the $(x-1)$-st row of $T) = b_{1,x-1} + b_{2,x-1} + \cdots + b_{i,x-1}$ (by (11.563), applied to $u = i$ and $j = x - 1$), this rewrites as follows:

\[b_{1,x-1} + b_{2,x-1} + \cdots + b_{i,x-1} \leq y - \left\lfloor \frac{1}{\mu_{x-1}} \right\rfloor - \mu_{x-1} < y - \mu_{x-1}.
\]

Hence, $\mu_{x-1} + (b_{1,x-1} + b_{2,x-1} + \cdots + b_{i,x-1}) < y$. This contradicts (11.570). This contradiction shows that our assumption was wrong. Hence, the entries of $T$ increase strictly top-to-bottom in columns. Since we also know that the entries of $T$ increase weakly left-to-right in rows, and that $T$ is a filling of the Ferrers diagram of $\lambda/\mu$, this yields that $T$ is a column-strict tableau of shape $\lambda/\mu$. Besides, we already know that all $(i, j) \in \{1, 2, 3, \ldots\}^2$ satisfy (2.52). Hence, Assertion $A$ holds (with the $\lambda$ and $T$ that we have constructed above). We thus have proved the implication $B \Rightarrow A$.

Now that both implications $A \Rightarrow B$ and $B \Rightarrow A$ are proven, we conclude that Assertions $A$ and $B$ are equivalent. Exercise 2.94(a) is solved.

(b) Since $T$ is a column-strict tableau, we know that the entries of $T$ increase weakly left-to-right in rows and increase strictly top-to-bottom along columns. Hence, if $c$ and $d$ are two cells of $T$ such that the cell $c$ lies southeast \footnote{A cell $(r, c)$ is said to lie southeast of a cell $(r', c')$ if and only if we have $r \geq r'$ and $c \geq c'$.} of the cell $d$, then

\[(11.571) \quad (\text{the entry of } T \text{ in cell } c) \geq (\text{the entry of } T \text{ in cell } d).
\]

For the same reason, if $c$ and $d$ are two cells of $T$ such that the cell $c$ lies southeast of the cell $d$ but not on the same row as $d$, then

\[(11.572) \quad (\text{the entry of } T \text{ in cell } c) > (\text{the entry of } T \text{ in cell } d).
\]

We shall first prove the equivalence of Assertions $D$ and $G$:

\textbf{Proof of the equivalence $D \iff G$:} For every $(i, j) \in \{1, 2, 3, \ldots\}^2$, let $b_{i,j}$ be the number of all entries $j$ in the $i$-th row of $T$. (This is not a typo; we don’t want the number of all entries $i$ in the $j$-th row of $T$.)

It is clear that $b_{i,j} = 0$ for all but finitely many pairs $(i, j)$. Hence, we can apply Exercise 2.94(a) to $\emptyset$ instead of $\mu$. We conclude that the following two assertions $A'$ and $B'$ are equivalent:\footnote{Note that we denote by $\nu$ and $S$ the variables that have been called $\lambda$ and $T$ in Exercise 2.94(a), since in our current situation the letters $\lambda$ and $T$ already have different meanings.}

- **Assertion $A'$:** There exist a partition $\nu$ and a column-strict tableau $S$ of shape $\nu/\emptyset$ such that all $(i, j) \in \{1, 2, 3, \ldots\}^2$ satisfy $b_{i,j} = (\text{the number of all entries } i \text{ in the } j\text{-th row of } S)$.

- **Assertion $B'$:** The inequality

\[\emptyset_{j+1} + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) \leq \emptyset_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j})
\]

holds for all $(i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}$.
Now, it is easy to see that Assertion $\mathcal{A}'$ is equivalent to Assertion $\mathcal{G}$ \Footnote{Proof. Assertion $\mathcal{A}'$ is equivalent to the following Assertion $\mathcal{A}''$:}

- **Assertion $\mathcal{A}''$:** There exists a column-strict tableau $S$ whose shape is a partition such that all $(i, j) \in \{1, 2, 3, \ldots\}^2$ satisfy
  
  \[ b_{i,j} = (\text{the number of all entries } i \text{ in the } j\text{-th row of } S). \]

(Indeed, Assertion $\mathcal{A}'$ is equivalent to Assertion $\mathcal{A}''$ because a column-strict tableau whose shape is a partition is the same thing as a column-strict tableau of shape $\nu/\nu$ with $\nu$ being a partition.)

Recall that $b_{i,j} = (\text{the number of all entries } j \text{ in the } i\text{-th row of } T)$ for all $(i, j) \in \{1, 2, 3, \ldots\}^2$ (by the definition of $b_{i,j}$).

Hence, Assertion $\mathcal{A}''$ is equivalent to the following Assertion $\mathcal{A}'''$:

- **Assertion $\mathcal{A}'''$:** There exists a column-strict tableau $S$ whose shape is a partition such that all $(i, j) \in \{1, 2, 3, \ldots\}^2$ satisfy
  
  \[ (\text{the number of all entries } j \text{ in the } i\text{-th row of } T) = (\text{the number of all entries } i \text{ in the } j\text{-th row of } S). \]

Assertion $\mathcal{A}'''$ is equivalent to the following Assertion $\mathcal{A}''''$:

- **Assertion $\mathcal{A}''''$:** There exists a column-strict tableau $S$ whose shape is a partition such that all $(i, j) \in \{1, 2, 3, \ldots\}^2$ satisfy
  
  \[ (\text{the number of all entries } i \text{ in the } j\text{-th row of } T) = (\text{the number of all entries } j \text{ in the } i\text{-th row of } S). \]

(Indeed, Assertion $\mathcal{A}''''$ is obtained from Assertion $\mathcal{A}'''$ by substituting $(j, i)$ for the index $(i, j)$.)

Assertion $\mathcal{A}''''$ is obviously equivalent to Assertion $\mathcal{G}$.

Altogether, we have obtained the chain of equivalences $\mathcal{A} \iff \mathcal{A}' \iff \mathcal{A}'' \iff \mathcal{A}''' \iff \mathcal{A}''' \iff \mathcal{G}$. Thus, we know that Assertion $\mathcal{A}'$ is equivalent to Assertion $\mathcal{G}$, qed.

\Footnote{Proof. Every $(i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}$ satisfies}

\[ \sum_{u=1}^{\infty} (b_{u,j}) = \sum_{u=0}^{i} (\text{the number of all entries } j \text{ in the } u\text{-th row of } T) \]

(11.573)

and

\[ \sum_{u=1}^{\infty} (b_{u,j+1}) = \sum_{u=0}^{i} (\text{the number of all entries } j+1 \text{ in the first } i+1 \text{ rows of } T) \]

(11.574)

(by (11.573), applied to $(i, j+1)$ instead of $(i, j)$). Hence, Assertion $\mathcal{B}'$ is equivalent to the following Assertion $\mathcal{B}''$: \Footnote{This Assertion $\mathcal{B}''$, in turn, is equivalent to the following Assertion $\mathcal{B}'''$:}

- **Assertion $\mathcal{B}''$:** The inequality
  
  \[ (\text{the number of all entries } j+1 \text{ in the first } i+1 \text{ rows of } T) \leq (\text{the number of all entries } j \text{ in the first } i \text{ rows of } T) \]

  holds for all $(i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}$.

This Assertion $\mathcal{B}''$, in turn, is equivalent to the following Assertion $\mathcal{B}'''$: \Footnote{(Indeed, Assertion $\mathcal{B}'''$ is obtained from Assertion $\mathcal{B}''$ by substituting $(j-1, i)$ for the index $(i, j)$.) But Assertion $\mathcal{B}'''$ is clearly equivalent to Assertion $\mathcal{D}$.}

- **Assertion $\mathcal{B}'''$:** The inequality
  
  \[ (\text{the number of all entries } i+1 \text{ in the first } j \text{ rows of } T) \leq (\text{the number of all entries } i \text{ in the first } j-1 \text{ rows of } T) \]

  holds for all $(i, j) \in \{1, 2, 3, \ldots\} \times \{1, 2, 3, \ldots\}$.

We thus have found the chain of equivalences $\mathcal{B}' \iff \mathcal{B}'' \iff \mathcal{B}''' \iff \mathcal{D}$. Thus, Assertion $\mathcal{B}'$ is equivalent to Assertion $\mathcal{D}$, qed.
We have the following equivalences:

\[ C \iff (C_i \text{ holds for every } i \in \{1, 2, 3, \ldots\}) \]

(since \(\text{cont} (T|_{\text{cols} \geq j})\) is a partition if and only if every \(i \in \{1, 2, 3, \ldots\}\) satisfies \((\text{cont} (T|_{\text{cols} \geq j})))_i \geq (\text{cont} (T|_{\text{cols} \geq j})))_{i+1}\), and

\[ D \iff (D_i \text{ holds for every } i \in \{1, 2, 3, \ldots\}) \]

(obviously), and

\[ E \iff (E_i \text{ holds for every } i \in \{1, 2, 3, \ldots\}) \]

(since \(\text{cont} (T|_S)\) is a partition if and only if every \(i \in \{1, 2, 3, \ldots\}\) satisfies \((\text{cont} (T|_S)))_i \geq (\text{cont} (T|_S)))_{i+1}\), and

\[ F \iff (F_i \text{ holds for every } i \in \{1, 2, 3, \ldots\}) \]

(by the definition of a Yamanouchi word). Hence, in order to prove the equivalence \(C \iff D \iff E \iff F\), it is enough to show that for every \(i \in \{1, 2, 3, \ldots\}\), we have an equivalence \(C_i \iff D_i \iff E_i \iff F_i\). So let us do this now.

Let \(i \in \{1, 2, 3, \ldots\}\). We need to prove the equivalence \(C_i \iff D_i \iff E_i \iff F_i\). We shall achieve this by proving the implications \(C_i \implies E_i, E_i \implies F_i, F_i \implies D_i, D_i \implies C_i\).

**Proof of the implication** \(C_i \implies E_i\): Assume that Assertion \(C_i\) holds. Let \(S\) be an NE-set of \(T\). We will prove that \((\text{cont} (T|_S)))_i \geq (\text{cont} (T|_S)))_{i+1}\)

Assume the contrary. Thus, \((\text{cont} (T|_S)))_i < (\text{cont} (T|_S)))_{i+1}\). In other words, \((\text{cont} (T|_S)))_{i+1} > (\text{cont} (T|_S)))_i\).

Recall that

\[ (\text{cont} (T|_S)))_i = \left|(T|_S)^{-1} (i)\right| = (\text{the number of entries } i \text{ in } T|_S) \]

and

\[ (\text{cont} (T|_S)))_{i+1} = \left|(T|_S)^{-1} (i+1)\right| = (\text{the number of entries } i+1 \text{ in } T|_S) \]

Hence,

\[ (\text{the number of entries } i+1 \text{ in } T|_S) = (\text{cont} (T|_S)))_{i+1} > (\text{cont} (T|_S)))_i \]

\[ = (\text{the number of entries } i \text{ in } T|_S) \geq 0 \]

Hence, there exists at least one entry \(i + 1 \text{ in } T|_S\). In other words, there exists at least one cell \(c \in S\) such that the entry of \(T\) in \(c\) equals \(i + 1\). Let \(d\) be the **leftmost** such cell \(c\) (or one of the leftmost, if there are several of them\(^{495}\)). Thus, we have \(d \in S\), and the entry of \(T\) in \(d\) equals \(i + 1\). Let \(j\) be the column in which this cell \(d\) lies. Assertion \(C_i\) then yields \((\text{cont} (T|_{\text{cols} \geq j})))_i \geq (\text{cont} (T|_{\text{cols} \geq j})))_{i+1}\). It is rather clear that \((\text{cont} (T|_{\text{cols} \geq j})))_{i+1} \geq (\text{cont} (T|_S)))_{i+1}\).

\[^{495}\text{A moment’s thought reveals that there cannot be several of them, but we don’t actually need to think about this.}\]

\[^{496}\text{Proof. We have defined } d \text{ as the leftmost cell } c \in S \text{ such that the entry of } T \text{ in } c \text{ equals } i + 1. \text{ Let } d \text{ be the leftmost such cell } c \text{ (or one of the leftmost, if there are several of them). Thus, we have } d \in S, \text{ and the entry of } T \text{ in } d \text{ equals } i + 1. \text{ Let } j \text{ be the column in which this cell } d \text{ lies. Assertion } C_i \text{ then yields } (\text{cont} (T|_{\text{cols} \geq j})))_i \geq (\text{cont} (T|_{\text{cols} \geq j})))_{i+1}. \text{ It is rather clear that } (\text{cont} (T|_{\text{cols} \geq j})))_{i+1} \geq (\text{cont} (T|_S)))_{i+1}. \]

\[ \{c \in S \mid \text{the entry of } T \text{ in } c \text{ equals } i + 1\} \]

is a subset of the set

\[ \{c \in S \mid \text{the entry of } T \text{ in } c \text{ equals } i + 1\} \]

As a consequence,

\[ |\{c \in S \mid \text{the entry of } T \text{ in } c \text{ equals } i + 1\}| \leq |\{c \in S \mid \text{the entry of } T \text{ in } c \text{ equals } i + 1\}| \]

\[ = (\text{the number of entries } i + 1 \text{ in } T|_{\text{cols} \geq j}) = \left|(T|_{\text{cols} \geq j})^{-1} (i + 1)\right| = (\text{cont} (T|_{\text{cols} \geq j})))_{i+1} \]

(because \((\text{cont} (T|_{\text{cols} \geq j})))_{i+1} = \left|(T|_{\text{cols} \geq j})^{-1} (i + 1)\right| \) \(\text{by the definition of } \text{cont} (T|_{\text{cols} \geq j})\)). Since

\[ |\{c \in S \mid \text{the entry of } T \text{ in } c \text{ equals } i + 1\}| \]

\[ = (\text{the number of entries } i + 1 \text{ in } T|_S) = (\text{cont} (T|_S)))_{i+1} \]

this rewrites as \((\text{cont} (T|_S)))_{i+1} \leq (\text{cont} (T|_{\text{cols} \geq j})))_{i+1}. \) Hence, \((\text{cont} (T|_{\text{cols} \geq j})))_{i+1} \geq (\text{cont} (T|_S)))_{i+1}, \text{ qed.} \]
We shall now show that \((\text{cont}(T|_S))_i \geq (\text{cont}(T|_{\text{cols} \geq j}))_i\). Indeed, let \(c\) be a cell of \(T|_{\text{cols} \geq j}\) such that the entry of \(T\) in \(c\) equals \(i\). We shall show that \(c \in S\).

Write the cell \(d\) as \(d = (x, j)\) for some positive integer \(x\) (this is possible, since \(d\) lies in column \(j\)). Write the cell \(c\) as \(c = (x', y')\) for two positive integers \(x'\) and \(y'\). Then, \(y'\) is the column in which the cell \(c\) lies. Since this column is one of the columns \(j, j + 1, j + 2, \ldots\) (because \(c\) is a cell of \(T|_{\text{cols} \geq j}\)), this yields that \(y' \in \{j, j + 1, j + 2, \ldots\}\), so that \(y' \geq j\).

If \(x' \geq x\), then the cell \((x', y')\) lies southeast of the cell \((x, j)\) (since \(x' \geq x\) and \(y' \geq j\)). Since \((x', y') = c\) and \((x, j) = d\), this rewrites as: If \(x' \geq x\), then the cell \(c\) lies southeast of the cell \(d\). Hence, if \(x' \geq x\), then (11.571) yields

\[
\text{(the entry of } T \text{ in cell } c) \geq \text{(the entry of } T \text{ in cell } d) = i + 1,
\]

which contradicts (the entry of \(T\) in cell \(c\)) = \(i < i + 1\). Therefore, we cannot have \(x' \geq x\). We thus have \(x' < x\).

Since \(x' < x\) and \(y' \geq j\), the cell \((x', y')\) lies northeast of the cell \((x, j)\). In other words, the cell \(c\) lies northeast of the cell \(d\) (since \((x', y') = c\) and \((x, j) = d\)). Since \(c\) is a cell of \(T\) and since \(d \in S\), this yields that \(c \in S\) as well (since \(S\) is an NE-set).

Now, forget that we fixed \(c\). We thus have shown that if \(c\) is a cell of \(T|_{\text{cols} \geq j}\) such that the entry of \(T\) in \(c\) equals \(i\), then \(c \in S\). Hence, the set \[
\{c \text{ is a cell of } T|_{\text{cols} \geq j} \mid \text{the entry of } T \text{ in } c \text{ equals } i\}
\]
is a subset of the set \[
\{c \in S \mid \text{the entry of } T \text{ in } c \text{ equals } i\}.
\]

As a consequence,

\[
|\{c \text{ is a cell of } T|_{\text{cols} \geq j} \mid \text{the entry of } T \text{ in } c \text{ equals } i\}| \\
\leq |\{c \in S \mid \text{the entry of } T \text{ in } c \text{ equals } i\}| \\
= (\text{the number of entries } i \text{ in } T|_S) = (\text{cont}(T|_S))_i.
\]

Since

\[
|\{c \text{ is a cell of } T|_{\text{cols} \geq j} \mid \text{the entry of } T \text{ in } c \text{ equals } i\}| \\
= (\text{the number of entries } i \text{ in } T|_{\text{cols} \geq j}) = \left|(T|_{\text{cols} \geq j})^{-1}(i)\right| = (\text{cont}(T|_{\text{cols} \geq j}))_i
\]

(because the definition of \(\text{cont}(T|_{\text{cols} \geq j})\) shows that \(\left|(T|_{\text{cols} \geq j})^{-1}(i)\right| = \left|(T|_{\text{cols} \geq j})^{-1}(i)\right|\)), this rewrites as

\[
(\text{cont}(T|_{\text{cols} \geq j}))_i \leq (\text{cont}(T|_S))_i.\]

In other words, \((\text{cont}(T|_S))_i \geq (\text{cont}(T|_{\text{cols} \geq j}))_i\). Altogether,

\[
(\text{cont}(T|_S))_i \geq (\text{cont}(T|_{\text{cols} \geq j}))_i \geq (\text{cont}(T|_{\text{cols} \geq j}))_{i+1} \geq (\text{cont}(T|_S))_{i+1},
\]

which contradicts \((\text{cont}(T|_S))_i < (\text{cont}(T|_S))_{i+1}\). This contradiction shows that our assumption was wrong. Hence, \((\text{cont}(T|_S))_i \geq (\text{cont}(T|_S))_{i+1}\).

Now, forget that we fixed \(S\). We thus have proven that for every NE-set \(S\) of \(T\), we have \((\text{cont}(T|_S))_i \geq (\text{cont}(T|_S))_{i+1}\). Thus, Assertion \(E_i\) holds. This proves the implication \(E_i \Rightarrow F_i\).

Proof of the implication \(E_i \Rightarrow F_i\): Assume that Assertion \(E_i\) holds. Let \(v\) be a prefix of the Semitic reading word of \(T\). We shall prove that there are at least as many \(i\)'s among the letters of \(v\) as there are \((i + 1)'s\) among them.

For every \(i \in \{1, 2, 3, \ldots\}\), let \(r_i\) be the word obtained by reading the \(i\)-th row of \(T\) from right to left. Then, the Semitic reading word of \(T\) is the concatenation \(r_1r_2r_3\cdots\) (according to its definition). Hence, every prefix of this Semitic reading word must have the form \(r_{i_1}r_{i_2}\cdots r_{i_k}\) for some \(k \in \{0, 1, 2, \ldots\}\) and some prefix \(s\) of \(r_{i+k}\). In particular, \(v\) must have this form (since \(v\) is a prefix of the Semitic reading word of \(T\)). In other words, there exists some \(k \in \{0, 1, 2, \ldots\}\) and some prefix \(s\) of \(r_{i+k}\) such that \(v = r_1r_2\cdots r_{i+k}\). Consider this \(k\) and this \(s\).

Let \(\ell\) be the length of \(s\). Since \(s\) is a prefix of \(r_{i+k}\) and has length \(\ell\), it is evident that the word \(s\) consists of the first \(\ell\) letters of \(r_{i+k}\). These first \(\ell\) letters of \(r_{i+k}\) are the rightmost \(\ell\) entries of the \((k+1)\)-st row of \(T\) (since \(r_{i+k}\) is the word obtained by reading the \((k+1)\)-st row of \(T\) from right to left). Thus, the word \(s\) consists of the rightmost \(\ell\) entries of the \((k+1)\)-st row of \(T\). Hence, the word \(r_1r_2\cdots r_{i+k}\) consists of all entries of the first \(k\) rows of \(T\) and the rightmost \(\ell\) entries of the \((k+1)\)-st row of \(T\). In other words, the
letters of the word \( r_1 r_2 \cdots r_k \) are precisely all entries of the first \( k \) rows of \( T \) and the rightmost \( \ell \) entries of the \((k + 1)\)-st row of \( T \).

Let \( S \) be the set which consists of all cells of the first \( k \) rows of \( T \) and the rightmost \( \ell \) cells of the \((k + 1)\)-st row of \( T \). Then, \( S \) is an NE-set of \( T \), and therefore we have \((\text{cont} (T|S))_i \geq (\text{cont} (T|S))_{i+1} \) (by Assertion \( \mathcal{E}_i \)).

Now, \((\text{cont} (T|S))_i = |(T|S)^{-1} (i)| \) (by the definition of \( \text{cont} (T|S) \)), so that \((\text{cont} (T|S))_i \) is the number of entries \( i \) in \( T|S \). In other words, \((\text{cont} (T|S))_i \) is the number of \( i \)'s among the entries of \( T|S \). But since the entries of \( T|S \) are precisely the letters of \( v \), this rewrites as follows: \((\text{cont} (T|S))_i \) is the number of \( i \)'s among the letters of \( v \). So we have

\[
(\text{cont} (T|S))_i = \text{the number of } i \text{'s among the letters of } v \).
\]

The same argument, with \( i + 1 \) in place of \( i \), shows that

\[
(\text{cont} (T|S))_{i+1} = \text{the number of } (i + 1) \text{'s among the letters of } v \).
\]

Hence,

\[
\text{the number of } i \text{'s among the letters of } v = (\text{cont} (T|S))_i \geq (\text{cont} (T|S))_{i+1} = \text{the number of } (i + 1) \text{'s among the letters of } v \).
\]

In other words, there are at least as many \( i \)'s among the letters of \( v \) as there are \((i + 1)\)'s among them.

Now, forget that we fixed \( v \). We thus have shown that for every prefix \( v \) of the Semitic reading word of \( T \), there are at least as many \( i \)'s among the letters of \( v \) as there are \((i + 1)\)'s among them. In other words, Assertion \( \mathcal{F}_i \) holds. We thus have shown the implication \( \mathcal{E}_i \Rightarrow \mathcal{F}_i \).

**Proof of the implication \( \mathcal{F}_i \Rightarrow \mathcal{D}_i \):** Assume that Assertion \( \mathcal{F}_i \) holds.

For every \( i \in \{1, 2, 3, \ldots \} \), let \( r_i \) be the word obtained by reading the \( i \)-th row of \( T \) from right to left. Then, the Semitic reading word of \( T \) is the concatenation \( r_1 r_2 r_3 \cdots \) (according to its definition).

Now, let \( j \) be a positive integer. The word \( r_j \) is the word obtained by reading the \( j \)-th row of \( T \) from right to left. Hence, the letters of this word \( r_j \) are in (weakly) decreasing order (since the entries of \( T \) increase weakly left-to-right in rows). Thus, the subword of \( r_j \) consisting of all letters \( > i \) in \( r_j \) is a prefix of \( r_j \). Denote this prefix by \( s \). By the definition of \( s \), all letters of \( s \) are \( > i \), so that

\[
\text{(the number of } i \text{'s among the letters of } s \text{)} = 0.
\]

Also, by the definition of \( s \), the word \( s \) consists of all letters \( > i \) in \( r_j \). As a consequence,

\[
\text{(the number of } (i + 1) \text{'s among the letters of } s \text{)} = \text{(the number of } (i + 1) \text{'s among the letters of } r_j \text{)}.
\]

Since \( s \) is a prefix of \( r_j \), it is clear that the word \( r_1 r_2 \cdots r_{j-1} s \) is a prefix of the word \( r_1 r_2 r_3 \cdots \). In other words, the word \( r_1 r_2 \cdots r_{j-1} s \) is a prefix of the Semitic reading word of \( T \) (since the Semitic reading word of \( T \) is the concatenation \( r_1 r_2 r_3 \cdots \)). Hence, there are at least as many \( i \)'s among the letters of \( r_1 r_2 \cdots r_{j-1} s \) as there are \((i + 1)\)'s among them (by Assertion \( \mathcal{S}_i \), applied to \( v = r_1 r_2 \cdots r_{j-1} s \)). In other words,

\[
\text{(the number of } i \text{'s among the letters of } r_1 r_2 \cdots r_{j-1} s \text{)} \\
\geq \text{(the number of } (i + 1) \text{'s among the letters of } r_1 r_2 \cdots r_{j-1} s \text{)}.
\]

\(^{497}\text{Proof.}\) The set \( S \) consists of all cells of the first \( k \) rows of \( T \) and the rightmost \( \ell \) cells of the \((k + 1)\)-st row of \( T \). Hence, the entries of \( T|S \) are precisely all entries of the first \( k \) rows of \( T \) and the rightmost \( \ell \) entries of the \((k + 1)\)-st row of \( T \). Comparing this with the fact that the letters of the word \( r_1 r_2 \cdots r_k \) are precisely all entries of the first \( k \) rows of \( T \) and the rightmost \( \ell \) entries of the \((k + 1)\)-st row of \( T \), we conclude the following: The entries of \( T|S \) are precisely the letters of \( r_1 r_2 \cdots r_k \). In other words, the entries of \( T|S \) are precisely the letters of \( v \) (since \( v = r_1 r_2 \cdots r_k \)), qed.
Since
\[
\sum_{k=1}^{j-1} \left( \text{(the number of } i \text{'s among the letters of } r_1 r_2 \cdots r_{j-1} s \right) \]
= \sum_{k=1}^{j-1} \left( \text{(the number of } i \text{'s among the letters of } r_k \right) + \left( \text{(the number of } i \text{'s among the letters of } s \right) \]
= \sum_{k=1}^{j-1} \left( \text{(the number of } i \text{'s among the letters of } r_k \right) \]
= \sum_{k=1}^{j-1} \left( \text{(the number of entries } i \text{ in the } k \text{-th row of } T \right) \]
= \sum_{k=1}^{j-1} \left( \text{(the number of entries } i \text{ in the } k \text{-th row of } T \right) \]
and
\[
\sum_{k=1}^{j-1} \left( \text{(the number of } (i + 1) \text{'s among the letters of } r_1 r_2 \cdots r_{j-1} s \right) \]
= \sum_{k=1}^{j-1} \left( \text{(the number of } (i + 1) \text{'s among the letters of } r_k \right) + \left( \text{(the number of } (i + 1) \text{'s among the letters of } s \right) \]
= \sum_{k=1}^{j-1} \left( \text{(the number of } (i + 1) \text{'s among the letters of } r_k \right) + \left( \text{(the number of } (i + 1) \text{'s among the letters of } r_j \right) \]
= \sum_{k=1}^{j} \left( \text{(the number of } (i + 1) \text{'s among the letters of } r_k \right) \]
= \sum_{k=1}^{j} \left( \text{(the number of entries } i + 1 \text{ in the } k \text{-th row of } T \right) \]
= \text{(the number of entries } i + 1 \text{ in the first } j \text{ rows of } T), \]
this rewrites as follows:
\[
\text{(the number of entries } i \text{ in the first } j - 1 \text{ rows of } T) \]
\[\geq \text{(the number of entries } i + 1 \text{ in the first } j \text{ rows of } T). \]
In other words, the number of entries \(i + 1\) in the first \(j\) rows of \(T\) is \(\leq\) to the number of entries \(i\) in the first \(j - 1\) rows of \(T\).

Now, forget that we fixed \(j\). We thus have shown that for every positive integer \(j\), the number of entries \(i + 1\) in the first \(j\) rows of \(T\) is \(\leq\) to the number of entries \(i\) in the first \(j - 1\) rows of \(T\). In other words, Assertion \(D_i\) holds. Thus, we have shown the implication \(F_i \implies D_i\).

Proof of the implication \(D_i \implies C_i\): Assume that Assertion \(D_i\) holds. We want to prove that Assertion \(C_i\) holds. In other words, we want to prove that, for every positive integer \(j\), we have \((\text{cont } (T|_{\text{cols} \geq j}))_i \geq (\text{cont } (T|_{\text{cols} \geq j}))_{i+1}\).

In order to prove this, we assume the contrary. That is, there exists a positive integer \(j\) such that we don’t have \((\text{cont } (T|_{\text{cols} \geq j}))_i \geq (\text{cont } (T|_{\text{cols} \geq j}))_{i+1}\). Consider the minimal such \(j\). Then, we don’t have \((\text{cont } (T|_{\text{cols} \geq j}))_i \geq (\text{cont } (T|_{\text{cols} \geq j}))_{i+1}\); but
\[(11.575) \text{ for every } k \in \{1, 2, \ldots, j - 1\}, \text{ we do have } (\text{cont } (T|_{\text{cols} \geq k}))_i \geq (\text{cont } (T|_{\text{cols} \geq k}))_{i+1}. \]
Applying \((11.575)\) to \(k = j - 1\), we obtain:
\[(11.576) \text{ if } j - 1 \text{ is positive, then we do have } (\text{cont } (T|_{\text{cols} \geq j - 1}))_i \geq (\text{cont } (T|_{\text{cols} \geq j - 1}))_{i+1}. \]
We don’t have \((\text{cont } (T|_{\text{cols} \geq j}))_i \geq (\text{cont } (T|_{\text{cols} \geq j}))_{i+1}\). Thus, \[(11.577) (\text{cont } (T|_{\text{cols} \geq j}))_i < (\text{cont } (T|_{\text{cols} \geq j}))_{i+1}. \]
In other words, \((\text{cont} \, (T|_{\text{cols} \geq j}))_{i+1} > (\text{cont} \, (T|_{\text{cols} \geq j}))_i\). Recall that
\[
(\text{cont} \, (T|_{\text{cols} \geq j}))_i = (\text{the number of entries } i \text{ in } T|_{\text{cols} \geq j})
\]
and
\[
(\text{cont} \, (T|_{\text{cols} \geq j}))_{i+1} = (\text{the number of entries } i + 1 \text{ in } T|_{\text{cols} \geq j}).
\]

Hence,
\[
(\text{the number of entries } i + 1 \text{ in } T|_{\text{cols} \geq j}) = (\text{cont} \, (T|_{\text{cols} \geq j}))_{i+1} > (\text{cont} \, (T|_{\text{cols} \geq j}))_i = (\text{the number of entries } i \text{ in } T|_{\text{cols} \geq j}).
\]

Hence, there exists at least one entry \(i + 1\) in \(T|_{\text{cols} \geq j}\). In other words, there exists at least one cell \(c\) of \(T|_{\text{cols} \geq j}\) such that the entry of \(T|_{\text{cols} \geq j}\) in \(c\) equals \(i + 1\). In other words, there exists at least one cell \(c\) of \(T|_{\text{cols} \geq j}\) such that the entry of \(T\) in \(c\) equals \(i + 1\) (since the entry of \(T|_{\text{cols} \geq j}\) in \(c\), when it is defined, equals the entry of \(T\) in \(c\)). Let \(r\) be the \textbf{bottommost} row which contains such a cell \(c\), and let \((x, y)\) be the \textbf{leftmost} such cell \(c\) on this row. Thus, \((x, y)\) is a cell of \(T|_{\text{cols} \geq j}\) such that the entry of \(T\) in \((x, y)\) equals \(i + 1\). Also, the cell \((x, y)\) lies in the \(r\)-th row; that is, \(r = x\).

The cell \((x, y)\) must lie in one of the columns \(j, j + 1, j + 2, \ldots\) (since it is a cell of \(T|_{\text{cols} \geq j}\)), so that we have \(y \geq j\). Also, \((x, y)\) is a cell of \(T\) (since it is a cell of \(T|_{\text{cols} \geq j}\)).

Applying Assertion \(D_i\), to \(x\) instead of \(j\), we see that the number of entries \(i + 1\) in the first \(x\) rows of \(T\) is \(\leq\) to the number of entries \(i\) in the first \(x - 1\) rows of \(T\). In other words,
\[
(\text{the number of entries } i + 1 \text{ in the first } x \text{ rows of } T)\]
\[
\leq (\text{the number of entries } i \text{ in the first } x - 1 \text{ rows of } T).
\]

It is easy to see that
\[
(\text{cont} \, (T|_{\text{cols} \geq j}))_{i+1} \leq (\text{the number of entries } i + 1 \text{ in the first } x \text{ rows of } T)
\]

\[498\]

We shall now show that
\[
(\text{the number of entries } i \text{ in the first } x - 1 \text{ rows of } T) \leq (\text{cont} \, (T|_{\text{cols} \geq j}))_i.
\]

Indeed, let \(c\) be a cell of \(T\) such that the entry of \(T\) in \(c\) equals \(i\) and such that \(c\) lies in one of the first \(x - 1\) rows of \(T\). We shall show that \(c\) is a cell of \(T|_{\text{cols} \geq j}\).

Assume the contrary. Thus, \(c\) is not a cell of \(T|_{\text{cols} \geq j}\).

Let us write the cell \(c\) as \(c = (x', y')\) for some integers \(x'\) and \(y'\). Then, the cell \(c\) lies in row \(x'\), and thus this row \(x'\) must be one of the first \(x - 1\) rows of \(T\) (since we know that \(c\) lies in one of the first \(x - 1\) rows of \(T\)). In other words, \(x' \leq x - 1\). Hence, \(x' \leq x - 1 < x\) and thus \(x > x'\) and \(x \neq x'\) and \(x \geq x'\).

\[498\text{Proof.}\] Let \(c\) be a cell of \(T|_{\text{cols} \geq j}\) such that the entry of \(T\) in \(c\) equals \(i + 1\). Since \(r\) is the bottommost row which contains such a cell \(c\), it is clear that the cell \(c\) must be in the \(r\)-th row or in a row further north. In other words, the cell \(c\) must lie in one of the first \(r\) rows of \(T\). Since \(r = x\), this rewrites as follows: The cell \(c\) must lie in one of the first \(x\) rows of \(T\).

Now, let us forget that we fixed \(c\). We thus have proven that if \(c\) is a cell of \(T|_{\text{cols} \geq j}\) such that the entry of \(T\) in \(c\) equals \(i + 1\), then the cell \(c\) must lie in one of the first \(x\) rows of \(T\). Hence, the set
\[
\{c \text{ is a cell of } T|_{\text{cols} \geq j} \mid \text{the entry of } T \text{ in } c \text{ equals } i + 1\}
\]
is a subset of the set
\[
\{c \text{ is a cell of } T \mid \text{the entry of } T \text{ in } c \text{ equals } i + 1; \text{ the cell } c \text{ lies in one of the first } x \text{ rows of } T\}.
\]

Hence,
\[
|\{c \text{ is a cell of } T|_{\text{cols} \geq j} \mid \text{the entry of } T \text{ in } c \text{ equals } i + 1\}| \leq |\{c \text{ is a cell of } T \mid \text{the entry of } T \text{ in } c \text{ equals } i + 1; \text{ the cell } c \text{ lies in one of the first } x \text{ rows of } T\}|
\]
\[
= (\text{the number of entries } i + 1 \text{ in the first } x \text{ rows of } T).
\]

But
\[
(\text{cont} \, (T|_{\text{cols} \geq j}))_{i+1} = (\text{the number of entries } i + 1 \text{ in } T|_{\text{cols} \geq j})
\]
\[
\leq |\{c \text{ is a cell of } T|_{\text{cols} \geq j} \mid \text{the entry of } T \text{ in } c \text{ equals } i + 1\}|
\]
\[
\leq (\text{the number of entries } i + 1 \text{ in the first } x \text{ rows of } T),
\]

qed.
The cell $c$ lies in the $y'$-th column (since $c = (x', y')$). Hence, if $y' \geq j$, then the cell $c$ lies in one of the columns $j, j + 1, j + 2, \ldots$ and therefore is a cell of $T_{\text{cols} \geq j}$, which contradicts our assumption that $c$ is not a cell of $T_{\text{cols} \geq j}$. Thus, we cannot have $y' \geq j$. We therefore have $y' < j$. That is, $y' \leq j - 1$ (since $y'$ and $j$ are integers), so that $j - 1 \geq y'$.

Now, the cell $(x, y)$ lies southeast of the cell $(x, j - 1)$ (since $x \geq x$ and $y \geq j \geq j - 1$), which in turn lies southeast of the cell $(x', y')$ (since $x \geq x'$ and $j - 1 \geq y'$). Since both cells $(x, y)$ and $(x', y')$ are cells of $T$ (indeed, $(x, y)$ is known to be a cell of $T$, and $(x', y') = c$ is also a cell of $T$), this yields that the intermediate cell $(x, j - 1)$ is also a cell of $T$. 499 Thus, $j - 1$ is a positive integer. Moreover, since the cell $(x, y)$ lies southeast of the cell $(x, j - 1)$, we can apply (11.571) to $(x, y)$ and $(x, j - 1)$ instead of $c$ and $d$. We thus obtain

$$(\text{the entry of } T \text{ in cell } (x, y)) \geq (\text{the entry of } T \text{ in cell } (x, j - 1)).$$

Since (the entry of $T$ in cell $(x, y)) = i + 1$, this rewrites as

$$(11.582) \quad i + 1 \geq (\text{the entry of } T \text{ in cell } (x, j - 1)).$$

But the cell $(x, j - 1)$ lies southeast of the cell $(x', y')$ and not on the same row as $(x', y')$ (since $x \neq x'$). Since $(x', y') = c$, this rewrites as follows: The cell $(x, j - 1)$ lies southeast of the cell $c$ and not on the same row as $c$. Hence, (11.572) (applied to $(x, j - 1)$ and $c$ instead of $c$ and $d$) yields

$$(\text{the entry of } T \text{ in cell } (x, j - 1)) > (\text{the entry of } T \text{ in cell } c) = i,$$

so that

$$(\text{the entry of } T \text{ in cell } (x, j - 1)) \geq i + 1$$

(since the entry of $T$ in cell $(x, j - 1)$ and the number $i$ are integers). Combined with (11.582), this yields

$$(\text{the entry of } T \text{ in cell } (x, j - 1)) = i + 1.$$

Thus, the number $i + 1$ appears in the $(j - 1)$-th column of $T$ (since the cell $(x, j - 1)$ lies in the $(j - 1)$-th column of $T$). In other words,

$$(11.583) \quad (\text{the number of entries } i + 1 \text{ in the } (j - 1)-\text{th column of } T) \geq 1.$$

On the other hand, the entries of $T$ increase strictly top-to-bottom along columns. In particular, in every given column of $T$, the entries are pairwise distinct. Applied to the $(j - 1)$-th column, this shows that the entries of the $(j - 1)$-th column of $T$ are pairwise distinct. In particular, for every $k \in \{1, 2, 3, \ldots\}$, the number $k$ appears at most once in the $(j - 1)$-th column of $T$. Applying this to $k = i$, we conclude that the number $i$ appears at most once in the $(j - 1)$-th column of $T$. That is,

$$(11.584) \quad (\text{the number of entries } i \text{ in the } (j - 1)-\text{th column of } T) \leq 1.$$

Now, for every $k \in \{1, 2, 3, \ldots\}$ and $h \in \{1, 2, 3, \ldots\}$, we have

$$\left(\text{cont } (T_{\text{cols} \geq k})\right)_h = \left(\left(T_{\text{cols} \geq k}\right)^{-1}(h)\right) = (\text{the number of entries } h \text{ in } T_{\text{cols} \geq k}),$$

(11.585)

and

$$= (\text{the number of entries } h \text{ in the columns } k, k + 1, k + 2, \ldots \text{ of } T).$$

Applying this to $k = j$ and $h = i$, we obtain

$$(11.586) \quad (\text{cont } (T_{\text{cols} \geq j}))_i = (\text{the number of entries } i \text{ in the columns } j, j + 1, j + 2, \ldots \text{ of } T).$$

499 Here, we are using the following fact: If $\alpha$, $\beta$ and $\gamma$ are three cells such that $\alpha$ lies southeast of the cell $\beta$, which in turn lies southeast of the cell $\gamma$, and if $\alpha$ and $\gamma$ are cells of $T$, then $\beta$ is also a cell of $T$. This can be easily derived from the fact that $T$ has shape $\lambda/\mu$. 
But applying (11.585) to \( k = j - 1 \) and \( h = i \), we obtain

\[
\begin{align*}
\left( \text{cont} \left( T_{\text{cols} \geq j-1} \right) \right)_i &= \text{(the number of entries } i \text{ in the columns } j-1, j, j+1, \ldots \text{ of } T) \\
&= \left( \text{the number of entries } i \text{ in the } (j-1)\text{-th column of } T \right)_{\leq i} \\
&\quad + \left( \text{the number of entries } i \text{ in the columns } j, j+1, j+2, \ldots \text{ of } T \right)
\end{align*}
\]

(by (11.584))

\[
\leq 1 + \left( \text{cont} \left( T_{\text{cols} \geq j} \right) \right)_i < 1 + \left( \text{cont} \left( T_{\text{cols} \geq j} \right) \right)_{i+1}.
\]

(by (11.586))

\[
\left( \text{cont} \left( T_{\text{cols} \geq j-1} \right) \right)_i \leq \left( 1 + \left( \text{cont} \left( T_{\text{cols} \geq j} \right) \right)_{i+1} \right) - 1 = \left( \text{cont} \left( T_{\text{cols} \geq j} \right) \right)_{i+1},
\]

so that

\[
\left( \text{cont} \left( T_{\text{cols} \geq j} \right) \right)_{i+1} \geq \left( \text{cont} \left( T_{\text{cols} \geq j-1} \right) \right)_i \geq \left( \text{cont} \left( T_{\text{cols} \geq j-1} \right) \right)_{i+1}
\]

(by (11.576))

\[
= \left( \text{the number of entries } i + 1 \text{ in the columns } j - 1, j, j + 1, \ldots \text{ of } T \right)
\]

(by (11.585), applied to \( k = j - 1 \) and \( h = i + 1 \))

\[
= \left( \text{the number of entries } i + 1 \text{ in the } (j - 1)\text{-th column of } T \right)
\]

(by (11.583))

\[
+ \left( \text{the number of entries } i + 1 \text{ in the columns } j, j + 1, j + 2, \ldots \text{ of } T \right)
\]

(since \( \left( \text{cont} \left( T_{\text{cols} \geq j} \right) \right)_{i+1} = \left( \text{the number of entries } i+1 \text{ in the columns } j, j+1, j+2, \ldots \text{ of } T \right) \)

(by (11.585), applied to \( k = j \) and \( h = i+1 \))

\[
\geq 1 + \left( \text{cont} \left( T_{\text{cols} \geq j} \right) \right)_{i+1} > \left( \text{cont} \left( T_{\text{cols} \geq j} \right) \right)_{i+1},
\]

which is absurd. This contradiction proves that our assumption was wrong, and thus we have proven that \( c \) is a cell of \( T_{\text{cols} \geq j} \).

Now, let us forget that we fixed \( c \). We thus have proven that if \( c \) is a cell of \( T \) such that the entry of \( T \) in \( c \) equals \( i \) and such that \( c \) lies in one of the first \( x - 1 \) rows of \( T \), then \( c \) is a cell of \( T_{\text{cols} \geq j} \). Hence, the set

\[
\{ c \text{ is a cell of } T \mid \text{ the entry of } T \text{ in } c \text{ equals } i; \text{ the cell } c \text{ lies in one of the first } x - 1 \text{ rows of } T \}
\]

is a subset of the set

\[
\{ c \text{ is a cell of } T_{\text{cols} \geq j} \mid \text{ the entry of } T \text{ in } c \text{ equals } i \}.
\]

Hence,

\[
|\{ c \text{ is a cell of } T \mid \text{ the entry of } T \text{ in } c \text{ equals } i; \text{ the cell } c \text{ lies in one of the first } x - 1 \text{ rows of } T \}| \\
\leq |\{ c \text{ is a cell of } T_{\text{cols} \geq j} \mid \text{ the entry of } T \text{ in } c \text{ equals } i \}| \\
= \left( \text{the number of entries } i \text{ in } T_{\text{cols} \geq j} \right) = \left( \text{cont} \left( T_{\text{cols} \geq j} \right) \right)_i \\
(\text{by (11.578)})
\]

Since

\[
|\{ c \text{ is a cell of } T \mid \text{ the entry of } T \text{ in } c \text{ equals } i; \text{ the cell } c \text{ lies in one of the first } x - 1 \text{ rows of } T \}| \\
= \left( \text{the number of entries } i \text{ in the first } x - 1 \text{ rows of } T \right),
\]

this rewrites as

\[
\left( \text{the number of entries } i \text{ in the first } x - 1 \text{ rows of } T \right) \leq \left( \text{cont} \left( T_{\text{cols} \geq j} \right) \right)_i.
\]
Thus, (11.581) is proven. Now, (11.581) becomes

\[
\begin{aligned}
\text{(the number of entries } i \text{ in the first } x - 1 \text{ rows of } T) \\
\leq (\text{cont } (T|_{\text{cols} \geq j}))_{i+1} & < (\text{cont } (T|_{\text{cols} \geq j}))_{i+1} \quad \text{(by (11.577))} \\
\leq (\text{the number of entries } i + 1 \text{ in the first } x \text{ rows of } T) & \quad \text{(by (11.580))},
\end{aligned}
\]

which contradicts (11.579). This contradiction proves that our assumption was wrong. Hence, Assertion \(C_i\) must hold. We thus have proven the implication \(D_i \implies C_i\).

We have now proven the four implications \(C_i \implies E_i, E_i \implies F_i, F_i \implies D_i\) and \(D_i \implies C_i\). Combining them, we obtain the equivalence \(C_i \iff D_i \iff E_i \iff F_i\).

Now, let us forget that we fixed \(i\). We thus have proven the equivalence \(C_i \iff D_i \iff E_i \iff F_i\) for every \(i \in \{1, 2, 3, \ldots\}\). As we know, this shows that we have the equivalence \(C \iff D \iff E \iff F\), and this finishes the solution of Exercise 2.94(b).

11.67. Solution to Exercise 2.96. Solution to Exercise 2.96. (a) The solution to Exercise 2.96(a) is a rather straightforward adaptation of the solution of Exercise 2.94(b) that we gave above. We leave the details to the reader (who can also look them up in the LaTeX source code of this file, where they appear in a "commented out" environment starting after this sentence).

(b) Exercise 2.96(a) yields the equivalence \(C^{(\kappa)} \iff D^{(\kappa)} \iff E^{(\kappa)} \iff F^{(\kappa)} \iff G^{(\kappa)}\). It thus remains to prove the equivalence \(G^{(\kappa)} \iff H^{(\kappa)}\). In order to do so, we must prove the implications \(G^{(\kappa)} \implies H^{(\kappa)}\) and \(H^{(\kappa)} \implies G^{(\kappa)}\).

The implication \(H^{(\kappa)} \implies G^{(\kappa)}\) is obvious (since we can just take the \(S\) whose existence is guaranteed by Assertion \(H^{(\kappa)}\), and set \(\zeta = \tau\)). It thus remains to prove the implication \(G^{(\kappa)} \implies H^{(\kappa)}\).

Proof of the implication \(G^{(\kappa)} \implies H^{(\kappa)}\): Assume that Assertion \(G^{(\kappa)}\) holds. We want to prove that Assertion \(H^{(\kappa)}\) holds.

We have assumed that Assertion \(G^{(\kappa)}\) holds. In other words, there exist a partition \(\zeta\) and a column-strict tableau \(S\) of shape \(\zeta/\kappa\) which satisfies the following property: For any positive integers \(i\) and \(j\),

\[
(11.587) \quad \left( \begin{array}{c}
\text{the number of entries } i \text{ in the } j\text{-th row of } T \\
\text{the number of entries } j \text{ in the } i\text{-th row of } S
\end{array} \right) = k_i - k_i.
\]

Consider this \(\zeta\) and this \(S\).

For every \(i \in \{1, 2, 3, \ldots\}\), we have

\[
(\text{cont } T)_i = |T^{-1}(i)| \quad \text{(by the definition of cont } T) \\
= (\text{the number of entries } i \text{ in } T) \\
= \sum_{j=1}^{\infty} \left( \begin{array}{c}
\text{the number of entries } i \text{ in the } j\text{-th row of } T \\
\text{the number of entries } j \text{ in the } i\text{-th row of } S
\end{array} \right) \quad \text{(by (11.587))} \\
= \sum_{j=1}^{\infty} (\text{the number of entries } j \text{ in the } i\text{-th row of } S) \\
= (\text{the length of the } i\text{-th row of the skew partition } \zeta/\kappa) \\
= \zeta_i - k_i.
\]

Hence, for every \(i \in \{1, 2, 3, \ldots\}\), we have

\[
(\kappa + \text{cont } T)_i = \kappa_i + (\text{cont } T)_i = \kappa_i + (\zeta_i - k_i) = \zeta_i.
\]

In other words, \(\kappa + \text{cont } T = \zeta\), so that \(\zeta = \kappa + \text{cont } T = \tau\). But \(S\) is a column-strict tableau of shape \(\zeta/\kappa\). In other words, \(S\) is a column-strict tableau of shape \(\tau/\kappa\) (since \(\zeta = \tau\)).
We thus have constructed a column-strict tableau $S$ of shape $\tau/\kappa$ such that for any positive integers $i$ and $j$, the property (11.587) holds. Therefore, Assertion $H^{(\kappa)}$ holds. We thus have proven the implication $G^{(\kappa)} \implies H^{(\kappa)}$. As we have said, this finishes the solution of Exercise 2.96(b).

11.68. Solution to Exercise 2.97. Solution to Exercise 2.97. (a) If $T$ is a column-strict tableau of shape $\lambda/\mu$, then we have the following chain of logical equivalences:

$(\text{for all } j \in \{1, 2, 3, \ldots \}, \text{ the weak composition } \kappa + \text{cont}(T_{|\text{cols} \geq j}))$ is a partition
\[
\iff \left( \text{Assertion } C^{(\kappa)} \text{ holds} \right) \quad \left( \text{because this is how we defined Assertion } C^{(\kappa)} \right)
\]

(11.588)
\[
\iff \left( \text{the five equivalent assertions } C^{(\kappa)}, D^{(\kappa)}, E^{(\kappa)}, F^{(\kappa)} \text{ and } G^{(\kappa)} \text{ hold} \right)
\]

(because the five assertions $C^{(\kappa)}, D^{(\kappa)}, E^{(\kappa)}, F^{(\kappa)}$ and $G^{(\kappa)}$ are equivalent (by Exercise 2.96(a))). We can apply (2.34) to $\nu = \kappa$. As a result, we see that
\[
s_{\kappa}s_{\lambda/\mu} = \sum_{T} s_{\kappa+\text{cont } T},
\]
where $T$ runs through all column-strict tableaux of shape $\lambda/\mu$ with the property that for all $j = 1, 2, 3, \ldots$ one has $\kappa + \text{cont}(T_{|\text{cols} \geq j})$ a partition. In other words,

\[
s_{\kappa}s_{\lambda/\mu} = \sum_{\substack{T \text{ is a column-strict tableau of shape } \lambda/\mu; \\
\text{the five equivalent assertions } C^{(\kappa)}, D^{(\kappa)}, E^{(\kappa)}, F^{(\kappa)} \text{ and } G^{(\kappa)} \text{ hold}}}
\]
\[
= \sum_{T} s_{\kappa+\text{cont } T},
\]

(11.590)

In other words,
\[
s_{\kappa}s_{\lambda/\mu} = \sum_{T} s_{\kappa+\text{cont } T},
\]
where the sum ranges over all column-strict tableaux $T$ of shape $\lambda/\mu$ satisfying the five equivalent assertions $C^{(\kappa)}, D^{(\kappa)}, E^{(\kappa)}, F^{(\kappa)}$ and $G^{(\kappa)}$ introduced in Exercise 2.96(a). This solves Exercise 2.97(a).

(b) If $T$ is a column-strict tableau of shape $\lambda/\mu$ such that the five equivalent assertions $C^{(\kappa)}, D^{(\kappa)}, E^{(\kappa)}, F^{(\kappa)}$ and $G^{(\kappa)}$ hold, then we have
\[
(s_{\kappa+\text{cont } T}, s_{\tau})_{\Lambda} = \delta_{\tau, \kappa+\text{cont } T}
\]

500 Proof of (11.591): We know that the basis $(s_{\lambda})_{\lambda \in \text{Par}}$ of $\Lambda$ is orthonormal with respect to the Hall inner product. In other words, we have
\[
(s_{\alpha}, s_{\beta})_{\Lambda} = \delta_{\alpha, \beta}
\]

for any $\alpha \in \text{Par}$ and $\beta \in \text{Par}$.

Now, let $T$ be a column-strict tableau of shape $\lambda/\mu$ such that the five equivalent assertions $C^{(\kappa)}, D^{(\kappa)}, E^{(\kappa)}, F^{(\kappa)}$ and $G^{(\kappa)}$ hold. Then, Assertion $C^{(\kappa)}$ holds (since the five equivalent assertions $C^{(\kappa)}, D^{(\kappa)}, E^{(\kappa)}, F^{(\kappa)}$ and $G^{(\kappa)}$ hold). In other words, for every positive integer $j$, the weak composition $\kappa + \text{cont}(T_{|\text{cols} \geq j})$ is a partition. Applying this to $j = 1$, we conclude that the weak composition $\kappa + \text{cont}(T_{|\text{cols} \geq 1})$ is a partition. But $T_{|\text{cols} \geq 1}$ is the subtableau which is the restriction of $T$ to the union of its columns 1, 2, 3, \ldots. In other words, $T_{|\text{cols} \geq 1}$ is the whole tableau $T$. In other words, $T_{|\text{cols} \geq 1} = T$. 

500
But if $T$ is a column-strict tableau of shape $\lambda/\mu$ satisfying $\tau = \kappa + \text{cont } T$, then we have the following chain of logical equivalences:

$$
(\text{the five equivalent assertions } C^{(\kappa)}, D^{(\kappa)}, E^{(\kappa)}, F^{(\kappa)} \text{ and } G^{(\kappa)} \text{ hold})

\quad \iff \quad
(\text{the six equivalent assertions } C^{(\kappa)}, D^{(\kappa)}, E^{(\kappa)}, F^{(\kappa)}, G^{(\kappa)} \text{ and } H^{(\kappa)} \text{ hold})

$$

(11.593)

(because the six assertions $C^{(\kappa)}, D^{(\kappa)}, E^{(\kappa)}, F^{(\kappa)}, G^{(\kappa)}$ and $H^{(\kappa)}$ are equivalent (by Exercise 2.96(b))).

Now, Proposition 2.73(i) (applied to $A = A$) shows that every $a \in \Lambda$, $f \in \Lambda$ and $g \in \Lambda$ satisfy $(g, f^\perp(a))_\Lambda = (f, g(a))_\Lambda$. We can apply this to $f = s_\kappa$, $g = s_\lambda/\mu$ and $a = s_\tau$. As a result, we obtain $(s_\lambda/\mu, s_\kappa, s_\tau)_\Lambda = (s_\kappa s_\lambda/\mu, s_\tau)_\Lambda$. Since $s_\kappa$ $(s_\tau) = s_\tau/\kappa$, this rewrites as follows:

$$
(s_\lambda/\mu, s_\tau/\kappa)_\Lambda = (s_\kappa s_\lambda/\mu, s_\tau)_\Lambda.

$$

(11.594)

Thus,

$$
(s_\lambda/\mu, s_\tau/\kappa)_\Lambda = \left(\sum_{T \text{ is a column-strict tableau of shape } \lambda/\mu; \text{ the five equivalent assertions } C^{(\kappa)}, D^{(\kappa)}, E^{(\kappa)}, F^{(\kappa)} \text{ and } G^{(\kappa)} \text{ hold}} s_{\kappa + \text{cont } T, s_\tau}\right)_\Lambda = \left(\delta_{\tau, \kappa + \text{cont } T}\right)_\Lambda

$$

(by (11.590))

$$
= \sum_{T} \delta_{\tau, \kappa + \text{cont } T} + \sum_{T} \delta_{\tau, \kappa + \text{cont } T}

$$

$$
\text{(since the Hall inner product is } k\text{-bilinear)}

$$

$$
= \sum_{T} \delta_{\tau, \kappa + \text{cont } T}

$$

$$
\text{if } \tau = \kappa + \text{cont } T \quad \text{and} \quad \delta_{\tau, \kappa + \text{cont } T} = 1

$$

$$
= \sum_{T} 0

$$

$$
\text{if } \tau \neq \kappa + \text{cont } T

$$

$$
\sum_{T} 1

$$

Now, recall that $\kappa + \text{cont } (T|_{\text{col} \geq 1})$ is a partition. In other words, $\kappa + \text{cont } T$ is a partition (since $T|_{\text{col} \geq 1} = T$). In other words, $\kappa + \text{cont } T \in \text{Par}$. Also, $\tau \in \text{Par}$ (since $\tau$ is a partition). Thus, (11.592) (applied to $\alpha = \kappa + \text{cont } T$ and $\beta = \tau$) shows that $(s_{\kappa + \text{cont } T, s_\tau})_\Lambda = \delta_{\tau, \kappa + \text{cont } T}$. This proves (11.591).
\[ T \text{ is a column-strict tableau of shape } \lambda/\mu; \]
\[ \tau = \kappa + \text{cont } T; \]
the six equivalent assertions
\[ C(\kappa), D(\kappa), E(\kappa), F(\kappa), G(\kappa) \text{ and } H(\kappa) \text{ hold} \]
\[ = (\text{the number of all column-strict tableaux } T \text{ of shape } \lambda/\mu \text{ such that} \]
\[ \tau = \kappa + \text{cont } T \text{ and such that the six equivalent} \]
assertions \[ C(\kappa), D(\kappa), E(\kappa), F(\kappa), G(\kappa) \text{ and } H(\kappa) \text{ hold} \]
\[ = (\text{the number of all column-strict tableaux } T \text{ of shape } \lambda/\mu \text{ satisfying} \]
\[ \tau = \kappa + \text{cont } T \text{ and also satisfying the six equivalent} \]
assertions \[ C(\kappa), D(\kappa), E(\kappa), F(\kappa), G(\kappa) \text{ and } H(\kappa) \].

This solves Exercise 2.97(b).

11.69. **Solution to Exercise 2.98.** Solution to Exercise 2.98. (a) We shall first prove that

(11.595) if \( N \) has Jordan type \( \lambda \), then every \( k \in \mathbb{N} \) satisfies \( \dim(\ker(N^k)) = (\lambda^t_1 + \lambda^t_2 + \cdots + \lambda^t_k). \)

**Proof of (11.595):** Assume that \( N \) has Jordan type \( \lambda \). Let \( k \in \mathbb{N} \).
Since \( N \) has Jordan type \( \lambda \), the Jordan normal form of \( N \) has Jordan blocks of sizes \( \lambda_1, \lambda_2, \lambda_3, \ldots \). In other words, \( N \) is similar to the block-diagonal matrix \( J_\lambda := \begin{pmatrix} J_{\lambda_1} & 0 & 0 & \cdots & 0 \\ 0 & J_{\lambda_2} & 0 & \cdots & 0 \\ 0 & 0 & J_{\lambda_3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \), where \( J_m \) is the \( m \times m \) matrix \( \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \) for every \( m \in \mathbb{N} \). (This block-diagonal matrix \( J_\lambda \) is finite, since only finitely many \( \lambda_p \) are nonzero.) We can thus WLOG assume that \( N \) is this matrix \( J_\lambda \) (because replacing \( N \) by a matrix similar to \( N \) changes neither the dimension \( \dim(\ker(N^k)) \) nor the Jordan type of \( N \)). Assume this. Thus, \( N = J_\lambda = \begin{pmatrix} J_{\lambda_1} & 0 & 0 & \cdots \\ 0 & J_{\lambda_2} & 0 & \cdots \\ 0 & 0 & J_{\lambda_3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \), so that

\[ N^k = \begin{pmatrix} J_{\lambda_1} & 0 & 0 & \cdots \\ 0 & J_{\lambda_2} & 0 & \cdots \\ 0 & 0 & J_{\lambda_3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}^k = \begin{pmatrix} J_{\lambda_1}^k & 0 & 0 & \cdots \\ 0 & J_{\lambda_2}^k & 0 & \cdots \\ 0 & 0 & J_{\lambda_3}^k & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \]

and therefore

\[ \ker(N^k) = \ker \begin{pmatrix} J_{\lambda_1}^k & 0 & 0 & \cdots \\ 0 & J_{\lambda_2}^k & 0 & \cdots \\ 0 & 0 & J_{\lambda_3}^k & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \cong \bigoplus_{p \geq 1} \ker(J_{\lambda_p}^k), \]
whence
\[
(11.596) \quad \dim(\ker(N^k)) = \dim \left( \bigoplus_{p \geq 1} \ker\left(J^k_{\lambda_p}\right) \right) = \sum_{p \geq 1} \dim(\ker\left(J^k_{\lambda_p}\right)) = \sum_{p \geq 1} \min\{\lambda_p, k\}.
\]

We are now going to prove that the right hand side of this equality is \((\lambda^i)_1 + (\lambda^i)_2 + \cdots + (\lambda^i)_k\). Indeed, let us use the Iverson bracket notation: For every assertion \(A\), we let \([A]\) denote the integer
\[
\begin{cases} 
1, & \text{if } A \text{ is true;} \\
0, & \text{if } A \text{ is false}. 
\end{cases}
\]
This integer is called the truth value of \(A\).

Any two nonnegative integers \(u\) and \(v\) satisfy \(\min\{u, v\} = \sum_{i=1}^v [i \leq u]\). Thus, \(\min\{\lambda_p, k\} = \sum_{i=1}^k [i \leq \lambda_p]\) for every \(p \in \{1, 2, 3, \ldots\}\). Hence, (11.596) becomes
\[
\dim(\ker(N^k)) = \sum_{p \geq 1} \min\{\lambda_p, k\} = \sum_{p \geq 1} \sum_{i=1}^k [i \leq \lambda_p] = \sum_{p \geq 1} \sum_{i=1}^k [i \leq \lambda_p].
\]

This proves (11.595).

With (11.595), we have proven one direction of the equivalence that Exercise 2.98(a) requires us to prove. To prove the other direction, we need to show that
(11.597) if every \(k \in \mathbb{N}\) satisfies \(\dim(\ker(N^k)) = (\lambda^i)_1 + (\lambda^i)_2 + \cdots + (\lambda^i)_k\), then \(N\) has Jordan type \(\mu\).

Proof of (11.597): Assume that every \(k \in \mathbb{N}\) satisfies \(\dim(\ker(N^k)) = (\lambda^i)_1 + (\lambda^i)_2 + \cdots + (\lambda^i)_k\). Let \(\mu\) be the Jordan type of \(N\). Then, (11.595) (applied to \(\mu\) instead of \(\lambda\)) shows that every \(k \in \mathbb{N}\) satisfies \(\dim(\ker(N^k)) = (\mu^i)_1 + (\mu^i)_2 + \cdots + (\mu^i)_k\). Hence, every \(k \in \mathbb{N}\) satisfies
\[
(11.598) \quad (\mu^i)_1 + (\mu^i)_2 + \cdots + (\mu^i)_k = \dim(\ker(N^k)) = (\lambda^i)_1 + (\lambda^i)_2 + \cdots + (\lambda^i)_k.
\]
Applying (11.598) to \(k-1\) instead of \(k\), and subtracting the result from (11.598), we obtain
\[
(\mu^i)_k = (\lambda^i)_k \quad \text{for every positive integer } k.
\]

Hence, \(\mu^i = \lambda^i\), so that \(\mu = \lambda\). Thus, \(N\) has Jordan type \(\mu = \lambda\). This proves (11.597), and thus the solution of Exercise 2.98(a) is complete.

Before we come to the solution of Exercise 2.98(b), let us show a linear-algebraic lemma:

**Lemma 11.12.** Let \(W\) be a finite-dimensional \(K\)-vector space. Let \(A\) and \(B\) be \(K\)-vector subspaces of \(W\). Let \(f \in \operatorname{End} W\) be such that \(f(A) \subset A\) and \(f(B) \subset B\). For any \((i, j) \in \mathbb{N}^2\), let us define a nonnegative integer \(w_{i,j}\) by \(w_{i,j} = \dim\left(\left(f^j\right)^{-1}(A) \cap \left(f^j\right)^{-1}(B)\right)\).

(a) For any \((i, j) \in \{1, 2, 3, \ldots\}^2\), we have \(w_{i,j} + w_{i-1,j-1} \geq w_{i,j-1} + w_{i-1,j}\).
(b) For any \((i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}\), we have \(w_{i+1,j+1} - w_{i+1,j} \leq w_{i,j} - w_{i,j-1}\).
(c) For any \((i, j) \in \{1, 2, 3, \ldots\} \times \mathbb{N}\), we have \(w_{i+1,j+1} - w_{i,j+1} \leq w_{i,j} - w_{i-1,j}\).
(d) If \(i \in \{1, 2, 3, \ldots\}\) and \(j \in \mathbb{N}\) are such that \(i > \dim W\), then \(w_{i,j} = w_{i-1,j}\).
(e) If \(j \in \{1, 2, 3, \ldots\}\) and \(i \in \mathbb{N}\) are such that \(j > \dim W\), then \(w_{i,j} = w_{i,j-1}\).
(f) Assume that \(f\) is nilpotent. For every \(j \in \mathbb{N}\), we have \(\sum_{i=1}^\infty (w_{i,j} - w_{i-1,j}) = \dim\left(\left(f^j\right)^{-1}(B)\right) - \dim\left(\left(A \cap \left(f^j\right)^{-1}(B)\right)\right)\). (In particular, the sum \(\sum_{i=1}^\infty (w_{i,j} - w_{i-1,j})\) converges with respect to the discrete topology, i.e., all but finitely many of its terms are 0.)
(g) Assume that \( f \) is nilpotent. For every \( i \in \mathbb{N} \), we have
\[
\sum_{j=1}^{\infty} (w_{i,j} - w_{i,j-1}) = \dim \left( (f^i)^{-1} (A) \right) - \dim \left( (f^i)^{-1} (A) \cap B \right).
\]
(In particular, the sum \( \sum_{j=1}^{\infty} (w_{i,j} - w_{i,j-1}) \) converges with respect to the discrete topology, i.e., all but finitely many of its terms are 0.)

Proof of Lemma 11.12. Notice first that
\[
\left( \sum_{i=1}^{\infty} \right)^{-1} (A) = \text{id}^{-1} (A) = A \text{ and } A = (f^0)^{-1} (A) \subset (f^1)^{-1} (A) \subset (f^2)^{-1} (A) \subset \cdots \text{ (since } f (A) \subset A \text{ and } B = (f^0)^{-1} (B) \subset (f^1)^{-1} (B) \subset (f^2)^{-1} (B) \subset \cdots \text{ (similarly)).}
\]

(a) Let \((i, j) \in \{1, 2, 3, \ldots \}^2\). We need to prove that \( w_{i,j} + w_{i-1,j} - w_{i,j-1} \) converges with respect to the discrete topology. Due to (11.599),
\[
w_{i,j} - w_{i,j-1} = \dim \left( (f^i)^{-1} (A) \cap (f^j)^{-1} (B) \right) - \dim \left( (f^i)^{-1} (A) \cap (f^j)^{-1} (B) \right)
\]

(11.599)
\[
- \dim \left( \left( (f^i)^{-1} (A) \cap (f^j)^{-1} (B) \right) / \left( (f^i)^{-1} (A) \cap (f^j)^{-1} (B) \right) \right)
\]
\[
\text{(by the definitions of } w_{i,j} \text{ and } w_{i,j-1})
\]

and
\[
w_{i-1,j} - w_{i-1,j-1} = \dim \left( \left( (f^{i-1})^{-1} (A) \cap (f^j)^{-1} (B) \right) / \left( (f^{i-1})^{-1} (A) \cap (f^j)^{-1} (B) \right) \right)
\]

(by the same argument as (11.599), only with \( i - 1 \) instead of \( i \)), this is equivalent to showing that
\[
\dim \left( \left( (f^i)^{-1} (A) \cap (f^j)^{-1} (B) \right) / \left( (f^i)^{-1} (A) \cap (f^j)^{-1} (B) \right) \right)
\]
\[
\geq \dim \left( \left( (f^{i-1})^{-1} (A) \cap (f^j)^{-1} (B) \right) / \left( (f^{i-1})^{-1} (A) \cap (f^j)^{-1} (B) \right) \right).
\]

This will clearly be achieved if we can construct a \( \mathbb{K} \)-linear injection
\[
\left( (f^i)^{-1} (A) \cap (f^j)^{-1} (B) \right) / \left( (f^{i-1})^{-1} (A) \cap (f^j)^{-1} (B) \right)
\]
\[
\rightarrow \left( (f^i)^{-1} (A) \cap (f^j)^{-1} (B) \right) / \left( (f^i)^{-1} (A) \cap (f^{i-1})^{-1} (B) \right).
\]

Here is how to construct it: Let \( \iota \) denote the canonical inclusion \((f^i)^{-1} (A) \cap (f^j)^{-1} (B) \rightarrow (f^i)^{-1} (A) \cap (f^j)^{-1} (B)\). Then, \( \iota \) restricts to an inclusion \((f^{i-1})^{-1} (A) \cap (f^j)^{-1} (B) \rightarrow (f^{i-1})^{-1} (A) \cap (f^j)^{-1} (B)\), and so gives rise to a map
\[
\left( (f^{i-1})^{-1} (A) \cap (f^j)^{-1} (B) \right) / \left( (f^{i-1})^{-1} (A) \cap (f^j)^{-1} (B) \right)
\]
\[
\rightarrow \left( (f^i)^{-1} (A) \cap (f^j)^{-1} (B) \right) / \left( (f^i)^{-1} (A) \cap (f^{i-1})^{-1} (B) \right).
\]

This map is injective because \( \iota^{-1} \left( (f^{i-1})^{-1} (A) \cap (f^j)^{-1} (B) \right) = (f^{i-1})^{-1} (A) \cap (f^j)^{-1} (B) \), and thus we have found our \( \mathbb{K} \)-linear injection. Lemma 11.12(a) is proven.

(b) Let \((i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots \} \). We need to prove that \( w_{i+1,j+1} - w_{i+1,j} \leq w_{i,j} - w_{i,j-1} \). Due to (11.599) and due to
\[
w_{i+1,j+1} - w_{i+1,j} = \dim \left( \left( (f^{i+1})^{-1} (A) \cap (f^{j+1})^{-1} (B) \right) / \left( (f^{i+1})^{-1} (A) \cap (f^j)^{-1} (B) \right) \right)
\]

(this follows by the same arguments as (11.599), only with \( i \) and \( j \) replaced by \( i + 1 \) and \( j + 1 \)), this is equivalent to showing that
\[
\dim \left( \left( (f^{i+1})^{-1} (A) \cap (f^{j+1})^{-1} (B) \right) / \left( (f^{i+1})^{-1} (A) \cap (f^j)^{-1} (B) \right) \right)
\]
\[
\leq \dim \left( \left( (f^i)^{-1} (A) \cap (f^j)^{-1} (B) \right) / \left( (f^{i-1})^{-1} (A) \cap (f^{j-1})^{-1} (B) \right) \right).
\]
This will clearly be achieved if we can construct a \( K \)-linear injection
\[
\left( (f^{i+1})^{-1}(A) \cap (f^{i+1})^{-1}(B) \right) / \left( (f^{i+1})^{-1}(A) \cap (f^{j})^{-1}(B) \right)
\rightarrow \left( (f^{i})^{-1}(A) \cap (f^{j})^{-1}(B) \right) / \left( (f^{i})^{-1}(A) \cap (f^{j-1})^{-1}(B) \right).
\]

Here is how this can be done: The map \( f \) restricts to a map \( \varphi : (f^{i+1})^{-1}(A) \cap (f^{j+1})^{-1}(B) \rightarrow (f^{i})^{-1}(A) \cap (f^{j})^{-1}(B) \), which further restricts to a map \( (f^{i+1})^{-1}(A) \cap (f^{j})^{-1}(B) \rightarrow (f^{i})^{-1}(A) \cap (f^{j-1})^{-1}(B) \). Hence, \( \varphi \) gives rise to a map
\[
\left( (f^{i+1})^{-1}(A) \cap (f^{j+1})^{-1}(B) \right) / \left( (f^{i+1})^{-1}(A) \cap (f^{j})^{-1}(B) \right)
\rightarrow \left( (f^{i})^{-1}(A) \cap (f^{j-1})^{-1}(B) \right),
\]
which is injective because
\[
\varphi^{-1} \left( (f^{i})^{-1}(A) \cap (f^{j-1})^{-1}(B) \right) = f^{-1} \left( (f^{i})^{-1}(A) \cap (f^{j-1})^{-1}(B) \right)
\]
\[
\geq f^{-1} \left( (f^{i})^{-1}(A) \right) \cap f^{-1} \left( (f^{j-1})^{-1}(B) \right) = (f^{i+1})^{-1}(A) \cap (f^{j})^{-1}(B).
\]
Thus, we have found our \( K \)-linear injection. Lemma 11.12(b) is proven.

(c) The proof of Lemma 11.12(c) is analogous to our above proof of Lemma 11.12(b) (one merely has to interchange \( A \) with \( B \) and \( i \) with \( j \)).

(d) We have
\[
(f^{0})^{-1}(A) \subset (f^{1})^{-1}(A) \subset \cdots \subset (f^{\dim W+1})^{-1}(A)
\]
and therefore
\[
\dim \left((f^{0})^{-1}(A)\right) \leq \dim \left((f^{1})^{-1}(A)\right) \leq \cdots \leq \dim \left((f^{\dim W+1})^{-1}(A)\right).
\]
This latter chain of inequalities contains \( \dim W + 1 \) inequality signs, but only at most \( \dim W \) of them can be strict (because each \( \dim \left((f^{k})^{-1}(A)\right) \) is an integer between 0 and \( \dim W \) inclusive\(^{501}\), and a sequence of \( \dim W + 2 \) integers between 0 and \( \dim W \) cannot strictly increase). Thus, at least one of the inequality signs is an equality. That is, there exists an \( I \in \{1, 2, \ldots, \dim W + 1\} \) such that \( \dim \left((f^{I-1})^{-1}(A)\right) = \dim \left((f^{I})^{-1}(A)\right) \). Consider this \( I \).

Since \( \dim \left((f^{I-1})^{-1}(A)\right) = \dim \left((f^{I})^{-1}(A)\right) \) and \( (f^{I-1})^{-1}(A) \subset (f^{I})^{-1}(A) \), we must have
\[
(f^{I-1})^{-1}(A) = (f^{I})^{-1}(A).
\]
(11.600)

From here, it is easy to see that
\[
(f^{i-1})^{-1}(A) = (f^{i})^{-1}(A) \quad \text{for every } i \in \{1, 2, 3, \ldots\} \text{ satisfying } i > \dim W.
\]

\(^{501}\)Since \( (f^{k})^{-1}(A) \) is a subspace of the finite-dimensional \( K \)-vector space \( W \).

\(^{502}\)Proof. Let \( i \in \{1, 2, 3, \ldots\} \) be such that \( i > \dim W \). Thus, \( i \geq \dim W + 1 \). But \( I \leq \dim W + 1 \) (since \( I \in \{1, 2, \ldots, \dim W + 1\} \)) and thus \( i \geq \dim W + 1 \geq I \). Hence, there exists some \( k \in \mathbb{N} \) such that \( i = I + k \). Consider this \( k \). Since \( i = I + k \), we have \( i - 1 = I + k - 1 = (I - 1) + k \), so that \( f^{i-1} = f^{I+1} \circ f^{k} \) and thus \( (f^{i-1})^{-1}(A) = (f^{I+1} \circ f^{k})^{-1}(A) = (f^{k})^{-1} \left((f^{I-1})^{-1}(A)\right) \). Similarly, \( (f^{I-1})^{-1}(A) = (f^{I})^{-1} \left((f^{I-1})^{-1}(A)\right) \). Now,
\[
(f^{I-1})^{-1}(A) = (f^{I})^{-1} \left((f^{I-1})^{-1}(A)\right) = (f^{k})^{-1} \left((f^{I-1})^{-1}(A)\right).
\]

\[\text{qed.}\]
Thus, for every $i \in \{1, 2, 3, \ldots \}$ satisfying $i > \dim W$, we have

\[
w_{i-1,j} = \dim \left( \frac{(f^{i-1})^{-1}(A) \cap (f^j)^{-1}(B)}{= (f^i)^{-1}(A)} \right)
\]

(by the definition of $w_{i-1,j}$)

\[
= \dim \left( (f^i)^{-1}(A) \cap (f^j)^{-1}(B) \right) = w_{i,j} \quad \text{for all } j \in \mathbb{N}.
\]

Thus, if $i \in \{1, 2, 3, \ldots \}$ and $j \in \mathbb{N}$ are such that $i > \dim W$, then $w_{i-1,j} = w_{i-1,j}$. This proves Lemma 11.12(d).

(e) The proof of Lemma 11.12(e) is analogous to our above proof of Lemma 11.12(d) (one merely has to interchange $A$ with $B$ and $i$ with $j$).

(f) Let $j \in \mathbb{N}$. We have

\[
\sum_{i=1}^{\infty} (w_{i,j} - w_{i-1,j}) = \sum_{i=1}^{\dim W} (w_{i,j} - w_{i-1,j}) + \sum_{i=\dim W+1}^{\infty} \left( \underbrace{w_{i,j}}_{=w_{i-1,j}} - w_{i-1,j} \right)
\]

(by Lemma 11.12(d), since $i > \dim W$)

\[
= \sum_{i=1}^{\dim W} (w_{i,j} - w_{i-1,j}) + \sum_{i=\dim W+1}^{\infty} \left( \underbrace{w_{i,j}}_{=0} - w_{i-1,j} \right) = \sum_{i=1}^{\dim W} (w_{i,j} - w_{i-1,j}) + \sum_{i=\dim W+1}^{\infty} 0
\]

\[
\quad = \sum_{i=1}^{\dim W} (w_{i,j} - w_{i-1,j}) = w_{\dim W,j} - w_{0,j} \quad \text{(by the telescope principle)}.
\]

Now, we are going to prove that $w_{\dim W,j} = \dim \left( (f^j)^{-1}(B) \right)$ and $w_{0,j} = \dim \left( A \cap (f^j)^{-1}(B) \right)$.

Indeed, it is well-known that if $g$ is any nilpotent endomorphism of a finite-dimensional vector space $V$, then $g^{\dim V} = 0$. Applied to $g = f$ and $V = W$, this yields $f^{\dim W} = 0$, and thus $f^{\dim W}(W) = 0(W) = 0 \subset A$, whence $W \subset (f^{\dim W})^{-1}(A)$, so that $(f^{\dim W})^{-1}(A) = W$. Hence,

\[
(f^{\dim W})^{-1}(A) \cap (f^j)^{-1}(B) = W \cap (f^j)^{-1}(B) = (f^j)^{-1}(B)
\]

(since $(f^j)^{-1}(B) \subset W$). Now, the definition of $w_{\dim W,j}$ yields

\[
w_{\dim W,j} = \dim \left( \frac{(f^{\dim W})^{-1}(A) \cap (f^j)^{-1}(B)}{=(f^j)^{-1}(B)} \right) = \dim \left( (f^j)^{-1}(B) \right).
\]

Also, the definition of $w_{0,j}$ yields

\[
w_{0,j} = \dim \left( \frac{(f^0)^{-1}(A) \cap (f^j)^{-1}(B)}{= (f^j)^{-1}(B)} \right) = \dim \left( \frac{id^{-1}(A) \cap (f^j)^{-1}(B)}{= A} \right) = \dim \left( A \cap (f^j)^{-1}(B) \right).
\]

Hence, (11.601) becomes

\[
\sum_{i=1}^{\infty} (w_{i,j} - w_{i-1,j}) = w_{\dim W,j} - w_{0,j} = \dim \left( (f^j)^{-1}(B) \right) - \dim \left( A \cap (f^j)^{-1}(B) \right).
\]

This proves Lemma 11.12(f).

(g) The proof of Lemma 11.12(g) is analogous to our above proof of Lemma 11.12(f) (one merely has to interchange $A$ with $B$ and $i$ with $j$).
(b) Assume that $Z$ is a subring of $k$.
For any $(i,j) \in \mathbb{N}^2$, let us define a nonnegative integer $a_{i,j}$ by
\[
a_{i,j} = \dim \left( (f^j)^{-1} (U) \cap \ker (f^i) \right).
\]
Furthermore, for every $(i,j) \in \{1,2,3,\ldots\}^2$, let us define an integer $b_{i,j}$ by
\[
b_{i,j} = a_{i,j} - a_{i,j-1} - a_{i-1,j} + a_{i-1,j-1}.
\]
We first observe that
\[
a_{i,0} = 0 \quad \text{for every } i \in \mathbb{N}.
\]
Furthermore,
\[
a_{0,j} = (\mu^j)_1 + (\mu^j)_2 + \cdots + (\mu^j)_j \quad \text{for every } j \in \mathbb{N}.
\]
Thus,
\[
a_{0,j} - a_{0,j-1} = (\mu^j)_j \quad \text{for every } j \in \{1,2,3,\ldots\}.
\]
Moreover, every $(i,j) \in \mathbb{N}^2$ satisfies
\[
a_{i,j} = \dim \left( (f^j)^{-1} (U) \cap \ker (f^i) \right) = \dim \left( (f^j)^{-1} (U) \cap (f^j)^{-1} (0) \right).
\]

---

**Proof of (11.602):**
For every $i \in \mathbb{N}$, the definition of $a_{i,0}$ yields $a_{i,0} = \dim \left( (f^i)^{-1} (U) \cap \ker (f^0) \right) = 0$, qed.

**Proof of (11.603):**
Let $j \in \mathbb{N}$. We can represent the endomorphism $f | U$ of $U$ as a $k \times k$-matrix $G \in \mathbb{K}^{k \times k}$ for $k = \dim U$. This $k \times k$-matrix $G$ is nilpotent (since $f | U$ is nilpotent) and has Jordan type $\mu$ (since $f | U$ has Jordan type $\mu$). But Exercise 2.98(a) (applied to $G$ and $\mu$ instead of $n$, $N$ and $\lambda$) shows that $G$ has Jordan type $\mu$ if and only if every $k \in \mathbb{N}$ satisfies $\dim (\ker (G^k)) = (\mu^k)_1 + (\mu^k)_2 + \cdots + (\mu^k)_k$. Since we know that $G$ has Jordan type $\mu$, we thus conclude that every $k \in \mathbb{N}$ satisfies $\dim (\ker (G^k)) = (\mu^k)_1 + (\mu^k)_2 + \cdots + (\mu^k)_k$. Applied to $k = j$, this yields $\dim (\ker (G^j)) = (\mu^j)_1 + (\mu^j)_2 + \cdots + (\mu^j)_j$.

Since the matrix $G$ represents the endomorphism $f | U$, we have
\[
\dim (\ker (G^j)) = \dim \left( \ker (f | U^j) \right) = \dim (U \cap \ker (f^j)).
\]
Compared with $\dim (\ker (G^j)) = (\mu^j)_1 + (\mu^j)_2 + \cdots + (\mu^j)_j$, this yields
\[
\dim (U \cap \ker (f^j)) = (\mu^j)_1 + (\mu^j)_2 + \cdots + (\mu^j)_j.
\]
Now, the definition of $a_{0,j}$ yields
\[
a_{0,j} = \dim \left( \left( \left( f^0 \right)^{-1} \right) (U) \cap \ker (f^j) \right) = \dim \left( \left( \left( f \right)^{-1} \right) (U) \cap \ker (f^j) \right) = \dim (U \cap \ker (f^j)) = (\mu^j)_1 + (\mu^j)_2 + \cdots + (\mu^j)_j.
\]
This proves (11.603).

**Proof of (11.604):**
For every $j \in \{1,2,3,\ldots\}$, we have
\[
a_{0,j} - a_{0,j-1} = \left( (\mu^j)_1 + (\mu^j)_2 + \cdots + (\mu^j)_j \right) - \left( (\mu^j)_1 + (\mu^j)_2 + \cdots + (\mu^j)_{j-1} \right) = (\mu^j)_j,
\]
qed.
Hence, Lemma 11.12(a) (applied to \(W = V\), \(A = U\), \(B = 0\) and \(w_{i,j} = a_{i,j}\)) shows that for any \((i, j) \in \{1, 2, 3, \ldots\}^2\), we have \(a_{i,j} + a_{i-1,j-1} \geq a_{i,j-1} + a_{i-1,j}\). In other words, for any \((i, j) \in \{1, 2, 3, \ldots\}^2\), we have \(a_{i,j} - a_{i,j-1} - a_{i-1,j} + a_{i-1,j-1} \geq 0\). In other words, for any \((i, j) \in \{1, 2, 3, \ldots\}^2\), we have

\[
b_{i,j} \geq 0
\]

(since \(b_{i,j} = a_{i,j} - a_{i,j-1} - a_{i-1,j} + a_{i-1,j-1}\)). Thus, \(b_{i,j}\) is a nonnegative integer for every two positive integers \(i\) and \(j\). Moreover,

\[
b_{i,j} = 0 \quad \text{for all but finitely many pairs } (i, j).
\]

Also,

\[
\left(\mu^t\right)_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j}) = a_{i,j} - a_{i,j-1} \quad \text{for every } (i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}.
\]

Now, Exercise 2.94(a) (applied to \(\mu^t\) instead of \(\mu\)) yields that the following two assertions are equivalent:

- **Assertion \(A'\):** There exist a partition \(\gamma\) and a column-strict tableau \(T\) of shape \(\gamma/\mu^t\) such that all \((i, j) \in \{1, 2, 3, \ldots\}^2\) satisfy

\[
b_{i,j} = (\text{the number of all entries } i \text{ in the } j\text{-th row of } T).
\]

- **Assertion \(B'\):** The inequality

\[
\left(\mu^t\right)_j + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) \leq \left(\mu^t\right)_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j})
\]

holds for all \((i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}\).

We shall now prove that Assertion \(B'\) holds. Indeed, any \((i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}\) satisfies

\[
\left(\mu^t\right)_j + (b_{1,j+1} + b_{2,j+1} + \cdots + b_{i+1,j+1}) = a_{i+1,j+1} - a_{i+1,j} \quad (\text{by } (11.608), \text{applied to } i+1 \text{ and } j+1 \text{ instead of } i \text{ and } j)
\]

\[
\leq a_{i,j+1} - a_{i,j} - a_{i-1,j+1} + a_{i-1,j} \quad (\text{by Lemma 11.12(b) (applied to } W = V, A = U, B = 0 \text{ and } w_{i,j} = a_{i,j}))
\]

\[
= \left(\mu^t\right)_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j}) \quad (\text{by } (11.608)).
\]

**Proof of (11.606):** Lemma 11.12(d) (applied to \(W = V, A = U, B = 0\) and \(w_{i,j} = a_{i,j}\)) shows that if \(i \in \{1, 2, 3, \ldots\}\) and \(j \in \mathbb{N}\) are such that \(i > \dim V\), then

\[
a_{i,j} = a_{i-1,j}.
\]

Lemma 11.12(e) (applied to \(W = V, A = U, B = 0\) and \(w_{i,j} = a_{i,j}\)) shows that if \(j \in \{1, 2, 3, \ldots\}\) and \(i \in \mathbb{N}\) satisfy \(j > \dim V\), then \(a_{i,j} = a_{i,j-1}\).

Now, let \((i, j) \in \{1, 2, 3, \ldots\}^2\). If \(i > \dim V\), then

\[
b_{i,j} = \frac{a_{i,j}}{a_{i-1,j}} - \frac{a_{i,j-1}}{a_{i-1,j-1}} = \frac{a_{i,j} - a_{i-1,j} - a_{i-1,j-1} = a_{i,j-1} - a_{i-1,j-1} - a_{i-1,j} + a_{i-1,j-1} = 0}.
\]

Similarly, \(b_{i,j} = 0\) if \(j > \dim V\). Hence, we see that \(b_{i,j} = 0\) if we have \(i > \dim V\) or \(j > \dim V\) (or both). Thus, \(b_{i,j} = 0\) for all but finitely many pairs \((i, j)\) (because all but finitely many pairs \((i, j)\) satisfy \(i > \dim V\) or \(j > \dim V\)). This proves (11.606).

**Proof of (11.608):** Let \((i, j) \in \mathbb{N} \times \{1, 2, 3, \ldots\}\). Then,

\[
b_{1,j} + b_{2,j} + \cdots + b_{i,j} = \sum_{k=1}^{i} b_{k,j} = \sum_{k=1}^{i} \frac{a_{k,j} - a_{k,j-1} - a_{k-1,j} + a_{k-1,j-1}}{(a_{k,j} - a_{k-1,j}) - (a_{k-1,j} - a_{k-1,j-1})}\]

\[
= \sum_{k=1}^{i} \frac{(a_{k,j} - a_{k,j-1}) - (a_{k-1,j} - a_{k-1,j-1})}{(a_{k-1,j} - a_{k-1,j-1})}\]

\[
= (\mu^t)_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j}) = a_{i,j} - a_{i,j-1} - (a_{0,j} - a_{0,j-1}) \quad (\text{by } (11.604)).
\]

and thus \((\mu^t)_j + (b_{1,j} + b_{2,j} + \cdots + b_{i,j}) = a_{i,j} - a_{i,j-1}\), qed.
Thus, Assertion $B'$ holds. Since Assertions $A'$ and $B'$ are equivalent, this yields that Assertion $A'$ also holds. In other words, there exist a partition $\gamma$ and a column-strict tableau $T$ of shape $\gamma/\mu^t$ such that all $(i, j) \in \{1, 2, 3, \ldots\}^2$ satisfy (11.609). Let us consider this $\gamma$ and this $T$.

We shall soon see that $\gamma = \lambda^t$ and $\text{cont } T = \nu^t$. Let us first prepare for this. We have

(11.611) \[ \dim \left( (f^j)^{-1}(0) \right) = (\lambda^t)_1 + (\lambda^t)_2 + \cdots + (\lambda^t)_j \quad \text{for every } j \in \mathbb{N}. \]

509 From here, it is easy to see that

(11.612) \[ \dim \left( (f^j)^{-1}(0) \right) - \dim \left( (f^{j-1})^{-1}(0) \right) = (\lambda^t)_j \quad \text{for every } j \in \{1, 2, 3, \ldots\}. \]

510 Now, for every $j \in \{1, 2, 3, \ldots\}$, we have

\[
(\text{the number of all entries in the } j\text{-th row of } T) = \gamma_j - (\mu^t)_j
\]

---

509 Proof of (11.611): Let $j \in \mathbb{N}$. We can represent the endomorphism $f$ of $V$ as an $n \times n$-matrix $F \in \mathbb{K}^{n \times n}$ for $n = \text{dim } V$. This $n \times n$-matrix $F$ is nilpotent (since $f$ is nilpotent) and has Jordan type $\lambda$ (since $f$ has Jordan type $\lambda$). But Exercise 2.98(a) (applied to $N = F$) shows that $F$ has Jordan type $\lambda$ if and only if every $k \in \mathbb{N}$ satisfies $\dim (\ker (F^k)) = (\lambda^t)_1 + (\lambda^t)_2 + \cdots + (\lambda^t)_k$.

Since we know that $F$ has Jordan type $\lambda$, we thus conclude that every $k \in \mathbb{N}$ satisfies $\dim (\ker (F^k)) = (\lambda^t)_1 + (\lambda^t)_2 + \cdots + (\lambda^t)_k$.

Applied to $k = j$, this yields $\dim (\ker (F^j)) = (\lambda^t)_1 + (\lambda^t)_2 + \cdots + (\lambda^t)_j$.

Since the matrix $F$ represents the endomorphism $f$, we have

\[
\dim (\ker (F^j)) = \dim \left( \ker \left( (f^j)^{-1}(0) \right) \right) = \dim \left( (f^j)^{-1}(0) \right).
\]

Compared with $\dim (\ker (F^j)) = (\lambda^t)_1 + (\lambda^t)_2 + \cdots + (\lambda^t)_j$, this yields

\[
\dim \left( (f^j)^{-1}(0) \right) = (\lambda^t)_1 + (\lambda^t)_2 + \cdots + (\lambda^t)_j,
\]

and thus (11.611) is proven.

510 Proof of (11.612): Let $j \in \{1, 2, 3, \ldots\}$. Then,

\[
\dim \left( (f^j)^{-1}(0) \right) - \dim \left( (f^{j-1})^{-1}(0) \right) = (\lambda^t)_1 + (\lambda^t)_2 + \cdots + (\lambda^t)_j
\]

(by (11.611))

\[
= (\lambda^t)_1 + (\lambda^t)_2 + \cdots + (\lambda^t)_{j-1} + (\lambda^t)_j
\]

(by (11.611), applied to $j-1$ instead of $j$)

\[
= (\lambda^t)_1 + (\lambda^t)_2 + \cdots + (\lambda^t)_j - (\lambda^t)_1 + (\lambda^t)_2 + \cdots + (\lambda^t)_{j-1}
= (\lambda^t)_j,
\]

so that (11.612) is proven.
(since \( T \) has shape \( \gamma/\mu^t \)), so that

\[
\gamma_j - (\mu^t)_j = \text{(the number of all entries in the } j\text{-th row of } T) = \sum_{i=1}^{\infty} (\text{the number of all entries } i \text{ in the } j\text{-th row of } T) = \sum_{i=1}^{\infty} b_{i,j} = a_{i,j} - a_{i-1,j-1} + a_{i-1,j-1} - a_{i-1,j-1} \quad \text{(by (11.609))}
\]

\[
= \sum_{i=1}^{\infty} \left( a_{i,j} - a_{i-1,j-1} + a_{i-1,j-1} - a_{i-1,j-1} \right) = \sum_{i=1}^{\infty} \left( (a_{i,j} - a_{i-1,j-1}) - (a_{i-1,j-1} - a_{i-1,j-1}) \right)
\]

\[
= \sum_{i=1}^{\infty} (a_{i,j} - a_{i-1,j}) - \sum_{i=1}^{\infty} (a_{i,j-1} - a_{i-1,j-1})
\]

\[
= \dim \left( (f^j)^{-1}(0) \right) - \dim \left( U \cap (f^j)^{-1}(0) \right) \quad \text{(by Lemma 11.12(f), applied to } W=V, A=U, B=0 \text{ and } w_{i,j}=a_{i,j} \text{)}
\]

\[
= \left( \dim \left( (f^j)^{-1}(0) \right) - \dim \left( U \cap (f^j)^{-1}(0) \right) \right) - \left( \dim \left( (f^j)^{-1}(0) \right) - \dim \left( U \cap (f^j)^{-1}(0) \right) \right)
\]

\[
= (\lambda')_j - \left( \dim \left( U \cap (f^j)^{-1}(0) \right) - \dim \left( U \cap (f^j)^{-1}(0) \right) \right)
\]

Adding this equality to the equality

\[
(\mu^t)_j = \underbrace{a_{0,j}}_{\text{by (11.604)}}, \quad \underbrace{a_{0,j-1}}_{\text{by the definition of } a_{0,j}} - \underbrace{a_{0,j-1}}_{\text{by the definition of } a_{0,j-1}} = \dim \left( (f^0)^{-1}(U) \cap \ker(f^j) \right) = \dim \left( (f^0)^{-1}(U) \cap \ker f^j \right)
\]

\[
= \dim \left( \left( \frac{f^0}{=\id} \right)^{-1}(U) \cap \ker f^j \right) - \dim \left( \left( \frac{f^0}{=\id} \right)^{-1}(U) \cap \ker f^j \right)
\]

\[
= \dim \left( \left( \id \right)^{-1}(U) \cap \ker f^j \right) - \dim \left( \left( \id \right)^{-1}(U) \cap \ker f^j \right)
\]

\[
= \dim \left( \left( \id \right)^{-1}(U) \cap \ker f^j \right) - \dim \left( \left( \id \right)^{-1}(U) \cap \ker f^j \right)
\]

we obtain \( \gamma_j = (\lambda')_j \) for every \( j \in \{1, 2, 3, \ldots \} \). Thus, \( \gamma = \lambda' \). Hence, \( T \) is a column-strict tableau of shape \( \lambda' / \mu^t \) (since \( T \) is a column-strict tableau of shape \( \gamma / \mu^t \)).

Next, let us prove that \( \cont T = \nu^t \). Indeed, we first notice that

\[
(11.613) \quad \dim \left( (f^j)^{-1}(U) \right) = (\dim U) + (\nu^t)_1 + (\nu^t)_2 + \cdots + (\nu^t)_i \quad \text{for every } i \in \mathbb{N}.
\]
As a consequence of this, we have

$$(11.614) \quad \dim \left( (f^i)^{-1}(U) \right) - \dim \left( (f^{i-1})^{-1}(U) \right) = (\nu^i)_i \quad \text{for every } i \in \{1, 2, 3, \ldots \}.$$

Proof of (11.613): Let $i \in \mathbb{N}$. Recall that the nilpotent endomorphism $\overline{T}$ of the quotient space $V/U$ (induced by $f \in \text{End} V$) has Jordan type $\nu$. We can represent this nilpotent endomorphism $\overline{T}$ of $V/U$ as an $\ell \times \ell$-matrix $H \in \mathbb{K}^{\ell \times \ell}$ for $\ell = \dim (V/U)$. This $\ell \times \ell$-matrix $H$ is nilpotent (since $\overline{T}$ is nilpotent) and has Jordan type $\nu$ (since $\overline{T}$ has Jordan type $\nu$). But Exercise 2.98(a) (applied to $N = H$) shows that $H$ has Jordan type $\nu$ if and only if every $k \in \mathbb{N}$ satisfies $\dim (\ker (H^k)) = (\nu^i)_1 + (\nu^i)_2 + \cdots + (\nu^i)_k$. Since we know that $H$ has Jordan type $\nu$, we thus conclude that every $k \in \mathbb{N}$ satisfies $\dim (\ker (H^k)) = (\nu^i)_1 + (\nu^i)_2 + \cdots + (\nu^i)_k$. Applied to $k = i$, this yields $\dim (\ker (H^i)) = (\nu^i)_1 + (\nu^i)_2 + \cdots + (\nu^i)_i$.

Since the matrix $H$ represents the endomorphism $\overline{T}$, we have

$$\dim \left( \ker (\overline{T}^i) \right) = \dim (\ker (H^i)) = (\nu^i)_1 + (\nu^i)_2 + \cdots + (\nu^i)_i.$$  

But $(f^i)^{-1}(U)$ is a $\mathbb{K}$-vector subspace of $V$ containing $U$. Thus, $(f^i)^{-1}(U)/U$ is canonically a $\mathbb{K}$-vector subspace of $V/U$. Moreover, this subspace $(f^i)^{-1}(U)/U$ is precisely the kernel $\ker (\overline{T}^i)$ (this is straightforward to check). Hence,

$$\dim \left( (f^i)^{-1}(U)/U \right) = \dim \left( \ker (\overline{T}^i) \right) = (\nu^i)_1 + (\nu^i)_2 + \cdots + (\nu^i)_i.$$  

Since

$$\dim \left( (f^i)^{-1}(U)/U \right) = \dim \left( (f^i)^{-1}(U) \right) - \dim U,$$

this rewrites as

$$\dim \left( (f^i)^{-1}(U) \right) - \dim U = (\nu^i)_1 + (\nu^i)_2 + \cdots + (\nu^i)_i.$$  

Adding $\dim U$ to both sides of this equality yields (11.613).

Proof of (11.614): Let $i \in \{1, 2, 3, \ldots \}$. Then,

$$\underbrace{\dim \left( (f^i)^{-1}(U) \right)}_{= (\dim U) + (\nu^i)_1 + (\nu^i)_2 + \cdots + (\nu^i)_i \text{ (by (11.613))}} - \underbrace{\dim \left( (f^{i-1})^{-1}(U) \right)}_{= (\dim U) + (\nu^i)_1 + (\nu^i)_2 + \cdots + (\nu^i)_{i-1} \text{ (by (11.613), applied to } i-1 \text{ instead of } i)}$$

$$= (\dim U) + (\nu^i)_1 + (\nu^i)_2 + \cdots + (\nu^i)_i - (\dim U) + (\nu^i)_1 + (\nu^i)_2 + \cdots + (\nu^i)_{i-1}) = (\nu^i)_i.$$  

Thus, (11.614) is proven.
Now, for every \( i \in \{1, 2, 3, \ldots \} \), we have

\[
(\text{cont } T)_i = |T^{-1}(i)| = \text{(the number of entries } i \text{ in } T)
\]

\[
= \sum_{j=1}^{\infty} \text{(the number of all entries } i \text{ in the } j\text{-th row of } T) = \sum_{j=1}^{\infty} b_{i,j}
\]

(by (11.609))

\[
= \sum_{j=1}^{\infty} (a_{i,j} - a_{i,j-1} - a_{i-1,j} + a_{i-1,j-1}) = \sum_{j=1}^{\infty} ((a_{i,j} - a_{i,j-1}) - (a_{i-1,j} - a_{i-1,j-1}))
\]

\[
= \sum_{j=1}^{\infty} (a_{i,j} - a_{i,j-1}) - \sum_{j=1}^{\infty} (a_{i-1,j} - a_{i-1,j-1})
\]

(by Lemma 11.12(g), applied to \( W = V', A = U, B = 0 \) and \( w_{i,j} = a_{i,j} \))

\[
= \dim \left( \left( f^i \right)^{-1}(U) \right) - \dim \left( \left( f^i \right)^{-1}(U) \cap 0 \right) = \dim \left( \left( f^{i-1} \right)^{-1}(U) \right) - \dim \left( \left( f^{i-1} \right)^{-1}(U) \cap 0 \right)
\]

(by Lemma 11.12(f), applied to \( V, U, 0, w_{i,j} = a_{i,j} \) and \( j-1 \) instead of \( W, A, B, w_{i,j} \) and \( i \))

\[
= \dim \left( \left( f^i \right)^{-1}(U) \right) - \dim \left( \left( f^i \right)^{-1}(U) \right)
\]

\[
= \dim \left( \left( f^{i-1} \right)^{-1}(U) \right) - \dim \left( \left( f^{i-1} \right)^{-1}(U) \right)
\]

(by (11.614)).

Hence, \( \text{cont } T = \nu^i \).

We shall next see that for every positive integer \( j \), the weak composition \( \text{cont}(T)_{\text{cub} \geq j} \) is a partition. Here, and in the following, we are using the notations of Exercise 2.94.

We first notice that every \( i \in \{1, 2, 3, \ldots \} \) and \( j \in \mathbb{N} \) satisfy

\[
(11.615) \quad \text{(the number of entries } i \text{ in the first } j \text{ rows of } T) = a_{i,j} - a_{i-1,j}.
\]
Now, for every positive integers \( j \) and \( i \), we have

\[
\begin{align*}
&= a_{i+1,j} - a_{i,j} \quad \text{(by (11.615), applied to } i+1 \text{ instead of } i) \\
&\leq a_{i,j-1} - a_{i-1,j-1} \quad \text{(by Lemma 11.12(c), applied to } V, U, 0, a_{i,j} \text{ and } j-1 \text{ instead of } W, A, B, w_{i,j} \text{ and } j) \\
&= (\text{the number of entries } i \text{ in the first } j-1 \text{ rows of } T)
\end{align*}
\]

(since (11.615) (applied to } j-1 \text{ instead of } j) \text{ yields } (\text{the number of entries } i \text{ in the first } j-1 \text{ rows of } T) = a_{i,j-1} - a_{i-1,j-1}). \text{ In other words, for every positive integers } j \text{ and } i, \text{ the number of entries } i+1 \text{ in the first } j \text{ rows of } T \text{ is } \leq \text{ to the number of entries } i \text{ in the first } j-1 \text{ rows of } T. \text{ In other words, Assertion } D \text{ of Exercise 2.94(b) (with } \lambda \text{ and } \mu \text{ replaced by } \lambda' \text{ and } \mu') \text{ is satisfied. Hence, Assertion } C \text{ of Exercise 2.94(b) (with } \lambda \text{ and } \mu \text{ replaced by } \lambda' \text{ and } \mu') \text{ is satisfied as well (because Exercise 2.94(b) yields that these Assertions } C \text{ and } D \text{ are equivalent). \text{ In other words, for every positive integer } j, \text{ the weak composition } \text{cont} (T|_{\text{cols } \geq j}) \text{ is a partition.}

Now, let us forget that we defined } \nu \text{ and } T. \text{ We thus have found a column-strict tableau } T \text{ of shape } \lambda'/\mu' \text{ with } \text{cont } T' = \nu' \text{ which has the property that for every positive integer } j, \text{ the weak composition } \text{cont} (T|_{\text{cols } \geq j}) \text{ is a partition. But we know that the number of such tableaux is } c_{\mu',\nu'}^\lambda \text{ (by Corollary 2.58, applied to } \lambda', \mu' \text{ and } \nu' \text{ instead of } \lambda, \mu \text{ and } \nu). \text{ Hence, this number } c_{\mu',\nu'}^\lambda \text{ must be } \neq 0 \text{ (because we have found such a tableau } T). \text{ So we have proven that } c_{\mu',\nu'}^\lambda \neq 0. \text{ Now, Exercise 2.69(c) yields } c_{\mu,\nu}^\lambda = c_{\mu',\nu'}^\lambda \neq 0, \text{ so that Exercise 2.98(b) is solved.}

11.70. **Solution to Exercise 3.7.** Solution to Exercise 3.7. Let } A \text{ be as in Proposition 3.3, and assume that } k = \mathbb{Q}. \text{ We have } p \cap I^2 = 0 \text{ (by Proposition 3.3 (b)). Thus, Lemma 3.5 yields that } A \text{ is commutative. Since } A \text{ is self-dual, this yields that } A \text{ is cocommutative. Hence, Exercise 1.65 (d) shows that the } k \text{-algebra } A \text{ is generated by the } k \text{-submodule } p. \text{ In other words, } A \text{ is the } k \text{-subalgebra of } A \text{ generated by the } k \text{-submodule } p. \text{ Since the } k \text{-subalgebra of } A \text{ generated by the } k \text{-submodule } p \text{ is } \sum_{n \geq 0} p^n \text{ (this is because generally, if } B \text{ is a } k \text{-algebra, and } U \text{ is a } k \text{-submodule of } B, \text{ then the } k \text{-subalgebra of } B \text{ generated by the } k \text{-submodule } U

---

\[513\] *Proof of (11.615):* Let \( i \in \{1, 2, 3, \ldots \} \) and \( j \in \mathbb{N}. \) We have

\[
\begin{align*}
&= \sum_{k=1}^{j} (\text{the number of all entries } i \text{ in the } k \text{-th row of } T) \\
&= \sum_{k=1}^{j} \left( b_{i,k} \right) \\
&\quad \text{(since (11.609) (applied to } (i,k) \text{ instead of } (i,j) \text{) yields } b_{i,k} \text{)} \\
&= \sum_{k=1}^{j} \left( a_{i,k} - a_{i-1,k-1} - a_{i-1,k} + a_{i,k-1} \right) \\
&\quad \text{(by the definition of } b_{i,k}) \\
&= \sum_{k=1}^{j} \left( a_{i,k} - a_{i-1,k} - a_{i-1,k-1} \right) \\
&\quad \text{(by the telescope principle)}
\end{align*}
\]

\[
\begin{align*}
&= \sum_{k=1}^{j} \left( a_{i,k} - a_{i-1,k} - a_{i-1,k-1} \right) \\
&= (a_{i,j} - a_{i-1,j}) - (a_{i,k} - a_{i-1,k}) - \left( \begin{array}{c}
\frac{a_{i,0}}{0} - \frac{a_{i-1,0}}{0} \\
\text{(by (11.602))}
\end{array} \right) \quad \text{(by the telescope principle)}
\end{align*}
\]

This proves (11.615).
is \( \sum_{n \geq 0} U^n \), this rewrites as follows: \( A \) is \( \sum_{n \geq 0} p^n \). Hence,

\[
A = \sum_{n \geq 0} p^n = p^0 + p^1 + \sum_{n \geq 2} p^n = k \cdot 1_A + p + \sum_{n \geq 2} p \cdot p^{n-2}
\]

\[
\subset k \cdot 1_A + p + \sum_{n \geq 2} I \cdot I \cdot p^{n-2} \subset k \cdot 1_A + p + \sum_{n \geq 2} I \subset k \cdot 1_A + p + I^2.
\]

Now, let \( a \in I \). Then, \( \epsilon (a) = 0 \) (by the definition of \( I \)). But we know that \( a \in I \subset A \subset k \cdot 1_A + p + I^2 \).

Hence, there exist some \( \lambda \in k \) and some \( a' \in p + I^2 \) such that \( a = \lambda \cdot 1_A + a' \). Since \( a' \in p + I^2 \subset I \), we have \( \epsilon (a') = 0 \). Now, \( \epsilon (a) = 0 \), so that

\[
0 = \epsilon \left( \frac{a}{\lambda \cdot 1_A + a'} \right) = \epsilon \left( \frac{\lambda \cdot 1_A + a'}{1} \right) = \lambda \epsilon (1_A) + \epsilon (a') = \lambda,
\]

so that \( \lambda = 0 \) and thus \( a = \frac{\lambda \cdot 1_A + a'}{1} = a' \in p + I^2 \). We thus have shown that every \( a \in I \) satisfies \( a \in p + I^2 \). Hence, \( I \subset p + I^2 \). Combined with \( p + I^2 \subset I \), this yields \( I = p + I^2 \), and thus \( I = p \oplus I^2 \) (since \( p \cap I^2 = 0 \)). This proves Proposition 3.3(c).

11.71. Solution to Exercise 3.10. Solution to Exercise 3.10. Corollary 2.45(a) yields that \((h_\lambda)_{\lambda \in \text{Par}}\) and \((m_\lambda)_{\lambda \in \text{Par}}\) are dual bases with respect to the Hall inner product on \( \Lambda \). In other words,

\[
(h_\lambda, m_\mu) = \delta_{\lambda, \mu} \quad \text{for any partitions } \lambda \text{ and } \mu.
\]

Thus, every nonempty partition \( \lambda \) satisfies \((h_\lambda, m_\emptyset) = \delta_{\lambda, \emptyset} = 0 \). Hence, every nonempty partition \( \lambda \) satisfies

\[
(11.616) \quad \begin{pmatrix} h_\lambda, \frac{1}{m_\emptyset} \end{pmatrix} = (h_\lambda, m_\emptyset) = 0.
\]

Recall the following fundamental fact from linear algebra: If \( k \) is a commutative ring, if \( A \) is a \( k \)-module, if \((\cdot, \cdot) : A \times A \to k\) is a symmetric \( k \)-bilinear form on \( A \), and if \((u_\lambda)_{\lambda \in L}\) and \((v_\lambda)_{\lambda \in L}\) are two \( k \)-bases of \( A \) which are dual to each other with respect to the form \((\cdot, \cdot)\) (where \( L \) is some indexing set), then every \( a \in A \) satisfies

\[
(11.617) \quad a = \sum_{\lambda \in L} (u_\lambda, a) v_\lambda.
\]

We can apply this fact to \( A = \Lambda, L = \text{Par}, (u_\lambda)_{\lambda \in L} = (h_\lambda)_{\lambda \in \text{Par}} \) and \((v_\lambda)_{\lambda \in L} = (m_\lambda)_{\lambda \in \text{Par}}\) (since the bases \((h_\lambda)_{\lambda \in \text{Par}}\) and \((m_\lambda)_{\lambda \in \text{Par}}\) of \( \Lambda \) are dual to each other with respect to the Hall inner product \((\cdot, \cdot)\)). As a result, we obtain that every \( a \in \Lambda \) satisfies

\[
(11.618) \quad a = \sum_{\lambda \in \text{Par}} (h_\lambda, a) m_\lambda.
\]

Now, let us solve the exercise. We need to show that for every \( a \in \Lambda \), the element \( a \) of \( \Lambda \) is primitive if and only if \( a \) lies in the \( k \)-linear span of \( p_1, p_2, p_3, \ldots \). The “if” direction of this statement is obvious\(^{514}\). Hence, we only need to prove the “only if” statement. In other words, we need to prove that if \( a \) is primitive, then \( a \) lies in the \( k \)-linear span of \( p_1, p_2, p_3, \ldots \). So let us assume that \( a \) is primitive. That is, \( \Delta (a) = 1 \otimes a + a \otimes 1 \).

For every \( x \in \Lambda \) and \( y \in \Lambda \), we have (using the Sweedler notation)

\[
x^+ (a) = \sum_{(a)} (x, a_1) a_2 = (x, 1) a + (x, a) 1 \quad \text{since } \sum_{(a)} a_1 \otimes a_2 = \Delta (a) = 1 \otimes a + a \otimes 1
\]

\(^{514}\)because Proposition 2.24(i) shows that each of \( p_1, p_2, p_3, \ldots \) is primitive, and therefore every element of their \( k \)-linear span is also primitive
and therefore

\[
(xy, a) = \left( y, \frac{x^a}{(x, a)_{a + (x, a)_{1}}} \right) \quad \text{(by Proposition 2.73(i))}
\]

(11.619) \quad = (y, (x, 1)_{a} + (x, a)_{1}) = (x, 1) (y, a) + (x, a) (y, 1) = (x, 1) (y, a) + (y, 1) (x, a).

Using this, we can easily obtain \((1, a) = 0\) \(^{515}\).

Now, using (11.619), it is easy to see that

(11.620) \quad (h_{\lambda}, a) = 0 \quad \text{for every partition } \lambda \text{ satisfying } \ell(\lambda) \geq 2.

\(^{516}\) Now, (11.618) becomes

\[
a = \sum_{\lambda \in \text{Par}} (h_{\lambda}, a) m_{\lambda} = \sum_{\lambda \in \text{Par}; \ell(\lambda)=0} (h_{\lambda}, a) m_{\lambda} + \sum_{\lambda \in \text{Par}; \ell(\lambda)=1} (h_{\lambda}, a) m_{\lambda} + \sum_{\lambda \in \text{Par}; \ell(\lambda) \geq 2 \text{ (by (11.620))}} (h_{\lambda}, a) m_{\lambda}
\]

\[
= \left( h_{\emptyset}, a \right) m_{\emptyset} + \sum_{n \geq 1} (h_{(n)}, a) m_{(n)} + \sum_{\lambda \in \text{Par}; \ell(\lambda) \geq 2} 0 m_{\lambda} = (1, a) m_{\emptyset} + \sum_{n \geq 1} (h_{(n)}, a) p_{n}
\]

\[
= \sum_{n \geq 1} (h_{(n)}, a) p_{n}.
\]

Thus, \(a\) lies in the \(k\)-linear span of \(p_{1}, p_{2}, p_{3}, \ldots\). This completes the solution to Exercise 3.10.


(a) This is straightforward: Let \(g_{1}\) and \(g_{2}\) be two elements of \(G\) belonging to the same conjugacy class. Thus, \(g_{1}\) and \(g_{2}\) are conjugate. In other words, there exists some \(x \in G\) such that \(g_{1} = xg_{2}x^{-1}\). Consider this \(x\).

\(^{515}\) Indeed, applying (11.619) to \(x = 1\) and \(y = 1\), we obtain \((1, a) = (1, 1) (1, a) + (1, 1) (1, a) = (1, a) + (1, a) = 2 (1, a)\), so that \((1, a) = 0\), qed.

\(^{516}\) Proof of (11.620): Let \(\lambda\) be a partition satisfying \(\ell(\lambda) \geq 2\). Write the partition \(\lambda\) in the form \(\lambda = (\lambda_{1}, \lambda_{2}, \ldots, \lambda_{\ell})\) with \(\ell = \ell(\lambda)\). Then, \(\lambda_{1}, \lambda_{2}, \ldots, \lambda_{\ell}\) are positive integers. Let \(\nu\) be the partition \((\lambda_{2}, \lambda_{3}, \ldots, \lambda_{\ell})\); then, \(\nu\) is nonempty (since \(\ell = \ell(\lambda) \geq 2\)), so that \((h_{\nu}, 1) = 0\) (by (11.616)), applied to \(\nu\) instead of \(\lambda\). Also, \((h_{(\lambda_{1}), 1}) = 0\) (by (11.616)), applied to \((\lambda_{1})\) instead of \(\lambda\).

By the definition of \(h_{\lambda}\), we have \(h_{\lambda} = h_{\lambda_{1}} h_{\lambda_{2}} \ldots h_{\lambda_{\ell}} = h_{\lambda_{1}} h_{\lambda_{2}} h_{\lambda_{3}} \ldots h_{\lambda_{\ell}} = h_{((\lambda_{1}), 1)} h_{\nu}\). Thus,

\[
\left( \frac{h_{\lambda}}{h_{(\lambda_{1}), 1} h_{\nu}}, a \right) = \left( h_{(\lambda_{1}), 1} h_{\nu}, a \right) = (1, a) = \frac{(h_{(\lambda_{1}), 1}) h_{\nu} + (h_{\nu}, 1) (h_{(\lambda_{1}), 1}) a}{0} = 0 + 0 = 0.
\]

This proves (11.620).
By the definition of $\text{Ind}_{H}^{G} f$, we have

$$
\left(\text{Ind}_{H}^{G} f\right)(g_1) = \frac{1}{|H|} \sum_{k \in G: k g_1 k^{-1} \in H} f(k g_1 k^{-1}) = \frac{1}{|H|} \sum_{k \in G: k x g_2 x^{-1} k^{-1} \in H} f(k x g_2 x^{-1} k^{-1}) \quad \text{(since } g_1 = x g_2 x^{-1})
$$

$$
= \frac{1}{|H|} \sum_{k \in G: k x g_2 (k x)^{-1} \in H} f(k x g_2 (k x)^{-1}) \quad \text{(since } x^{-1} k^{-1} = (k x)^{-1})
$$

$$
= \frac{1}{|H|} \sum_{k \in G: k g_2 k^{-1} \in H} f(k g_2 k^{-1})
$$

(here, we have substituted $k$ for $k x$ in the sum, because the map $G \to G, k \to k x$ is a bijection). Compared with

$$
\left(\text{Ind}_{H}^{G} f\right)(g_2) = \frac{1}{|H|} \sum_{k \in G: k g_2 k^{-1} \in H} f(k g_2 k^{-1})
$$

(this follows from the definition of $\text{Ind}_{H}^{G} f$), this yields $\left(\text{Ind}_{H}^{G} f\right)(g_1) = \left(\text{Ind}_{H}^{G} f\right)(g_2)$.

Now, forget that we fixed $g_1$ and $g_2$. We thus have proven that if $g_1$ and $g_2$ are two elements of $G$ belonging to the same conjugacy class, then $\left(\text{Ind}_{H}^{G} f\right)(g_1) = \left(\text{Ind}_{H}^{G} f\right)(g_2)$. In other words, the map $\text{Ind}_{H}^{G} f$ is constant on conjugacy classes. Hence, $\text{Ind}_{H}^{G} f$ is a class function on $G$. In other words, $\text{Ind}_{H}^{G} f \in R_{C}(G)$. This solves part (a) of the exercise.

(b) We have $G = \bigcup_{j \in J} H j$. Thus, every $k \in G$ can be uniquely written in the form $h j$ for some $j \in J$ and $h \in H$. Hence,

$$
\sum_{k \in G: k g k^{-1} \in H} f(k g k^{-1}) = \sum_{j \in J: h \in H: (h j) g (h j)^{-1} \in H} f(h j g (h j)^{-1}) = \sum_{j \in J: h \in H: h j g j^{-1} h^{-1} \in H} f(h j g j^{-1} h^{-1})
$$

(by under the condition that $h \in H$, the relation $h j g j^{-1} h^{-1} \in H$ is equivalent to the relation $j g j^{-1} \in H$)

$$
= \sum_{j \in J: h \in H: j g j^{-1} \in H} f(h j g j^{-1} h^{-1}) = \sum_{j \in J: h \in H: j g j^{-1} \in H} f(j g j^{-1})
$$

(since $f$ is a class function on $H$, $h j g j^{-1} h^{-1}$ is $H$-conjugate to $j g j^{-1}$)

$$
= \sum_{j \in J: h \in H: j g j^{-1} \in H} f(j g j^{-1}) \quad \text{for some } C
$$

$$
= |H| \sum_{j \in J: j g j^{-1} \in H} f(j g j^{-1})
$$

(11.621)

Now, the definition of $\text{Ind}_{H}^{G} f$ yields

$$
\left(\text{Ind}_{H}^{G} f\right)(g) = \frac{1}{|H|} \sum_{k \in G: k g k^{-1} \in H} f(k g k^{-1}) = \sum_{j \in J: j g j^{-1} \in H} f(j g j^{-1})
$$

This solves part (b) of the exercise.
11.73. **Solution to Exercise 4.2.** Solution to Exercise 4.2. We have $CG \otimes_{CH} CH \cong CG$ as $(CG, CH)$-bimodules, and thus also as $(CG, CI)$-bimodules\(^{517}\).

By the definition of induction, we have $\text{Ind}^H_I U = CH \otimes_{CI} U$. But by the definition of induction, we also have

\[
\text{Ind}^G_H \text{Ind}^H_I U = CG \otimes_{CH} \text{Ind}^H_I U = CG \otimes_{CH} (CH \otimes_{CI} U)
\]

\[
= (CG \otimes_{CH} CH) \otimes_{CI} U
\]

\[
\cong CG \otimes_{CI} U
\]

(by the associativity of the tensor product)

\[
\cong CG \otimes_{CI} U = \text{Ind}^G_I U
\]

(since $\text{Ind}^G_I U = CG \otimes_{CI} U$ (by the definition of induction)). This solves Exercise 4.2.

11.74. **Solution to Exercise 4.3.** Solution to Exercise 4.3. The solution rests upon the following general fact:

**Proposition 11.13.** Let $k$ be a commutative ring. Let $A$, $B$, $C$, $A'$, $B'$ and $C'$ be six $k$-algebras. Let $P$ be an $(A, B)$-bimodule\(^{518}\). Let $Q$ be a $(B, C)$-bimodule. Let $P'$ be an $(A', B')$-bimodule. Let $Q'$ be a $(B', C')$-bimodule. Then,

\[
(P \otimes P') \otimes_{B \otimes B'} (Q \otimes Q') \cong (P \otimes_B Q) \otimes (P' \otimes_{B'} Q')
\]

as $(A \otimes A', C \otimes C')$-bimodules. Here, all $\otimes$ signs without subscript stand for $\otimes_k$.

This proposition is proven by straightforward (repeated) application of the universal property of the tensor product. (Of course, the $(A \otimes A', C \otimes C')$-bimodule isomorphism $(P \otimes P') \otimes_{B \otimes B'} (Q \otimes Q') \to (P \otimes_B Q) \otimes (P' \otimes_{B'} Q')$ sends every $(p \otimes p') \otimes_{B \otimes B'} (q \otimes q')$ to $(p \otimes_B q) \otimes (p' \otimes_{B'} q')$.) We leave all details to the reader.

Now, let us come to the solution of Exercise 4.3.\(^{519}\) We want to prove the isomorphism

\[
\text{Ind}_{H_1 \times H_2}^{G_1 \times G_2} (U_1 \otimes U_2) \cong \left( \text{Ind}_{H_1}^{G_1} U_1 \right) \otimes \left( \text{Ind}_{H_2}^{G_2} U_2 \right)
\]

of $\mathbb{C}[G_1 \times G_2]$-modules. Recalling the definition of Ind, we rewrite this as

\[
\mathbb{C}[G_1 \times G_2] \otimes_{\mathbb{C}[H_1 \times H_2]} (U_1 \otimes U_2) \cong (\text{Ind}_{H_1}^{G_1 \circ H_2} U_1) \otimes (\text{Ind}_{H_2}^{G_2 \circ H_2} U_2).
\]

But Proposition 11.13 (applied to $k = \mathbb{C}$, $A = CG_1$, $B = CH_1$, $C = \mathbb{C}$, $A' = CG_2$, $B' = CH_2$, $C' = \mathbb{C}$, $P = CG_1$, $Q = U_1$, $P' = CG_2$ and $Q' = U_2$) yields

\[
(\text{Ind}_{H_1}^{G_1 \circ H_2} U_1) \otimes (\text{Ind}_{H_2}^{G_2 \circ H_2} U_2) \cong (\text{Ind}_{H_1}^{G_1} U_1) \otimes (\text{Ind}_{H_2}^{G_2} U_2)
\]

as $(CG_1 \otimes CG_2, \mathbb{C} \otimes \mathbb{C})$-bimodules (thus, as left $CG_1 \otimes CG_2$-modules). In order to derive (11.622) from this isomorphism, we need to realize that:

- there exist algebra isomorphisms $\mathbb{C}[G_1 \times G_2] \to CG_1 \otimes CG_2$ and $\mathbb{C}[H_1 \times H_2] \to CH_1 \otimes CH_2$ which commute with the canonical inclusion maps $\mathbb{C}[H_1 \times H_2] \to \mathbb{C}[G_1 \times G_2]$ and $CH_1 \otimes CH_2 \to CH_1 \otimes CH_2$;
- when we identify $\mathbb{C}[H_1 \times H_2]$ with $CH_1 \otimes CH_2$ along the isomorphism $\mathbb{C}[H_1 \times H_2] \to CH_1 \otimes CH_2$, the $CH_1 \otimes CH_2$-module $U_1 \otimes U_2$ becomes exactly the $\mathbb{C}[H_1 \times H_2]$-module $U_1 \otimes U_2$;
- when we identify $\mathbb{C}[G_1 \times G_2]$ with $CG_1 \otimes CG_2$ along the isomorphism $\mathbb{C}[G_1 \times G_2] \to CG_1 \otimes CG_2$, the $CG_1 \otimes CG_2$-module $(CG_1 \otimes CH_1 U_1) \otimes (CG_2 \otimes CH_2 U_2)$ becomes exactly the $\mathbb{C}[G_1 \times G_2]$-module $(CG_1 \otimes CH_1 U_1) \otimes (CG_2 \otimes CH_2 U_2)$.

\(^{517}\)Since the right $CI$-module structures on $CG \otimes_{CH} CH$ and on $CG$ are obtained from the respective right $CH$-module structures by restriction

\(^{518}\)As usual, we understand the notion of a bimodule to be defined over $k$; that is, the left $A$-module structure and the right $B$-module structure of an $(A, B)$-bimodule must restrict to one and the same $k$-module structure.

\(^{519}\)The following solution involves some handwaving: We are going to use certain isomorphisms to identify $\mathbb{C}[G_1 \times G_2]$ with $CG_1 \otimes CG_2$ and to identify $\mathbb{C}[H_1 \times H_2]$ with $CH_1 \otimes CH_2$. See the solution of Exercise 4.15 for an example of how to avoid this kind of handwaving. (Actually, Exercise 4.14(b) shows that Exercise 4.15 is a generalization of Exercise 4.3.)
These facts are all trivial to verify (of course, the isomorphism \( \mathbb{C} [G_1 \times G_2] \to \mathbb{C} G_1 \otimes \mathbb{C} G_2 \) is given by sending every \( t_{(g_1, g_2)} \) to \( t_g \otimes t_g \), and similarly for the other isomorphism). Thus, (11.622) holds, and the exercise is solved.

11.75. **Solution to Exercise 4.4.** Solution to Exercise 4.4. In the following, we will write \( g \) for the element \( t_g \) of \( CG \) whenever \( g \) is an element of \( G \). This is a relatively common abuse of notation, and it is harmless because the map \( G \to CG, \ g \to t_g \) is an injective homomorphism of multiplicative monoids (so \( t_{gh} = t_g t_h \) and \( t_1 = 1 \), which means that we won’t run into ambiguities denoting \( t_g \) by \( g \)) and because every \( CG \)-module \( M \), every \( m \in M \) and every \( g \in G \) satisfy \( gm = t_g m \).

Recall that \( \text{Ind}^G_U \) is defined as the \( CG \)-module \( \mathbb{C} G \otimes \mathbb{C} H U \), where \( \mathbb{C} G \) is regarded as a \((CG, CH)\)-bimodule.

The \( \mathbb{C} \)-vector space \( \mathbb{C} G \) is endowed with both a \((CG, CH)\)-bimodule structure (this is the structure used in the definition of \( \text{Ind}^G_U \)) and a \((CH, CG)\)-bimodule structure (this is the structure used in the statement of Exercise 4.4). Thus, \( \mathbb{C} G \) has a \( CG \)-module structure, a right \( CH \)-module structure, a left \( CH \)-module structure, and a \( CG \)-module structure. All four of these structures are simply given by multiplication inside \( \mathbb{C} G \) (since \( \mathbb{C} H \) is a \( \mathbb{C} \)-subalgebra of \( \mathbb{C} G \)). Therefore, notations like \( xy \) with \( x \) and \( y \) being two elements of \( \mathbb{C} G \) or \( CH \) will never be ambiguous: While they might be interpreted in different ways, all of the possible interpretations will produce identical results.

We recall that \( \text{Hom}_{\mathbb{C} H} (\mathbb{C} G, U) \) is the left \( CG \)-module consisting of all left \( CH \)-module homomorphisms from \( \mathbb{C} G \) to \( U \). This uses only the \((CH, CG)\)-bimodule structure on \( \mathbb{C} G \) that was used in the statement of Exercise 4.4 (but not the \((CG, CH)\)-bimodule structure that was used in the definition of \( \text{Ind}^G_U \)).

Let \( J \) be a system of distinct representatives for the right \( H \)-cosets in \( G \). Then, \( G = \bigcup_{j \in J} H j \).

We now define a map \( \alpha : \text{Hom}_{\mathbb{C} H} (\mathbb{C} G, U) \to \mathbb{C} G \otimes \mathbb{C} H U \) by setting

\[
\alpha (f) = \sum_{j \in J} j^{-1} \otimes \mathbb{C} H f (j) \quad \text{for all } f \in \text{Hom}_{\mathbb{C} H} (\mathbb{C} G, U).
\]

This \( \alpha \) is a map \( \text{Hom}_{\mathbb{C} H} (\mathbb{C} G, U) \to \text{Ind}^G_U \) (since \( \text{Ind}^G_U = \mathbb{C} G \otimes \mathbb{C} H U \)). It is easy to see that this map \( \alpha \) does not depend on the choice of \( J \). (In fact, if \( j_1 \) and \( j_2 \) are two elements of \( G \) lying in the same right \( H \)-coset, and if \( f \in \text{Hom}_{\mathbb{C} H} (\mathbb{C} G, U) \), then it is easy to see that \( j_1^{-1} \otimes \mathbb{C} H f (j_1) = j_2^{-1} \otimes \mathbb{C} H f (j_2) \).) In other words, if \( J' \) is any system of distinct representatives for the right \( H \)-cosets in \( G \) (which may and may not not be equal to \( J \)), then

\[
\alpha (f) = \sum_{j \in J'} j^{-1} \otimes \mathbb{C} H f (j) \quad \text{for all } f \in \text{Hom}_{\mathbb{C} H} (\mathbb{C} G, U).
\]

Notice that if \( g \in G \) is arbitrary, then \( J g = \{ j g \mid j \in J \} \) is also a system of distinct representatives for the right \( H \)-cosets in \( G \).

We will show that \( \alpha \) is a \( CG \)-module isomorphism.

First, let us prove that \( \alpha \) is a left \( CG \)-module homomorphism. In fact, any \( f \in \text{Hom}_{\mathbb{C} H} (\mathbb{C} G, U) \) and \( g \in G \) satisfy

\[
\alpha (gf) = \sum_{j \in J} j^{-1} \otimes \mathbb{C} H (gf)(j) = \sum_{j \in J} g (j g)^{-1} \otimes \mathbb{C} H f (jg) = \sum_{j \in J} g j^{-1} \otimes \mathbb{C} H f (j) = g \cdot \alpha (f) \quad \text{(by (11.623), applied to } J' = Jg) \,
\]

The map \( \alpha \) is thus a homomorphism of left \( G \)-sets. Since \( \alpha \) is furthermore \( \mathbb{C} \)-linear, this yields that \( \alpha \) is a left \( CG \)-module homomorphism.

We now are going to construct an inverse for \( \alpha \). This will be more cumbersome.
For every $g \in G$ and every $p \in CG$, we denote by $\epsilon_g(p)$ the $g$-coordinate of $p$ with respect to the basis $G$ of the $C$-vector space $CG$. By the definition of “coordinate”, we have

\begin{equation}
q = \sum_{g \in G} \epsilon_g(q) \quad \text{for every } q \in CG.
\end{equation}

For every $g \in G$, we have defined a map $\epsilon_g : CG \to C$ (because we have defined an element $\epsilon_g(p)$ for every $p \in CG$). This map $\epsilon_g$ is $C$-linear. Here are some simple properties of this map:

- For every $g \in G$ and $h \in G$, we have

\begin{equation}
\epsilon_g(h) = \delta_{g,h}.
\end{equation}

- We have

\begin{equation}
\epsilon_1(pq) = \epsilon_1(qp) \quad \text{for all } p \in CG \text{ and } q \in CG.
\end{equation}

- Moreover,

\begin{equation}
\epsilon_1(g^{-1}q) = \epsilon_g(q) \quad \text{for every } g \in G \text{ and } q \in CG.
\end{equation}

Now, fix $q \in CG$ and $u \in U$. We let $f_{q,u}$ be the map $CG \to U$ defined by

\begin{equation}
f_{q,u}(p) = \sum_{h \in H} \epsilon_1(hpq) h^{-1}u \quad \text{for every } p \in CG.
\end{equation}

It is obvious that this map $f_{q,u}$ is $C$-linear. We will show that $f_{q,u}$ is a left $CH$-module homomorphism.

The map $f_{q,u}$ is a homomorphism of left $H$-sets.\(^{523}\) Since $f_{q,u}$ is furthermore $C$-linear, this yields that $f_{q,u}$ is a left $CH$-module homomorphism. Hence, $f_{q,u} \in \text{Hom}_{CH}(CG,U)$.

Now, forget that we fixed $q$ and $u$. We thus have defined a map $f_{q,u} \in \text{Hom}_{CH}(CG,U)$ for every $q \in CG$ and $u \in U$. It is easy to see that this map $f_{q,u}$ depends $C$-linearly on each of $q$ and $u$. Now, define a map $\tilde{\beta} : CG \times U \to \text{Hom}_{CH}(CG,U)$ by

\[ \tilde{\beta}(q,u) = f_{q,u} \quad \text{for every } (q,u) \in CG \times U. \]

\(520\)Proof. Let $g \in G$ and $h \in G$. Then, $\epsilon_g(h)$ is defined as the $g$-coordinate of $h$ with respect to the basis $G$ of the $C$-vector space $CG$. This $g$-coordinate is precisely 1 (since $h$ is an element of this basis $G$). Thus, $\epsilon_g(h) = 1$, qed.

\(521\)Proof. Let $p \in CG$ and $q \in CG$. We need to prove the equality (11.626). This equality is $C$-linear in each of $p$ and $q$, and thus we can WLOG assume that both $p$ and $q$ belong to the basis $G$ of the $C$-vector space $CG$. Assume this. Then, $pq$ and $qp$ belong to $G$ as well. Hence, (11.625) yields $\epsilon_1(pq) = \delta_{1,pq}$ and $\epsilon_1(qp) = \delta_{1,qp}$. But $p$ and $q$ are elements of the group $G$. Hence, we have 1 = $pq$ if and only if 1 = $qp$ (because both of these statements are equivalent to $q = p^{-1}$). Therefore, $\delta_{1,pq} = \delta_{1,qp}$, so that $\epsilon_1(pq) = \delta_{1,pq} = \delta_{1,qp} = \epsilon_1(qp)$. This proves (11.626).

\(522\)Proof. Let $g \in G$ and $q \in CG$. We need to prove the equality (11.627). This equality is $C$-linear in $q$, and thus we can WLOG assume that $q$ belongs to the basis $G$ of the $C$-vector space $CG$. Assume this. Then, $g^{-1}q \in G$ as well. Hence, (11.625) yields $\epsilon_1(g^{-1}q) = \delta_{1,g^{-1}q}$ and $\epsilon_g(q) = \delta_{g,q}$. But $g$ and $q$ are elements of the group $G$. Hence, we have 1 = $g^{-1}q$ if and only if $g = q$. Therefore, $\delta_{1,g^{-1}q} = \delta_{g,q}$, so that $\epsilon_1(g^{-1}q) = \delta_{1,g^{-1}q} = \delta_{g,q} = \epsilon_g(q)$. This proves (11.627).

\(523\)Proof. In fact, for every $h' \in H$ and every $p \in CG$, we have

\[ f_{q,u}(h'p) = \sum_{h \in H} \epsilon_1(hh'pq) h^{-1}u \quad \text{by the definition of } f_{q,u} \]

\[ = \sum_{h \in H} \epsilon_1(h(h')^{-1}h'pq) \left( h(h')^{-1}h^{-1}u \right) \]

\[ = \sum_{h \in H} \epsilon_1(hpq) h'h^{-1}u = h' \cdot \sum_{h \in H} \epsilon_1(hpq) h^{-1}u = h' \cdot f_{q,u}(p). \]

In other words, $f_{q,u}$ is a homomorphism of left $H$-sets, qed.
Then, $\tilde{\beta}$ is a $\mathbb{C}$-bilinear map (because $\tilde{\beta}(q,u) = f_{q,u}$ depends $\mathbb{C}$-linearly on each of $q$ and $u$). We are now going to prove that the map $\tilde{\beta}$ is $\mathcal{CH}$-bilinear with respect to the right $\mathcal{CH}$-module structure on $\mathcal{CG}$ and the left $\mathcal{CH}$-module structure on $U$.

In fact, every $h' \in H, q \in \mathcal{CG}$ and $u \in U$ satisfy

$$\tilde{\beta}(q,h'u) = \tilde{\beta}(qh',u). \quad (11.629)$$

$524$ As a consequence of this, every $r \in \mathcal{CH}, q \in \mathcal{CG}$ and $u \in U$ satisfy

$$\tilde{\beta}(q,ru) = \tilde{\beta}(qr,u) \quad (by \ (11.630))$$

(by $\mathbb{C}$-bilinearity of $\tilde{\beta}$). In other words, the map $\tilde{\beta}$ is $\mathcal{CH}$-bilinear with respect to the right $\mathcal{CH}$-module structure on $\mathcal{CG}$ and the left $\mathcal{CH}$-module structure on $U$. Hence, by the universal property of the tensor product, we conclude that there exists a unique $\mathbb{C}$-linear map $\beta : \mathcal{CG} \otimes_{\mathcal{CH}} U \rightarrow \text{Hom}_{\mathcal{CH}}(\mathcal{CG},U)$ such that every $(q,u) \in \mathcal{CG} \times U$ satisfies

$$\beta(q \otimes_{\mathcal{CH}} u) = \tilde{\beta}(q,u). \quad (11.631)$$

Consider this map $\beta$. Clearly, every $q \in \mathcal{CG}$ and $u \in U$ satisfy

$$\beta(q \otimes_{\mathcal{CH}} u) = \tilde{\beta}(q,u) = f_{q,u} \quad (by \ the \ definition \ of \ \tilde{\beta}). \quad (11.632)$$

Hence, every $q \in \mathcal{CG}, u \in U$ and $p \in \mathcal{CG}$ satisfy

$$\beta(q \otimes_{\mathcal{CH}} u) = f_{q,u} \quad (by \ (11.628)). \quad (11.633)$$

We shall now show that the maps $\alpha$ and $\beta$ are mutually inverse. To do so, we will show that $\alpha \circ \beta = \text{id}$ and $\beta \circ \alpha = \text{id}.$

$524$ Proof of (11.629): Let $h' \in H, q \in \mathcal{CG}$ and $u \in U$. Then, the map $H \rightarrow H, h \mapsto h'h$ is a bijection (since $H$ is a group). Now, let $p \in \mathcal{CG}$. The definition of $\tilde{\beta}$ yields $\tilde{\beta}(q,h'u) = f_{q,h'u}$. Hence,

$$\begin{align*}
\tilde{\beta}(q,h'u) &= f_{q,h'u}(p) \\
&= \sum \epsilon_1(h'pq) \frac{(h'h)^{-1}}{h'\epsilon_1(h'h')^{-1}} h'u \\
&= \sum \epsilon_1(h'pq) \frac{(h'h)^{-1}}{h'\epsilon_1(h'h')^{-1}} h'u \\
&= \sum \epsilon_1(h'pq) \frac{(h'h)^{-1}}{h'\epsilon_1(h'h')^{-1}} h'u \\
&= \sum \epsilon_1(h'pq) h^{-1}u.
\end{align*} \quad (11.634)$$

But the definition of $\tilde{\beta}$ also yields $\tilde{\beta}(qh',u) = f_{qh',u}$. Hence,

$$\begin{align*}
\tilde{\beta}(qh',u) &= f_{qh',u}(p) \\
&= \sum \epsilon_1(hpqh') h^{-1}u \\
&= \left(\tilde{\beta}(q,h'u)\right)(p) \quad (by \ (11.630)).
\end{align*} \quad (11.635)$$

Now, forget that we fixed $p$. We have thus proven that $\left(\tilde{\beta}(qh',u)\right)(p) = \left(\tilde{\beta}(q,h'u)\right)(p)$ for every $p \in \mathcal{CG}$. In other words, $\tilde{\beta}(qh',u) = \tilde{\beta}(q,h'u)$, so that $\tilde{\beta}(q,h'u) = \tilde{\beta}(qh',u)$. This proves (11.629).
Let us first show that $\alpha \circ \beta = \text{id}$. In fact, every $q \in CG$ and $u \in U$ satisfy

$$(\alpha \circ \beta) (q \otimes_{CH} u) = \alpha (\beta (q \otimes_{CH} u)) = \sum_{j \in J} j^{-1} \otimes_{CH} \left( \beta (q \otimes_{CH} u) (j) \right)$$

(by the definition of $\alpha$)

$$= \sum_{j \in J} j^{-1} \otimes_{CH} \left( \sum_{h \in H} \epsilon_1 (hj) h^{-1} u \right) = \sum_{j \in J} \sum_{h \in H} \epsilon_1 (hj) j^{-1} \otimes_{CH} h^{-1} u$$

(by (11.633), applied to $p=j$)

$$= \sum_{j \in J} \sum_{h \in H} \epsilon_1 (hj) j^{-1} \otimes_{CH} u = \sum_{j \in J} \sum_{h \in H} \epsilon_1 (hj) (hj)^{-1} \otimes_{CH} u$$

(here, we substituted $g$ for $hj$ in the second sum, since the map $H \rightarrow Hj$, $h \mapsto hj$ is a bijection (because $G$ is a group))

$$= \sum_{g \in G} \epsilon_g (g) g^{-1} \otimes_{CH} u = \sum_{g \in G} \epsilon_1 (g^{-1} q) (g^{-1})^{-1} \otimes_{CH} u$$

(here, we substituted $g^{-1}$ for $g$ in the sum, since the map $G \rightarrow G$, $g \mapsto g^{-1}$ is a bijection (since $G$ is a group))

$$= \sum_{g \in G} \epsilon_g (g) g \otimes_{CH} u = \sum_{g \in G} \epsilon_g (g) g \otimes_{CH} u$$

(by (11.627))

$$= q \otimes_{CH} u = \text{id} (q \otimes_{CH} u).$$

Thus, the two maps $\alpha \circ \beta$ and $\text{id}$ are equal on every pure tensor. Since these two maps are $\mathbb{C}$-linear, this yields that $\alpha \circ \beta = \text{id}$.

Next, we are going to show that $\beta \circ \alpha = \text{id}$.

Let $f \in \text{Hom}_{CH} (CG, U)$. Let $p \in CG$. The map $f$ is left $CH$-linear (since $f \in \text{Hom}_{CH} (CG, U)$), hence $\mathbb{C}$-linear. We have $\alpha (f) = \sum_{j \in J} j^{-1} \otimes_{CH} f (j)$ (by the definition of $\alpha$). Applying the map $\beta$ to this equality, we obtain

$$\beta (\alpha (f)) = \beta \left( \sum_{j \in J} j^{-1} \otimes_{CH} f (j) \right) = \sum_{j \in J} \beta (j^{-1} \otimes_{CH} f (j))$$
yields that the maps \( \alpha \) isomorphism is functorial in \( G \). Since \( \alpha \) isomorphism (as we already know that \( \beta \) words, \( -G j \)).

This isomorphism is functorial with respect to \( U \). This solves Exercise 4.4.

Remark: In our solution, we explicitly constructed a \( CG \)-module isomorphism \( \alpha : \text{Hom}_{CH} (CG, U) \cong CG \otimes_{CH} U = \text{Ind}^G_H U \) as left \( CG \)-modules. This solves Exercise 4.4.

Now, forget that we fixed \( p \). We thus have proven that \( (\beta (\alpha (f))) (p) = f (p) \) for every \( p \in CG \). In other words, \( \beta (\alpha (f)) = f \). Hence, \( (\beta \circ \alpha) (f) = \beta (\alpha (f)) = f = \text{id} (f) \).

Since we have shown this for every \( f \), we can thus conclude that \( \beta \circ \alpha = \text{id} \). Combined with \( \alpha \circ \beta = \text{id} \), this yields that the maps \( \alpha \) and \( \beta \) are mutually inverse. Hence, the map \( \alpha \) is invertible, and thus a left \( CG \)-module isomorphism (as we already know that \( \alpha \) is a left \( CG \)-module homomorphism). Hence, \( \text{Hom}_{CH} (CG, U) \cong CG \otimes_{CH} U = \text{Ind}^G_H U \) as left \( CG \)-modules. This solves Exercise 4.4.

Remark: In our solution, we explicitly constructed a \( CG \)-module isomorphism \( \alpha : \text{Hom}_{CH} (CG, U) \rightarrow \text{Ind}^G_H U \). This isomorphism is functorial with respect to \( U \). It is also independent on the choice of \( J \) (this is not immediately clear from its definition, but it can be shown very easily, by observing that the tensor \( j^{-1} \otimes_{CH} f (j) \) for \( j \in G \) depends only on the coset \( HJ \) and not on \( j \) itself). One might ask whether this isomorphism is functorial in \( G \) and \( H \); but to make sense of this question, one has to define the category with respect to which this functoriality is to be understood. I don’t know a good answer.
11.76. Solution to Exercise 4.6. Solution to Exercise 4.6. Consider the $\mathbb{C}G$-module $\text{Hom}_{\mathbb{C}H}(\mathbb{C}G, V)$ defined as in Exercise 4.4. Then, Exercise 4.4 (applied to $V$ instead of $U$) yields that this module $\text{Hom}_{\mathbb{C}H}(\mathbb{C}G, V)$ is isomorphic to $\text{Ind}_{H}^{G} V$. That is, $\text{Ind}_{H}^{G} V \cong \text{Hom}_{\mathbb{C}H}(\mathbb{C}G, V)$. Hence, $\text{Hom}_{\mathbb{C}G} \left( U, \text{Ind}_{H}^{G} V \right) \cong \text{Hom}_{\mathbb{C}G}(U, \text{Hom}_{\mathbb{C}H}(\mathbb{C}G, V))$. But (4.8) (applied to $R = \mathbb{C}G$, $S = \mathbb{C}H$, $A = \mathbb{C}G$, $B = U$ and $C = V$) yields $\text{Hom}_{\mathbb{C}H}(\mathbb{C}G \otimes_{\mathbb{C}G} U, V) \cong \text{Hom}_{\mathbb{C}G}(U, \text{Hom}_{\mathbb{C}H}(\mathbb{C}G, V))$. Altogether, we thus have

$$\text{Hom}_{\mathbb{C}G} \left( U, \text{Ind}_{H}^{G} V \right) \cong \text{Hom}_{\mathbb{C}G}(U, \text{Hom}_{\mathbb{C}H}(\mathbb{C}G, V)) \cong \text{Hom}_{\mathbb{C}H}( \mathbb{C}G \otimes_{\mathbb{C}G} U, V).$$

It remains to see what the $\mathbb{C}H$-module $\mathbb{C}G \otimes_{\mathbb{C}G} U$ is. By the basic properties of tensor products, the $\mathbb{C}H$-module $\mathbb{C}G \otimes_{\mathbb{C}G} U$ is isomorphic to the $\mathbb{C}H$-module obtained by restricting the action of $\mathbb{C}G$ on the $\mathbb{C}G$-module $U$ to the subalgebra $\mathbb{C}H$. Since this latter $\mathbb{C}H$-module is precisely $\text{Res}_{H}^{G} U$ (in fact, this is how $\text{Res}_{H}^{G} U$ was defined!), this rewrites as follows: The $\mathbb{C}H$-module $\mathbb{C}G \otimes_{\mathbb{C}G} U$ is isomorphic to $\text{Res}_{H}^{G} U$. In other words, $\mathbb{C}G \otimes_{\mathbb{C}G} U \cong \text{Res}_{H}^{G} U$. Now,

$$\text{Hom}_{\mathbb{C}G} \left( U, \text{Ind}_{H}^{G} V \right) \cong \text{Hom}_{\mathbb{C}G} \left( \mathbb{C}G \otimes_{\mathbb{C}G} U, V \right) \cong \text{Hom}_{\mathbb{C}H} \left( \text{Res}_{H}^{G} U, V \right).$$

This solves the exercise.

11.77. Solution to Exercise 4.9. Solution to Exercise 4.9. In the following, a “Hom” symbol without a subscript means “$\text{Hom}_{\mathbb{C}}$” (rather than “$\text{Hom}_{\mathbb{C}G}$” or whatever other meaning this symbol could possibly have in the context).

If $G$ is a group and if $M$ and $N$ are two $\mathbb{C}G$-modules, then $\text{Hom}(M, N)$ becomes a $\mathbb{C}G$-module, with $G$ acting as follows: If $g \in G$ and $f \in \text{Hom}(M, N)$, then $t_{g}f$ is the $\mathbb{C}$-linear map $M \to N$ sending every $m \in M$ to $t_{g}f \left( t_{g^{-1}}m \right)$. This $\mathbb{C}G$-module structure is precisely the one we know from Remark 1.33.

Now, it is well-known (and straightforward to verify) that every two $\mathbb{C}G$-modules $M$ and $N$ satisfy

$$\text{Hom}_{\mathbb{C}G}(M, N) = \left( \text{Hom}(M, N) \right)^{G}$$

(where we regard $\text{Hom}_{\mathbb{C}G}(M, N)$ as a $\mathbb{C}$-vector subspace of $\text{Hom}(M, N)$ because every $\mathbb{C}G$-linear map $M \to N$ is a $\mathbb{C}$-linear map $M \to N$).

Let us now come to the solution of the exercise.

(a) Let $\psi$ denote the $\mathbb{C}$-linear map

$$\text{Hom}(V_{1}, W_{1}) \otimes \text{Hom}(V_{2}, W_{2}) \to \text{Hom}(V_{1} \otimes V_{2}, W_{1} \otimes W_{2})$$

sending each tensor $f \otimes g$ to the tensor product $f \otimes g$ of homomorphisms. This map $\psi$ is completely independent of $G_{1}$ and $G_{2}$ (it is defined whenever $V_{1}$, $V_{2}$, $W_{1}$ and $W_{2}$ are four $\mathbb{C}$-vector spaces) and is a vector space isomorphism (this is a basic fact from linear algebra, relying only on the finite-dimensionality of $V_{1}, V_{2}, W_{1}$ and $W_{2}$). But we can regard $\text{Hom}(V_{1}, W_{1})$ as a $\mathbb{C}G_{1}$-module, $\text{Hom}(V_{2}, W_{2})$ as a $\mathbb{C}G_{2}$-module and $\text{Hom}(V_{1} \otimes V_{2}, W_{1} \otimes W_{2})$ as a $\mathbb{C}G_{1} \times G_{2}$-module. Then, the map $\psi$ is a homomorphism of $\mathbb{C}G_{1} \times G_{2}$-modules$^{525}$. Hence, this map $\psi$ must be an isomorphism of $\mathbb{C}G_{1} \times G_{2}$-modules (being a vector space isomorphism). As a consequence, it sends the $G_{1} \times G_{2}$-fixed space of its domain to the $G_{1} \times G_{2}$-fixed space of its target:

$$\psi \left( (\text{Hom}(V_{1}, W_{1}) \otimes \text{Hom}(V_{2}, W_{2}))^{G_{1} \times G_{2}} \right) = (\text{Hom}(V_{1} \otimes V_{2}, W_{1} \otimes W_{2}))^{G_{1} \times G_{2}}.$$

$^{525}$This is easy to verify by checking that $t_{(h_{1}, h_{2})}(\psi(f \otimes g)) = \psi(t_{h_{1}}f \otimes t_{h_{2}}g)$ for all $(h_{1}, h_{2}) \in G_{1} \times G_{2}$, $f \in \text{Hom}(V_{1}, W_{1})$ and $g \in \text{Hom}(V_{2}, W_{2})$. 

It is also worth noting that in our solution, $\mathbb{C}$ could be replaced by any commutative ring. We used neither that $\mathbb{C}$ is a field, nor that $\mathbb{C}$ has characteristic 0.
Since
\[ (\text{Hom}(V_1, W_1) \otimes \text{Hom}(V_2, W_2))^{G_1 \times G_2} = \text{Hom}_{CG_1}(V_1, W_1) \otimes \text{Hom}_{CG_2}(V_2, W_2) \]
by (11.634),

applied to \(G = G_1, M = V_1\) and \(N = W_1\)
and applied to \(G = G_2, M = V_2\) and \(N = W_2\)

\[ \left( \begin{array}{c}
\text{by } (11.635), \text{ applied to } K_1 = G_1, K_2 = G_2, \\
U_1 = \text{Hom}(V_1, W_1) \text{ and } U_2 = \text{Hom}(V_2, W_2)
\end{array} \right) \]

\[ = \text{Hom}_{CG_1}(V_1, W_1) \otimes \text{Hom}_{CG_2}(V_2, W_2), \]

this rewrites as
\[ \psi(\text{Hom}_{CG_1}(V_1, W_1) \otimes \text{Hom}_{CG_2}(V_2, W_2)) = (\text{Hom}(V_1 \otimes V_2, W_1 \otimes W_2))^{G_1 \times G_2}. \]

Hence, the isomorphism \(\psi\) restricts to an isomorphism from \(\text{Hom}_{CG_1}(V_1, W_1) \otimes \text{Hom}_{CG_2}(V_2, W_2)\) to
\((\text{Hom}(V_1 \otimes V_2, W_1 \otimes W_2))^{G_1 \times G_2}\). This restriction is precisely the \(C\)-linear map
\[ \text{Hom}_{CG_1}(V_1, W_1) \otimes \text{Hom}_{CG_2}(V_2, W_2) \rightarrow \text{Hom}_{[G_1 \times G_2]}(V_1 \otimes V_2, W_1 \otimes W_2) \]
sending each tensor \(f \otimes g\) to the tensor product \(f \otimes g\) of homomorphisms (i.e., the map alleged to be an
isomorphism in the statement of the exercise). So we know now that this map is an isomorphism. This
solves Exercise 4.9(a).

(b) Let \(G_1\) and \(G_2\) be two groups. Let \(V_i\) and \(W_i\) be finite-dimensional \(CG_i\)-modules for every \(i \in \{1, 2\}\). Exercise 4.9(a) provides a vector space isomorphism from \(\text{Hom}_{CG_1}(V_1, W_1) \otimes \text{Hom}_{CG_2}(V_2, W_2)\) to
\(\text{Hom}_{[G_1 \times G_2]}(V_1 \otimes V_2, W_1 \otimes W_2)\). As a consequence,
\[ \dim_C(\text{Hom}_{CG_1}(V_1, W_1) \otimes \text{Hom}_{CG_2}(V_2, W_2)) \]
(11.635)

Applying (4.1) to \(G_1, V_1\) and \(W_1\) instead of \(G, V_1\) and \(W_2\), we obtain
\[ (\chi_{V_1}, \chi_{W_1})_{G_1} = \dim_C(\text{Hom}_{CG_1}(V_1, W_1)). \]

Applying (4.1) to \(G_2, V_2\) and \(W_2\) instead of \(G, V_1\) and \(W_2\), we obtain
\[ (\chi_{V_2}, \chi_{W_2})_{G_2} = \dim_C(\text{Hom}_{CG_2}(V_2, W_2)) \]
(11.636)

Applying (4.1) to \(G_1 \times G_2, V_1 \otimes V_2\) and \(W_1 \otimes W_2\) instead of \(G, V_1\) and \(W_2\), we obtain
\[ (\chi_{V_1}, \chi_{W_1})_{G_1 \times G_2} = \dim_C(\text{Hom}_{[G_1 \times G_2]}(V_1 \otimes V_2, W_1 \otimes W_2)) \]
\[ = \dim_C(\text{Hom}_{CG_1}(V_1, W_1) \otimes \text{Hom}_{CG_2}(V_2, W_2)) \]
\[ = \dim_C(\text{Hom}_{CG_1}(V_1, W_1)) \cdot \dim_C(\text{Hom}_{CG_2}(V_2, W_2)) \]
(11.635)

\[ = (\chi_{V_1}, \chi_{W_1})_{G_1} \cdot (\chi_{V_2}, \chi_{W_2})_{G_2}. \]

This proves (4.2). Thus, Exercise 4.9(b) is solved.

11.78. Solution to Exercise 4.10. Solution to Exercise 4.10. If \(A\) and \(B\) are two algebras, \(P\) is a \((B, A)\)-bimodule and \(Q\) is a left \(B\)-module, then \(\text{Hom}_B(P, Q)\) is a left \(A\)-module. Consequently, \(\text{Hom}_{CG}(CG, U)\) is
a left \(CG\)-module, and \(\text{Hom}_{CG[H/(H \cap K)]}(C[G/K], U^{H \cap K})\) is a left \(C[G/K]\)-module. Exercise 4.4 yields that
the \(CG\)-module \(\text{Hom}_{CG}(CG, U)\) is isomorphic to \(\text{Ind}_H^G U\). In other words,
\[ \text{Hom}_{CG}(CG, U) \cong \text{Ind}_H^G U \]
(11.638)\]
as \(CG\)-modules.

Also, Exercise 4.4 (applied to \(G/K, H/(H \cap K)\) and \(U^{H \cap K}\) instead of \(G, H\) and \(U\)) yields that
the \(C[G/K]\)-module \(\text{Hom}_{CG[H/(H \cap K)]}(C[G/K], U^{H \cap K})\) is isomorphic to \(\text{Ind}^{G/K}_H(U^{H \cap K})\). In other words,
\[ \text{Hom}_{CG[H/(H \cap K)]}(C[G/K], U^{H \cap K}) \cong \text{Ind}^{G/K}_H(U^{H \cap K}) \]
(11.639)\]
as \(C[G/K]\)-modules.
But we also have

\[(11.640) \quad (\text{Hom}_C (\mathbb{C}G, U))^K \cong \text{Hom}_{\mathbb{C}[H/(H \cap K)]} (\mathbb{C}[G/K], U^{H \cap K}) \quad \text{as } \mathbb{C}[G/K]-\text{modules.}\]

Here is just a brief sketch of the proof of (11.640): Let \(\pi\) be the canonical projection \(G \to G/K\), and let \(\mathbb{C}[\pi]\) be the \(\mathbb{C}\)-linear map \(\mathbb{C}G \to \mathbb{C}[G/K]\) obtained by \(\mathbb{C}\)-linearly extending \(\pi\). Clearly, \(\mathbb{C}[\pi]\) is a surjective \(\mathbb{C}\)-linear \(\mathbb{C}\)-algebra homomorphism, and we have

\[(11.641) \quad \ker (\mathbb{C}[\pi]) = \langle g - g' \mid g \in G, \ g' \in G, \ gK = g'K \rangle\]

\[(11.642) \quad = \langle g - gk \mid g \in G, \ k \in K \rangle\]

\[(11.643) \quad = \langle g - g' \mid g \in G, \ g' \in G, \ Kg = Kg' \rangle\]

\[(11.644) \quad = \langle g - kg \mid g \in G, \ k \in K \rangle\]

(where the \(\langle\rangle\) brackets stand for “\(\mathbb{C}\)-span”).

Let \(f\) be an element of \((\text{Hom}_C (\mathbb{C}G, U))^K\). Then, \(f \in (\text{Hom}_C (\mathbb{C}G, U))^K \subset \text{Hom}_C (\mathbb{C}G, U)\) is \(\mathbb{C}\)-linear, and it can easily be shown that \(\ker (\mathbb{C}[\pi]) \subset \ker f\) \(^{526}\). Hence, the map \(f\) factors through the surjective map \(\mathbb{C}[\pi]\). The resulting map \(\mathbb{C}[G/K] \to U\) is a \(\mathbb{C}\)-module homomorphism (since \(\mathbb{C}[\pi]\) and \(f\) were both \(\mathbb{C}\)-linear), and it is easy to see that its image (i.e., the image of \(f\)) is contained in \(U^{H \cap K}\) \(^{527}\). Hence, this map factors through the canonical inclusion \(U^{H \cap K} \to U\), leaving behind a map \(\mathbb{C}[G/K] \to U^{H \cap K}\) which we denote by \(\Phi (f)\). This resulting map \(\Phi (f)\) turns out to be \(\mathbb{C}[H/(H \cap K)]\)-linear \(^{528}\), thus belongs to \(\text{Hom}_{\mathbb{C}[H/(H \cap K)]} (\mathbb{C}[G/K], U^{H \cap K})\). Since this holds for every \(f\), we thus obtain a \(\mathbb{C}\)-linear map

\[\text{Hom}_C (\mathbb{C}G, U)^K \to \text{Hom}_{\mathbb{C}[H/(H \cap K)]} (\mathbb{C}[G/K], U^{H \cap K}), \quad f \mapsto \Phi (f).\]

This map is invertible\(^{529}\) and \(\mathbb{C}[G/K]\)-linear\(^{530}\), therefore an isomorphism of \(\mathbb{C}[G/K]\)-modules. This proves (11.640). All steps that were left to the reader are straightforward.

Now,

\[
\begin{pmatrix}
\text{Ind}_H^G U \\
\cong \text{Hom}_C (\mathbb{C}G, U) \text{ as } \mathbb{C}\text{-modules (by (11.638))}
\end{pmatrix}^K
\cong (\text{Hom}_C (\mathbb{C}G, U))^K
\cong \text{Hom}_{\mathbb{C}[H/(H \cap K)]} (\mathbb{C}[G/K], U^{H \cap K}) \quad \text{(by (11.640))}
\cong \text{Ind}_{H/(H \cap K)}^{G/K} (U^{H \cap K}) \quad \text{(by (11.639))}
\]

as \(\mathbb{C}[G/K]\)-modules. This solves Exercise 4.10.

\(^{526}\) Indeed, \(f \in (\text{Hom}_C (\mathbb{C}G, U))^K\). Hence, every \(k \in K\) satisfies \(kf = f\). Thus, every \(g \in G\) and \(k \in K\) satisfy \((kf)(g) = f(g)\). But since \((kf)(g) = f(gk)\) (by the definition of the action of \(G\) on \(\text{Hom}_C (\mathbb{C}G, U)\)), this becomes \(f(gk) = f(g)\), so that \(f(g - gk) = 0\). The map \(f\) therefore annihilates \(g - gk\) for all \(g \in G\) and \(k \in K\). Due to (11.642), this yields \(f(\ker (\mathbb{C}[\pi])) = 0\), so that \(\ker (\mathbb{C}[\pi]) \subset \ker f\), qed.

\(^{527}\) Proof. We want to show that the image of \(f\) is contained in \(U^{H \cap K}\).

For this, it clearly suffices to prove that \(f(g) \in U^{H \cap K}\) for every \(g \in G\). So fix \(g \in G\). Let \(k \in H \cap K\). Then, \(k \cdot f (g) = f(kg)\) (since \(f\) is \(\mathbb{C}\)-linear and \(k \in H\)) and \(f(g) = f(kg)\) (since \(k \in K\), so that (11.644) yields \(g - kg \in \ker (\mathbb{C}[\pi]) \subset \ker f\) and thus \(f(g - kg) = 0\)). Hence, \(k \cdot f (g) = f(kg) = f(g)\). Since this has been proven for all \(k \in H \cap K\), we thus have \(f(g) \in U^{H \cap K}\), qed.

\(^{528}\) This is straightforward to see (everything in sight is \(\mathbb{C}\)-linear).

\(^{529}\) In fact, defining the inverse is very easy (just send every map \(g \in \text{Hom}_{\mathbb{C}[H/(H \cap K)]} (\mathbb{C}[G/K], U^{H \cap K})\) to the composition \(\mathbb{C}[G/K] \to U^{H \cap K}\) via \(\text{inclusion}\)). Checking that these maps are mutually inverse is also straightforward.

\(^{530}\) Indeed, it is easier to check that its inverse is \(\mathbb{C}[G/K]\)-linear (this can be proven by straightforward computations).
11.79. **Solution to Exercise 4.11.** Solution to Exercise 4.11. In the following, we will write $g$ for the element $t_g$ of $CG$ whenever $g$ is an element of $G$. This is a relatively common abuse of notation, and it is harmless because the map $G \rightarrow CG$, $g \mapsto t_g$ is an injective homomorphism of multiplicative monoids (so $t_{gh} = t_g t_h$ and $t_1 = 1$, which means that we won’t run into ambiguities denoting $t_g$ by $g$) and because every $CG$-module $M$, every $m \in M$ and every $g \in G$ satisfy $gm = t_g m$. We will do the same abuse of notation for elements of $H$ and of $K$.

Inflation does not change the underlying $\mathbb{C}$-vector space of a representation. Thus, $\text{Infl}_{G/K}^G \text{Ind}_{H/K}^H V = \text{Ind}_{H/K}^H V$ as $\mathbb{C}$-vector spaces. For the same reason, $\text{Infl}_{H/K}^H V = V$ as $\mathbb{C}$-vector spaces.

Let $\pi_G$ be the canonical projection map $G \rightarrow G/K$. This gives rise to a surjective $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\pi_G] : CG \rightarrow \mathbb{C}[G/K]$ (which sends every $g \in G$ to $\pi_G(g)$). We now define a $\mathbb{C}$-linear map

$$\beta : CG \otimes_{CH} \text{Infl}_{H/K}^H V \rightarrow CG \otimes_{C[H/K]} V,$$

$$s \otimes_{CH} v \mapsto (\mathbb{C}[\pi_G]) (s) \otimes_{C[H/K]} v.$$

This is easily seen to be well-defined (using the universal property of the tensor product and the observation that every $t \in CH$ satisfies $(\mathbb{C}[\pi_G])(t) \in C[H/K]$).

We also want to define a map $\alpha$ in the opposite direction, but this will require some more work. First, for every $g \in G$, we define a $\mathbb{C}$-linear map

$$a_g : V \rightarrow CG \otimes_{CH} \text{Infl}_{H/K}^H V,$$

$$v \mapsto g \otimes_{CH} v.$$

It is easily seen that if $g_1$ and $g_2$ are two elements of $G$ satisfying $\pi_G(g_1) = \pi_G(g_2)$, then $a_{g_1} = a_{g_2}$. In other words, the map $a_g$ depends only on $\pi_G(g)$ rather than on $g$ itself. Thus, we can define a $\mathbb{C}$-linear map $\tilde{a}_p : V \rightarrow CG \otimes_{CH} \text{Infl}_{H/K}^H V$ for every $p \in G/K$ by choosing any $g \in G$ satisfying $\pi_G(g) = p$, and then setting $\tilde{a}_p = a_g$; the resulting map $\tilde{a}_p$ does not depend on the choice of $g$.

Hence, we have defined a $\mathbb{C}$-linear map $\tilde{a}_p : V \rightarrow CG \otimes_{CH} \text{Infl}_{H/K}^H V$ for every $p \in G/K$. In other words, we have defined an element $\tilde{a}_p$ of $\text{Hom}_C \left(V, CG \otimes_{CH} \text{Infl}_{H/K}^H V\right)$ for every $p \in G/K$. Hence, we can define a $\mathbb{C}$-linear map

$$A : \mathbb{C}[G/K] \rightarrow \text{Hom}_C \left(V, CG \otimes_{CH} \text{Infl}_{H/K}^H V\right),$$

$$p \mapsto \tilde{a}_p \quad \text{for every } p \in G/K$$

(because in order to define a $\mathbb{C}$-linear map from the $\mathbb{C}$-vector space $\mathbb{C}[G/K]$, it is enough to assign its values on the basis $G/K$). Using this map $A$, we can now define a $\mathbb{C}$-linear map

$$\alpha : \mathbb{C}[G/K] \otimes_{C[H/K]} V \rightarrow CG \otimes_{CH} \text{Infl}_{H/K}^H V,$$

$$s \otimes_{C[H/K]} v \mapsto (A(s))(v).$$

---

531** Indeed, let $g_1$ and $g_2$ be two elements of $G$ satisfying $\pi_G(g_1) = \pi_G(g_2)$. Then, $g_1 \in g_2 K$, so that there exists some $k \in K$ such that $g_1 = g_2 k$. Consider this $k$. Then, every $v \in V$ satisfies

$$a_{g_1} (v) = g_1 \otimes_{CH} v \quad \text{(by the definition of } a_{g_1})$$

$$= g_2 k \otimes_{CH} v = g_2 \otimes_{CH} v \quad \text{(since } k \in K \text{ acts trivially on } \text{Infl}_{H/K}^H V)$$

$$= g_2 \otimes_{CH} v = a_{g_2} (v) \quad \text{(by the definition of } a_{g_2}).$$

Thus, $a_{g_1} = a_{g_2}$, qed.
It is again easy to check that this is well-defined\footnote{In order to prove this well-definedness, we have to check that $\alpha \circ \beta = \text{id}$ (indeed, this only needs to be proven on tensors of the form $g \otimes_{CH} v$ for $g \in G$ and $v \in \text{Inf}^H_{H/K} V$, because the $\mathbb{C}$-vector space $\mathbb{C}G \otimes_{CH} \text{Inf}^H_{H/K} V$ is spanned by such tensors; but on such tensors it is very easy to check) and $\beta \circ \alpha = \text{id}$ (using a similar argument). Thus, the maps $\alpha$ and $\beta$ are mutually inverse, and therefore $\beta$ is invertible. We can regard $\beta$ as a map from $\text{Ind}^G_H \mathbb{C}G \otimes_{CH} \text{Inf}^H_{H/K} V$ to $\text{Ind}^G_H \mathbb{C}G \otimes_{CH} \mathbb{C}[G/K]$, since $\text{Ind}^G_H \mathbb{C}G \otimes_{CH} \text{Inf}^H_{H/K} V = \mathbb{C}G \otimes_{CH} \text{Inf}^H_{H/K} V$ and $\text{Ind}^G_H \mathbb{C}G \otimes_{CH} \mathbb{C}[G/K] = \mathbb{C}[G/K] \otimes_{CH} \mathbb{C}[H/K]$ as $\mathbb{C}$-vector spaces, and it is easy to verify that $\beta$ becomes a $\mathbb{C}G$-module homomorphism when regarded this way. Thus, $\beta$ is an invertible $\mathbb{C}G$-module homomorphism from $\text{Ind}^G_H \mathbb{C}G \otimes_{CH} \text{Inf}^H_{H/K} V$ to $\text{Ind}^G_H \mathbb{C}G \otimes_{CH} \mathbb{C}[G/K]$, hence a $\mathbb{C}G$-module isomorphism.} \footnote{Let $v \in V$. Let $j \in K$. Then, the map $K \rightarrow K$, $s \mapsto js$ is a bijection (since $j \in K$ and since $K$ is a group). Now,}

\begin{equation}
\sum_{j \in K} j v = \frac{1}{|K|} \sum_{j \in K} j v \in V^K.
\end{equation}

\footnote{This vanishes on $IV_{V,K}$. In order to do so, we only need to check that $\psi(v - kv) = 0$ for all $k \in K$ and $v \in V$ (since $IV_{V,K}$ is spanned by all $v - kv$ with $k \in K$ and $v \in V$). But this follows from the fact that all $k \in K$}

Thus, we can define a map $\Phi : G/K \rightarrow \text{End}^G_K \mathbb{C}G$ such that $\Phi(g)(v) = g v$ for all $v \in V$. Hence, all that remains to be done is proving that $\Phi$ is bijective. We will prove that $\Phi$ is a $G$-module isomorphism.

\begin{equation}
\text{Ind}^G_H \mathbb{C}G \otimes_{CH} \text{Inf}^H_{H/K} V \cong \text{Inf}^G_H \mathbb{C}G \otimes_{CH} \text{Inf}^H_{H/K} V \cong \mathbb{C}G \otimes_{CH} \text{Inf}^H_{H/K} V \cong \mathbb{C}[G/K] \otimes_{CH} \mathbb{C}[H/K] \cong \mathbb{C}[G/K] \otimes_{CH} \mathbb{C}[H/K],
\end{equation}

as $\mathbb{C}$-vector spaces. This is an $\mathbb{C}G$-module isomorphism.

\footnote{We want to show that $\psi$ vanishes on $IV_{V,K}$. In order to do so, we only need to check that $\psi(v - kv) = 0$ for all $k \in K$ and $v \in V$ (since $IV_{V,K}$ is spanned by all $v - kv$ with $k \in K$ and $v \in V$). But this follows from the fact that all $k \in K$}

11.80. \textbf{Solution to Exercise 4.12.} Solution to Exercise 4.12. (a) It is clearly easy to show that $g(v - k v) \in IV_{V,K}$ for all $g \in G$, $k \in K$ and $v \in V$. But if $g \in G$, $k \in K$ and $v \in V$, then $g k$ has the form $g k = k' g$ for some $k' \in K$ (since $K \triangleleft G$), and thus we have $g(v - k v) = g v - g k v = g v - k' g v \in IV_{V,K}$ (by the definition of $IV_{V,K}$, since $k' \in K$ and $g v \in V$). Hence, part (a) of the exercise is solved.

(b) For every $v \in V$, let $\pi$ denote the projection of $v$ onto the quotient space $V/IV_{V,K} = V_K$.

We can define a map $\Phi : V^K \rightarrow V_K$ by sending every $v \in V^K$ to $\pi v \in V_K$. This map $\Phi$ is easily seen to be a $G$-module isomorphism $\text{Ind}_{\mathbb{C}G}^{\mathbb{C}G} (V^K) \rightarrow V_K$. If we can show that $\Phi$ is also bijective, then it will follow that $\Phi$ is a $\mathbb{C}G$-module isomorphism $\text{Ind}_{\mathbb{C}G}^{\mathbb{C}G} (V^K) \rightarrow V_K$, whence part (b) of the exercise will be solved. Hence, all that remains to be done is proving that $\Phi$ is bijective.

Let us construct an inverse map to $\Phi$. First, let us notice that every $v \in V$ satisfies

\begin{equation}
\frac{1}{|K|} \sum_{j \in K} j v \in V^K.
\end{equation}

Thus, we can define a map $\psi : V \rightarrow V^K$ by sending every $v \in V$ to $\frac{1}{|K|} \sum_{j \in K} j v$. This map $\psi$ is $\mathbb{C}$-linear and vanishes on $IV_{V,K}$.

\footnote{We want to show that $\psi$ vanishes on $IV_{V,K}$. In order to do so, we only need to check that $\psi(v - kv) = 0$ for all $k \in K$ and $v \in V$ (since $IV_{V,K}$ is spanned by all $v - kv$ with $k \in K$ and $v \in V$). But this follows from the fact that all $k \in K$}

\begin{equation}
\sum_{j \in K} j v = \frac{1}{|K|} \sum_{j \in K} j v.
\end{equation}

\footnote{We want to show that $\psi$ vanishes on $IV_{V,K}$. In order to do so, we only need to check that $\psi(v - kv) = 0$ for all $k \in K$ and $v \in V$ (since $IV_{V,K}$ is spanned by all $v - kv$ with $k \in K$ and $v \in V$). But this follows from the fact that all $k \in K$}

\begin{equation}
\sum_{j \in K} j v = \frac{1}{|K|} \sum_{j \in K} j v.
\end{equation}

\footnote{We want to show that $\psi$ vanishes on $IV_{V,K}$. In order to do so, we only need to check that $\psi(v - kv) = 0$ for all $k \in K$ and $v \in V$ (since $IV_{V,K}$ is spanned by all $v - kv$ with $k \in K$ and $v \in V$). But this follows from the fact that all $k \in K$}
Thus, $\Psi$ and $K$ convention: Whenever $11.81$. Solution to Exercise 4.14. Solution to Exercise 4.14. In the following, we will use the following convention: Whenever $K$ is a group, and $k$ is an element of $K$, we shall write $k$ for the element $t_k$ of $\mathbb{C}K$. This is a relatively common abuse of notation, and it is harmless because the map $K \to \mathbb{C}K$, $k \mapsto t_k$ is an injective homomorphism of multiplicative monoids (so $t_{gh} = t_g t_h$ and $t_1 = 1$, which means that we won’t run into ambiguities denoting $t_k$ by $k$) and because every $\mathbb{C}K$-module $M$, every $m \in M$ and every $k \in K$ satisfy $km = t_k m$.

Proof. Let $w \in V_K$. Then, there exists some $v \in V$ such that $w = \tau$. Consider this $v$. Since $\Psi$ was defined as a quotient of the map $\psi$, we have $\Psi(\tau) = \psi(v)$. Now,

$$
(\Phi \circ \Psi) \left( \begin{array}{c} w \\ \tau \end{array} \right) = (\Phi \circ \Psi) (\tau) = \Phi \left( \begin{array}{c} \psi(v) \\ = \psi(v) \end{array} \right) = \Phi \left( \begin{array}{c} \psi(v) \\ \tau \end{array} \right) = \Phi \left( \begin{array}{c} 1 \\ j \in K \end{array} \right) jv
$$

(by the definition of $\Phi$)

$$
= \frac{1}{|K|} \sum_{j \in K} jv = \frac{1}{|K|} \sum_{j \in K} jv - \frac{1}{|K|} \sum_{j \in K} jkv = \frac{1}{|K|} \sum_{j \in K} jv - \frac{1}{|K|} \sum_{j \in K} jv
$$

(here, we have substituted $j$ for $jk$ in the second sum, since the map $K \to K$, $j \mapsto jk$ is a bijection (because $k \in K$ and because $K$ is a group))

$$
= 0.
$$

Proof. For every $v \in V_K$, we have

$$
(\Psi \circ \Phi) (v) = \Psi \left( \begin{array}{c} \Phi (v) \\ \tau \end{array} \right) = \Psi (\tau) = \psi(v)
$$

(by the definition of $\Phi$)

$$
= \frac{1}{|K|} \sum_{j \in K} jv
$$

(since $v \in V_K$ and $j \in K$)

$$
= \frac{1}{|K|} \sum_{j \in K} v = \frac{1}{|K|} |K| \cdot v = v.
$$

Thus, $\Psi \circ \Phi = \text{id}$, qed.
Let us first notice a trivial fact: If $K$ is a finite group, and if $f: K \to \mathbb{C}$ is any function, then we have the following equivalence:

\[
(f \in R_C(K)) \iff (f \text{ is a class function on } K)
\]

\[
\iff (f \text{ is constant on } K\text{-conjugacy classes})
\]

\[
(11.645) \iff (\text{any two conjugate elements } k \text{ and } k' \text{ of } K \text{ satisfy } f(k) = f(k')).
\]

(a) Let $f \in R_C(H)$.

We can apply (11.645) to $K = H$. As a consequence, we obtain the following equivalence:

\[
(f \in R_C(H)) \iff (\text{any two conjugate elements } k \text{ and } k' \text{ of } H \text{ satisfy } f(k) = f(k')).
\]

Hence,

\[
(11.646) \text{ any two conjugate elements } k \text{ and } k' \text{ of } H \text{ satisfy } f(k) = f(k')
\]

(because we know that $f \in R_C(H)$).

Now, let $g$ and $g'$ be two conjugate elements of $G$. Then, there exists a $p \in G$ such that $g' = pgp^{-1}$ (since $g$ and $g'$ are conjugate). Consider this $p$. 
Applying the map $\text{Ind}_\rho f$ to both sides of the equality $g' = pgp^{-1}$, we obtain

$$\left(\text{Ind}_\rho f\right) (g') = \left(\text{Ind}_\rho f\right) (pgp^{-1}) = \frac{1}{|H|} \sum_{(h,k) \in H \times G; k\rho(h)k^{-1} = pgp^{-1}} f(h) = \frac{1}{|H|} \sum_{h \in H} \sum_{k\rho(h)k^{-1} = pgp^{-1}} f(h)$$

(by the definition of $\text{Ind}_\rho f$)

$$= \frac{1}{|H|} \sum_{h \in H} \sum_{k\rho(h)(pk)^{-1} = pgp^{-1}} f(h) = \frac{1}{|H|} \sum_{h \in H} \sum_{pk\rho(h)k^{-1} = pgp^{-1}} f(h)$$

(here, we have substituted $pk$ for $k$ in the sum, since the map $G \to G$, $k \mapsto pk$ is a bijection)

$$= \frac{1}{|H|} \sum_{h \in H} \sum_{k\rho(h)(pk)^{-1} = pgp^{-1}} f(h) = \frac{1}{|H|} \sum_{h \in H} \sum_{pk\rho(h)k^{-1} = pgp^{-1}} f(h)$$

(since $pk\rho(h)k^{-1} = pgp^{-1}$ is equivalent to $pk\rho(h)k^{-1} = pg$)

$$= \frac{1}{|H|} \sum_{h \in H} f(h) = (\text{Ind}_\rho f) (g)$$

(since the definition of $\text{Ind}_\rho f$ yields $(\text{Ind}_\rho f) (g) = \frac{1}{|H|} \sum_{(h,k) \in H \times G; k\rho(h)k^{-1} = g} f(h)$).

Let us now forget that we fixed $g$ and $g'$. We thus have shown that any two conjugate elements $g$ and $g'$ of $G$ satisfy $(\text{Ind}_\rho f) (g) = (\text{Ind}_\rho f) (g')$. Renaming $g$ and $g'$ as $k$ and $k'$ in this statement, we obtain the following: Any two conjugate elements $k$ and $k'$ of $G$ satisfy $(\text{Ind}_\rho f) (k) = (\text{Ind}_\rho f) (k')$.

But (11.645) (applied to $G$ and $\text{Ind}_\rho f$ instead of $K$ and $f$) yields the following equivalence:

$$(\text{Ind}_\rho f \in R_C (G)) \iff (\text{any two conjugate elements } k \text{ and } k' \text{ of } G \text{ satisfy } (\text{Ind}_\rho f) (k) = (\text{Ind}_\rho f) (k')).$$

Thus, $\text{Ind}_\rho f \in R_C (G)$ (because we know that any two conjugate elements $k$ and $k'$ of $G$ satisfy $(\text{Ind}_\rho f) (k) = (\text{Ind}_\rho f) (k')$). This solves Exercise 4.14(a).

(b) Let us first introduce an elementary (but apocryphal) notion from linear algebra: the notion of finite dual generating systems.

**Definition 11.14.** Let $K$ be a commutative ring. Let $V$ be a $K$-module. A **finite dual generating system** for $V$ means a triple $(I, (a_i)_{i \in I}, (f_i)_{i \in I})$, where

- $I$ is a finite set;
- $(a_i)_{i \in I}$ is a family of elements of $V$;
- $(f_i)_{i \in I}$ is a family of elements of $V^*$ (where $V^*$ means $\text{Hom}_K (V, K)$)

such that every $v \in V$ satisfies $v = \sum_{i \in I} f_i (v) a_i$. 


In the following, we shall only use finite dual generating systems in the case when $\mathbb{K}$ is a field; nevertheless, they are more useful in the general case.

The first question one might ask about finite dual generating systems for $V$ is when they exist. The answer is very simple when $\mathbb{K}$ is a field:

**Proposition 11.15.** Let $\mathbb{K}$ be a field. Let $V$ be a $\mathbb{K}$-vector space. Then, a finite dual generating system for $V$ exists if and only if the vector space $V$ is finite-dimensional.

**Proof of Proposition 11.15.** Proposition 11.15 is an “if and only if” statement. Hence, in order to prove Proposition 11.15, it is sufficient to verify the following two claims:

**Claim 1:** If a finite dual generating system for $V$ exists, then the vector space $V$ is finite-dimensional.

**Claim 2:** If the vector space $V$ is finite-dimensional, then a finite dual generating system for $V$ exists.

Let us now prove these two claims.

**Proof of Claim 1.** Assume that a finite dual generating system for $V$ exists. Let $(I, (a_i)_{i \in I}, (f_i)_{i \in I})$ be such a finite dual generating system for $V$.

We know that $I$ is a finite set, that $(a_i)_{i \in I}$ is a family of elements of $V$, and that every $v \in V$ satisfies $v = \sum_{i \in I} f_i(v) a_i$. (Indeed, this is part of what it means for $(I, (a_i)_{i \in I}, (f_i)_{i \in I})$ to be a finite dual generating system for $V$.)

Now, every $v \in V$ satisfies $v = \sum_{i \in I} f_i(v) a_i \in \sum_{i \in I} \mathbb{K} a_i$. Thus, $V \subset \sum_{i \in I} \mathbb{K} a_i$. Combined with the (obvious) inclusion $\sum_{i \in I} \mathbb{K} a_i \subset V$, this yields $V = \sum_{i \in I} \mathbb{K} a_i$. But the vector space $\sum_{i \in I} \mathbb{K} a_i$ is finite-dimensional (since $I$ is a finite set). In other words, the vector space $V$ is finite-dimensional (since $V = \sum_{i \in I} \mathbb{K} a_i$). This proves Claim 1.

**Proof of Claim 2.** Assume that $V$ is a finite-dimensional vector space. Then, $V$ has a finite basis. Let $(e_i)_{i \in I}$ be such a basis. Thus, $I$ is a finite set, and $(e_i)_{i \in I}$ is a basis of the $\mathbb{K}$-vector space $V$. Let $(e^*_i)_{i \in I}$ be the basis of $V^*$ dual to the basis $(e_i)_{i \in I}$ of $V$. (This is well-defined, since $V$ is finite-dimensional.)

We know that $(e^*_i)_{i \in I}$ is the basis of $V^*$ dual to the basis $(e_i)_{i \in I}$ of $V$. Thus,

\[(11.647) \quad e^*_i \left( \sum_{j \in I} \lambda_j e_j \right) = \lambda_i \quad \text{for all} \ i \in I \text{ and } \ (\lambda_j)_{j \in I} \in \mathbb{K}^I.\]

(Indeed, this is one of the ways to define a dual basis.)

Now, every $v \in V$ satisfies $v = \sum_{i \in I} e^*_i(v) e_i \quad \text{537.}$

So we know that $(I, (e_i)_{i \in I}, (e^*_i)_{i \in I})$ is a triple such that

- $I$ is a finite set;
- $(e_i)_{i \in I}$ is a family of elements of $V$;
- $(e^*_i)_{i \in I}$ is a family of elements of $V^*$ (where $V^*$ means $\text{Hom}_\mathbb{K}(V, \mathbb{K})$)

such that every $v \in V$ satisfies $v = \sum_{i \in I} e^*_i(v) e_i$. In other words, $(I, (e_i)_{i \in I}, (e^*_i)_{i \in I})$ is a finite dual generating system for $V$ (by the definition of a “finite dual generating system”). Thus, a finite dual generating system for $V$ exists (namely, $\left( I, (e_i)_{i \in I}, (e^*_i)_{i \in I} \right)$). This proves Claim 2.

Now, both Claim 1 and Claim 2 are proven. Hence, the proof of Proposition 11.15 is complete. \(\square\)

\text{537} Proof. Let $v \in V$. Then, we can write $v$ in the form $v = \sum_{i \in I} \lambda_i e_i$ for some family $(\lambda_i)_{i \in I} \in \mathbb{K}^I$ (since $(e_i)_{i \in I}$ is a basis of $V$). Consider this family $(\lambda_i)_{i \in I}$. Now, $v = \sum_{i \in I} \lambda_i e_i = \sum_{i \in I} \lambda_j e_j$ (here, we have renamed the summation index $i$ as $j$ in the sum). For each $i \in I$, we now have

\[ e^*_i \left( \sum_{j \in I} \lambda_j e_j \right) = e^*_i \left( \sum_{j \in I} \lambda_j e_j \right) = \lambda_i \quad \text{(by (11.647)).} \]

Thus, $\sum_{i \in I} e^*_i(v) e_i = \sum_{i \in I} \lambda_i e_i$. Compared with $v = \sum_{i \in I} \lambda_i e_i$, this yields $v = \sum_{i \in I} e^*_i(v) e_i$, qed.
[Remark: The notion of finite dual generating system for $V$ is more versatile than the notion of a finite basis of $V$. One difference between these notions is that all bases of $V$ have the same size, while the set $I$ in a finite dual generating system $(I, (a_i)_{i \in I}, (f_i)_{i \in I})$ of $V$ can have any (finite) size $\geq \dim V$. Another difference manifests itself in the general setting when $K$ is a commutative ring, not necessarily a field. In this generality, a finite basis of $V$ exists if and only if $V$ is a finite free $K$-module (this is the definition of a finite free $K$-module), whereas a finite dual generating system for $V$ exists if and only if $V$ is a finitely generated projective $K$-module. Projective $K$-modules are a more frequent occurrence in commutative algebra than free $K$-modules, and in the absence of a finite basis, a finite dual generating system is the thing that comes closest to allowing “computing in a basis”.]

One significant application of finite dual generating systems is computing traces of endomorphisms:

**Proposition 11.16.** Let $K$ be a field. Let $V$ be a finite-dimensional $K$-vector space. Let $(I, (a_i)_{i \in I}, (f_i)_{i \in I})$ be a finite dual generating system for $V$. Let $T : V \to V$ be a $K$-linear map. Then,

$$\text{trace } T = \sum_{i \in I} f_i(Ta_i).$$

Proposition 11.16 can be easily proven directly, but let us take a slight detour and derive it from the following more general fact:

**Proposition 11.17.** Let $K$ be a commutative ring. Let $V$ be a $K$-module. Let $(I, (a_i)_{i \in I}, (f_i)_{i \in I})$ be a finite dual generating system for $V$. Let $(J, (b_j)_{j \in J}, (g_j)_{j \in J})$ be a further finite dual generating system for $V$. Let $T : V \to V$ be a $K$-linear map. Then,

$$\sum_{i \in I} f_i(Ta_i) = \sum_{j \in J} g_j(Tb_j).$$

**Proof of Proposition 11.17.** We know that $(I, (a_i)_{i \in I}, (f_i)_{i \in I})$ is a finite dual generating system for $V$. In other words, $(I, (a_i)_{i \in I}, (f_i)_{i \in I})$ is a triple such that

- $I$ is a finite set;
- $(a_i)_{i \in I}$ is a family of elements of $V$;
- $(f_i)_{i \in I}$ is a family of elements of $V^*$ (where $V^*$ means $\text{Hom}_K(V, K)$)

such that every $v \in V$ satisfies

$$(11.648) \quad v = \sum_{i \in I} f_i(v) a_i$$

We know that $(J, (b_j)_{j \in J}, (g_j)_{j \in J})$ is a finite dual generating system for $V$. In other words, $(J, (b_j)_{j \in J}, (g_j)_{j \in J})$ is a triple such that

- $J$ is a finite set;
- $(b_j)_{j \in J}$ is a family of elements of $V$;
- $(g_j)_{j \in J}$ is a family of elements of $V^*$ (where $V^*$ means $\text{Hom}_K(V, K)$)

such that every $v \in V$ satisfies

$$(11.649) \quad v = \sum_{j \in J} g_j(v) b_j.$$
Now,

\[ \sum_{i \in I} f_i \begin{pmatrix} T \alpha_i \\ \vdots \\ T \beta_j \end{pmatrix} = \sum_{i \in I} g_j (T a_i) b_j \]

(by (11.649), applied to \( v = T a_i \))

\[ = \sum_{i \in I} f_i \left( \sum_{j \in J} g_j (T a_i) b_j \right) = \sum_{j \in J} g_j (T a_i) \left( \sum_{i \in I} f_i (b_j) \right) \]

(since the map \( f_i \) is \( \mathbb{K} \)-linear)

\[ = \sum_{j \in J} \sum_{i \in I} f_i (b_j) g_j (T a_i) . \]

Compared with

\[ \sum_{j \in J} g_j \begin{pmatrix} \vdots \\ b_j \end{pmatrix} = \sum_{i \in I} f_i (b_j) a_i \]

(by (11.648), applied to \( v = b_j \))

\[ = \sum_{j \in J} g_j \left( \sum_{i \in I} f_i (b_j) a_i \right) = \sum_{j \in J} \sum_{i \in I} f_i (b_j) g_j (T a_i) \]

(since the map \( T \) is \( \mathbb{K} \)-linear)

this yields \( \sum_{i \in I} f_i (T a_i) = \sum_{j \in J} g_j (T b_j) \). This proves Proposition 11.17. \( \square \)

\textbf{Proof of Proposition 11.16.} The vector space \( V \) has a finite basis (since it is finite-dimensional). Let \( (e_1, e_2, \ldots, e_n) \) be such a basis. Let \( (m_{i,j})_{1 \leq i, j \leq n} \in \mathbb{K}^{n \times n} \) be the matrix which represents the map \( T : V \to V \) with respect to this basis \( (e_1, e_2, \ldots, e_n) \) of \( V \). Then,

(11.650) \[ T e_j = \sum_{i=1}^{n} m_{i,j} e_i \quad \text{for every } j \in \{1, 2, \ldots, n\} \]

(due to the definition of “the matrix which represents the map \( T : V \to V \) with respect to this basis \( (e_1, e_2, \ldots, e_n) \) of \( V \)).

Let \( (e_1^*, e_2^*, \ldots, e_n^*) \) be the basis of \( V^* \) dual to the basis \( (e_1, e_2, \ldots, e_n) \) of \( V \). (This is well-defined, since \( V \) is finite-dimensional.) Thus,

(11.651) \[ e_k^* \left( \sum_{i=1}^{n} \lambda_i e_i \right) = \lambda_k \quad \text{for all } k \in \{1, 2, \ldots, n\} \text{ and } (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n . \]

(This follows immediately from the definition of a “dual basis”.) Now, every \( k \in \{1, 2, \ldots, n\} \) and \( j \in \{1, 2, \ldots, n\} \) satisfy

(11.652) \[ e_k^* (T e_j) = m_{k,j} . \]
Now, every \( v \in V \) satisfies \( v = \sum_{i=1}^{n} e_i^+ (v) e_i \). So we know that \( \{1, 2, \ldots, n\}, (e_i)_{i \in \{1, 2, \ldots, n\}}, (e_i^+)_{i \in \{1, 2, \ldots, n\}} \) is a triple such that

- \( \{1, 2, \ldots, n\} \) is a finite set;
- \( (e_i)_{i \in \{1, 2, \ldots, n\}} \) is a family of elements of \( V \);
- \( (e_i^+)_{i \in \{1, 2, \ldots, n\}} \) is a family of elements of \( V^* \) (where \( V^* \) means \( \text{Hom}_K (V, K) \)) such that every \( v \in V \) satisfies \( v = \sum_{i=1}^{n} e_i^+ (v) e_i \). In other words,

\[
\{1, 2, \ldots, n\}, (e_i)_{i \in \{1, 2, \ldots, n\}}, (e_i^+)_{i \in \{1, 2, \ldots, n\}}
\]

is a finite dual generating system for \( V \). Therefore, Proposition 11.17 (applied to \( (J, \{b_j\}_{j \in J}, \{g_j\}_{j \in J}) = \{1, 2, \ldots, n\}, (e_i)_{i \in \{1, 2, \ldots, n\}}, (e_i^+)_{i \in \{1, 2, \ldots, n\}} \)) yields

\[
\sum_{i \in I} f_i (T a_i) = \sum_{j=1}^{n} e_j^+ (T_e_j) = n \sum_{j=1}^{n} m_{j,j}
\]

(by (11.652), applied to \( k=j \))

\[
(11.653)
\]

(here, we renamed the summation index \( j \) as \( i \)).

But recall that if \( \mathcal{G} \) is any endomorphism of the \( K \)-vector space \( V \), then the trace of \( \mathcal{G} \) equals the trace of any matrix which represents \( \mathcal{G} \) with respect to a basis of \( V \). Applying this to \( \mathcal{G} = T \), we conclude that the trace of \( T \) equals the trace of any matrix which represents \( T \) with respect to a basis of \( V \). In particular, the trace of \( T \) equals the trace of the matrix \( (m_{i,j})_{1 \leq i,j \leq n} \) (since \( (m_{i,j})_{1 \leq i,j \leq n} \)) is a matrix which represents \( T \) with respect to the basis \( (e_1, e_2, \ldots, e_n) \) of \( V \). In other words, trace \( T \) is trace \( (m_{i,j})_{1 \leq i,j \leq n} \) (by the definition of trace \( (m_{i,j})_{1 \leq i,j \leq n} \)). Compared with (11.653), this yields trace \( T = \sum_{i \in I} f_i (T a_i) \). This proves Proposition 11.16.

[Remark: Proposition 11.17 can be used to define the trace of an endomorphism of a finitely generated projective \( K \)-module when \( K \) is a commutative ring. Indeed, if \( K \) is a commutative ring and if \( V \) is a finitely generated projective \( K \)-module, and if \( T : V \rightarrow V \) is a \( K \)-linear map, then the trace \( \text{trace} (T) \) of \( T \) can be defined as \( \sum_{i \in I} f_i (T a_i) \), where \( (I, \{a_i\}_{i \in I}, \{f_i\}_{i \in I}) \) is a finite dual generating system for \( V \). This notion of trace is well-defined and generalizes the classical notion from linear algebra (which is defined only for

---

538 Proof of (11.652): Fix \( k \in \{1, 2, \ldots, n\} \) and \( j \in \{1, 2, \ldots, n\} \). Applying the map \( e_k^+ \) to both sides of (11.650), we obtain

\[
e_k^+ (T e_j) = e_k^+ \left( \sum_{i=1}^{n} m_{i,j} e_i \right) = m_{k,j}
\]

(by (11.651), applied to \( \lambda_i = m_{i,j} \)).

This proves (11.652).

539 Proof. Let \( v \in V \). Then, we can write \( v \) in the form \( v = \sum_{i=1}^{n} \lambda_i e_i \) for some \( n \)-tuple \((\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n \) (since \((e_1, e_2, \ldots, e_n) \) is a basis of \( V \)). Consider this \( n \)-tuple \((\lambda_1, \lambda_2, \ldots, \lambda_n) \). Now, \( v = \sum_{i=1}^{n} \lambda_i e_i \). Hence, for each \( k \in \{1, 2, \ldots, n\} \), we have

\[
e_k^+ \left( \sum_{i=1}^{n} \lambda_i e_i \right) = \lambda_k
\]

(by (11.651)).

Renaming \( k \) as \( i \) in this statement, we obtain the following: For each \( i \in \{1, 2, \ldots, n\} \), we have \( e_i^+ (v) = \lambda_i \). Thus,

\[
\sum_{i \in \{1, 2, \ldots, n\}} e_i^+ (v) e_i = \sum_{i=1}^{n} \lambda_i e_i
\]

Compared with \( v = \sum_{i=1}^{n} \lambda_i e_i \), this yields \( v = \sum_{i \in \{1, 2, \ldots, n\}} e_i^+ (v) e_i \), qed.

540 Indeed, this is how the trace of \( \mathcal{G} \) is defined.

541 because Proposition 11.17 shows that \( \sum_{i \in I} f_i (T a_i) \) does not depend on the choice of \((I, \{a_i\}_{i \in I}, \{f_i\}_{i \in I}) \)
finitely generated free $K$-modules). But we will not concern ourselves with these generalizations, since our exercise deals only with representations of groups over a field.

After all these preparations, we finally come to the actual solution of Exercise 4.14(b). Let $U$ be any finite-dimensional $\mathbb{C}H$-module. We want to prove $\chi_{\text{Ind}_H} = \text{Ind}_\rho \chi_U$.

For every $g \in G$, we define a $\mathbb{C}$-linear map $g^* : \mathbb{C}G \to \mathbb{C}$ by
\[(11.654)\quad (g^* (k) = \delta_{g,k} \quad \text{for all } k \in G).
\]
(This is well-defined, since $(k)_{k \in G}$ is a basis of the $\mathbb{C}$-vector space $\mathbb{C}G$.) Then,
\[(11.655)\quad \sum_{g \in G} g^* (\gamma) \cdot g = \gamma \quad \text{for every } \gamma \in \mathbb{C}G.
\]

Also,
\[(11.656)\quad (gr)^* (\gamma r) = g^* (\gamma) \quad \text{for all } g, r \in G \text{ and } \gamma \in \mathbb{C}G.
\]

\[543\]

Proposition 11.15 (applied to $K = \mathbb{C}$ and $V = U$) yields that a finite dual generating system for $U$ exists if and only if the vector space $U$ is finite-dimensional. Thus, a finite dual generating system for $U$ exists (since the vector space $U$ is finite-dimensional). Let us fix such a finite dual generating system for $U$, and denote it by $\left( J, (b_j)_{j \in J}, (g_j)_{j \in J} \right)$.

Hence, $\left( J, (b_j)_{j \in J}, (g_j)_{j \in J} \right)$ is a finite dual generating system for $U$. In other words, $\left( J, (b_j)_{j \in J}, (g_j)_{j \in J} \right)$ is a triple such that
- $J$ is a finite set;
- $\{b_j\}_{j \in J}$ is a family of elements of $U$;
- $\{g_j\}_{j \in J}$ is a family of elements of $U^*$ (where $U^*$ means $\text{Hom}_\mathbb{C} (U, \mathbb{C})$) such that every $v \in U$ satisfies
\[(11.657)\quad v = \sum_{j \in J} g_j (v) b_j.
\]

\[543\]

Proof of (11.655): Let $\gamma \in \mathbb{C}G$. We need to prove the equality (11.655). We notice that this equality is $\mathbb{C}$-linear in $\gamma$. Hence, we can WLOG assume that $\gamma \in \mathbb{C}$ (since $G$ is a basis of the $\mathbb{C}$-vector space $\mathbb{C}G$). Assume this. Now,
\[\sum_{g \in G} g^* (\gamma) \cdot g = \sum_{g \in G} \underbrace{g^* (\gamma)}_{= \delta_{g,\gamma}} \cdot g = \sum_{g \in G} \delta_{g,\gamma} \cdot g \quad (\text{by } (11.654), \text{ applied to } k=\gamma \text{ (since } \gamma \in \mathbb{C}G))
\]
\[= \sum_{g \in G; \; g=\gamma} \delta_{g,\gamma} \cdot g + \sum_{g \in G; \; g \neq \gamma} \delta_{g,\gamma} \cdot g = \sum_{g \in G; \; g=\gamma} 1 \cdot g + \sum_{g \in G; \; g \neq \gamma} 0 \cdot g \quad (\text{since } g^* (\gamma) \text{ is a finite dual generating system for } U)
\]
\[= \sum_{g \in G; \; g=\gamma} 1 \cdot g = \sum_{g \in G; \; g=\gamma} g = \gamma \quad (\text{since } \gamma \in \mathbb{C}G).
\]
This proves (11.655).

\[544\]

Proof of (11.656): Let $g \in G, r \in G$ and $\gamma \in \mathbb{C}G$. We need to prove the equality (11.656). Since this equality is $\mathbb{C}$-linear in $\gamma$ (because the maps $(gr)^*$ and $g^*$ are $\mathbb{C}$-linear), we can WLOG assume that $\gamma \in G$ (since $G$ is a basis of the $\mathbb{C}$-vector space $\mathbb{C}G$). Assume this.

We have $\gamma \in G$, and thus we can apply (11.654) to $k = \gamma$. We thus obtain $g^* (\gamma) = \delta_{g,\gamma}$. Also, $\gamma r \in \mathbb{C}G \subset G$. Hence, (11.654) (applied to $g r$ and $\gamma r$ instead of $g$ and $k$) yields $(gr)^* (\gamma r) = \delta_{gr,\gamma r}$.

Now, we have $g = \gamma$ if and only if $gr = \gamma r$ (because $G$ is a group). Now,
\[g^* (\gamma) = \delta_{g,\gamma} = \begin{cases} 1, & \text{if } g = \gamma; \\ 0, & \text{if } g \neq \gamma \end{cases} = \begin{cases} 1, & \text{if } gr = \gamma r; \\ 0, & \text{if } gr \neq \gamma r \end{cases} \quad (\text{since } g = \gamma \text{ if and only if } gr = \gamma r)
\]
\[= \delta_{gr,\gamma r} = (gr)^* (\gamma r).
\]
This proves (11.656).
We now endow \( CG \) with the \((CG, CH)\)-bimodule structure that was used to define \( \text{Ind}_p U \). We recall that \( \text{Ind}_p U = CG \otimes U \) (by the definition of \( \text{Ind}_p U \)).

Let us fix \( g \in G \). We define a map \( \tilde{F}_g : CG \times U \to U \) by setting

\[
\tilde{F}_g(\gamma, u) = \frac{1}{|H|} \sum_{h \in H} (g p(h))^* (\gamma) h u \quad \text{for all} \; (\gamma, u) \in CG \times U.
\]

The map \( \tilde{F}_g \) is \( CH \)-bilinear with respect to the right \( CH \)-module structure on \( CG \) and the left \( CH \)-module structure on \( U \).\(^{545}\)

According to the universal property of the tensor product, this yields that there exists a unique \( C \)-linear map \( F_g : CG \otimes CH U \to U \) such that

\[
(F_g (\gamma \otimes CH u) = \tilde{F}_g (\gamma, u) \quad \text{for all} \; (\gamma, u) \in CG \times U).
\]

Consider this map \( F_g \).

So we know that any \((\gamma, u) \in CG \times U\) satisfies

\[
(11.658) \quad F_g (\gamma \otimes CH u) = \tilde{F}_g (\gamma, u) = \frac{1}{|H|} \sum_{h \in H} (g p(h))^* (\gamma) h u.
\]

Let us now forget that we fixed \( g \). We thus have constructed a \( C \)-linear map \( F_g : CG \otimes CH U \to U \) for each \( g \in G \). We have shown that this map satisfies (11.658) for any \((\gamma, u) \in CG \times U \).

The set \( G \times J \) is finite (since the sets \( G \) and \( J \) are finite). Now, we define a family \((a_i)_{i \in G \times J} \) of elements of \( CG \otimes CH U \) by

\[
(a_{k,j}) = k \otimes CH b_j \quad \text{for all} \; (k, j) \in G \times J.
\]

Furthermore, we define a family \((f_i)_{i \in G \times J} \) of elements of \( (CG \otimes CH U)^* \) by

\[
(f_{k,j}) = g_j \circ F_k \quad \text{for all} \; (k, j) \in G \times J.
\]

(\( CG \otimes CH U \to U \))

\((k, j) \in G \times J \).

(\( CG \otimes CH U \to U \))

\((G \times J, (a_i)_{i \in G \times J}, (f_i)_{i \in G \times J}) \) is a finite dual generating system for \( CG \otimes CH U \).

---

\(^{545}\)Proof. We notice that \( \tilde{F}_g (\gamma, u) \) depends \( C \)-linearly on each of \( \gamma \) and \( u \) (for obvious reasons).

In other words, the map \( \tilde{F}_g \) is \( C \)-bilinear.

Now, let us fix \( \gamma \in CG, u \in U \) and \( \kappa \in CH \). We are going to prove the equality \( \tilde{F}_g (\gamma \kappa, u) = \tilde{F}_g (\gamma, \kappa u) \).

Since this equality is \( C \)-linear in \( \kappa \) (because the map \( \tilde{F}_g \) is \( C \)-bilinear), we can WLOG assume that \( \kappa \in H \) (since \( H \) is a basis of the \( C \)-vector space \( CH \)). Assume this.

We have \( \kappa \in H \). Thus, the map \( H \to H, h \to h \kappa \) is a bijection (since \( H \) is a group). Hence, we can substitute \( h \kappa \) for \( h \) in the sum \( \sum_{h \in H} (g p(h))^* (\gamma \rho(\kappa)) h u \). As a result, we obtain \( \sum_{h \in H} (g p(h))^* (\gamma \rho(\kappa)) h u = \sum_{h \in H} (g p(h))^* (\gamma (\rho(\kappa))) (h \kappa) u \).

Now, the definition of \( \tilde{F}_g (\gamma \kappa, u) \) yields

\[
\tilde{F}_g (\gamma \kappa, u) = \frac{1}{|H|} \sum_{h \in H} (g p(h))^* (\gamma (\rho(\kappa))) h u = \frac{1}{|H|} \sum_{h \in H} \left( g \rho(h \kappa) \right)^* (\gamma (\rho(\kappa))) h u = \frac{1}{|H|} \sum_{h \in H} (g p(h))^* (\gamma (h \kappa)) h u.
\]

\[
\tilde{F}_g (\gamma, \kappa u) = \frac{1}{|H|} \sum_{h \in H} (g p(h))^* (\gamma (h \kappa)) h u.
\]

This yields \( \tilde{F}_g (\gamma \kappa, u) = \tilde{F}_g (\gamma, \kappa u) \).

Now let us forget that we fixed \( \gamma, u \) and \( \kappa \). Thus, we have shown that \( \tilde{F}_g (\gamma \kappa, u) = \tilde{F}_g (\gamma, \kappa u) \) for all \( \gamma \in CG, u \in U \) and \( \kappa \in CH \). This yields that the map \( \tilde{F}_g \) is \( CH \)-bilinear with respect to the right \( CH \)-module structure on \( CG \) and the left \( CH \)-module structure on \( U \) (since we already know that this map \( \tilde{F}_g \) is \( C \)-bilinear). Qed.
Indeed, every \( v \in CG \otimes_{CH} U \) satisfies \( v = \sum_{i \in G \times J} f_i(v) a_i \). Thus, we know that \( (G \times J, (a_i)_{i \in G \times J} : (f_i)_{i \in G \times J}) \)
is a triple such that

- \( G \times J \) is a finite set;
- \( (a_i)_{i \in G \times J} \) is a family of elements of \( CG \otimes_{CH} U \);

---

**Proof.** Let \( v \in CG \otimes_{CH} U \). We need to prove the equality \( v = \sum_{i \in G \times J} f_i(v) a_i \). Since this equality is \( C \)-linear in \( v \), we can WLOG assume that \( v \) is a pure tensor (since the pure tensors in \( CG \otimes_{CH} U \) span the whole \( C \)-vector space \( CG \otimes_{CH} U \)). Assume this. Thus, \( v \) is a pure tensor. In other words, \( v = \gamma \otimes CH u \) for some \( \gamma \in CG \) and \( u \in U \). Consider these \( \gamma \) and \( u \).

We notice that for any \( w \in G \), the map \( G \to G, \ k \mapsto kw \) is a bijection (since \( G \) is a group). Hence, for any \( w \in G \), we have

\[
\sum_{k \in G} (kw)^* (\gamma) kw = \sum_{k \in G} k^* (\gamma) k \quad \begin{cases} \text{here, we substituted } k \text{ for } kw \text{ in the sum,} \\ \text{since the map } G \to G, \ k \mapsto kw \text{ is a bijection} \end{cases}
\]

\[
= \sum_{g \in G} g^* (\gamma) g \\ \text{(here, we renamed the summation index } k \text{ as } g) 
\]

\[
(11.659) \quad = \sum_{g \in G} g^* (\gamma) \cdot g = \gamma \quad \text{(by (11.655)).}
\]

Now, every \( h \in H \) satisfies

\[
\sum_{k \in G} (kp(h))^* (\gamma) = \sum_{k \in G} (kp(h))^* (\gamma) k \quad \begin{cases} \text{by the definition of the} \\ \text{right CH-module structure on } CG) \\ \text{since } h \in H \end{cases}
\]

\[
= \sum_{g \in G} (g \circ F_k) (v) k \otimes CH b_j
\]

\[
= \sum_{j \in J} (g_j \circ F_k) (v) b_j
\]

\[
(11.660) \quad = \gamma \quad \text{(by (11.659), applied to } w = \rho(h)).
\]

Now,

\[
\sum_{i \in G \times J} f_i(v) a_i = \sum_{i \in G \times J} f_i(k, j) (v) a(k, j) = \sum_{k \in G} k \otimes CH \left( \sum_{j \in J} g_j \circ F_k(v) b_j \right)
\]

\[
\text{(because } F_k(v) = \sum_{j \in J} g_j(F_k(v)) b_j \text{)}
\]

\[
\text{(by (11.657), applied to } F_k(v) \text{ instead of } v) \quad \text{(since the tensor product is } C \text{-bilinear)}
\]

\[
= \sum_{k \in CH} k \otimes CH \left( \frac{1}{|H|} \sum_{h \in H} (k \rho(h))^* (\gamma) hu \right)
\]

\[
= \frac{1}{|H|} \sum_{h \in H} (k \rho(h))^* (\gamma) hu \quad \text{(since the tensor product is } C \text{-bilinear)}
\]

\[
\frac{1}{|H|} \sum_{h \in H} (k \rho(h))^* (\gamma) hu = \frac{1}{|H|} \sum_{h \in H} \sum_{k \in G} (k \rho(h))^* (\gamma) k \otimes CH u
\]

\[
\text{(since } h \text{ can be moved past the } \otimes CH \text{ sign (since } h \in CH \subseteq CH) \quad \text{(by (11.660))}
\]

Thus, \( v = \sum_{i \in G \times J} f_i(v) a_i \) is proven, qed.
• \((f_i)_{i \in G \times J}\) is a family of elements of \((\mathbb{C}G \otimes_{CH} U)^*\) (where \((\mathbb{C}G \otimes_{CH} U)^*\) means \(\text{Hom}_{\mathbb{C}}(\mathbb{C}G \otimes_{CH} U, \mathbb{C})\)) such that every \(v \in \mathbb{C}G \otimes_{CH} U\) satisfies \(v = \sum_{i \in G \times J} f_i(v) a_i\). In other words, \((G \times J, (a_i)_{i \in G \times J}, (f_i)_{i \in G \times J})\) is a finite dual generating system for \(\mathbb{C}G \otimes_{CH} U\) (because this is precisely how a “finite dual generating system” was defined). In other words, \((G \times J, (a_i)_{i \in G \times J}, (f_i)_{i \in G \times J})\) is a finite dual generating system for \(\text{Ind}_p U\) (since \(\text{Ind}_p U = \mathbb{C}G \otimes_{CH} U\)).

Let us now notice that

\[\sum_{j \in J} g_j(hb_j) = \chi_U(h) \quad \text{for every } h \in H\]

Now, let us fix \(g \in G\). We want to compute \(\chi_{\text{Ind}_p U}(g)\). The definition of \(\chi_{\text{Ind}_p U}(g)\) yields \(\chi_{\text{Ind}_p U}(g) = \text{trace}(g : \text{Ind}_p U \to \text{Ind}_p U)\). But Proposition 11.16 (applied to \(K = \mathbb{C}, V = \text{Ind}_p U, I = G \times J\) and \(T = (g : \text{Ind}_p U \to \text{Ind}_p U)\)) yields

\[
\text{trace}(g : \text{Ind}_p U \to \text{Ind}_p U) = \sum_{i \in G \times J} f_i(ga_i)
\]

\[
= \sum_{(k,j) \in G \times J} f_{(k,j)} \left( \frac{(g : \text{Ind}_p U \to \text{Ind}_p U) a_j}{= ga_i} \right)
\]

\[
= \sum_{(k,j) \in G \times J} f_{(k,j)} \left( \frac{g \circ a_{(k,j)}}{= g \circ F_k} \right)
\]

\[
= \sum_{(k,j) \in G \times J} (g_j \circ F_k) \left( \frac{g (k \otimes_{CH} b_j)}{= gk \otimes_{CH} b_j} \right)
\]

\[
= \sum_{(k,j) \in G \times J} \left( \frac{g_j \circ F_k}{= g_j(F_k(gk \otimes_{CH} b_j))} \right)
\]

\[
= \sum_{(k,j) \in G \times J} \frac{F_k(gk \otimes_{CH} b_j)}{= 1 \sum_{h \in H} \frac{(k \rho(h))^* (gk) hb_j}{= 1 \sum_{h \in H} \frac{(k \rho(h))^* (gk) hb_j}{(by (11.658), applied to \(k, gk\) and \(b_j\) instead of \(g, \gamma\) and \(u\))}}}
\]

\[
= \sum_{(k,j) \in G \times J} \frac{g_j}{= \sum_{h \in H} \frac{1}{|H|} \sum_{h \in H} (k \rho(h))^* (gk) hb_j}{= \sum_{k \in G} \sum_{j \in J} \frac{1}{|H|} \sum_{h \in H} (k \rho(h))^* (gk) g_j(hb_j)}
\]

\[
= \sum_{k \in G} \sum_{j \in J} \frac{1}{|H|} \sum_{h \in H} (k \rho(h))^* (gk) g_j(hb_j).
\]

\[
(11.662)
\]

---

\[547\) Proof of (11.661): Let \(h \in H\). The definition of \(\chi_U(h)\) yields \(\chi_U(h) = \text{trace}(h : U \to U)\). But Proposition 11.16 (applied to \(\mathbb{C}, U, (J, (b_j))_{j \in J}, (g_j)_{j \in J}\) and \(h : U \to U\) instead of \(\mathbb{C}, U, (I, (a_i))_{i \in I}, (f_i)_{i \in I}\) and \(T\)) yields

\[
\text{trace}(h : U \to U) = \sum_{j \in J} g_j \left( \frac{(h : U \to U) b_j}{= hb_j} \right) = \sum_{j \in J} g_j(hb_j).
\]

Thus, \(\chi_U(h) = \text{trace}(h : U \to U) = \sum_{j \in J} g_j(hb_j)\). This proves (11.661).
But every \( k \in G \) satisfies

\[
\sum_{h \in H} (k \rho(h))^* (gk) \sum_{j \in J} g_j (hb_j) = \sum_{h \in H} \delta_{k \rho(h), gk} \chi_U (h)
\]

(by \((11.654), \text{ applied to } k \rho(h) \text{ and } gk \text{ instead of } g \text{ and } k\))

\[
= \sum_{h \in H} \delta_{k \rho(h), gk} \chi_U (h) = \sum_{h \in H; k \rho(h) = g} \chi_U (h) + \sum_{h \in H; k \rho(h) \neq g} \chi_U (h) = \sum_{h \in H; k \rho(h)^{-1} = g} \chi_U (h)
\]

(Exercise 4.14(b))

Now, recall that

\[
\chi_{\text{Ind}_\rho U} (g) = \text{trace} (g : \text{Ind}_\rho U \rightarrow \text{Ind}_\rho U)
\]

\[
= \frac{1}{|H|} \sum_{h \in H} (k \rho(h))^* (gk) \sum_{j \in J} g_j (hb_j) = \sum_{h \in H; k \rho(h)^{-1} = g} \chi_U (h)
\]

(by \((11.662)\))

\[
= \frac{1}{|H|} \sum_{h \in H; k \rho(h)^{-1} = g} \chi_U (h) = \frac{1}{|H|} \sum_{h \in H; k \rho(h)^{-1} = g} \chi_U (h)
\]

(by the definition of \( \text{Ind}_\rho \chi_U (g) \))

\[
(\text{Ind}_\rho \chi_U) (g) = \frac{1}{|H|} \sum_{(h,k) \in H \times G; k \rho(h)^{-1} = g} \chi_U (h)
\]

this yields \( \chi_{\text{Ind}_\rho U} (g) = (\text{Ind}_\rho \chi_U) (g) \).

Now, let us forget that we fixed \( g \). We thus have proven that \( \chi_{\text{Ind}_\rho U} (g) = (\text{Ind}_\rho \chi_U) (g) \) for every \( g \in G \).

In other words, \( \chi_{\text{Ind}_\rho U} = \text{Ind}_\rho \chi_U \). This solves Exercise 4.14(b).

(c) Assume that \( H \) is a subgroup of \( G \), and that \( \rho : H \rightarrow G \) is the inclusion map. We need to show that \( \text{Ind}_\rho f = \text{Ind}_H^G f \) for every \( f \in R_C (H) \).

We notice that the map \( G \rightarrow G, \ k \mapsto k^{-1} \) is a bijection (since \( G \) is a group).
Let $f \in R_G(H)$. Let $g \in G$. Then, the definition of $(\text{Ind}_\rho f)(g)$ yields

\[
(\text{Ind}_\rho f)(g) = \frac{1}{|H|} \sum_{(h,k) \in H \times G; \ k \rho(h) k^{-1} = g} f(h) = \frac{1}{|H|} \sum_{(h,k) \in H \times G; \ k h k^{-1} = g} f(h)
\]

(since $\rho(h) = h$ for every $h \in H$)

\[
= \frac{1}{|H|} \sum_{k \in G} \sum_{h : h = k^{-1} g k} f(h)
\]

(since every $k \in G$ satisfies either $k^{-1} g k \in H$ or $k^{-1} g k \notin H$ (but never both))

\[
= \frac{1}{|H|} \left( \sum_{k \in G; \ k^{-1} g k \in H} \sum_{h : h = k^{-1} g k} f(h) + \sum_{k \in G; \ k^{-1} g k \notin H} f(h) \right)
\]

(since there is no $h \in H$ satisfying $h = k^{-1} g k$ (because $k^{-1} g k \notin H$))

\[
= \frac{1}{|H|} \left( \sum_{k \in G; \ k^{-1} g k \in H} f(k^{-1} g k) + \sum_{k \in G; \ k^{-1} g k \notin H} \text{(empty sum)} \right) = \frac{1}{|H|} \left( \sum_{k \in G; \ k^{-1} g k \in H} f(k^{-1} g k) + \sum_{k \in G; \ k^{-1} g k \notin H} 0 \right)
\]

\[
= \frac{1}{|H|} \sum_{k \in G; \ k^{-1} g k \in H} f(k^{-1} g k)
\]

\[
= \frac{1}{|H|} \sum_{k \in G; \ k^{-1} g k \in H} \left( k^{-1} g k \right) = \frac{1}{|H|} \sum_{k \in G; \ k^{-1} g k \in H} (k^{-1})^{-1} g k^{-1} = k
\]

(here, we substituted $k^{-1}$ for $k$ in the sum, since the map $G \to G$, $k \mapsto k^{-1}$ is a bijection)

\[
= \frac{1}{|H|} \sum_{k \in G; \ k g k^{-1} \in H} f(k g k^{-1}) = \left( \text{Ind}_H^G f \right)(g)
\]
(since (4.4) yields \( \left( \text{Ind}_H^G f \right) (g) = \frac{1}{|H|} \sum_{k \in G} f (kgk^{-1}) \)).

Let us now forget that we fixed \( g \). We thus have proven that \( (\text{Ind}_\rho f) (g) = \left( \text{Ind}_H^G f \right) (g) \) for every \( g \in G \).

In other words, \( \text{Ind}_\rho f = \text{Ind}_H^G f \). This solves Exercise 4.14(c).

(d) Assume that \( H \) is a subgroup of \( G \), and that \( \rho : H \to G \) is the inclusion map. We need to show that \( \text{Ind}_\rho U = \text{Ind}_H^G U \) for every \( CH \)-module \( U \).

We notice that \( \rho : H \to G \) is the inclusion map. Hence, \( \mathbb{C}[\rho] : CH \to CG \) is also the inclusion map (since we identify \( CH \) with a \( \mathbb{C} \)-subalgebra of \( CG \) along this map \( \mathbb{C}[\rho] \)). Thus,

\[
\text{(11.664)} \quad (\mathbb{C}[\rho]) \eta = \eta \quad \text{for every } \eta \in CH.
\]

Let \( U \) be a \( CH \)-module. Both \( \text{Ind}_H^G U \) and \( \text{Ind}_\rho U \) are defined as \( CG \otimes CH U \) for some \( (CG, CH) \)-bimodule structure on \( CG \); however, their definitions differ at how this \( (CG, CH) \)-bimodule structure is defined. We are now going to prove that these two \( (CG, CH) \)-bimodule structures on \( CG \) are identical.

The definition of \( \text{Ind}_\rho U \) shows that we have

\[
\text{(11.665)} \quad \text{Ind}_\rho U = CG \otimes_{CH} U,
\]

where \( CG \) is regarded as a \( (CG, CH) \)-bimodule according to the following rule: The left \( CG \)-module structure on \( CG \) is plain multiplication inside \( CG \); the right \( CH \)-module structure on \( CG \) is induced by the \( \mathbb{C} \)-algebra homomorphism \( \mathbb{C}[\rho] : CH \to CG \) (thus, it is explicitly given by \( \gamma \eta = \gamma \cdot (\mathbb{C}[\rho]) \eta \) for all \( \gamma \in CG \) and \( \eta \in CH \)).

We denote this \( (CG, CH) \)-bimodule structure on \( CG \) as the first structure.

On the other hand, the definition of \( \text{Ind}_H^G U \) shows that we have

\[
\text{(11.666)} \quad \text{Ind}_H^G U = CG \otimes_{CH} U,
\]

where \( CG \) is regarded as a \( (CG, CH) \)-bimodule in the usual way (i.e., the left \( CG \)-module structure on \( CG \) is plain multiplication inside \( CG \), and the right \( CH \)-module structure on \( CG \) is also plain multiplication inside \( CG \) because \( CH \subset CG \)). We denote this \( (CG, CH) \)-bimodule structure on \( CG \) as the second structure.

The right hand sides of the equalities (11.665) and (11.666) appear identical, but so far we do not know if they actually mean the same thing, because the meanings of “\( CG \)” possibly differ. Namely, we have two \( (CG, CH) \)-bimodule structures on the \( \mathbb{C} \)-vector space \( CG \): the first structure (used in (11.665)) and the second structure (used in (11.666)). These two structures clearly have the same left \( CG \)-module structure. But they also have the same right \( CH \)-module structure, because every \( \gamma \in CG \) and \( \eta \in CH \) satisfy

\[
\text{(the result of the right action of } \eta \text{ on } \gamma \text{ according to the first structure)}
\]

\[
= \gamma \cdot \left( \mathbb{C}[\rho] \right) \eta = \gamma \cdot \eta
\]

(by \( (11.664) \))

\[
= \text{(the result of the right action of } \eta \text{ on } \gamma \text{ according to the second structure)}.
\]

Hence, the first structure and the second structure are identical. Thus, the right hand sides of the equalities (11.665) and (11.666) really mean the same thing. Thus, comparing the equalities (11.665) and (11.666), we obtain \( \text{Ind}_\rho U = \text{Ind}_H^G U \) as left \( CG \)-modules. This solves Exercise 4.14(d).

(c) Assume that \( \bar{G} = H/K \) for some normal subgroup \( K \) of \( H \). Let \( \rho : H \to G \) be the projection map. We want to prove that \( \text{Ind}_\rho f = f^K \) for every \( f \in R_\bar{C}(H) \).

Let \( f \in R_\bar{C}(H) \). Applying (11.645) to \( H \) instead of \( K \), we obtain the following equivalence:

\[
(f \in R_\bar{C}(H)) \Leftrightarrow \text{any two conjugate elements } k \text{ and } k' \text{ of } H \text{ satisfy } f(k) = f(k') \, .
\]

Hence,

(11.667) \( \text{any two conjugate elements } k \text{ and } k' \text{ of } H \text{ satisfy } f(k) = f(k') \) (since we know that \( f \in R_\bar{C}(H) \)). Thus,

(11.668) \( f(y^{-1}zy) = f(z) \) for all \( z \in H \) and \( y \in H \) (because if \( z \in H \) and \( y \in H \), then \( y^{-1}zy \) and \( z \) are two conjugate elements of \( H \)).
Let \( g \in G \). Then, \( g \in G = H/K \). Hence, there exists an \( x \in H \) such that \( g = xK \). Consider this \( x \).
The map \( \rho : H \to G \) is the projection map from \( H \) to \( G = H/K \), and thus sends every \( h \in H \) to the coset \( hK \in G \). In other words, \( \rho(h) = hK \) for every \( h \in H \). Applied to \( h = x \), this yields \( \rho(x) = xK = g \). Of course, \( \text{Ker} \rho = K \) (since \( \rho \) is the projection map from \( H \) to \( H/K \)).

We have \( |G| = |H/K| = |H : K| = |H|/|K| \).

We make another simple observation: If \( h \in H \) and \( y \in H \), then we have the following logical equivalence:

\[
(11.669) \quad \left( \rho(h) = (\rho(y))^{-1} gp(y) \right) \iff (h \in y^{-1}xK) .
\]

The definition of \( f^K \) yields \( f^K(xK) = \frac{1}{|K|} \sum_{k \in K} f(xk) \). Hence,

\[
(11.670) \quad f^K \left( \frac{g}{xK} \right) = f^K(xK) = \frac{1}{|K|} \sum_{k \in K} f(xk) .
\]

But the definition of \( \left( \text{Ind}_{\rho} f \right)(g) \) yields

\[
\left( \text{Ind}_{\rho} f \right)(g) = \frac{1}{|H|} \sum_{(h,k) \in H \times G; \quad \rho(h)k^{-1} = g} f(h) = \frac{1}{|H|} \sum_{(h,k) \in H \times G; \quad \rho(h)=k^{-1}gk} f(h) = \frac{1}{|H|} \sum_{h \in H; \quad \rho(h)=p^{-1}gp} f(h)
\]

\[
(11.671) \quad = \frac{1}{|H|} \sum_{h \in H; \quad \rho(h)=k^{-1}gk} f(h) = \frac{1}{|H|} \sum_{h \in H; \quad \rho(h)=p^{-1}g} f(h)
\]

Proof of (11.669): Let \( h \in H \) and \( y \in H \). Recall that the map \( \rho : H \to G \) is the projection map from \( H \) to \( G = H/K \).

Thus, \( \rho(w) = wK \) for every \( w \in H \).

We have the following chain of equivalences:

\[
\left( \rho(h) = (\rho(y))^{-1} gp(y) \right) \iff \left( \rho(y) \rho(h) = gp(y) \right) \iff \left( \rho(y) \rho(h)(\rho(y))^{-1} = g \right) = \rho(yh^{-1}) = \rho(x) \quad (\text{since } \rho \text{ is a group homomorphism})
\]

\[
\iff \left( \rho(yh^{-1}) = \rho(x) \right) \iff (yhy^{-1}K = xK) \iff (yhy^{-1} \in xK) \iff (yh \in xK) \iff (h \in y^{-1}xK).
\]

This proves (11.669).
(here, we renamed the index $k$ as $p$ in the first sum). But every $p \in G$ satisfies
\[ \sum_{h \in H; \rho(h) = p \cdot g} f(h) = \sum_{k \in K} f(xk) \]
549. Hence, (11.671) becomes
\[ (\text{Ind}_\rho f)(g) = \frac{1}{|H|} \sum_{p \in G} \sum_{h \in H; \rho(h) = p \cdot g} f(h) = \frac{1}{|H|} \sum_{p \in G} \sum_{k \in K} f(xk) = \frac{1}{|H|} \cdot \sum_{k \in K} f(xk) = \frac{1}{|K|} \cdot \sum_{k \in K} f(xk) = f^K(g) \]
(by (11.670)).
Let us now forget that we fixed $g$. We thus have shown that $(\text{Ind}_\rho f)(g) = f^K(g)$ for every $g \in G$. In other words, Ind$_\rho f = f^K$. This solves Exercise 4.14(e).

(f) Assume that $G = H/K$ for some normal subgroup $K$ of $H$. Let $\rho : H \to G$ be the projection map. We want to prove that Ind$_\rho U \cong U^K$ for every $C$-module $U$.

The map $\rho$ is the projection map from $H$ to $G = H/K$. Thus, the map $\rho$ is surjective and has kernel Ker $\rho = K$. Furthermore,
\[ (11.673) \quad |\rho^{-1}(g)| = |K| \quad \text{for every } g \in G. \]
550

Let $U$ be a $C$-module. Recall that $U^K$ is a $C[H/K]$-module, thus a $C$-$G$-module (since $H/K = G$).

Recall that Ind$_\rho U$ is defined as the $C$-$G$-module $C[G] \otimes_{CH} U$, where $C[G]$ is regarded as a $C[G,C]$-bimodule according to the following rule: The left $C$-$G$-module structure on $C[G]$ is plain multiplication inside $C[G]$; the right $C$-$H$-module structure on $C[G]$ is induced by the $C$-algebra homomorphism $C[H] : CH \to C[G]$ (thus, it is explicitly given by $\gamma \eta = \gamma \cdot (C[H]) \eta$ for all $\gamma \in C[G]$ and $\eta \in CH$). From now on, we regard $C[G]$ as endowed with this $(C[G], C[H])$-bimodule structure.

549 Proof. Let $p \in G$. Then, $p \in G = H/K$. Hence, there exists a $y \in H$ such that $p = yK$. Consider this $y$.

Recall that $\rho(h) = hK$ for every $h \in H$. Applying this to $h = y$, we obtain $\rho(y) = yK = p$.

We notice that the map $K \to y^{-1}xKy$, $k \mapsto y^{-1}xky$ is a bijection (since $H$ is a group).

Now, for every $h \in H$, we have the following equivalence:
\[ (\rho(h) = p \cdot g) \iff (\rho(h) = (\rho(y))^{-1}gp) \quad (\text{since } p = \rho(y)) \]
(11.672) \quad \iff \quad (h \in y^{-1}xKy) \quad (\text{according to (11.669))}.

Now,
\[ \sum_{h \in H; \rho(h) = p \cdot g} f(h) = \sum_{h \in y^{-1}xKy} f(h) = \sum_{k \in K} f(y^{-1}xky) \]
\[ (\text{because for every } h \in H, \text{ the statement } \rho(h) = p \cdot g \text{ is equivalent to } h \in y^{-1}xKy) \]
\[ (\text{in case of } \rho(h) = y^{-1}xKy) \]
(here, we substituted $y^{-1}xky$ for $h$ in the sum, since the map $K \to y^{-1}xKy$, $k \mapsto y^{-1}xky$ is a bijection). Thus,
\[ \sum_{h \in H; \rho(h) = p \cdot g} f(h) = \sum_{k \in K} f(xk), \]
by (11.668), applied to $z = xk$.

qed.
550 Proof of (11.673): Let $g \in G$. The map $\rho$ is surjective. Hence, $\rho(H) = G$. Thus, $g \in G = \rho(H)$. Therefore, there exists some $x \in H$ such that $g = \rho(x)$. Let us fix such an $x$.

We know that $H$ is a group. Hence, the map $K \to xK$, $k \mapsto xk$ is a bijection. Thus, the sets $K$ and $xK$ are in bijection. Therefore, $|xK| = |K|$. 

We define a map $\alpha : U^K \to CG \otimes_{CH} U$ by setting
$$\alpha (u) = 1 \otimes_{CH} u \quad \text{for every } u \in U^K.$$  
This $\alpha$ is a map $U^K \to \text{Ind}_{\rho} U$ (since $\text{Ind}_{\rho} U = CG \otimes_{CH} U$) and is $\mathbb{C}$-linear (since $\alpha (u) = 1 \otimes_{CH} u$ depends $\mathbb{C}$-linearly on $u$).

We will show that $\alpha$ is a $CG$-module isomorphism.

The $CG$-module structure on $U^K$ has the property that

$$\rho (h) \cdot v = hv \quad \text{for any } h \in H \text{ and } v \in U^K.$$  

551

Now, $\alpha$ is a $CG$-module homomorphism. We will eventually construct an inverse to $\alpha$; but first we need to prepare.

But for every $y \in H$, we have the following logical equivalence:

$$(y \in \rho^{-1} (g)) \iff \left( \rho (y) = g \underset{= \rho (x)}{\longrightarrow} \rho (x) \right) \iff (\rho (y) = \rho (x)) \iff \left( \frac{\rho (y) \cdot (\rho (x))^{-1} = 1}{\rho (y) \rho (x) = 1} \right) \iff \left( \frac{yz = 1}{(\text{since } \rho \text{ is a group homomorphism})} \right)$$

Hence, $\rho^{-1} (g) = xK$, so that $|\rho^{-1} (g)| = |xK| = |K|$. This proves (11.673).

Proof of (11.674): The map $\rho$ is the projection map from $H$ to $H/K$. Thus, the map $\rho$ sends every $h \in H$ to the coset $hK \in H/K$. In other words,

$$\rho (h) = hK \quad \text{for every } h \in H,$$

where $hK$ means the coset $hK \in H/K$. But by the definition of the $\mathbb{Z} [H/K]$-module structure on $U^K$, we have

$$(hK) \cdot v = hv \quad \text{for any } h \in H \text{ and } v \in U^K,$$

where $hK$ means the coset $hK \in H/K$. Thus, any $h \in H$ and $v \in U^K$ satisfy

$$\rho (h) \underset{= hK}{\longrightarrow} \cdot v = (hK) \cdot v = hv,$$

where $hK$ means the coset $hK \in H/K$. This proves (11.674).

Proof. Let $u \in U^K$ and $g \in G$. We have $g \in G = \rho (H)$ (since the map $\rho : H \to G$ is surjective). Thus, $g = \rho (y)$ for some $y \in H$. Let us consider this $y$.

Applying (11.674) to $h = y$ and $v = u$, we obtain $\rho (y) \cdot u = yu$. Thus, $\underset{= \rho (y)}{u} \overset{\rho (y)}{\longrightarrow} u = \rho (y) \cdot u = yu.$

Now, the definition of $\alpha (gu)$ yields

$$\alpha (gu) = 1 \otimes_{CH} \underset{= yu}{\overset{\rho (y)}{\longrightarrow}} gu = 1 \otimes_{CH} yu = 1y \otimes_{CH} u$$

(here, we moved the $y$ past the tensor sign; this is allowed because $y \in H \subset CH$).

By the definition of the left $CG$-module structure on $CG$, we have $g \cdot 1 = g1 = g$.

By the definition of the right $\mathbb{C}H$-module structure on $CG$, we have $\gamma \eta = \gamma \cdot (\mathbb{C} [\rho] \eta)$ for all $\gamma \in CG$ and $\eta \in CH$.

Applying this to $\gamma = 1$ and $\eta = y$, we obtain $1y = 1 \cdot (\mathbb{C} [\rho] y)$. But $y \in H$ and thus $(\mathbb{C} [\rho] y) = \rho (y) = g$. Therefore, $1y = 1 \cdot (\mathbb{C} [\rho] y = 1 \cdot g = g = g \cdot 1$.

Now, $\alpha (gu) = 1y \otimes_{CH} u = g \cdot 1 \otimes_{CH} u$. Compared with $g \cdot \alpha (u) = g \cdot (1 \otimes_{CH} u) = g1 \otimes_{CH} u$, this yields $\alpha (gu) = g \cdot \alpha (u)$.

Now, let us forget that we fixed $u$ and $g$. We thus have shown that $\alpha (gu) = g \cdot \alpha (u)$ for all $u \in U^K$ and $g \in G$. Thus, the map $\alpha$ is a homomorphism of $G$-sets. Since the map $\alpha$ is also $\mathbb{C}$-linear (for obvious reasons), this yields that $\alpha$ is a $CG$-module homomorphism. Qed.
For every \( g \in G \) and every \( p \in \mathbb{C}G \), we denote by \( \epsilon_g (p) \) the \( g \)-coordinate of \( p \) with respect to the basis \( G \) of the \( \mathbb{C} \)-vector space \( \mathbb{C}G \). By the definition of “coordinate”, we have

\[
\epsilon_g (q) = \sum_{g \in G} \epsilon_g (q) g \quad \text{for every} \quad q \in \mathbb{C}G.
\]

(11.676)

For every \( g \in G \), we have defined a map \( \epsilon_g : \mathbb{C}G \to \mathbb{C} \) (because we have defined an element \( \epsilon_g (p) \) for every \( p \in \mathbb{C}G \)). This map \( \epsilon_g \) is \( \mathbb{C} \)-linear. We notice some basic properties of these maps:

- For every \( g \in G \) and \( h \in G \), we have

\[
\epsilon_g (h) = \delta_{g,h}.
\]

(11.677)

- We have

\[
\epsilon_1 (pq) = \epsilon_1 (qp) \quad \text{for all} \quad p \in \mathbb{C}G \quad \text{and} \quad q \in \mathbb{C}G.
\]

(11.678)

- We have

\[
\epsilon_1 (g^{-1}q) = \epsilon_g (q) \quad \text{for every} \quad g \in G \quad \text{and} \quad q \in \mathbb{C}G.
\]

(11.679)

Now, for every \((q,u) \in \mathbb{C}G \times U\), we have

\[
\sum_{h \in H} \epsilon_1 (\rho (h) q) h^{-1} u \in U^K.
\]

(11.680)\footnote{This equality has been proven in our solution of Exercise 4.4. (Namely, it appeared there as (11.625).)}

Hence, we can define a map \( \beta : \mathbb{C}G \times U \to U^K \) by setting

\[
\beta (q,u) = \sum_{h \in H} \epsilon_1 (\rho (h) q) h^{-1} u \quad \text{for every} \quad (q,u) \in \mathbb{C}G \times U.
\]

Consider this map \( \beta \). Then, \( \beta \) is a \( \mathbb{C} \)-bilinear map (because \( \beta (q,u) = \sum_{h \in H} \epsilon_1 (\rho (h) q) h^{-1} u \) depends \( \mathbb{C} \)-linearly on each of \( q \) and \( u \)). We are now going to prove that the map \( \beta \) is \( \mathbb{C} \)-bilinear with respect to the right \( CH \)-module structure on \( \mathbb{C}G \) and the left \( CH \)-module structure on \( U \).

In fact, every \( h' \in H \), \( q \in \mathbb{C}G \) and \( u \in U \) satisfy

\[
\beta (q,h'u) = \beta (qh',u).
\]

(11.681)\footnote{This equality has been proven in our solution of Exercise 4.4. (Namely, it appeared there as (11.626).)}

\footnote{Proof. Let \((q,u) \in \mathbb{C}G \times U \). Let \( x = \sum_{h \in H} \epsilon_1 (\rho (h) q) h^{-1} u \). We shall show that \( x \in U^K \).

Let \( k \in K \). Then, the map \( H \to H \), \( h \mapsto hk \) is a bijection (since \( H \) is a group). Now, multiplying both sides of the equality \( x = \sum_{h \in H} \epsilon_1 (\rho (h) q) h^{-1} u \) with \( k \) from the left, we obtain

\[
kx = k \sum_{h \in H} \epsilon_1 (\rho (h) q) h^{-1} u = \sum_{h \in H} \epsilon_1 (\rho (h) q) kh^{-1} u
\]

\[
= \sum_{h \in H} \epsilon_1 \left( \frac{\rho (hk)}{\rho (h) \rho (k)} \right) q \left( \begin{array}{c} k (hk)^{-1} u \\ -k^{-1}h^{-1} \end{array} \right) \quad \text{(since \( \rho \) is a group homomorphism)}
\]

\[
= \sum_{h \in H} \epsilon_1 \left( \frac{\rho (h) \rho (k)}{\rho (k) \rho (k)} \right) q \left( \begin{array}{c} k^{-1}h^{-1} u \\ -1 \end{array} \right) = \sum_{h \in H} \epsilon_1 (\rho (h) q) h^{-1} u = x.
\]

Let us now forget that we fixed \( k \). We thus have shown that \( kx = x \) for every \( k \in K \). In other words, \( x \) is an element \( y \in U \) satisfying \( ky = y \) for every \( k \in K \). Hence,

\[
x \in \{ y \in U \mid ky = y \text{ for every } k \in K \} = U^K.
\]

Thus, \( \sum_{h \in H} \epsilon_1 (\rho (h) q) h^{-1} u = x \in U^K \), qed.}
As a consequence of this, we can see that every \( r \in CH, q \in CG \) and \( u \in U \) satisfy
\[
(11.682) \quad \bar{\beta}(q, ru) = \bar{\beta}(qr, u).
\]

In other words, the map \( \bar{\beta} \) is \( CH \)-bilinear with respect to the right \( CH \)-module structure on \( CG \) and the left \( CH \)-module structure on \( U \) (since we already know that \( \bar{\beta} \) is \( C \)-bilinear). Hence, by the universal property of the tensor product, we conclude that there exists a unique \( C \)-linear map \( \beta : CG \otimes_{CH} U \rightarrow U^K \) such that every \((q, u) \in CG \times U\) satisfies
\[
(11.683) \quad \beta(q \otimes_{CH} u) = \bar{\beta}(q, u).
\]
Consider this map \( \beta \). Clearly, every \( q \in CG \) and \( u \in U \) satisfy
\[
(11.684) \quad \beta(q \otimes_{CH} u) = \bar{\beta}(q, u) = \sum_{h \in H} \varepsilon_1(\rho(h)q) h^{-1}u \quad \text{(by the definition of \( \bar{\beta} \))}.
\]

We shall now show that the maps \( \alpha \) and \( \frac{1}{|K|} \beta \) are mutually inverse. To do so, we will show that
\[
\alpha \circ \beta = |K| \text{id} \quad \text{and} \quad \beta \circ \alpha = |K| \text{id}.
\]

Let us first notice that
\[
(11.685) \quad 1 \otimes_{CH} h^{-1}u = (\rho(h))^{-1} \otimes_{CH} u \quad \text{for any} \ h \in H \ \text{and} \ u \in U.
\]

---

**Proof of (11.680):** Let \( h' \in H, q \in CG \) and \( u \in U \). Then, the map \( H \rightarrow H, h \mapsto h'h \) is a bijection (since \( H \) is a group). The definition of \( \bar{\beta} \) yields
\[
\bar{\beta}(q, h') = \sum_{h \in H} \varepsilon_1(\rho(h)q) h^{-1}h'u = \sum_{h \in H} \varepsilon_1(\rho(h)q) h^{-1}(h')^{-1}h'u
\]
\[
= \sum_{h \in H} \varepsilon_1(\rho(h)q) h^{-1}(h')^{-1}h'u
\]
\[
= \sum_{h \in H} \varepsilon_1(\rho(h)q) \rho(h')(h')^{-1}h'u
\]
\[
= \sum_{h \in H} \varepsilon_1(\rho(h)q) \rho(h')(h')^{-1}h'u
\]
\[
= \sum_{h \in H} \varepsilon_1(\rho(h)q) \rho(h')(h')^{-1}h'u
\]
\[
= \bar{\beta}(q, h') u.
\]

But the definition of the right \( CH \)-module structure on \( CG \) yields \( \gamma \eta = \gamma \cdot (\mathbb{C}[\rho]) \eta \) for all \( \gamma \in CG \) and \( \eta \in CH \). Applying this to \( \gamma = q \) and \( \eta = h' \), we obtain \( qh' = q \cdot (\mathbb{C}[\rho]) (h') \). But \( h' \in H \) and thus \( (\mathbb{C}[\rho]) (h') = \rho(h') \). Hence, \( qh' = q \cdot (\mathbb{C}[\rho]) (h') = q \rho(h') \).

But the definition of \( \bar{\beta} \) also yields
\[
\bar{\beta}(qh', u) = \sum_{h \in H} \varepsilon_1(\rho(h)qh') h^{-1}u = \sum_{h \in H} \varepsilon_1(\rho(h)q \rho(h')) h^{-1}u
\]
\[
= \bar{\beta}(q, h') u \quad \text{(by (11.681))}.
\]

In other words, \( \bar{\beta}(q, h') = \bar{\beta}(qh', u) \). This proves (11.680).

**Proof of (11.682):** Let \( r \in CH, q \in CG \) and \( u \in U \). We need to prove the equality (11.682). But this equality is \( C \)-linear in \( r \). Hence, we can WLOG assume that \( r \) belongs to the basis \( H \) of the \( C \)-vector space \( CH \). Assume this. Then, (11.682) follows from (11.680) (applied to \( h' = r \)).

**Proof of (11.685):** Let \( h \in H \) and \( u \in U \). We have \( h \in H \). Since \( H \) is a group, this yields \( h^{-1} \in H \subset CH \). Hence, we can move the \( h^{-1} \) past the tensor sign in \( 1 \otimes_{CH} h^{-1}u \). We thus obtain \( 1 \otimes_{CH} h^{-1}u = 1h^{-1} \otimes_{CH} u \).

But the definition of the right \( CH \)-module structure on \( CG \) yields \( \eta \gamma = \gamma \cdot (\mathbb{C}[\rho]) \eta \) for all \( \gamma \in CG \) and \( \eta \in CH \). Applying this to \( \gamma = 1 \) and \( \eta = h^{-1} \), we obtain \( 1h^{-1} = 1 \cdot (\mathbb{C}[\rho]) (h^{-1}) \). But \( h^{-1} \in H \) and thus \( (\mathbb{C}[\rho]) (h^{-1}) = \rho(h^{-1}) = (\rho(h))^{-1} \) (since \( \rho \)}
Now, we are going to show that $\alpha \circ \beta = |K| \text{id}$. In fact, let $q \in CG$ and $u \in U$ be arbitrary. Then,

$$(\alpha \circ \beta)(q \otimes_{CH} u) = \alpha(\beta(q \otimes_{CH} u)) = 1 \otimes_{CH} \beta(q \otimes_{CH} u)$$

(by the definition of $\alpha$)

$$= 1 \otimes_{CH} \left( \sum_{h \in H} \epsilon_1(\rho(h)q)h^{-1}u \right)$$

(by (11.684))

$$= \sum_{g \in G} \sum_{h \in H; \rho(h) = g} \epsilon_1(\rho(h)q)\left(\rho(h)\right)^{-1} \otimes_{CH} u$$

(by (11.685))

$$= \sum_{g \in G} \left| \rho^{-1}(g) \right| \epsilon_1(\rho^{-1}(g)q)g^{-1} \otimes_{CH} u = |K| \epsilon_1(\rho^{-1}(g)q)g^{-1} \otimes_{CH} u$$

(by (11.673))

$$= |K| \left( \sum_{g \in G} \epsilon_1(\rho^{-1}(g)q)g^{-1} \right) \otimes_{CH} u = |K| \left( \sum_{g \in G} \epsilon_1(\rho^{-1}q)g^{-1} \right) \otimes_{CH} u$$

(by (11.676))

$$= |K| q \otimes_{CH} u = |K| \text{id}(q \otimes_{CH} u).$$

Now, let us forget that we fixed $q$ and $u$. We have shown that $(\alpha \circ \beta)(q \otimes_{CH} u) = |K| \text{id}(q \otimes_{CH} u)$ for all $q \in CG$ and $u \in U$. In other words, the two maps $\alpha \circ \beta: CG \otimes_{CH} U \to CG \otimes_{CH} U$ and $|K| \text{id}: CG \otimes_{CH} U \to CG \otimes_{CH} U$ are equal to each other on each pure tensor. Since these two maps are $C$-linear, this yields that these two maps $\alpha \circ \beta: CG \otimes_{CH} U \to CG \otimes_{CH} U$ and $|K| \text{id}: CG \otimes_{CH} U \to CG \otimes_{CH} U$ must be identical. In other words, $\alpha \circ \beta = |K| \text{id}$.

Next, we are going to show that $\beta \circ \alpha = |K| \text{id}$.

We first notice that

$$(11.685) \quad h^{-1}u = (\rho(h))^{-1}u \quad \text{for every } h \in H \text{ and } u \in U^K.$$

is a group homomorphism). Hence, $1h^{-1} = 1 \cdot (C[\rho](h^{-1})) = 1 \cdot (\rho(h))^{-1} = (\rho(h))^{-1}$. Now, $1 \otimes_{CH} h^{-1}u = (\rho(h))^{-1} \otimes_{CH} u$. This proves (11.685).

560 Proof of (11.686): Let $h \in H$ and $u \in U^K$. Applying (11.674) to $h^{-1}$ and $u$ instead of $h$ and $v$, we obtain $\rho(h^{-1}) \cdot u = h^{-1}u$. But $\rho$ is a group homomorphism, and thus we have $(\rho(h))^{-1} = (h^{-1})$. Hence, $\rho(h^{-1}) \cdot u = \rho(h^{-1})u = (\rho(h^{-1})$ $\cdot u = h^{-1}u$. This proves (11.686).
Let $u \in U^K$. The definition of $\alpha (u)$ yields $\alpha (u) = 1 \otimes_{CH} u$. Applying the map $\beta$ to this equality, we obtain

$$\beta (\alpha (u)) = \beta (1 \otimes_{CH} u) = \sum_{h \in H} \sum_{g \in G} \epsilon_1 (\rho (h) 1) \frac{h^{-1}u}{\epsilon_1 (\rho (h) 1) \frac{h^{-1}u}{\rho (h) = g}} = \sum_{g \in G} |K| \epsilon_1 (g 1) g^{-1}u$$

(by (11.684), applied to $q = 1$)

$$= \sum_{g \in G} \left[ \frac{\rho^{-1} (g)}{|K|} \right] \epsilon_1 (g 1) g^{-1}u = \sum_{g \in G} |K| \epsilon_1 (g 1) g^{-1}u$$

(by (11.673))

$$= |K| \sum_{g \in G} \epsilon_1 (g 1) g^{-1}u = |K| \sum_{g \in G} \epsilon_1 (g^{-1} 1) (g^{-1})^{-1}u$$

(here, we substituted $g^{-1}$ for $g$ in the sum, since the map $G \rightarrow G, \ g \mapsto g^{-1}$ is a bijection (since $G$ is a group))

$$= |K| \sum_{g \in G} \epsilon_1 (g^{-1} 1) \epsilon_1 (g) (1) \epsilon_1 (g) (1) = |K| \sum_{g \in G} \epsilon_1 (g) (1) g$$

(by (11.679), applied to $q = 1$)

$$= |K| \frac{u}{\epsilon_1 (1)} = |K| \text{id} (u).$$

Now, forget that we fixed $u$. We thus have shown that $(\beta \circ \alpha) (u) = |K| \text{id} (u)$ for every $u \in U^K$. In other words, $\beta \circ \alpha = |K| \text{id}$.

The equalities $\alpha \circ \beta = \frac{1}{|K|} = 1 = \frac{1}{|K|} \text{id}$ id and $\frac{1}{|K|} \text{id} = \frac{1}{|K|} \epsilon_1 (1) g$ show that the maps $\alpha$ and $\beta$ are mutually inverse. Hence, the map $\alpha$ is invertible.

Now, we know that the map $\alpha$ is an invertible $CG$-module homomorphism. Hence, $\alpha$ is a $CG$-module isomorphism. Therefore, there exists a $CG$-module isomorphism $U^K \rightarrow CG \otimes CH U$ (namely, $\alpha$). Therefore, $U^K \cong CG \otimes CH U = \text{Ind}_p U$ as $CG$-modules. This solves Exercise 4.14(f).

[Remark: There is another solution of Exercise 4.14(f), which uses the result of Exercise 4.12(b). Yet another solution of Exercise 4.14(f) relies on Exercise 4.14(i), and will be given after the solution of the latter.]

(g) Let $\alpha \in R_C (H)$ and $\beta \in R_C (G)$. We notice that

$$\text{(11.687)} \quad \text{any two conjugate elements } k \text{ and } k' \text{ of } G \text{ satisfy } \beta (k) = \beta (k')$$

$561$.

$561$Proof of (11.687): We can apply (11.645) to $K = G$ and $f = \beta$. As a consequence, we obtain the following equivalence:

$$(\beta \in R_C (G))$$

$$\iff \text{(any two conjugate elements } k \text{ and } k' \text{ of } G \text{ satisfy } \beta (k) = \beta (k') \text{).}$$

Hence, any two conjugate elements $k$ and $k'$ of $G$ satisfy $\beta (k) = \beta (k')$ (because we know that $\beta \in R_C (G)$).
Let us first prove (4.17). The definition of \( \langle \text{Ind}_\rho \alpha, \beta \rangle_G \) yields

\[
\langle \text{Ind}_\rho \alpha, \beta \rangle_G = \frac{1}{|G|} \sum_{g \in G} (\text{Ind}_\rho \alpha)(g) \beta(g^{-1}) = \frac{1}{|H|} \sum_{(h,k) \in H \times G; \ k \rho(h) k^{-1} = g} \alpha(h) \beta(g^{-1})
\]

(by the definition of \( \text{Ind}_\rho \alpha \))

\[
= \frac{1}{|G|} \frac{1}{|H|} \sum_{g \in G} \sum_{(h,k) \in H \times G; \ k \rho(h) k^{-1} = g} \alpha(h) \beta(g^{-1}) = \sum_{(h,k) \in H \times G} \sum_{g \in G; \ k \rho(h) k^{-1} = g} \alpha(h) \beta(g^{-1})
\]

\[
(11.688)
\]

But for every \((h,k) \in H \times G\), we have

\[
\sum_{g \in G; \ k \rho(h) k^{-1} = g} \alpha(h) \beta(g^{-1}) = \alpha(h) \beta((k \rho(h) k^{-1})^{-1})
\]

\[
(11.689)
\]
Thus, (11.688) becomes

\[ (\text{Ind}_\rho \alpha, \beta)_G = \frac{1}{|G|} \frac{1}{|H|} \sum_{(h,k) \in H \times G} \sum_{g \in G; k \rho(h) k^{-1} = g} \alpha(h) \beta \left( (k \rho(h) k^{-1})^{-1} \right) \]

\[ = \frac{1}{|G|} \frac{1}{|H|} \sum_{h \in H} \sum_{k \in G} \alpha(h) \left( \text{Res}_\rho \beta \right)(h^{-1}) = \frac{1}{|G|} \frac{1}{|H|} \sum_{h \in H} |G| \cdot \alpha(h) \left( \text{Res}_\rho \beta \right)(h^{-1}) \]

\[ = \frac{1}{|H|} \sum_{h \in H} \alpha(h) \left( \text{Res}_\rho \beta \right)(h^{-1}). \]

Compared with

\[ \langle \alpha, \text{Res}_\rho \beta \rangle_H = \frac{1}{|H|} \sum_{g \in H} \alpha(g) \left( \text{Res}_\rho \beta \right)(g^{-1}) = \text{(by the definition of } \langle \alpha, \text{Res}_\rho \beta \rangle_H \text{)} \]

\[ = \frac{1}{|H|} \sum_{h \in H} \alpha(h) \left( \text{Res}_\rho \beta \right)(h^{-1}) \]

(here, we renamed the summation index \( g \) as \( h \)),

this yields \( (\text{Ind}_\rho \alpha, \beta)_G = \langle \alpha, \text{Res}_\rho \beta \rangle_H \). Thus, (4.17) is proven.

---

**Proof of (11.689):** Fix a \((h, k) \in H \times G\). Thus, \( h \in H \) and \( k \in G \). Now, there exists only one \( g \in G \) such that \( k \rho(h) k^{-1} = g \) (namely, \( g = k \rho(h) k^{-1} \)). Hence, the sum \( \sum_{g \in G; k \rho(h) k^{-1} = g} \alpha(h) \beta \left( (k \rho(h) k^{-1})^{-1} \right) \) has only one addend. Hence, this sum simplifies as follows:

\[ \sum_{g \in G; k \rho(h) k^{-1} = g} \alpha(h) \beta \left( (k \rho(h) k^{-1})^{-1} \right) \]

\[ = \alpha(h) \beta \left( (k \rho(h) k^{-1})^{-1} \right) = \alpha(h) \beta \left( k^{-1} \left( (\rho(h))^{-1} \right)^{-1} \right) = \alpha(h) \beta \left( k \rho(h) k^{-1} \right). \]  

(11.690)

But \( \rho(h^{-1}) \) and \( k \rho(h^{-1}) k^{-1} \) are two conjugate elements of \( G \). Hence, (11.687) (applied to \( \rho(h^{-1}) \) and \( k \rho(h^{-1}) k^{-1} \) instead of \( k \) and \( k' \)) yields \( \beta \left( \rho(h^{-1}) \right) = \beta \left( k \rho(h^{-1}) k^{-1} \right) \). Thus,

\[ \beta \left( (k \rho(h) k^{-1})^{-1} \right) = \beta \left( \rho(h^{-1}) \right) = \frac{\beta \circ \rho}{\text{Res}_\rho \beta} \]  

(hence, \( \text{Res}_\rho \beta \) is defined as \( \beta \circ \rho \)).

Hence, (11.690) becomes

\[ \sum_{g \in G; k \rho(h) k^{-1} = g} \alpha(h) \beta \left( (k \rho(h) k^{-1})^{-1} \right) = \alpha(h) \beta \left( k \rho(h^{-1}) k^{-1} \right) = \alpha(h) \left( \text{Res}_\rho \beta \right)(h^{-1}). \]

This proves (11.689).
Let us now prove (4.16). (This proof will be very much similar to the proof of (4.17) above, but in a few places even easier.) The definition of \((\text{Ind}_{\rho} \alpha, \beta)_G\) yields

\[
(\text{Ind}_{\rho} \alpha, \beta)_G = \frac{1}{|G|} \sum_{g \in G} (\text{Ind}_{\rho} \alpha)(g) \beta(g) = \frac{1}{|H|} \sum_{(h,k) \in H \times G; \ k \rho(h) k^{-1} = g} \frac{1}{|G|} \sum_{g \in G} \alpha(h) \beta(g).
\]

(by the definition of \(\text{Ind}_{\rho} \alpha\))

\[
= \frac{1}{|G|} \frac{1}{|H|} \sum_{g \in G} \sum_{(h,k) \in H \times G; \ k \rho(h) k^{-1} = g} \alpha(h) \beta(g)
\]

\[
= \sum_{(h,k) \in H \times G} \sum_{g \in G; \ k \rho(h) k^{-1} = g} \alpha(h) \beta(k \rho(h) k^{-1}).
\]

(11.691)

But for every \((h,k) \in H \times G\), we have

\[
\sum_{g \in G; \ k \rho(h) k^{-1} = g} \alpha(h) \beta(k \rho(h) k^{-1}) = \alpha(h) (\text{Res}_{\rho} \beta)(h)
\]

(11.692)
Thus, \((11.691)\) becomes

\[
(\text{Ind}_p \alpha, \beta)_G = \frac{1}{|G| |H|} \sum_{(h,k) \in H \times G} \alpha(h) \beta(k \rho(h) k^{-1})
\]

Compared with

\[
(\alpha, \text{Res}_p \beta)_G = \frac{1}{|G|} \sum_{g \in G} \alpha(g) \beta(\text{Res}_p \beta)(g)
\]

(by the definition of \((\alpha, \text{Res}_p \beta)_G\))

\[
= \frac{1}{|H|} \sum_{h \in H} \alpha(h) \beta(\text{Res}_p \beta)(h)
\]

(here, we renamed the summation index \(g\) as \(h\)).

this yields \((\text{Ind}_p \alpha, \beta)_G = (\alpha, \text{Res}_p \beta)_H\). Thus, \((4.16)\) is proven. The solution of Exercise 4.14(g) is thus complete.

(h) Let \(U\) be a \(CH\)-module, and let \(V\) be a \(CG\)-module.

Recall that \(\text{Ind}_p U\) is defined as the \(CG\)-module \(CG \otimes_C U\), where \(CG\) is regarded as a \((CG, CH)\)-bimodule according to the following rule: The left \(CG\)-module structure on \(CG\) is plain multiplication inside \(CG\); the right \(CH\)-module structure on \(CG\) is induced by the \(C\)-algebra homomorphism \(\mathbb{C} [\rho] : CH \to CG\) (thus, it is explicitly given by \(\gamma \eta = \gamma \cdot (C[\rho])\eta\) for all \(\gamma \in CG\) and \(\eta \in CH\)). From now on, we regard \(CG\) as endowed with this \((CG, CH)\)-bimodule structure.

Now, \((4.8)\) (applied to \(R = CH, S = CG, A = CG, B = U\) and \(C = V\)) yields

\[(11.694) \quad \text{Hom}_{CG}(CG \otimes_CH U, V) \cong \text{Hom}_{CH}(U, \text{Hom}_{CG}(CG, V)).\]

We shall now prove that \(\text{Hom}_{CG}(CG, V) \cong \text{Res}_p V\) as left \(CH\)-modules.

Indeed, a fundamental fact in abstract algebra says the following: If \(A\) is a \(C\)-algebra, and if \(M\) is a left \(A\)-module, then there exists a \(C\)-vector space isomorphism \(\Xi : \text{Hom}_A (A, M) \to M\) which satisfies \((\Xi (f)) = f(1)\) for every \(f \in \text{Hom}_A (A, M)\). Applying this fact to \(A = CG\) and \(M = V\), we conclude that

563 Proof of \((11.692)\): Fix a \((h, k) \in H \times G\). Thus, \(h \in H\) and \(k \in G\). Now, there exists only one \(g \in G\) such that \(k \rho(h) k^{-1} = g\) (namely, \(g = k \rho(h) k^{-1}\)). Hence, the sum \(\sum_{g \in G, k \rho(h) k^{-1} = g} \alpha(h) \beta(\text{Res}_p \beta)(k \rho(h) k^{-1})\) has only one addend. Hence, this sum simplifies as follows:

\[
(11.693) \quad \sum_{g \in G, k \rho(h) k^{-1} = g} \alpha(h) \beta(\text{Res}_p \beta)(k \rho(h) k^{-1}) = \alpha(h) \beta(\text{Res}_p \beta)(k \rho(h) k^{-1}).
\]

But \(\rho(h)\) and \(k \rho(h) k^{-1}\) are two conjugate elements of \(G\). Hence, \((11.687)\) (applied to \(\rho(h)\) and \(k \rho(h) k^{-1}\) instead of \(k\) and \(k'\)) yields \(\beta(\rho(h)) = \beta(k \rho(h) k^{-1})\). Thus,

\[
\beta(k \rho(h) k^{-1}) = \beta(\rho(h)) = \frac{\beta \circ \rho}{\text{Res}_p \beta} \quad (\text{since } \text{Res}_p \beta \text{ is defined as } \beta \circ \rho)
\]

Hence, \((11.693)\) becomes

\[
(11.694) \quad \sum_{g \in G, k \rho(h) k^{-1} = g} \alpha(h) \beta(\text{Res}_p \beta)(k \rho(h) k^{-1}) = \alpha(h) \beta(\text{Res}_p \beta)(k \rho(h) k^{-1}) = \alpha(h) \beta(\text{Res}_p \beta)(k \rho(h) k^{-1})
\]

This proves \((11.692)\).
there exists a $\mathbb{C}$-vector space isomorphism $\Xi : \text{Hom}_{CG}(\mathbb{C}G, V) \to V$ which satisfies $\Xi(f) = f(1)$ for every $f \in \text{Hom}_{CG}(\mathbb{C}G, V)$. Consider this $\Xi$.

The map $\Xi$ is a homomorphism of $H$-sets from $\text{Hom}_{CG}(\mathbb{C}G, V)$ to $\text{Res}_p V$. Therefore, $\Xi$ is a $\mathbb{C}H$-module homomorphism from $\text{Hom}_{CG}(\mathbb{C}G, V)$ to $\text{Res}_p V$ (since $\Xi$ is a $\mathbb{C}$-linear map). Consequently, $\Xi$ is a $\mathbb{C}H$-module isomorphism from $\text{Hom}_{CG}(\mathbb{C}G, V)$ to $\text{Res}_p V$ (since $\Xi$ is a $\mathbb{C}$-vector space isomorphism). Therefore, $\text{Hom}_{CG}(\mathbb{C}G, V) \cong \text{Res}_p V$ as $\mathbb{C}H$-modules. Now,

$$\text{Hom}_{CG} \left( \underbrace{\text{Ind}_p U}_{=CG \otimes_{\mathbb{C}H} U} , V \right) = \text{Hom}_{CG} \left( \mathbb{C}G \otimes_{\mathbb{C}H} U, V \right)$$

$$\cong \text{Hom}_{\mathbb{C}H} \left( U, \underbrace{\text{Hom}_{CG}(\mathbb{C}G, V)}_{\cong \text{Res}_p V \text{ as } \mathbb{C}H\text{-modules}} \right) \quad \text{(by (11.694))}$$

$$\cong \text{Hom}_{\mathbb{C}H} \left( U, \text{Res}_p V \right).$$

This solves Exercise 4.14(h).

(i) The following solution will mostly be an imitation of the solution of Exercise 4.4.

Let $U$ be any $\mathbb{C}H$-module. Recall that $\text{Ind}_p U$ is defined as the $\mathbb{C}G$-module $\mathbb{C}G \otimes_{\mathbb{C}H} U$, where $\mathbb{C}G$ is regarded as a $(\mathbb{C}G, \mathbb{C}H)$-bimodule according to the following rule: The left $\mathbb{C}G$-module structure on $\mathbb{C}G$ is plain multiplication inside $\mathbb{C}G$; the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ is induced by the $\mathbb{C}$-algebra homomorphism $\mathbb{C}[\rho] : \mathbb{C}H \to \mathbb{C}G$ (thus, it is explicitly given by $\gamma \eta = \gamma \cdot (\mathbb{C}[\rho]) \eta$ for all $\gamma \in \mathbb{C}G$ and $\eta \in \mathbb{C}H$). From now on, we regard $\mathbb{C}G$ as endowed with this $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure (besides the $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure that was introduced in the statement of Exercise 4.14(i)).

The $\mathbb{C}$-vector space $\mathbb{C}G$ is thus endowed with a left $\mathbb{C}H$-module structure (which is part of the $(\mathbb{C}G, \mathbb{C}H)$-bimodule structure) and with a right $\mathbb{C}H$-module structure (which is part of the $(\mathbb{C}H, \mathbb{C}G)$-bimodule structure). These two structures, combined, form a $(\mathbb{C}H, \mathbb{C}H)$-bimodule structure. This allows us to write expressions like $xyz$ with $x \in \mathbb{C}H$, $y \in \mathbb{C}G$ and $z \in \mathbb{C}H$, without having to disambiguate whether they mean $(xy)z$ or $x(yz)$.

We recall that $\text{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)$ is the left $\mathbb{C}G$-module consisting of all left $\mathbb{C}H$-module homomorphisms from $\mathbb{C}G$ to $U$. This uses only the $(\mathbb{C}H, \mathbb{C}H)$-bimodule structure on $\mathbb{C}G$ that was introduced in the definition of $\text{Ind}_p U$.

Proof. The map $\Xi$ is clearly a map from $\text{Hom}_{CG}(\mathbb{C}G, V)$ to $V$, therefore a map from $\text{Hom}_{CG}(\mathbb{C}G, V)$ to $\text{Res}_p V$ (since $\text{Res}_p V = V$ as sets).

Recall that the $\mathbb{C}H$-module structure on $\text{Res}_p V$ is given by

$$h \cdot v = \rho(h) \cdot v \quad \text{for every } h \in H \text{ and } v \in V. \quad (11.695)$$

Now, let $f \in \text{Hom}_{CG}(\mathbb{C}G, V)$ and $h \in H$. We are going to prove that $\Xi(h \cdot f) = h \cdot \Xi(f)$, where $h \cdot \Xi(f)$ is computed in the $\mathbb{C}H$-module $\text{Res}_p V$.

Indeed, the definition of the left $\mathbb{C}H$-module structure on $\text{Hom}_{CG}(\mathbb{C}G, V)$ yields $(\eta \cdot \alpha)(p) = \alpha(\eta p)$ for every $\eta \in \mathbb{C}H$, $\alpha \in \text{Hom}_{CG}(\mathbb{C}G, V)$ and $p \in \mathbb{C}G$. Applying this to $\eta = h$, $\alpha = f$ and $p = 1$, we obtain $h \cdot f(1) = f(1h)$ (since $h \in H \subset \mathbb{C}H$). But the definition of $\Xi$ yields $\Xi(h \cdot f) = (h \cdot f)(1) = f(1h)$.

Recall that the right $\mathbb{C}H$-module structure on $\mathbb{C}G$ is given by the equality $\gamma \eta = \gamma \cdot (\mathbb{C}[\rho]) \eta$ for all $\gamma \in \mathbb{C}G$ and $\eta \in \mathbb{C}H$. Applying this equality to $\gamma = 1$ and $\eta = h$, we obtain $1h = 1 \cdot (\mathbb{C}[\rho]) (h)$ (since $h \in H \subset \mathbb{C}H$). But $h \in H$ and thus $\mathbb{C}[\rho]) (h) = \rho(h)$. Thus, $1h = 1 \cdot (\mathbb{C}[\rho]) (h) = (\mathbb{C}[\rho]) (h) = \rho(h) = \rho(h) \cdot 1$. Now, $\Xi(h \cdot f) = f \left( \underbrace{1h}_{=\rho(h) \cdot 1} \right) = f(\rho(h) \cdot 1) = \rho(h) \cdot f(1)$ (since $f$ is a $\mathbb{C}G$-module homomorphism (because $f \in \text{Hom}_{CG}(\mathbb{C}G, V)$) and since $\rho(h) \in G \subset \mathbb{C}G$).

On the other hand, (11.695) (applied to $v = f(1)$) yields $h \cdot f(1) = \rho(h) \cdot f(1)$. Compared with $\Xi(h \cdot f) = \rho(h) \cdot f(1)$, this yields $\Xi(h \cdot f) = h \cdot f(1)$.

But the definition of $\Xi$ yields $\Xi(f) = f(1)$. Thus, $f(1) = \Xi(f)$, so that $\Xi(h \cdot f) = h \cdot \Xi(f)$.

Now, let us forget that we fixed $f$ and $h$. We thus have shown that $\Xi(h \cdot f) = h \cdot \Xi(f)$ for every $f \in \text{Hom}_{CG}(\mathbb{C}G, V)$ and $h \in H$. In other words, $\Xi$ is a homomorphism of $H$-sets from $\text{Hom}_{CG}(\mathbb{C}G, V)$ to $\text{Res}_p V$, qed.
Recall that the right \( \mathbb{C}H \)-module structure on \( \mathbb{C}G \) is given by

\[
(11.696) \quad \gamma \eta = \gamma \cdot (\mathbb{C} \rho) \eta \quad \text{for all } \gamma \in \mathbb{C}G \text{ and } \eta \in \mathbb{C}H.
\]

For every \( \gamma \in \mathbb{C}G \) and \( \eta \in H \), we have

\[
(11.697) \quad \gamma \eta = \gamma \cdot (\mathbb{C} \rho(\eta)) \eta \quad \text{(by (11.696) (since } \eta \in H \subseteq \mathbb{C}H)\text{))}
\]

\[\text{(since } \eta \in H)\]

On the other hand, the left \( \mathbb{C}H \)-module structure on \( \mathbb{C}G \) is induced by the \( \mathbb{C} \)-algebra homomorphism \( \mathbb{C} \rho : \mathbb{C}H \to \mathbb{C}G \). In other words, this structure is given by

\[
(11.698) \quad \eta \gamma = (\mathbb{C} \rho(\eta)) \gamma \quad \text{for all } \gamma \in \mathbb{C}G \text{ and } \eta \in \mathbb{C}H.
\]

For every \( \gamma \in \mathbb{C}G \) and \( \eta \in H \), we have

\[
(11.699) \quad \eta \gamma = (\mathbb{C} \rho(\eta)) \gamma \quad \text{(by (11.698) (since } \eta \in H \subseteq \mathbb{C}H)\text{))}
\]

\[\text{(since } \eta \in H)\]

The map \( \rho \) is a group homomorphism. Hence, \( \rho(H) \) is a subgroup of \( G \). Let us denote this subgroup by \( \overline{H} \). Thus, \( \overline{H} = \rho(H) \).

Let \( J \) be a system of distinct representatives for the right \( \overline{H} \)-cosets in \( G \). Then, \( G = \bigcup_{j \in J} \overline{H} j \).

For every \( g \in G \), define a map \( \mathcal{R}_g : J \to J \) as follows: Let \( i \in J \). Then, \( ig \in G = \bigcup_{j \in J} \overline{H} j \). Thus, there exists a unique \( j \in J \) such that \( ig \in \overline{H} j \). Define \( \mathcal{R}_g(i) \) to be this \( j \). Hence, we have defined \( \mathcal{R}_g(i) \) for every \( i \in J \). Thus, we have defined a map \( \mathcal{R}_g : J \to J \).

For every \( g \in G \) and \( i \in J \), we have

\[
(11.700) \quad ig \in \overline{H} \cdot \mathcal{R}_g(i)
\]

(because \( \mathcal{R}_g(i) \) is defined as the \( j \in J \) satisfying \( ig \in \overline{H} j \)).

It is easy to see that for every \( g \in G \), the map \( \mathcal{R}_g : J \to J \) is a bijection.\(^{566}\) We have

\[
(11.702) \quad j^{-1} \otimes_{\mathbb{C}H} f(\mathcal{R}_g(j))^{-1} \otimes_{\mathbb{C}H} f(\mathcal{R}_g(j)) \quad \text{in } \mathbb{C}G \otimes_{\mathbb{C}H} U
\]

\(^{566}\)\text{Proof.} Fix \( g \in G \).

The sets \( \overline{H} j \) for all \( j \in J \) are pairwise disjoint (since \( \bigcup_{j \in J} \overline{H} j \) is well-defined). In other words, if \( i \) and \( i' \) are two elements of \( J \) such that \( \overline{H} i \) and \( \overline{H} i' \) are not disjoint, then

\[
(11.701) \quad i = i'.
\]

Let us now prove that \( \mathcal{R}_g \) is injective.

Let \( i \) and \( i' \) be two elements of \( J \). Assume that \( \mathcal{R}_g(i) = \mathcal{R}_g(i') \). We will show that \( i = i' \).

From (11.700), we have \( ig \in \overline{H} \cdot \mathcal{R}_g(i) \). Hence, there exists some \( h \in \overline{H} \) such that \( ig = h \cdot \mathcal{R}_g(i) \). Consider this \( h \). We have

\[
\overline{H} \cdot ig = \overline{H}h \cdot \mathcal{R}_g(i) = \overline{H} \cdot \mathcal{R}_g(i).
\]

The same argument (for \( i' \) instead of \( i \)) yields \( \overline{H}i'g = \overline{H} \cdot \mathcal{R}_g(i') \). Hence, \( \overline{H}ig = \overline{H} \cdot \mathcal{R}_g(i) = \overline{H} \cdot \mathcal{R}_g(i') = \overline{H}i'g \).

Thus, \( \overline{H}i'g = \overline{H}i'gg^{-1} = \overline{H}i'g \), so that \( \overline{H}i'g = \overline{H}gg^{-1} = \overline{H}i \). Thus, the sets \( \overline{H}i \) and \( \overline{H}i' \) are not disjoint (because

\[
(\overline{H}i) \cap (\overline{H}i') = (\overline{H}i) \cap (\overline{H}i) = \overline{H}i \neq \emptyset.
\]

Therefore, \( i = i' \) (by (11.701)).

Now, forget that we fixed \( i \) and \( i' \). We thus have shown that if \( i \) and \( i' \) are two elements of \( J \) such that \( \mathcal{R}_g(i) = \mathcal{R}_g(i') \), then \( i = i' \). In other words, the map \( \mathcal{R}_g \) is injective. But the set \( J \) is finite (since it is a subset of the finite set \( G \)). Hence, \( \mathcal{R}_g \) is a map from a finite set (namely, \( J \)) to itself. Since we know that \( \mathcal{R}_g \) is injective, this yields that \( \mathcal{R}_g \) is surjective (because any injective map from a finite set to itself must be surjective). Hence, \( \mathcal{R}_g \) is bijective (since \( \mathcal{R}_g \) is injective and surjective), that is, a bijection, qed.
for every $g \in G$, every $j \in J$ and every $f \in \text{Hom}_{CH}(CG, U)$.\footnote{Proof of (11.702): Let $g \in G$, $j \in J$ and $f \in \text{Hom}_{CH}(CG, U)$. Applying (11.700) to $i = j$, we obtain $jg \in \overline{H} \cdot \mathcal{R}_g(j)$. In other words, there exists an $h \in \overline{H}$ such that $jg = h \cdot \mathcal{R}_g(j)$. Consider this $h$. We have}

We now define a map $\alpha : \text{Hom}_{CH}(CG, U) \to CG \otimes CH U$ by setting

$$\alpha(f) = \sum_{j \in J} j^{-1} \otimes_{CH} f(j) \quad \text{for all } f \in \text{Hom}_{CH}(CG, U).$$

This $\alpha$ is a map $\text{Hom}_{CH}(CG, U) \to \text{Ind}_p U$ (since $\text{Ind}_p U = CG \otimes CH U$).

We will show that $\alpha$ is a $CH$-module isomorphism.

First, let us prove that $\alpha$ is a left $CG$-module homomorphism. In fact, any $f \in \text{Hom}_{CH}(CG, U)$ and $g \in G$ satisfy

$$\alpha(gf) = \sum_{j \in J} j^{-1} \otimes_{CH} (gf)(j) \quad \text{(by the definition of $\alpha(gf)$)}$$

$$= \sum_{j \in J} \left[ j^{-1} \otimes_{CH} f(jg) \right] = \sum_{j \in J} g \cdot (\mathcal{R}_g(j))^{-1} \otimes_{CH} f(\mathcal{R}_g(j)) = \sum_{j \in J} g \cdot j^{-1} \otimes_{CH} f(j)$$

$$= g \cdot \sum_{j \in J} j^{-1} \otimes_{CH} f(j) = g \cdot \alpha(f).$$

The map $\alpha$ is thus a homomorphism of left $G$-sets. Since $\alpha$ is furthermore $C$-linear, this yields that $\alpha$ is a left $CG$-module homomorphism.

We now are going to construct an inverse for $\alpha$. This will be more cumbersome.

For every $g \in G$ and every $p \in CG$, we denote by $\epsilon_g(p)$ the $g$-coordinate of $p$ with respect to the basis $G$ of the $C$-vector space $CG$. By the definition of “coordinate”, we have

$$q = \sum_{g \in G} \epsilon_g(q) g \quad \text{for every } q \in CG. \quad (11.704)$$

For every $g \in G$, we have defined a map $\epsilon_g : CG \to C$ (because we have defined an element $\epsilon_g(p)$ for every $p \in CG$). This map $\epsilon_g$ is $C$-linear. We record some properties of these maps:}

\[ j^{-1} \equiv_{h \cdot \mathcal{R}_g(j)} h^{-1} \cdot \mathcal{R}_g(j) = j^{-1} \cdot (\mathcal{R}_g(j))^{-1} \]

\[ j^{-1} h = j^{-1} jg = g \cdot (\mathcal{R}_g(j))^{-1} \]

But $h \in \overline{H} = \rho(H)$. Hence, there exists a $h' \in H$ such that $h = \rho(h')$. Consider this $h'$.

Applying (11.699) to $\eta = h'$ and $\gamma = \mathcal{R}_g(j)$, we obtain $h' \cdot \mathcal{R}_g(j) = \rho(h') \cdot \mathcal{R}_g(j) = h \cdot \mathcal{R}_g(j) = jg$ (since $jg = h \cdot \mathcal{R}_g(j)$).

But the map $\alpha$ is left $CH$-linear (since $f \in \text{Hom}_{CG}(CG, U)$). Thus, $f(h' \cdot \mathcal{R}_g(j)) = h' \cdot f(\mathcal{R}_g(j))$ (since $h' \in H \subset CH$).

Since $h' \cdot \mathcal{R}_g(j) = jg$, this rewrites as $f(jg) = h' \cdot f(\mathcal{R}_g(j))$. Hence,

\[ (11.703) \]

\[ j^{-1} \otimes_{CH} f(jg) = j^{-1} \otimes_{CH} h' \cdot f(\mathcal{R}_g(j)) = j^{-1} h' \otimes_{CH} f(\mathcal{R}_g(j)) \]

(here, we have moved $h'$ past the $\otimes_{CH}$ sign, since $h' \in H \subset CH$).

On the other hand, applying (11.697) to $\eta = j^{-1}$ and $\gamma = h'$, we obtain $j^{-1} h' = j^{-1} \cdot \rho(h') = j^{-1} h = g \cdot (\mathcal{R}_g(j))^{-1}$. Hence,

\[ (11.703) \]

becomes

\[ j^{-1} \otimes_{CH} f(jg) = g \cdot (\mathcal{R}_g(j))^{-1} \otimes_{CH} f(\mathcal{R}_g(j)). \]

This proves (11.702).
In other words, we have
\[ \epsilon_g (h) = \delta_{g,h}. \]

We have
\[ \epsilon_1 (pq) = \epsilon_1 (qp) \quad \text{for all } p \in \mathbb{C}G \text{ and } q \in \mathbb{C}G. \]

We have
\[ \epsilon_1 (g^{-1}q) = \epsilon_g (q) \quad \text{for every } g \in G \text{ and } q \in \mathbb{C}G. \]

Now, fix \( q \in \mathbb{C}G \) and \( u \in U \). We let \( f_{q,u} \) be the map \( \mathbb{C}G \to U \) defined by
\[ f_{q,u} (p) = \sum_{h \in H} \epsilon_1 (\rho (h) pq) h^{-1} u \quad \text{for every } p \in \mathbb{C}G. \]

It is obvious that this map \( f_{q,u} \) is \( \mathbb{C} \)-linear. We will show that \( f_{q,u} \) is a left \( \mathcal{C} \)-module homomorphism.

The map \( f_{q,u} \) is a homomorphism of left \( \mathcal{H} \)-sets.\(^{571}\) Since \( f_{q,u} \) is furthermore \( \mathbb{C} \)-linear, this yields that \( f_{q,u} \) is a left \( \mathcal{C} \)-module homomorphism. Hence, \( f_{q,u} \in \text{Hom}_\mathcal{C} (\mathcal{C}G, U) \).

Now, forget that we fixed \( q \) and \( u \). We thus have defined a map \( f_{q,u} \in \text{Hom}_\mathcal{C} (\mathcal{C}G, U) \) for every \( q \in \mathbb{C}G \) and \( u \in U \). It is easy to see that this map \( f_{q,u} \) depends \( \mathbb{C} \)-linearly on each of \( q \) and \( u \). Now, define a map \( \tilde{\beta} : \mathcal{C}G \times U \to \text{Hom}_\mathcal{C} (\mathcal{C}G, U) \) by
\[ \tilde{\beta} (q, u) = f_{q,u} \quad \text{for every } (q, u) \in \mathcal{C}G \times U. \]

\(^{568}\)This equality has been proven in our solution of Exercise 4.4. (Namely, it appeared there as (11.625).)

\(^{569}\)This equality has been proven in our solution of Exercise 4.4. (Namely, it appeared there as (11.626).)

\(^{570}\)This equality has been proven in our solution of Exercise 4.4. (Namely, it appeared there as (11.627).)

\(^{571}\)Proof. In fact, for every \( h' \in H \) and every \( p \in \mathbb{C}G \), we have

\[ f_{q,u} (h'p) = \sum_{h \in H} \epsilon_1 \left( \rho (h) \underbrace{(h'p)}_{=\rho (h') \frac{\rho (h') - 1}{\rho (h)' - 1}} \right) q h^{-1} u \quad \text{(by the definition of } f_{q,u}) \]

\[ = \sum_{h \in H} \epsilon_1 (\rho (h) \rho (h') \cdot pq) h^{-1} u = \sum_{h \in H} \epsilon_1 \left( \underbrace{\rho (h') \rho (h) - 1}_{=\rho (h)'(\rho (h)) - 1} \right) q \left( h (h')^{-1} \right)^{-1} u \quad \text{(since } \rho \text{ is a group homomorphism)} \]

\[ \text{here, we have substituted } h (h')^{-1} \text{ for } h \text{ in the sum, because the map } H \to H, h \mapsto h (h')^{-1} \text{ is a bijection (since } H \text{ is a group and since } h' \in H) \]

\[ = \sum_{h \in H} \epsilon_1 \left( \rho (h) \left( \rho (h') \frac{\rho (h') - 1}{\rho (h)' - 1} \right) \frac{\rho (h) \rho (h') - 1}{\rho (h)' - 1} \right) h^{-1} u \]

\[ = \sum_{h \in H} \epsilon_1 (\rho (h) pq) h^{-1} u = \sum_{h \in H} \epsilon_1 (\rho (h) pq) h^{-1} u = h' \cdot f_{q,u} (p). \]

In other words, \( f_{q,u} \) is a homomorphism of left \( \mathcal{H} \)-sets, qed.
Then, \( \tilde{\beta} \) is a \( \mathbb{C} \)-bilinear map (because \( \tilde{\beta}(q,u) = f_{q,u} \) depends \( \mathbb{C} \)-linearly on each of \( q \) and \( u \)). We are now going to prove that the map \( \tilde{\beta} \) is \( CH \)-bilinear with respect to the right \( CH \)-module structure on \( CG \) and the left \( CH \)-module structure on \( U \).

In fact, every \( h' \in H \), \( q \in CG \) and \( u \in U \) satisfy

\[
\tilde{\beta}(q,h'u) = \tilde{\beta}(qh',u).
\]  

As a consequence of this, we can see that every \( r \in CH \), \( q \in CG \) and \( u \in U \) satisfy

\[
\tilde{\beta}(q,ru) = \tilde{\beta}(qr,u).
\]

In other words, the map \( \tilde{\beta} \) is \( CH \)-bilinear with respect to the right \( CH \)-module structure on \( CG \) and the left \( CH \)-module structure on \( U \) (since we already know that \( \tilde{\beta} \) is \( \mathbb{C} \)-bilinear). Hence, by the universal property of the tensor product, we conclude that there exists a unique \( \mathbb{C} \)-linear map \( \beta : CG \otimes_{CH} U \to \text{Hom}_{CH}(CG, U) \) such that every \( (q,u) \in CG \times U \) satisfies

\[
\beta(q \otimes u) = \tilde{\beta}(q,u).
\]

Consider this map \( \beta \). Clearly, every \( q \in CG \) and \( u \in U \) satisfy

\[
\beta(q \otimes u) = \tilde{\beta}(q,u) = f_{q,u} \quad \text{(by the definition of \( \tilde{\beta} \)).}
\]

\[572\] Proof of (11.709): Let \( h' \in H \), \( q \in CG \) and \( u \in U \). Then, the map \( H \to H \), \( h \mapsto h'h \) is a bijection (since \( H \) is a group).

Now, let \( p \in CG \). The definition of \( \tilde{\beta} \) yields \( \tilde{\beta}(q,h'u) = f_{q,h'u} \). Hence,

\[
\begin{aligned}
\left( \tilde{\beta}(q,h'u) \right)(p) &= f_{q,h'u}(p) = \sum_{h \in H} \epsilon_1(\rho(h)pq)h^{-1}h'u \\
&= \sum_{h \in H} \epsilon_1(\rho(h)pq)h^{-1}(h'u) \\
&= \sum_{h \in H} \epsilon_1(\rho(h')\rho(h)q)h^{-1}(h'u) \\
&= \sum_{h \in H} \epsilon_1(\rho(h)pq\rho(h'))h^{-1}(h'u). \\
\end{aligned}
\]

But the definition of \( \tilde{\beta} \) also yields \( \tilde{\beta}(qh',u) = f_{qh',u} \). Hence,

\[
\left( \tilde{\beta}(qh',u) \right)(p) = f_{qh',u}(p) = \sum_{h \in H} \epsilon_1(\rho(h)p\rho(h')q)h^{-1}u. \\
\]

Now, forget that we fixed \( p \). We have thus proven that \( \left( \tilde{\beta}(qh',u) \right)(p) = \left( \tilde{\beta}(q,h'u) \right)(p) \) for every \( p \in CG \). In other words, \( \tilde{\beta}(qh',u) = \tilde{\beta}(q,h'u) \), so that \( \tilde{\beta}(q,h'u) = \tilde{\beta}(qh',u) \). This proves (11.709).

\[573\] Proof of (11.711): Let \( r \in CH \), \( q \in CG \) and \( u \in U \). We need to prove the equality (11.711). But this equality is \( \mathbb{C} \)-linear in \( r \). Hence, we can WLOG assume that \( r \) belongs to the basis \( H \) of the \( \mathbb{C} \)-vector space \( CH \). Assume this. Then, (11.711) follows from (11.709) (applied to \( h' = r \)).
Hence, every $q \in CG$, $u \in U$ and $p \in CG$ satisfy

$$\beta(q \otimes_{CH} u) = f_{q,u}(p) = \sum_{h \in H} \epsilon_{1}(\rho(h)pq)h^{-1}u$$

(by (11.708)).

Let $K = \text{Ker} \rho$. Thus, $K$ is the kernel of a group homomorphism out of $H$ (since $\rho$ is a group homomorphism out of $H$), and therefore a normal subgroup of $H$.

We notice that

$$\rho^{-1}(g) = |K| \quad \text{for every } g \in \overline{H}. \quad \text{(11.714)}$$

We shall now show that the maps $\alpha$ and $\frac{1}{|K|} \beta$ are mutually inverse. To do so, we will show that

$\alpha \circ \beta = |K| \text{id}$ and $\beta \circ \alpha = |K| \text{id}.$

\footnote{\textit{Proof of (11.714)}: Let $g \in \overline{H}$. Then, $g \in \overline{H} = \rho(H)$. Hence, $g = \rho(x)$ for some $x \in H$. Let us fix such a $x$.

We know that $H$ is a group. Hence, the map $K \rightarrow xK$, $k \mapsto xk$ is a bijection. Thus, the sets $K$ and $xK$ are in bijection. Therefore, $|xK| = |K|$.

But for every $y \in H$, we have the following logical equivalence:

$$y \in \rho^{-1}(g) \iff \left( \rho(y) = \underbrace{g}_{=\rho(x)} \right) \iff \left( \rho(y) = \rho(x) \right) \iff \left( \rho(y) \cdot (\rho(x))^{-1} = 1 \right) \iff \left( \rho(yx^{-1}) = 1 \right) \iff \left( yx^{-1} \in \text{Ker} \rho \right) \iff \left( yx^{-1} \in K \right) \iff \left( y \in xK \right).$$

Hence, $\rho^{-1}(g) = xK$, so that $|\rho^{-1}(g)| = |xK| = |K|$. This proves (11.714).}
Let us first show that $\alpha \circ \beta = \text{id}$. In fact, let $q \in \mathbb{C}G$ and $u \in U$ be arbitrary. Then,

$$(\alpha \circ \beta)(q \otimes_C H u) = \alpha(\beta(q \otimes_C H u)) = \sum_{j \in J} j^{-1} \otimes_C H = \sum_{h \in H} \epsilon_1(\rho(h)q)h^{-1}u$$

(by the definition of $\alpha$)

$$= \sum_{j \in J} \sum_{h \in H} \epsilon_1(\rho(h)q) j^{-1} h^{-1} \otimes_C H u$$

(by (11.697), applied to $\gamma = j^{-1}$ and $\eta = h^{-1}$ (since $h^{-1} \in H$ (since $h \in H$)) and thus $h^{-1} \in \mathbb{C} H$ (since $h \in H$))

$$= \sum_{j \in J} \sum_{g \in H} \epsilon_1(\rho(h)q) j^{-1} g^{-1} \otimes_C H u$$

(by (11.714))

$$= \left| K \right| \cdot \epsilon_1(gj)j^{-1} \cdot g^{-1} \otimes_C H u$$

(by (11.714))
we obtain

\[ \sum_{j \in J} \sum_{g \in \Pi} |K| \cdot \epsilon_1 (gj q) (gj)^{-1} \otimes_{\mathcal{C} \mathcal{H}} u = |K| \cdot \sum_{j \in J} \sum_{g \in \Pi} \epsilon_1 (gj q) (gj)^{-1} \otimes_{\mathcal{C} \mathcal{H}} u \]

\[ = |K| \cdot \sum_{j \in J} \sum_{g \in \Pi} \epsilon_1 (gq) g^{-1} \otimes_{\mathcal{C} \mathcal{H}} u \]

(since \( \bigcup_{j \in J} \Pi_j = G \))

\[ = |K| \cdot \sum_{g \in G} \epsilon_1 (gq) g^{-1} \otimes_{\mathcal{C} \mathcal{H}} u = |K| \cdot \sum_{g \in G} \epsilon_1 \left( (g^{-1}q) \right) (g^{-1})^{-1} \otimes_{\mathcal{C} \mathcal{H}} u \]

(here, we substituted \( g^{-1} \) for \( g \) in the sum, since the map \( G \to G, \ g \mapsto g^{-1} \) is a bijection (since \( G \) is a group))

\[ = |K| \cdot \sum_{g \in G} \epsilon_2 (q) g \otimes_{\mathcal{C} \mathcal{H}} u = |K| \cdot \left( \sum_{g \in G} \epsilon_2 (q) g \right) \otimes_{\mathcal{C} \mathcal{H}} u \]

(by (11.704))

\[ = |K| \cdot (q \otimes_{\mathcal{C} \mathcal{H}} u) = |K| \cdot \text{id} (q \otimes_{\mathcal{C} \mathcal{H}} u) = (|K| \cdot \text{id}) (q \otimes_{\mathcal{C} \mathcal{H}} u). \]

Now, let us forget that we fixed \( q \) and \( u \). We thus have shown that \( (\alpha \circ \beta) (q \otimes_{\mathcal{C} \mathcal{H}} u) = (|K| \cdot \text{id}) (q \otimes_{\mathcal{C} \mathcal{H}} u) \) for all \( q \in \mathbb{C}G \) and \( u \in \mathbb{C}H \). In other words, the two maps \( \alpha \circ \beta : \mathbb{C}G \otimes_{\mathcal{C} \mathcal{H}} U \to \mathbb{C}G \otimes_{\mathcal{C} \mathcal{H}} U \) and \( |K| \cdot \text{id} : \mathbb{C}G \otimes_{\mathcal{C} \mathcal{H}} U \to \mathbb{C}G \otimes_{\mathcal{C} \mathcal{H}} U \) are equal to each other on each pure tensor. Since these two maps are \( \mathbb{C} \)-linear, this yields that these two maps \( \alpha \circ \beta : \mathbb{C}G \otimes_{\mathcal{C} \mathcal{H}} U \to \mathbb{C}G \otimes_{\mathcal{C} \mathcal{H}} U \) and \( |K| \cdot \text{id} : \mathbb{C}G \otimes_{\mathcal{C} \mathcal{H}} U \to \mathbb{C}G \otimes_{\mathcal{C} \mathcal{H}} U \) must be identical. In other words, \( \alpha \circ \beta = |K| \cdot \text{id} \).

Next, we are going to show that \( \beta \circ \alpha = |K| \cdot \text{id} \).

Let \( f \in \text{Hom}_{\mathcal{C} \mathcal{H}} (\mathbb{C}G, U) \). Let \( p \in \mathbb{C}G \). The map \( f \) is left \( \mathcal{C} \mathcal{H} \)-linear (since \( f \in \text{Hom}_{\mathcal{C} \mathcal{H}} (\mathbb{C}G, U) \)), hence \( \mathbb{C} \)-linear. We have \( \alpha (f) = \sum_{j \in J} j^{-1} \otimes_{\mathcal{C} \mathcal{H}} f (j) \) (by the definition of \( \alpha \)). Applying the map \( \beta \) to this equality, we obtain

\[ \beta (\alpha (f)) = \beta \left( \sum_{j \in J} j^{-1} \otimes_{\mathcal{C} \mathcal{H}} f (j) \right) = \sum_{j \in J} \beta \left( j^{-1} \otimes_{\mathcal{C} \mathcal{H}} f (j) \right) \]
(since $\beta$ is a $\mathbb{C}$-linear map). Thus,

\[
\begin{align*}
&\sum_{j \in J} \beta(j^{-1} \otimes \mathbb{C} H f(j)) \\
= &\left( \sum_{j \in J} \beta\left(j^{-1} \otimes \mathbb{C} H f(j)\right) \right)(p) = \sum_{j \in J} \left( \sum_{j \in J} \beta\left(j^{-1} \otimes \mathbb{C} H f(j)\right) \right)(p) \\
= &\sum_{j \in J} \left( \sum_{h \in H} \epsilon_1\left(\rho(h) p j^{-1}\right) h^{-1} f(j)\right). \\
&= \sum_{j \in J} \sum_{h \in H} \epsilon_1\left(\rho(h) p j^{-1}\right) h^{-1} f(j) = \sum_{j \in J} \sum_{h \in H} \epsilon_1\left(\rho(h^{-1}) p j^{-1}\right) (h^{-1})^{-1} f(j) \\
&\text{here, we substituted } h^{-1} \text{ for } h \text{ in the second sum,} \\
&\text{since the map } H \rightarrow H, h \mapsto h^{-1} \text{ is a bijection} \\
&\text{(because } H \text{ is a group)} \\
&= \sum_{j \in J} \sum_{h \in H} \epsilon_1\left(\rho(h^{-1}) p j^{-1}\right) (j^{-1})^{-1} p f(j) \\
&\text{since } f \text{ is left } \mathbb{C} H \text{-linear and since } h \in H \subseteq H \\
&= \sum_{j \in J} \sum_{h \in H} \epsilon_1\left(\rho(h) p j^{-1}\right) h f(j) \\
&\text{by (11.713), applied to } h=g \\
&= \sum_{j \in J} \sum_{h \in H} \epsilon_1\left(\rho(h) j^{-1} p\right) f(\rho(h) j) \\
&\text{by (11.699), applied to } \eta=h \text{ and } \gamma=j \\
&= \sum_{j \in J} \sum_{h \in H} \epsilon_1\left(\rho(h) j^{-1} p\right) f(\rho(h) j) \\
&\text{by (11.714))} \\
&= \sum_{j \in J} \sum_{g \in H} \sum_{h \in H: p h=j} \epsilon_1\left(\rho(h) j^{-1} p\right) f(\rho(h) j) \\
&\text{(since every } h \in H \text{ satisfies } p h j=g) \\
&= \sum_{j \in J} \sum_{g \in H} \sum_{h \in H: p h=g} \epsilon_1\left(\rho(h) j^{-1} p\right) f(\rho(h) j) \\
&\text{by (11.714))} \\
&= \sum_{j \in J} \sum_{g \in H} \sum_{h \in H: p h=g} \epsilon_1\left(\rho(h) j^{-1} p\right) f(g j) \\
&\text{by (11.713), applied to } h=g \\
&= \sum_{j \in J} \sum_{g \in H} \sum_{h \in H: p h=g} \epsilon_1\left(\rho(h) j^{-1} p\right) f(\rho(h) j) \\
&\text{(by (11.714))} \\
&= \sum_{j \in J} \sum_{g \in H} \sum_{h \in H: p h=g} \epsilon_1\left(\rho(h) j^{-1} p\right) f(g j) \\
&\text{by (11.714))} \\
&= \sum_{j \in J} \sum_{g \in H} \sum_{h \in H: p h=g} \epsilon_1\left(\rho(h) j^{-1} p\right) f(g j) \\
&\text{(by (11.714))} \\
&= \sum_{j \in J} \sum_{g \in H} \sum_{h \in H: p h=g} \epsilon_1\left(\rho(h) j^{-1} p\right) f(g j). \\
\end{align*}
\]
\[
= |K| \cdot \sum_{j \in J} \sum_{g \in \mathcal{H}} \epsilon_1 \left( (gj)^{-1} p \right) f(gj) = |K| \cdot \sum_{j \in J} \sum_{g \in \mathcal{H} j} \epsilon_1 (g^{-1} p) f(g)
\]

(since \( \bigcup_{j \in J} \mathcal{H} j = \mathcal{G} \))

\[
(\text{here, we substituted } g \text{ for } gj \text{ in the second sum, since the map } \mathcal{H} \to \mathcal{H} j, g \to gj \text{ is a bijection (because } G \text{ is a group)})
\]

\[
= |K| \cdot \sum_{g \in G} \epsilon_1 (g^{-1} p) f(g) = |K| \cdot \sum_{g \in G} \epsilon_g(p) f(g) = |K| \cdot f\left( \sum_{g \in G} \epsilon_g(p) g \right).
\]

Compared with

\[
(|K| f)(p) = |K| \cdot f\left( \sum_{g \in G} \epsilon_g(p) g \right) = |K| \cdot f\left( \sum_{g \in G} \epsilon_g(p) g \right),
\]

this yields \((\beta (\alpha (f)))(p) = (|K| f)(p)\).

Now, forget that we fixed \(p\). We thus have proven that \((\beta (\alpha (f)))(p) = (|K| f)(p)\) for every \(p \in \mathbb{C}G\). In other words, \(\beta (\alpha (f)) = |K| f\). Hence, \((\beta \circ \alpha)(f) = \beta (\alpha (f)) = |K| f = (|K| \text{id})(f)\).

Now, forget that we fixed \(f\). We thus have shown that \((\beta \circ \alpha)(f) = (|K| \text{id})(f)\) for every \(f \in \text{Hom}_{\mathbb{C}H}(\mathbb{C}G, U)\). In other words, \(\beta \circ \alpha = |K| \text{id}\).

We now have \(\beta \circ \alpha = |K| \text{id}\) and thus \(\left( \frac{1}{|K|} \beta \right) \circ \alpha = \frac{1}{|K|} (\beta \circ \alpha) = \frac{1}{|K|} |K| \text{id} = \text{id}\).

On the other hand, recall that \(\alpha \circ \beta = |K| \text{id}\). Hence, \(\alpha \circ \left( \frac{1}{|K|} \beta \right) = \frac{1}{|K|} (\alpha \circ \beta) = \frac{1}{|K|} |K| \text{id} = \text{id}\).

Combining the equalities \(\left( \frac{1}{|K|} \beta \right) \circ \alpha = \text{id}\) and \(\alpha \circ \left( \frac{1}{|K|} \beta \right) = \text{id}\), we conclude that the maps \(\alpha\) and \(\frac{1}{|K|} \beta\) are mutually inverse. Hence, the map \(\alpha\) is invertible.

Now, we know that the map \(\alpha\) is an invertible left \(\mathbb{C}G\)-module homomorphism. Hence, \(\alpha\) is a left \(\mathbb{C}G\)-module isomorphism. Therefore, there exists a left \(\mathbb{C}G\)-module isomorphism \(\text{Hom}_{\mathbb{C}H}(\mathbb{C}G, U) \to \mathbb{C}G \otimes_{\mathbb{C}H} U\) (namely, \(\alpha\)). Therefore, \(\text{Hom}_{\mathbb{C}H}(\mathbb{C}G, U) \cong \mathbb{C}G \otimes_{\mathbb{C}H} U \cong \text{Ind}_p U\) as left \(\mathbb{C}G\)-modules. This solves Exercise 4.14(i).

[Remark: In our above solution of Exercise 4.14(i), we explicitly constructed a \(\mathbb{C}G\)-module isomorphism \(\alpha : \text{Hom}_{\mathbb{C}H}(\mathbb{C}G, U) \to \text{Ind}_p U\). This isomorphism is functorial with respect to \(U\). It is also independent on the choice of \(J\) (this is not immediately clear from its definition, but it can be shown very easily, by observing that the tensor \(j^{-1} \otimes_{\mathbb{C}H} f(j)\) for \(j \in G\) depends only on the coset \(\mathcal{H} j\) and not on \(j\) itself).]

[Remark: Exercise 4.14(i) allows us to give yet another solution to Exercise 4.14(f):]

Second solution to Exercise 4.14(f): Assume that \(G = H/K\) for some normal subgroup \(K\) of \(H\). Let \(\rho : H \to G\) be the projection map. We want to prove that \(\text{Ind}_p U \cong U^K\) for every \(\mathbb{C}H\)-module \(U\).

We know that \(\rho\) is the projection map from \(H\) to \(H/K\). Hence, \(\rho\) is a surjective group homomorphism and has kernel \(\text{Ker } \rho = K\).

Let \(U\) be a \(\mathbb{C}H\)-module. Recall that \(U^K\) is a \(\mathbb{C}[H/K]\)-module, thus a \(\mathbb{C}G\)-module (since \(H/K = G\)). This \(\mathbb{C}G\)-module structure has the property that

\[
(11.715) \quad \rho(h) \cdot v = hv \quad \text{for any } h \in H \text{ and } v \in U^K.
\]
Let us make $CG$ into a $(CH, CG)$-bimodule as in Exercise 4.14(i). From now on, we regard $CG$ as endowed with this $(CH, CG)$-bimodule structure. Then, the $CG$-module $\text{Hom}_C (CG, U)$ (defined as in Exercise 4.4 using the $(CH, CG)$-bimodule structure on $CG$) is isomorphic to $\text{Ind}_C U$ (because of Exercise 4.14(i)). In other words, $\text{Hom}_C (CG, U) \cong \text{Ind}_C U$ as left $CG$-modules.

We recall that $\text{Hom}_C (CG, U)$ is the left $CG$-module consisting of all left $CH$-module homomorphisms from $CG$ to $U$. This uses only the $(CH, CG)$-bimodule structure on $CG$ that was introduced in the statement of Exercise 4.14(i) (but not the $(CG, CH)$-bimodule structure on $CG$ that was introduced in the definition of $\text{Ind}_C U$).

We have $f(1) \in U^K$ for every $f \in \text{Hom}_C (CG, U)$ \footnote{Proof of (11.715): The proof of (11.715) is identical with the proof of (11.674) in the First solution to Exercise 4.14(f).}. Thus, we can define a map $\phi : \text{Hom}_C (CG, U) \to U^K$ by

\[
(\phi(f) = f(1) \quad \text{for all } f \in \text{Hom}_C (CG, U)).
\]

Consider this $\phi$. This map $\phi$ is clearly $C$-linear (since $f(1)$ depends $C$-linearly on $f$). Also, the map $\phi$ is surjective\footnote{Proof of (11.716): Let $u \in U^K$. We are going to construct a map $f \in \text{Hom}_C (CG, U)$ which satisfies $f(1) = u$. Indeed, let us define a map $f : CG \to U$ by (11.716)} and injective\footnote{Proof of (11.716): Let $u \in U^K$. We are going to construct a map $f \in \text{Hom}_C (CG, U)$ which satisfies $f(1) = u$. Indeed, let us define a map $f : CG \to U$ by (11.716)}. Hence, the map $\phi$ is bijective. Thus, $\phi$ is a $C$-vector space isomorphism (since $\phi$ is $C$-linear).

Now, let us forget that we fixed $k$. We thus have shown that $f(k) = f(1)$ for every $k \in K$. In other words, $f(1)$ is an element $y$ of $U$ which satisfies $ky = y$ for every $k \in K$. In other words,

\[
f(1) \in \{ y \in U \mid ky = y \text{ for every } k \in K \} = U^K.
\]

\text{qed.}

\footnote{Proof of (11.716): Let $u \in U^K$. We are going to construct a map $f \in \text{Hom}_C (CG, U)$ which satisfies $f(1) = u$. Indeed, let us define a map $f : CG \to U$ by (11.716). (This is well-defined, since $U^K$ is a $C$-module.) This map $f$ is $C$-linear (since $\gamma u$ depends $C$-linearly on $\gamma$). We shall now check that the map $f$ is a homomorphism of left $H$-sets.

Indeed, let $h \in H$ and $\gamma \in CG$. We have $f(h\gamma) = (h\gamma) u$ (by the definition of $f$).

On the other hand, (11.699) (applied to $\gamma = 1_{CG}$ and $\eta = k$) yields $k1_{CG} = \rho(k)1_{CG} = 1_{CG}1_{CG} = 1_{CG}$.

But $f \in \text{Hom}_C (CG, U)$. Hence, the map $f$ is left $CH$-linear. Thus, $f(k1_{CG}) = f(1_{CG})$ (since $k \in H \subset CH$). Since $k1_{CG} = 1_{CG}$, this rewrites as $f(1_{CG}) = kf(1) = f(1)$.}

\text{Now, let us forget that we fixed $k$ and $\gamma$. We thus have proven that $f(h\gamma) = hf(\gamma)$ for all $h \in H$ and $\gamma \in CG$. In other words, $f$ is a homomorphism of left $H$-sets. Hence, $f$ is a left $CH$-module homomorphism (since $f$ is $C$-linear).}

\text{In other words, $f \in \text{Hom}_C (CH, U)$. Now, the definition of $\phi$ yields $\phi(f) = f(1) = 1u$ (by (11.716), applied to $\gamma = 1$).}

\[
\phi(f) = f(1) = 1u \quad \text{by (11.716), applied to } \gamma = 1.
\]

Thus, $u = \phi \left( \begin{array}{c} 1 \\ \in \text{Hom}_C (CG, U) \end{array} \right) \in \phi (\text{Hom}_C (CG, U)).$

\text{Let us now forget that we fixed $u$. We thus have shown that $u \in \phi (\text{Hom}_C (CG, U))$ for every $u \in U^K$. In other words, $U^K \subset \phi (\text{Hom}_C (CG, U))$. This proves that the map $\phi$ is surjective. Qed.}

\footnote{Proof of (11.716): Let $f \in \ker \phi$. Then, $f \in \text{Hom}_C (CG, U)$ and $f(1) = 0$ (since $f \in \ker \phi$). The definition of $\phi$ yields $\phi(f) = f(1_{CG}) = 0$, so that $f(1_{CG}) = \phi(f) = 0$.}

Now, let $g \in G$ be arbitrary. Then, there exists some $h \in H$ such that $g = \rho(h)$ (since $\rho$ is surjective). Consider this $h$. Applying (11.699) to $\gamma = 1_{CG}$ and $\eta = h$, we obtain $h1_{CG} = \rho(h)1_{CG} = \rho(h) = g$. But the map $f$ is left $CH$-linear (since
Moreover, \( \phi \) is a homomorphism of left \( G \)-sets, and therefore a left \( CG \)-module homomorphism (since \( f \) is \( C \)-linear). Thus, \( \phi \) is a left \( CG \)-module isomorphism (since \( \phi \) is a \( C \)-vector space isomorphism). We thus have found a left \( CG \)-module isomorphism from \( \text{Hom}_{CG}(CG, U) \) to \( U^K \) (namely, \( \phi \)). Hence, \( U^K \cong \text{Hom}_{CG}(CG, U) \) as left \( CG \)-modules. Hence, \( U^K \cong \text{Hom}_{CG}(CG, U) \cong \text{Ind} \rho U \) as left \( CG \)-modules. This solves Exercise 4.14(f) again.

(j) Let \( U \) be a \( CG \)-module, and let \( V \) be a \( CH \)-module.

Let us make \( CG \) into a \((CH, CG)\)-bimodule as in Exercise 4.14(i). From now on, we regard \( CG \) as endowed with this \((CH, CG)\)-bimodule structure. Then, the \( CG \)-module \( \text{Hom}_{CH}(CH, V) \) (defined as in Exercise 4.4 using the \((CH, CG)\)-bimodule structure on \( CG \)) is isomorphic to \( \text{Ind} \rho V \) (because of Exercise 4.14(i), applied to \( V \) instead of \( U \)). In other words,

\[
\text{Hom}_{CG}(CG, V) \cong \text{Ind} \rho V \quad \text{as left } CG \text{-modules.}
\]

Now, (4.8) (applied to \( R = CG, S = CH, A = CG, B = U \) and \( C = V \)) yields

\[
\text{Hom}_{CG}(CG \otimes CG U, V) \cong \text{Hom}_{CG}(U, \text{Hom}_{CG}(CG, V)).
\]

We shall now prove that \( CG \otimes CG U \cong \text{Res} \rho U \) as left \( CH \)-modules.

Indeed, a fundamental fact in abstract algebra says the following: If \( A \) is a \( C \)-algebra, and if \( \mathfrak{M} \) is a left \( A \)-module, then there exists a \( C \)-vector space isomorphism \( \Xi : A \otimes \mathfrak{M} \to \mathfrak{M} \) which satisfies \( (a \otimes a \mathfrak{m}) = a \mathfrak{m} \) for every \( a \in A \) and \( \mathfrak{m} \in \mathfrak{M} \). Applying this fact to \( A = CG \) and \( \mathfrak{M} = U \), we conclude that there exists a \( C \)-vector space isomorphism \( \Xi : CG \otimes CG U \to U \) which satisfies \( (a \otimes \mathfrak{m}) = a \mathfrak{m} \) for every \( a \in CG \) and \( \mathfrak{m} \in U \). Consider this \( \Xi \).

The map \( \Xi \) is a \( C \)-vector space isomorphism. Hence, its inverse \( \Xi^{-1} \) exists and also is a \( C \)-vector space isomorphism, and therefore a \( C \)-linear map. It furthermore satisfies

\[
\Xi^{-1}(u) = 1 \otimes CG u \quad \text{for every } u \in U.
\]

---

**Proof.** Let \( g \in G \) and \( f \in \text{Hom}_{CH}(CH, U) \). The definition of the left \( CG \)-module structure on \( \text{Hom}_{CH}(CH, U) \) yields

\[
(g f)(1) = f \begin{pmatrix} 1 \end{pmatrix} = \phi \begin{pmatrix} 1 \end{pmatrix} = f(g).
\]

Now, the definition of \( \phi \) yields \( \phi(g f) = (g f)(1) = f(g) \).

But there exists some \( h \in H \) such that \( g = \rho(h) \) (since \( \rho \) is surjective). Consider this \( h \). The map \( f \) is left \( CH \)-linear (since \( f \in \text{Hom}_{CH}(CH, U) \)). Thus, we have \( f(\rho(h)CG) = h f(1_{CG}) \). But (11.69) (applied to \( \gamma = 1_{CG} \) and \( \eta = h \)) yields

\[
\rho(\mathfrak{m})_hCG = g \cdot 1_{CG} = g,
\]

so that \( g = h \). Hence, \( f \begin{pmatrix} 1 \end{pmatrix} = f(h \cdot 1_{CG}) = h f(1_{CG}). \)

The definition of \( \phi \) yields \( \phi(f) = f \begin{pmatrix} 1 \end{pmatrix} = \phi(1_{CG}) \), so that \( f(1_{CG}) = \phi(f) \). Now, \( f(g) = h f(1_{CG}) = h \phi(f) \). Hence,

\[
\phi(g f) = f(g) = h \phi(f).
\]

The equality (11.715) (applied to \( \phi(f) \) instead of \( \psi \)) yields \( \rho(h) \cdot \phi(f) = h \phi(f) \). Thus, \( h \phi(f) = \rho(h) \cdot \phi(f) = g \cdot \phi(f) = g \phi(f) \).

Hence, \( \phi(g f) = h \phi(f) = g \phi(f) \).

Now, let us forget that we fixed \( g \) and \( f \). We thus have proven that \( \phi(g f) = g \phi(f) \) for all \( g \in G \) and \( f \in \text{Hom}_{CH}(CH, U) \).

In other words, the map \( \phi \) is a homomorphism of left \( G \)-sets. Qed.

---

**Proof of (11.719):** Let \( u \in U \). Recall that \( \Xi(a \otimes CG m) = am \) for every \( a \in CG \) and \( m \in U \). Applying this to \( a = 1 \) and \( m = u \), we obtain \( \Xi(1 \otimes CG u) = 1 u = u \). Hence, \( 1 \otimes CG u = \Xi^{-1}(u) \) (since \( \Xi \) is invertible). This proves (11.719).
The map $\Xi^{-1}$ is a homomorphism of $H$-sets from $\text{Res}_\rho U$ to $CG \otimes_{CG} U$. Therefore, $\Xi^{-1}$ is a $CH$-module homomorphism from $\text{Res}_\rho U$ to $CG \otimes_{CG} U$ (since $\Xi^{-1}$ is a $C$-linear map). Consequently, $\Xi^{-1}$ is a $CH$-module isomorphism from $\text{Res}_\rho U$ to $CG \otimes_{CG} U$ (since $\Xi^{-1}$ is a $C$-vector space isomorphism). Therefore, $CG \otimes_{CG} U \cong \text{Res}_\rho U$ as $CH$-modules. Now, from (11.694), we obtain

$$\text{Hom}_{CG}(U, \text{Hom}_{CH}(CG, V)) \cong \text{Hom}_{CH}\left(\begin{array}{c}
CG \otimes_{CG} U \\
\cong \text{Res}_\rho U \text{ as } CH\text{-modules}
\end{array}, V\right)
\cong \text{Hom}_{CH}(\text{Res}_\rho U, V).$$

Thus,

$$\text{Hom}_{CH}(\text{Res}_\rho U, V) \cong \text{Hom}_{CG}\left(U, \text{Hom}_{CH}(CG, V) \cong \text{Hom}_{CG}(U, \text{Ind}_\rho V) \right).$$

This solves Exercise 4.14(i).

(k) *Alternative proof of (4.3)*: Let $G$ be a finite group, and let $H$ be a subgroup of $G$. Let $U$ be a finite-dimensional $CH$-module. We need to prove the identity (4.3).

Let $\rho$ denote the inclusion map $H \to G$. Clearly, $\rho$ is a group homomorphism. But Exercise 4.14(d) yields $\text{Ind}_\rho U = \text{Ind}^G_H U$. Hence, $\chi_{\text{Ind}_\rho U} = \chi_{\text{Ind}^G_H U}$, so that $\chi_{\text{Ind}_\rho U} = \chi_{\text{Ind}_\rho U} = \chi_{\text{Ind}^G_H U} = \chi_{\text{Ind}^G_H U}$ (by Exercise 4.14(h)). But Exercise 4.14(c) (applied to $f = \chi_U$) yields $\text{Ind}_\rho \chi_U = \text{Ind}^G_H \chi_U$. Thus, $\chi_{\text{Ind}_\rho U} = \chi_{\text{Ind}^G_H U}$. Thus, every $g \in G$ satisfies

$$\chi_{\text{Ind}_\rho U}(g) = \left(\text{Ind}^G_H \chi_U\right)(g) = \frac{1}{|H|} \sum_{k g k^{-1} \in H} \chi_U(k g k^{-1}) \quad \text{(by the definition of } \text{Ind}^G_H \chi_U).$$

This proves (4.3). Thus, Exercise 4.14(k) is solved.

(l) *Alternative proof of (4.12)*: Let $G$ be a finite group, and let $K$ be a normal subgroup of $G$. Let $V$ be a finite-dimensional $CG$-module. We need to prove the identity (4.12).

Let $\rho$ denote the projection map $G \to G/K$. Exercise 4.14(f) (applied to $G/K$, $G$ and $V$ instead of $G$, $H$ and $U$) yields $\text{Ind}_\rho V \cong V^K$ as $CGK/K$-modules. Hence, $\chi_{\text{Ind}_\rho V} = \chi_{V^K}$ (since isomorphic $CGK/K$-modules

---

581 Proof. The map $\Xi^{-1}$ is a map from $U$ to $CG \otimes_{CG} U$, therefore a map from $\text{Res}_\rho U$ to $CG \otimes_{CG} U$ (since $\text{Res}_\rho U = U$ as sets).

Recall that the $CH$-module structure on $\text{Res}_\rho U$ is given by

$$h \cdot v = \rho(h) \cdot v \quad \text{for every } h \in H \text{ and } v \in U.$$

Now, let $\eta \in \text{Res}_\rho U$ and $h \in H$. We are going to prove that $\Xi^{-1}(h \cdot v) = h \cdot \Xi^{-1}(v)$, where $h \cdot \Xi^{-1}(v)$ is computed in the $CH$-module $CG \otimes_{CG} U$.

Indeed, recall that the left $CH$-module structure is induced by the $CG$-algebra homomorphism $[\rho] : CH \to CG$. Thus, it is explicitly given by

$$\eta \cdot (\overline{\rho}) = \left(\overline{\rho}\right) \eta \cdot \gamma \quad \text{for all } \gamma \in CG \text{ and } \eta \in CH.$$

Applying this to $\eta = h \cdot \gamma = 1_{CG},$ we obtain $h \cdot \gamma = \frac{\overline{\rho} \gamma}{\gamma \rho(h)}$ (since $h \in H \subset CH$). But $h \in H$ and thus $\left(\overline{\rho}\right)(h) = \rho(h).$ Hence, $h \cdot \gamma = \frac{\overline{\rho} \gamma}{\gamma \rho(h)} = \rho(h) \cdot 1_{CG} = \rho(h) = 1_{CG} \cdot \rho(h).$ Now, $v \in \text{Res}_\rho U = U$. Hence, (11.719) applied to $u = v$ yields $\Xi^{-1}(v) = 1 \otimes_{CG} v$. Thus,

$$h \cdot \Xi^{-1}(v) = h \cdot (1 \otimes_{CG} v) = h \cdot (1 \otimes_{CG} v) = \frac{1}{1_{CG} \rho(h)} \otimes_{CG} v = \frac{1}{1_{CG} \rho(h)} \otimes_{CG} v = \rho(h) \cdot v \quad \text{(here, we have moved } \rho(h) \text{ past the tensor sign)}$$

$$\text{(this is allowed since } \rho(h) \in G \subset CG) \right).$$

Compared with $\Xi^{-1}(h \cdot v) = 1 \otimes_{CG} h \cdot v$ (by (11.719), applied to $u = h \cdot v$), this yields $\Xi^{-1}(h \cdot v) = h \cdot \Xi^{-1}(v)$.

Now, let us forget that we fixed $v$ and $h$. We thus have shown that $\Xi^{-1}(h \cdot v) = h \cdot \Xi^{-1}(v)$ for every $v \in \text{Res}_\rho U$ and $h \in H$. In other words, $\Xi^{-1}$ is a homomorphism of $H$-sets from $\text{Res}_\rho U$ to $CG \otimes_{CG} U$, qed.
have equal characters). Thus, \( \chi_{V^K} = \chi_{\text{Ind}_G V} = \text{Ind}_G \chi_V \) (by Exercise 4.14(b), applied to \( G/K, G \) and \( V \) instead of \( G, H \) and \( U \)). But Exercise 4.14(e) (applied to \( G/K, G, V \) and \( \chi_V \) instead of \( G, H, U \) and \( f \)) yields \( \text{Ind}_G \chi_V = (\chi_V)^{K} \). Hence, \( \chi_{V^K} = \text{Ind}_G \chi_V = (\chi_V)^{K} \). Thus, every \( g \in G \) satisfies

\[
\chi_{V^K}(gK) = (\chi_V)^{K}(gK) = \frac{1}{|K|} \sum_{k \in K} \chi_V(gk)
\]

(by the definition of \( (\chi_V)^{K} \)). This proves (4.12). Thus, Exercise 4.14(l) is solved.

[Remark: Most parts of Exercise 4.14 work in a far greater generality than they are stated in. Let us briefly survey the straightforward generalizations. We begin with the generalizations of the notions of induction, restriction, inflation and fixed point constructions defined in Chapter 4.1:

- Recall that if \( G \) is a finite group and \( H \) is a subgroup of \( G \), then we have defined a \( CG \)-module Ind\(^G\)\(_H\) \( U \) for every \( CH \)-module \( U \), and we have also defined a \( CH \)-module Res\(^G\)\(_H\) \( V \) for every \( CG \)-module \( V \). Both of these definitions are also valid when \( C \) is replaced by any commutative ring. Basic properties of induction and restriction (such as the equality (4.7), the statement of Exercise 4.2 and the statement of Exercise 4.3) are still true in this generality (and the proofs that we gave can be transferred to this generality with almost no changes). However, more advanced properties might fail in this generality, and some can not even be stated over a general commutative ring \( C \).]

- Recall that if \( G \) is a finite group and \( K \) is a normal subgroup of \( G \), then we have defined a \( CG \)-module Ind\(^G\)\(_K\) \( U \) for every \( [G/K] \)-module \( U \), and we have also defined a \( [G/K] \)-module V\(^K\) for every \( CG \)-module \( V \). Both of these definitions are also valid when \( C \) is replaced by any commutative ring. Some basic properties of inflation and fixed points (such as the isomorphisms (4.10) and (4.11)) are still correct in this generality (again, they can be proven in the same way as above), but some others are not (e.g., Exercise 4.12(b) is generally false\(^{583} \)).

- If \( G \) is a finite group, then we defined \( RC(G) \) as the \( C \)-vector space of all class functions \( G \to C \). This definition still applies when \( C \) is replaced by any commutative ring. (Of course, in this case, “\( C \)-vector space” will have to be replaced by “\( C \)-module”.)

- Recall that if \( G \) is a finite group and \( H \) is a subgroup of \( G \), then we have defined a class function Ind\(^G\)\(_H\) \( f \in RC(G) \) for every \( f \in RC(H) \), and we have also defined a class function Res\(^G\)\(_H\) \( f \in RC(H) \) for every \( f \in RC(G) \). The definition of Res\(^G\)\(_H\) \( f \) still works when \( C \) is replaced by any commutative ring. Our definition of Ind\(^G\)\(_H\) \( f \) cannot be reasonably interpreted in this generality (due to the denominator \( |H| \) in (4.4)); however, Exercise 4.1(b) can be used as an alternative definition of Ind\(^G\)\(_H\) \( f \) when \( C \) is replaced by any commutative ring. (Of course, the necessity of using Exercise 4.1(b) as a definition of Ind\(^G\)\(_H\) \( f \) entails that most of our proofs concerning Ind\(^G\)\(_H\) \( f \) no longer are valid in this generality, because they use the equality (4.4) as the definition of Ind\(^G\)\(_H\) \( f \). Even the fact that Ind\(^G\)\(_H\) \( f \) is well-defined needs to be proven anew, since it is not immediately obvious that the sum \( \sum_{j \in J} f(jgj^{-1}) \) in Exercise 4.1(b) is independent on \( J \).

Now, let us move on to generalizing Remark 4.13 and Exercise 4.14. Let \( G \) and \( H \) be two finite groups, and let \( \rho : H \to G \) be a group homomorphism.

- Remark 4.13 still holds if \( C \) is replaced by any commutative ring.

- In Exercise 4.14, we defined a map Ind\(_\rho\) \( f : G \to C \) for every \( f \in RC(H) \). This definition still works when \( C \) is replaced by any commutative ring in which \( |H| \) is invertible. There is an alternative definition which works in even greater generality: Namely, as long as \( |\ker \rho| \) is invertible\(^{584} \), we can

\(^{582}\) For instance, any properties of characters of modules can only be stated when these characters are well-defined. Since characters are defined as traces of certain endomorphisms of \( C \)-vector spaces, they are not automatically well-defined when \( C \) is replaced by a commutative ring. They are, however, well-defined if our modules are projective modules over this base ring; see the Remark after the proof of Proposition 11.16 for details.

\(^{583}\) However, Exercise 4.12(b) holds if we additionally assume that \( |K| \) is invertible in the ring that replaces \( C \). Again, this can be proven by repeating our solution of Exercise 4.12(b).

\(^{584}\) This is a weaker condition than \( |H| \) being invertible.
define \(\text{Ind}_f : G \to \mathbb{C}\) by the equality

\[
(\text{Ind}_f)(g) = \frac{1}{|\ker \rho|} \sum_{(h,k) \in H \times J; k \rho(h) k^{-1} = g} f(h),
\]

where \(J\) is a system of left coset representatives for \(G/\rho(H)\) (so that \(G = \bigcup_{j \in J} j \rho(H)\)). We leave it to the reader to show that this definition is still well-defined (i.e., that \(\sum_{(h,k) \in H \times J; k \rho(h) k^{-1} = g} f(h)\) is independent of the choice of \(J\)), and that it defines the same function \(\text{Ind}_f\) as the definition made in Exercise 4.14 when \(|K|\) is invertible in the base ring. To my knowledge, there exists no reasonable definition of \(\text{Ind}_f\) that completely avoids making any requirements on the base ring.

- In Exercise 4.14, we defined a \(CG\)-module \(\text{Ind}_f U\) for every \(CH\)-module \(U\). This definition is still valid when \(C\) is replaced by a commutative ring. (Notice the slightly surprising fact that if \(C\) is replaced by a commutative ring, then the \(\rho\)-induction of a \(CH\)-module is always well-defined, whereas the \(\rho\)-induction of a class function might not be.)

- Exercise 4.14(a) is still valid when \(C\) is replaced by any commutative ring in which \(|H|\) is invertible. And the solution that we gave still applies in this generality. More generally, if we define \(\text{Ind}_f\) by (11.721), then Exercise 4.14(a) is still valid when \(C\) is replaced by any commutative ring in which \(|\ker \rho|\) is invertible. However, this can no longer be proven by blindly copying our solution.

- Exercise 4.14(b) is still valid when \(C\) is replaced by any field in which \(|H|\) is invertible. Again, our solution is still valid in this generality. More generally, Exercise 4.14(b) is still valid when \(C\) is replaced by any field in which \(|\ker \rho|\) is invertible (as long as \(\text{Ind}_f\) is defined through (11.721)); but our solution is not sufficient to prove this. Even more generally, Exercise 4.14(b) holds whenever \(C\) is replaced by any commutative ring \(A\) in which \(|\ker \rho|\) is invertible, as long as \(U\) is assumed to be a finitely generated projective \(A\)-module. However, in order to make sense of this statement, one needs to know that \(\text{Ind}_f U\) is a finitely generated projective \(A\)-module as well, and one needs to know how the trace of an endomorphism of a finitely generated projective \(A\)-module is defined. We essentially did most of this in our solution above.

- Exercise 4.14(c) is still valid when \(C\) is replaced by any commutative ring in which \(|H|\) is invertible. Again, our solution is still valid in this generality. Again, if we define \(\text{Ind}_f\) by (11.721) and define \(\text{Ind}_f^H\) by Exercise 4.1(b), then Exercise 4.14(c) even holds when \(C\) is replaced by any arbitrary commutative ring. (The definition of \(\text{Ind}_f\) requires \(|\ker \rho|\) to be invertible, but this is automatically satisfied since \(\rho\) is injective.)

- Exercise 4.14(d) is still valid when \(C\) is replaced by any commutative ring whatsoever. Again, our solution is still valid in this generality.

---

585 We can also replace \(C\) by a commutative ring \(A\) rather than by a field; but then we need to replace “finite-dimensional \(CH\)-module \(U\)” by “\(AH\)-module \(U\) which is a finitely generated projective \(A\)-module” in the statement of the exercise. The “finitely generated projective \(A\)-module” condition is needed in order to make sure that \(\chi_U\) and \(\chi_{\text{Ind}_f U}\) are well-defined. (See the Remark after the proof of Proposition 11.16 for details.)

586 However, a certain modification of this solution does the trick. Namely, we replace the definition of \(\tilde{F}_g\) by

\[
\tilde{F}_g(\gamma, u) = \frac{1}{|\ker \rho|} \sum_{h \in H} (g \rho(h))^* (\gamma) h u \quad \text{for all } (\gamma, u) \in CG \times U.
\]

Then, we should pick a system \(J'\) of left coset representatives for \(G/\rho(H)\) (so that \(G = \bigcup_{j \in J'} j \rho(H)\)). (Notice that we do not call it \(J\) because the letter \(J\) already has a different meaning in this solution.) Then, we define a family \(\{a_{i}\}_{i \in J' \times J}\) of elements of \(CG \otimes_{CH} U\) by

\[
a_{(k,j)} = k \otimes_{CH} b_j \quad \text{for all } (k,j) \in J' \times J,
\]

and a family \(\{f_{i}\}_{i \in J' \times J}\) of elements of \((CG \otimes_{CH} U)^*\) by

\[
f_{(k,j)} = g_j \circ F_k \quad \text{for all } (k,j) \in J' \times J.
\]

(Notice that these families \(\{a_{i}\}_{i \in J' \times J}\) and \(\{f_{i}\}_{i \in J' \times J}\) are subfamilies of the families \(\{a_{i}\}_{i \in G \times J}\) and \(\{f_{i}\}_{i \in G \times J}\) from the original solution to Exercise 4.14(b).) It can then be shown that \((J' \times J, \{a_{i}\}_{i \in J' \times J}, \{f_{i}\}_{i \in J' \times J})\) is a finite dual generating system for \(\text{Ind}_f U\) (though the proof is somewhat different from the way this was shown in the solution to Exercise 4.14(b)). This allows us to use Proposition 11.16 to derive \(\chi_{\text{Ind}_f U} = \text{Ind}_f \chi_U\). All details are left to the reader.

587 The only missing link is the fact that an \(A\)-module is finitely generated and projective if and only if it has a finite dual generating system; this fact is easy to prove.
Exercise 4.14(e) is still valid when $C$ is replaced by any commutative ring in which $|H|$ is invertible. Again, our solution is still valid in this generality. Again, it is possible to prove Exercise 4.14(e) also when $C$ is replaced by any commutative ring in which $|\ker \rho|$ is invertible (as long as $\ind_{\rho} f$ is defined using (11.721)). (This is actually very easy to check – arguably even easier than our above solution of Exercise 4.14(e).)

Exercise 4.14(f) is still valid when $C$ is replaced by any commutative ring in which $|K|$ is invertible. Both solutions of Exercise 4.14(f) given above still remain valid in this generality.

The second claim of Exercise 4.14(g) (that is, the equality (4.17)) is still valid when $C$ is replaced by any commutative ring in which $|G|$ and $|H|$ is invertible. The first claim, a priori, makes no sense when $C$ is replaced by an arbitrary commutative ring, because the definition of the Hermitian forms $(\cdot,\cdot)_G$ and $(\cdot,\cdot)_H$ involves complex conjugation (which is only defined on $C$). However, it turns out that this complex conjugation can be replaced by any map from the base ring to itself and Exercise 4.14(g) remains valid.

Exercise 4.14(h) is still valid when $C$ is replaced by any commutative ring whatsoever. Again, our solution is still valid in this generality.

Exercise 4.14(i) is still valid when $C$ is replaced by any commutative ring in which $|\ker \rho|$ is invertible. Again, our solution is still valid in this generality (because if we set $K = \ker \rho$ as we did in our solution, then $K = |\ker \rho|$ is invertible).

Exercise 4.14(j) is still valid when $C$ is replaced by any commutative ring in which $|\ker \rho|$ is invertible.

Exercise 4.14(k) still gives a proof of the formula (4.3) when $C$ is replaced by any field in which $|H|$ is invertible.

Exercise 4.14(l) still gives a proof of the formula (4.12) when $C$ is replaced by any field in which $|G|$ and $|K|$ are invertible.

We actually tailored our above solutions to apply in a reasonably high generality (although we did not state Exercise 4.14 in this generality). Had we contented ourselves with solving no more and no less than what Exercise 4.14 demanded of us, we could have obtained much shorter solutions of some parts of it:

Exercise 4.14(b) can be obtained by combining parts (g) and (h) of Exercise 4.14 with the following two observations:

- Every $\mathbb{C}G$-module $V$ satisfies $\chi_{\res_{\rho} V} = \res_{\rho} V$. (This is obvious.)
- If $\alpha$ and $\beta$ are two elements of $\mathbb{R}_G (G)$ such that every finite-dimensional $\mathbb{C}G$-module $V$ satisfies $(\alpha, \chi_V)_G = (\beta, \chi_V)_G$, then $\alpha = \beta$. (This follows from the fact that $\mathbb{C}$-vector space $\mathbb{R}_G (G)$ and from the fact that the Hermitian form $(\cdot,\cdot)_G$ is nondegenerate.)

However, this quick argument only works when the base ring is $\mathbb{C}$. It cannot be easily generalized to the case when $C$ is replaced by any field in which $|H|$ is invertible.

Part (e) (at least in the case when $U$ is finite-dimensional) and part (f) of Exercise 4.14 can be derived from each other using the correspondence between irreducible representations and irreducible characters. Again, this argument is quick but hard to generalize.

11.82. **Solution to Exercise 4.15.** Solution to Exercise 4.15. We shall use the following fact:

---

588More precisely: Let $A$ be any commutative ring in which $|G|$ and $|H|$ are invertible. Fix an arbitrary map $\conj : A \to A$ (not necessarily linear). For every $a \in A$, let $\pi$ denote the image $\conj (a)$ of $a$ under this map. For any $f_1 \in R_A (G)$ and $f_2 \in R_A (G)$, define $(f_1, f_2)_G \in A$ by $(f_1, f_2)_G = \frac{1}{|G|} \sum_{g \in G} f_1 (g) \overline{f_2 (g)}$ (using the notation $\pi$ that we just introduced).

Similarly, define $(f_1, f_2)_H \in A$ for any $f_1 \in R_A (H)$ and $f_2 \in R_A (H)$. Then, Exercise 4.14(g) is still valid when $C$ is replaced by $A$. (And our solution to this exercise still applies.)

589We only sketch this argument; the details are left to the reader.
Proposition 11.18. Let $k$ be a commutative ring. Let $A$, $B$, $C$, $A'$, $B'$ and $C'$ be six $k$-algebras. Let $P$ be an $(A,B)$-bimodule$^{590}$. Let $Q$ be a $(B,C)$-bimodule. Let $P'$ be an $(A',B')$-bimodule. Let $Q'$ be a $(B',C')$-bimodule. Then,

$$(P \otimes P') \otimes_{B \otimes B'} (Q \otimes Q') \cong (P \otimes_B Q) \otimes (P' \otimes_{B'} Q')$$

as $(A \otimes A', C \otimes C')$-bimodules. Here, all $\otimes$ signs without subscript stand for $\otimes_k$.

Proposition 11.18 is identical with the Proposition 11.13 that appeared in our solution to Exercise 4.3. It is proven by straightforward (repeated) use of the universal property of the tensor product.

We have a canonical $C$-algebra isomorphism

$$\mathcal{f} : C [H_1 \times H_2] \rightarrow CH_1 \otimes CH_2$$

which satisfies

$$\mathcal{f} (t_{(h_1, h_2)}) = t_{h_1} \otimes t_{h_2} \quad \text{for every } (h_1, h_2) \in H_1 \times H_2.$$

We also have a canonical $C$-algebra isomorphism

$$\mathcal{g} : C [G_1 \times G_2] \rightarrow CG_1 \otimes CG_2$$

which satisfies

$$\mathcal{g} (t_{(g_1, g_2)}) = t_{g_1} \otimes t_{g_2} \quad \text{for every } (g_1, g_2) \in G_1 \times G_2.$$

Recall how the left $C [H_1 \times H_2]$-module $U_1 \otimes U_2$ is defined: Its left $C [H_1 \times H_2]$-module structure is given by

$$(t_{(h_1, h_2)}) (u_1 \otimes u_2) = t_{h_1} u_1 \otimes t_{h_2} u_2 \quad \text{for all } (h_1, h_2) \in H_1 \times H_2 \text{ and } (u_1, u_2) \in U_1 \times U_2.$$

We shall now recall the definitions of $\text{Ind}_{p_1} U_1$, $\text{Ind}_{p_2} U_2$ and $\text{Ind}_{p_1 \times p_2} (U_1 \otimes U_2)$:

- We defined $\text{Ind}_{p_1} U_1$ as the $CG_1$-module $CG_1 \otimes_{CH_1} U_1$, where $CG_1$ is regarded as a $(CG_1, CH_1)$-bimodule according to the following rule: The left $CG_1$-module structure on $CG_1$ is plain multiplication inside $CG_1$: the right $CH_1$-module structure on $CG_1$ is induced by the $C$-algebra homomorphism $C [p_1] : CH_1 \rightarrow CG_1$ (thus, it is explicitly given by $\gamma \eta = \gamma \cdot (C [p_1]) \eta$ for all $\gamma \in CG_1$ and $\eta \in CH_1$).
- We defined $\text{Ind}_{p_2} U_2$ as the $CG_2$-module $CG_2 \otimes_{CH_2} U_2$, where $CG_2$ is regarded as a $(CG_2, CH_2)$-bimodule in a similar fashion.
- We defined $\text{Ind}_{p_1 \times p_2} (U_1 \otimes U_2)$ as the $C [G_1 \times G_2]$-module $C [G_1 \times G_2] \otimes_{C[H_1 \times H_2]} (U_1 \otimes U_2)$, where $C [G_1 \times G_2]$ is regarded as a $(C [G_1 \times G_2], C [H_1 \times H_2])$-bimodule in a similar fashion.

Proposition 11.13 (applied to $k = C$, $A = CG_1$, $B = CH_1$, $C = C$, $A' = CG_2$, $B' = CH_2$, $C' = C$, $P = CG_1$, $Q = U_1$, $P' = CG_2$ and $Q' = U_2$) yields

$$(CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2) \cong (CG_1 \otimes_{CH_1} U_1) \otimes (CG_2 \otimes_{CH_2} U_2)$$

as $(CG_1 \otimes CG_2, C \otimes C)$-bimodules, hence also as left $CG_1 \otimes CG_2$-modules. Thus,

$$(CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2) \cong (CG_1 \otimes_{CH_1} U_1) \otimes (CG_2 \otimes_{CH_2} U_2)$$

equals (11.722)

$$(\text{Ind}_{p_1} U_1) \otimes (\text{Ind}_{p_2} U_2)$$

as left $CG_1 \otimes CG_2$-modules.

At this point, there is a quick way to finish the solution using handwaving: We use the $C$-algebra isomorphism $\mathcal{f} : C [H_1 \times H_2] \rightarrow CH_1 \otimes CH_2$ to identify the $C$-algebra $C [H_1 \times H_2]$ with the $C$-algebra $CH_1 \otimes CH_2$, and we use the $C$-algebra isomorphism $\mathcal{g} : C [G_1 \times G_2] \rightarrow CG_1 \otimes CG_2$ to identify the $C$-algebra $C [G_1 \times G_2]$ with the $C$-algebra $CG_1 \otimes CG_2$. It is “easy to see” that these two identifications “play nicely with each other and with the module structures on $U_1 \otimes U_2$” (this is the part where we wave our hands). Now, (11.722)
yields that \((CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2) \cong (\text{Ind}_{\rho_1} U_1) \otimes (\text{Ind}_{\rho_2} U_2)\) as left \(CG_1 \otimes CG_2\)-modules, and therefore also as left \(\mathbb{C}[G_1 \times G_2]\)-modules (since \(\mathbb{C}[G_1 \times G_2] = CG_1 \otimes CG_2\)). Now,

\[
\text{Ind}_{\rho_1 \times \rho_2} (U_1 \otimes U_2) = \mathbb{C}[G_1 \times G_2] \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2)
\]

\[(\text{since } \mathbb{C}[G_1 \times G_2] = CG_1 \otimes CG_2 \text{ and } \mathbb{C}[H_1 \times H_2] = CH_1 \otimes CH_2) \cong (\text{Ind}_{\rho_1} U_1) \otimes (\text{Ind}_{\rho_2} U_2)\]

as left \(\mathbb{C}[G_1 \times G_2]\)-modules. This solves Exercise 4.15 if you believe the handwaving I have done above.

The handwaving we have done is slightly questionable, since we have made two identifications which turned two isomorphisms into identities. In reality, they are merely isomorphisms, not identities, and it is not immediately clear that regarding them as identities will not lead to contradictions. (Indeed, this is the meaning of our vague claim that the two identifications “play nicely with each other”.)

Let us now show a way to formalize the above questionable argument. We shall focus on explaining how to do this proof in a clean fashion (in particular, we shall avoid identifying any things that are not already identical; instead, we will work with the isomorphisms \(\mathcal{G}\) and \(\mathcal{H}\) explicitly); we will leave straightforward computations and arguments to the reader.

We first state a general fact:

**Proposition 11.19.** Let \(k\) be a commutative ring. In the following, all \(\otimes\) signs without subscript stand for \(\otimes_k\).

Let \(A, B, C, A', B'\) and \(C'\) be six \(k\)-algebras. Let \(M\) be an \((A, B)\)-bimodule. Let \(N\) be a \((B, C)\)-bimodule. Let \(M'\) be an \((A', B')\)-bimodule. Let \(N'\) be a \((B', C')\)-bimodule. Let \(\beta : B \to B'\), \(\mu : M \to M'\) and \(\nu : N \to N'\) be three \(k\)-module homomorphisms. Assume that we have

\[
(\mu (mb)) = (\mu (m)) \beta (b) \quad \text{for all } b \in B \text{ and } m \in M
\]

and

\[
(\nu (bn)) = \beta (b) (\nu (n)) \quad \text{for all } b \in B \text{ and } n \in N.
\]

Then:

(a) There exists a unique \(k\)-module homomorphism

\[
\Omega : M \otimes_B N \to M' \otimes_{B'} N'
\]

which satisfies

\[
(\Omega (m \otimes_B n)) = (\mu (m)) \otimes_{B'} (\nu (n)) \quad \text{for all } (m, n) \in M \times N.
\]

We shall denote this homomorphism \(\Omega\) by \(\Omega_{\beta, \mu, \nu}\).

(b) If \(\beta, \mu\) and \(\nu\) are bijections, then \(\Omega_{\beta, \mu, \nu}\) is a \(k\)-module isomorphism.

(c) Let \(\alpha : A \to A'\) be a \(k\)-module homomorphism. Assume that

\[
(\mu (am)) = \alpha (a) (\mu (m)) \quad \text{for all } a \in A \text{ and } m \in M.
\]

Assume also that the \(k\)-module \(M'\) is endowed with a left \(A\)-module structure. Assume that this left \(A\)-module structure on \(M'\) and the right \(B'\)-module structure on \(M'\) together form an \((A, B')\)-bimodule structure on \(M'\). Thus, \(M' \otimes_{B'} N'\) becomes a left \(A\)-module. Assume furthermore that

\[
am = \alpha (a) m \quad \text{for every } a \in A \text{ and } m \in M'.
\]

592 Then, \(\Omega_{\beta, \mu, \nu}\) is a left \(A\)-module homomorphism.

The proof of Proposition 11.19 is straightforward593 and is left to the reader. We could also add a part (d) to Proposition 11.19, which would give a criterion for \(\Omega_{\beta, \mu, \nu}\) to be a right \(C\)-module homomorphism given an appropriate right \(C\)-module structure on \(N'\).

Let us now return to solving Exercise 4.15. We have four bimodules:

591 As usual, we understand the notion of a bimodule to be defined over \(k\); that is, the left \(A\)-module structure and the right \(B\)-module structure of an \((A, B)\)-bimodule must restrict to one and the same \(k\)-module structure.

592 Here the \(a (a) m\) on the right hand side is defined using the action of \(A'\) on the left \(A'\)-module \(M'\), whereas the \(am\) on the left hand side is defined using the action of \(A\) on the left \(A\)-module \(M'\).

593 For part (b), the inverse of \(\Omega_{\beta, \mu, \nu}\) is \(\Omega_{\beta^{-1}, \mu^{-1}, \nu^{-1}}\), of course.
• The \( \mathbb{C} \)-vector space \( CG_1 \otimes CG_2 \) is a \((CG_1 \otimes CG_2, CH_1 \otimes CH_2)\)-bimodule (because it is the tensor product of the \((CG_1, CH_1)\)-bimodule \( CG_1 \) with the \((CG_2, CH_2)\)-bimodule \( CG_2 \)). As a left \( CG_1 \otimes CG_2 \)-module, it is thus the tensor product of the left \( CG_1 \)-module \( CG_1 \) with the left \( CG_2 \)-module \( CG_2 \). Hence, its left \( CG_1 \otimes CG_2 \)-module structure is given by plain multiplication inside \( CG_1 \otimes CG_2 \). This is straightforward to check.

• The \( \mathbb{C} \)-vector space \( U_1 \otimes U_2 \) is a left \( CH_1 \otimes CH_2 \)-module (because it is the tensor product of the left \( CH_1 \)-module \( U_1 \) with the left \( CH_2 \)-module \( U_2 \)), and thus a \((CH_1 \otimes CH_2, \mathbb{C})\)-bimodule.

• The \( \mathbb{C} \)-vector space \( C\{G_1 \times G_2\} \) is a \((C\{G_1 \times G_2\}, C\{H_1 \times H_2\})\)-bimodule (as we already know).

• The \( \mathbb{C} \)-vector space \( U_1 \otimes U_2 \) is a left \( CH_1 \otimes CH_2 \)-module (since it is the tensor product of the left \( CH_1 \)-module \( U_1 \) with the left \( CH_2 \)-module \( U_2 \)), and thus a \((CH_1 \otimes CH_2, \mathbb{C})\)-bimodule.

We have

\[
(11.727) \quad (\mathfrak{S} (mb) = \mathfrak{S} (m) \mathfrak{S} (b) \quad \text{for all } b \in C \{H_1 \times H_2\} \text{ and } m \in C \{G_1 \times G_2\})
\]

and

\[
(11.728) \quad (id (bn) = \mathfrak{S} (b) id (n) \quad \text{for all } b \in C \{H_1 \times H_2\} \text{ and } n \in U_1 \otimes U_2).
\]

(In fact, both of these equalities are easily checked on basis elements and pure tensors.) Hence, we can apply Proposition 11.19(a) to \( k = \mathbb{C} \), \( A = C \{G_1 \times G_2\} \), \( B = C \{H_1 \times H_2\} \), \( C = \mathbb{C} \), \( A' = CG_1 \otimes CG_2 \), \( B' = CH_1 \otimes CH_2 \), \( C' = \mathbb{C} \), \( M = C \{G_1 \times G_2\} \), \( N = U_1 \otimes U_2 \), \( M' = CG_1 \otimes CG_2 \), \( N' = U_1 \otimes U_2 \), \( \beta = \mathfrak{S} \), \( \mu = \mathfrak{S} \) and \( \nu = id \). As a consequence, we conclude that there exists a unique \( \mathbb{C} \)-module homomorphism

\[
(11.729) \quad \Omega : C \{G_1 \times G_2\} \otimes_{C \{H_1 \times H_2\}} (U_1 \otimes U_2) \to (CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2)
\]

which satisfies

\[
(11.730) \quad \Omega (m \otimes_{C \{H_1 \times H_2\}} n) = \mathfrak{S} (m) \otimes_{CH_1 \otimes CH_2} id (n) \quad \text{for all } (m, n) \in C \{G_1 \times G_2\} \times (U_1 \otimes U_2).
\]

According to Proposition 11.19(a), this homomorphism \( \Omega \) is denoted by \( \Omega_{\mathfrak{S}, id} \).

The maps \( \mathfrak{S} \), \( \mathfrak{S} \) and \( id \) are \( \mathbb{C} \)-algebra isomorphisms, and therefore invertible. Hence, Proposition 11.19(b) (applied to \( k = \mathbb{C} \), \( A = C \{G_1 \times G_2\} \), \( B = C \{H_1 \times H_2\} \), \( C = \mathbb{C} \), \( A' = CG_1 \otimes CG_2 \), \( B' = CH_1 \otimes CH_2 \), \( C' = \mathbb{C} \), \( M = C \{G_1 \times G_2\} \), \( N = U_1 \otimes U_2 \), \( M' = CG_1 \otimes CG_2 \), \( N' = U_1 \otimes U_2 \), \( \beta = \mathfrak{S} \), \( \mu = \mathfrak{S} \) and \( \nu = id \) yields that \( \Omega_{\mathfrak{S}, id} \) is a \( \mathbb{C} \)-module isomorphism.

Next, we notice that the \( \mathbb{C} \)-vector space \( CG_1 \otimes CG_2 \) is a left \( C \{G_1 \times G_2\} \)-module (since it is the tensor product of the left \( CH_1 \)-module \( CG_1 \) with the left \( CH_2 \)-module \( CG_2 \)). This left \( C \{G_1 \times G_2\} \)-module structure is defined by the rule

\[
(11.731) \quad t_{(g_1, g_2)} (u_1 \otimes u_2) = t_{g_1} u_1 \otimes t_{g_2} u_2 \quad \text{for all } (g_1, g_2) \in G_1 \times G_2 \text{ and } (u_1, u_2) \in CG_1 \times CG_2.
\]

This left \( C \{G_1 \times G_2\} \)-module structure on \( CG_1 \otimes CG_2 \) and the right \( CH_1 \otimes CH_2 \)-module structure on \( CG_1 \otimes CG_2 \) are connected by the equality

\[
(11.732) \quad ((gm) h = g (mh) \quad \text{for all } g \in C \{G_1 \times G_2\}, m \in CG_1 \otimes CG_2 \text{ and } h \in CH_1 \otimes CH_2)
\]

(which is easily checked). Hence, these two structures together form an \((C \{G_1 \times G_2\}, CH_1 \otimes CH_2)\)-bimodule structure on \( CG_1 \otimes CG_2 \) (because both of these structures are \( \mathbb{C} \)-bilinear). Thus, the tensor product \( CG_1 \otimes CG_2 \otimes CH_1 \otimes CH_2 \) \( (U_1 \otimes U_2) \) becomes a left \( C \{G_1 \times G_2\} \)-module.

Next, it is easy to check that

\[
(11.733) \quad \mathfrak{S} (am) = \mathfrak{S} (a) \mathfrak{S} (m) \quad \text{for all } a \in C \{G_1 \times G_2\} \text{ and } m \in C \{G_1 \times G_2\}.
\]

It is also easy to see that

\[
(11.734) \quad am = \mathfrak{S} (a) m \quad \text{for every } a \in C \{G_1 \times G_2\} \text{ and } m \in CG_1 \otimes CG_2.
\]

Thus, we can apply Proposition 11.19(c) to \( k = \mathbb{C} \), \( A = C \{G_1 \times G_2\} \), \( B = C \{H_1 \times H_2\} \), \( C = \mathbb{C} \), \( A' = CG_1 \otimes CG_2 \), \( B' = CH_1 \otimes CH_2 \), \( C' = \mathbb{C} \), \( M = C \{G_1 \times G_2\} \), \( N = U_1 \otimes U_2 \), \( M' = CG_1 \otimes CG_2 \), \( N' = U_1 \otimes U_2 \), \( \beta = \mathfrak{S} \), \( \mu = \mathfrak{S} \), \( \nu = id \) and \( \alpha = \mathfrak{S} \). As a result, we obtain that \( \Omega_{\mathfrak{S}, id} \) is a left \( CG_1 \otimes CG_2 \)-module homomorphism (since \( \Omega_{\mathfrak{S}, id} \) is invertible). Thus,

\[
(11.735) \quad C \{G_1 \times G_2\} \otimes_{C \{H_1 \times H_2\}} (U_1 \otimes U_2) \cong (CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2)
\]

594 That is, the action of any \( a \in CG_1 \otimes CG_2 \) on any element \( m \) of the left \( CG_1 \otimes CG_2 \)-module \( CG_1 \otimes CG_2 \) equals the product of \( a \) with \( m \) in the \( \mathbb{C} \)-algebra \( CG_1 \otimes CG_2 \).
as left \(\mathbb{C} [G_1 \times G_2]\)-modules.

Now, let us recall the isomorphism (11.722). It is an isomorphism of left \(CG_1 \otimes CG_2\)-modules, and thus cannot be immediately combined with (11.733). However, it is easy to see that the corresponding isomorphism of left \(\mathbb{C} [G_1 \times G_2]\)-modules holds as well: Namely, we have

\[
(11.734) \quad (CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2) \cong (\text{Ind}_{\rho_1} U_1) \otimes (\text{Ind}_{\rho_2} U_2)
\]

as left \(\mathbb{C} [G_1 \times G_2]\)-modules. Before we prove this, let us make three auxiliary observations which connect the left \(\mathbb{C} [G_1 \times G_2]\)-module structures on the modules appearing in (11.734) with the left \(CG_1 \otimes CG_2\)-module structures on the same modules:

- We have

\[
(11.735) \quad t_{(g_1, g_2)} m = (t_{g_1} \otimes t_{g_2}) m \quad \text{for every } (g_1, g_2) \in G_1 \times G_2 \text{ and } m \in CG_1 \otimes CG_2
\]

(where the expression \(t_{(g_1, g_2)} m\) on the left hand side is defined using the left \(\mathbb{C} [G_1 \times G_2]\)-module structure on \(CG_1 \otimes CG_2\), whereas the expression \((t_{g_1} \otimes t_{g_2}) m\) on the right hand side is defined using the left \(CG_1 \otimes CG_2\)-module structure on \(CG_1 \otimes CG_2\)). This is easy to prove.

- We have

\[
(11.736) \quad t_{(g_1, g_2)} n = (t_{g_1} \otimes t_{g_2}) n \quad \text{for every } (g_1, g_2) \in G_1 \times G_2 \text{ and } n \in (CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2)
\]

(where the expression \(t_{(g_1, g_2)} n\) on the left hand side is defined using the left \(\mathbb{C} [G_1 \times G_2]\)-module structure on \((CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2)\), whereas the expression \((t_{g_1} \otimes t_{g_2}) n\) on the right hand side is defined using the left \(CG_1 \otimes CG_2\)-module structure on \((CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2)\)). This is easy to prove using (11.735).

- Finally, we have

\[
(11.737) \quad t_{(g_1, g_2)} m = (t_{g_1} \otimes t_{g_2}) m \quad \text{for every } (g_1, g_2) \in G_1 \times G_2 \text{ and } m \in (\text{Ind}_{\rho_1} U_1) \otimes (\text{Ind}_{\rho_2} U_2)
\]

(where the expression \(t_{(g_1, g_2)} m\) on the left hand side is defined using the left \(\mathbb{C} [G_1 \times G_2]\)-module structure on \((\text{Ind}_{\rho_1} U_1) \otimes (\text{Ind}_{\rho_2} U_2)\), whereas the expression \((t_{g_1} \otimes t_{g_2}) m\) on the right hand side is defined using the left \(CG_1 \otimes CG_2\)-module structure on \((\text{Ind}_{\rho_1} U_1) \otimes (\text{Ind}_{\rho_2} U_2)\)). Again, this is easy to check.

Now, it is easy to see that (11.734) holds\(^{595}\).

Now, (11.733) becomes

\[
\mathbb{C} [G_1 \times G_2] \otimes \mathbb{C}[H_1 \times H_2] (U_1 \otimes U_2) \\
\cong (CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2) \cong (\text{Ind}_{\rho_1} U_1) \otimes (\text{Ind}_{\rho_2} U_2) \quad \text{(by (11.734))}
\]

\(^{595}\)Proof of (11.734): From (11.722), we conclude that there exists an isomorphism

\[
T : (CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2) \rightarrow (\text{Ind}_{\rho_1} U_1) \otimes (\text{Ind}_{\rho_2} U_2)
\]

of left \(CG_1 \otimes CG_2\)-modules. We shall now show that this map \(T\) is an isomorphism of left \(\mathbb{C} [G_1 \times G_2]\)-modules as well. In order to do so, it is sufficient to show that the map \(T\) is a homomorphism of left \(G_1 \times G_2\)-sets (because we already know that \(T\) is \(\mathbb{C}\)-linear and invertible). In other words, it is sufficient to show that

\[
T(\alpha n) = a T(\alpha) \quad \text{for every } \alpha \in \mathbb{C} [G_1 \times G_2] \text{ and } n \in (CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2).
\]

Proof of (11.738): Let \(\alpha \in \mathbb{C} [G_1 \times G_2]\) and \(n \in (CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2)\). We need to prove the equality

\[
T(\alpha n) = a T(\alpha) \quad \text{for every } \alpha \in \mathbb{C} [G_1 \times G_2] \text{ and } n \in (CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2).
\]

Consider this \((g_1, g_2)\). We have

\[
\sum_{(g_1, g_2)} \alpha n = \sum_{(g_1, g_2)} t_{(g_1, g_2)} n = \sum_{(g_1, g_2)} (t_{g_1} \otimes t_{g_2}) n \quad \text{(by (11.736))}
\]

Applying the map \(T\) to both sides of this equality, we obtain

\[
T(\alpha n) = T(\sum_{(g_1, g_2)} t_{g_1} \otimes t_{g_2}) n = \sum_{(g_1, g_2)} t_{g_1} \otimes t_{g_2} T(n) \quad \text{(since } T\text{ is a homomorphism of left } CG_1 \otimes CG_2\text{-modules)}.
\]

Compared with

\[
T(\alpha n) = \sum_{(g_1, g_2)} T(n) = \sum_{(g_1, g_2)} (t_{g_1} \otimes t_{g_2}) T(n) = \sum_{(g_1, g_2)} (t_{g_1} \otimes t_{g_2}) T(n) \quad \text{(by (11.737), applied to } m = T(n))
\]

this yields \(T(\alpha n) = a T(\alpha)\). This proves (11.738).

Now, (11.738) shows that \(T\) is a homomorphism of left \(G_1 \times G_2\)-sets (because (11.738) holds), therefore a homomorphism of left \(\mathbb{C} [G_1 \times G_2]\)-modules (since \(T\) is \(\mathbb{C}\)-linear), and therefore an isomorphism of left \(\mathbb{C} [G_1 \times G_2]\)-modules (since \(T\) is invertible). Thus, \((CG_1 \otimes CG_2) \otimes_{CH_1 \otimes CH_2} (U_1 \otimes U_2) \cong (\text{Ind}_{\rho_1} U_1) \otimes (\text{Ind}_{\rho_2} U_2)\) as left \(\mathbb{C} [G_1 \times G_2]\)-modules. This proves (11.734).
as left $\mathbb{C}[G_1 \times G_2]$-modules. Therefore,
\[
\text{Ind}_{\rho_1 \times \rho_2} (U_1 \otimes U_2) = \mathbb{C}[G_1 \times G_2] \otimes_{\mathbb{C}[H_1 \times H_2]} (U_1 \otimes U_2) \cong (\text{Ind}_{\rho_1} U_1) \otimes (\text{Ind}_{\rho_2} U_2)
\]
as left $\mathbb{C}[G_1 \times G_2]$-modules. This finishes the solution of Exercise 4.15.

11.83. Solution to Exercise 4.16. Solution to Exercise 4.16. In the following, we will use the following
convention: Whenever $K$ is a group, and $k$ is an element of $K$, we shall write $k$ for the element $t_k$ of $\mathbb{C}K$.
This is a relatively common abuse of notation, and it is harmless because the map $K \to \mathbb{C}K$, $k \mapsto t_k$ is an
injective homomorphism of multiplicative monoids (so $t_{gh} = t_g t_h$ and $t_1 = 1$, which means that we won’t
run into ambiguities denoting $t_k$ by $k$) and because every $\mathbb{C}K$-module $M$, every $m \in M$ and every $k \in K$
satisfy $km = t_k m$.

We solve the four parts of Exercise 4.16 in the following order: first, part (c); then, part (d); then, part
(a); finally, part (b).

(c) The definition of $\text{Res}_{\rho} V$ yields $\text{Res}_{\rho} V = V$ as $\mathbb{C}$-vector spaces. Similarly, $\text{Res}_{\tau} \text{Res}_{\rho} V = \text{Res}_{\rho} V$ as
$\mathbb{C}$-vector spaces, and $\text{Res}_{\rho \circ \tau} V = V$ as $\mathbb{C}$-vector spaces. Thus, $\text{Res}_{\tau} \text{Res}_{\rho} V = \text{Res}_{\rho} V = V = \text{Res}_{\rho \circ \tau} V$ as $\mathbb{C}$-vector spaces. But our goal is to show that $\text{Res}_{\tau} \text{Res}_{\rho} V = \text{Res}_{\rho \circ \tau} V$ as $\mathbb{C}I$-modules. Thus, it suffices
to prove that the left $\mathbb{C}I$-module structures on $\text{Res}_{\tau} \text{Res}_{\rho} V$ and $\text{Res}_{\rho \circ \tau} V$ are identical. In other words, it
suffices to show that every $i \in \mathbb{C}I$ and $v \in V$ satisfy the equality

\[
(\text{the action of } i \in \mathbb{C}I \text{ on the element } v \text{ of the left } \mathbb{C}I\text{-module } \text{Res}_{\tau} \text{Res}_{\rho} V) \quad = \quad (\text{the action of } i \in \mathbb{C}I \text{ on the element } v \text{ of the left } \mathbb{C}I\text{-module } \text{Res}_{\rho \circ \tau} V) \quad = \quad (11.739)
\]

(Of course, both sides of the equality (11.739) could be rewritten as $iv$, but this notation is ambiguous,
because $i$ simultaneously belongs to two $\mathbb{C}I$-modules $\text{Res}_{\tau} \text{Res}_{\rho} V$ and $\text{Res}_{\rho \circ \tau} V$ which are not yet known to be identical.)

Proof of (11.739): Let $i \in \mathbb{C}I$ and $v \in V$. We need to prove the equality (11.739). Since this equality is
$\mathbb{C}$-linear in $i$, we can WLOG assume that $i$ belongs to the basis $I$ of the $\mathbb{C}$-vector space $\mathbb{C}I$. Assume this.
Now, the definition of the left $\mathbb{C}I$-module $\text{Res}_{\tau} \text{Res}_{\rho} V$ yields

\[
(\text{the action of } i \in \mathbb{C}I \text{ on the element } v \text{ of the left } \mathbb{C}I\text{-module } \text{Res}_{\tau} \text{Res}_{\rho} V) = (\text{the action of } \tau (i) \in \mathbb{C}H \text{ on the element } v \text{ of the left } \mathbb{C}H\text{-module } \text{Res}_{\rho} V)
\]
\[
= \left( \text{the action of } \rho (\tau (i)) \in \mathbb{C}G \text{ on the element } v \text{ of the left } \mathbb{C}G\text{-module } V \right)
\]
\[
(\text{by the definition of the left } \mathbb{C}H\text{-module } \text{Res}_{\rho} V)
\]
\[
= (\text{the action of } (\rho \circ \tau) (i) \in \mathbb{C}G \text{ on the element } v \text{ of the left } \mathbb{C}G\text{-module } V).\n\]

Compared with

\[
(\text{the action of } i \in \mathbb{C}I \text{ on the element } v \text{ of the left } \mathbb{C}I\text{-module } \text{Res}_{\rho \circ \tau} V)
\]
\[
= (\text{the action of } (\rho \circ \tau) (i) \in \mathbb{C}G \text{ on the element } v \text{ of the left } \mathbb{C}G\text{-module } V),\n\]
\[
(\text{by the definition of the left } \mathbb{C}I\text{-module } \text{Res}_{\rho \circ \tau} V),\n\]
this yields

\[
(\text{the action of } i \in \mathbb{C}I \text{ on the element } v \text{ of the left } \mathbb{C}I\text{-module } \text{Res}_{\tau} \text{Res}_{\rho} V) = (\text{the action of } i \in \mathbb{C}I \text{ on the element } v \text{ of the left } \mathbb{C}I\text{-module } \text{Res}_{\rho \circ \tau} V) \quad = \quad (11.739).\n\]

Now, we know that $\text{Res}_{\tau} \text{Res}_{\rho} V = \text{Res}_{\rho \circ \tau} V$ as $\mathbb{C}$-vector spaces. Thus, (11.739) shows that $\text{Res}_{\tau} \text{Res}_{\rho} V = \text{Res}_{\rho \circ \tau} V$ as left $\mathbb{C}I$-modules as well. Exercise 4.16(c) is thus solved.
(d) Let \( f \in R_C(G) \). The definition of \( \text{Res}_{\rho \tau} f \) yields \( \text{Res}_{\rho \tau} f = f \circ (\rho \circ \tau) \). But the definition of \( \text{Res}_\rho f \) yields \( \text{Res}_\rho f = f \circ \rho \). The definition of \( \text{Res}_\tau \), \( \text{Res}_\rho \) yields
\[
\text{Res}_\tau \text{Res}_\rho f = (\text{Res}_\rho f) \circ \tau = (f \circ \rho) \circ \tau = f \circ (\rho \circ \tau) = \text{Res}_{\rho \tau} f.
\]
This solves Exercise 4.16(d).

(a) Let \( U \) be any \( CI \)-module.

Recall that \( \text{Ind}_\tau U \) is defined as the \( CH \)-module \( CH \otimes_{CI} U \), where \( CH \) is regarded as a \( (CH, CI) \)-bimodule according to the following rule: The left \( CH \)-module structure on \( CH \) is plain multiplication inside \( CH \); the right \( CI \)-module structure on \( CH \) is induced by the \( \mathbb{C} \)-algebra homomorphism \( \mathbb{C} [\tau] : CI \to CH \) (thus, it is explicitly given by \( \gamma \eta = \gamma \cdot (\mathbb{C} [\tau]) \eta \) for all \( \gamma \in CH \) and \( \eta \in CI \).

Furthermore, \( \text{Ind}_\rho (\text{Ind}_\tau U) \) is defined as the \( CG \)-module \( CG \otimes_{CH} (\text{Ind}_\tau U) \), where \( CG \) is regarded as a \( (CG, CH) \)-bimodule according to the following rule: The left \( CG \)-module structure on \( CG \) is plain multiplication inside \( CG \); the right \( CH \)-module structure on \( CG \) is induced by the \( \mathbb{C} \)-algebra homomorphism \( \mathbb{C} [\rho] : CH \to CG \) (thus, it is explicitly given by \( \gamma \eta = \gamma \cdot (\mathbb{C} [\rho]) \eta \) for all \( \gamma \in CG \) and \( \eta \in CH \).

Finally, \( \text{Ind}_{\rho \tau} U \) is defined as the \( CG \)-module \( CG \otimes_{CI} U \), where \( CG \) is regarded as a \( (CG, CI) \)-bimodule according to the following rule: The left \( CG \)-module structure on \( CG \) is plain multiplication inside \( CG \); the right \( CI \)-module structure on \( CG \) is induced by the \( \mathbb{C} \)-algebra homomorphism \( \mathbb{C} [\rho \circ \tau] : CI \to CG \) (thus, it is explicitly given by \( \gamma \eta = \gamma \cdot (\mathbb{C} [\rho \circ \tau]) \eta \) for all \( \gamma \in CG \) and \( \eta \in CI \).

Thus, we have introduced a \( (CH, CI) \)-bimodule structure on \( CH \), a \( (CG, CH) \)-bimodule structure on \( CG \) and a \( (CG, CI) \)-bimodule structure on \( CG \). The left \( CG \)-module structure that underlies the \( (CG, CH) \)-bimodule structure on \( CG \) is identical with the left \( CG \)-module structure that underlies the \( (CG, CI) \)-bimodule structure on \( CG \) (because both of these left \( CG \)-module structures are defined to be plain multiplication inside \( CG \)). Therefore, we will not run into ambiguities if we write expressions such as \( ab \) for \( a \in CG \) and \( b \in CG \).

We shall now show that
\[
(11.740) \quad CG \otimes_{CH} CH \cong CG \quad \text{as \( (CG, CI) \)-bimodules}.
\]
(Here, on the left hand side, \( CG \) is regarded as a \( (CG, CH) \)-bimodule and \( CH \) is regarded as \( (CH, CI) \)-bimodule, whereas on the right hand side, \( CG \) is regarded as a \( (CG, CI) \)-bimodule.)

596 Proof of (11.740): There is clearly a unique \( \mathbb{C} \)-vector space isomorphism \( \Phi : CG \to CG \otimes_{CH} CH \) which satisfies
\[
(11.741) \quad \Phi(m) = m \otimes_{CH} 1_{CH} \quad \text{for all } m \in CG.
\]

Consider this \( \Phi \). It is straightforward to see that \( \Phi \) is a homomorphism of left \( G \)-sets and a homomorphism of right \( I \)-sets. Hence, \( \Phi \) is a homomorphism of \( (CG, CI) \)-bimodules (since \( \Phi \) is \( \mathbb{C} \)-linear), thus an isomorphism of \( (CG, CI) \)-bimodules (since \( \Phi \) is invertible). Therefore, there exists an isomorphism of \( (CG, CI) \)-bimodules from \( CG \) to \( CG \otimes_{CI} CH \) (namely, \( \Phi \)). In other words, \( CG \otimes_{CH} CH \cong CG \) as \( (CG, CI) \)-bimodules. This proves (11.740).

Now,
\[
\text{Ind}_\rho (\text{Ind}_\tau U) = CG \otimes_{CH} (\text{Ind}_\tau U) = CG \otimes_{CH} (CH \otimes_{CI} U) =CG \otimes_{CI} U \quad \text{(by the associativity of the tensor product)}
\]
\[
\cong CG \otimes_{CI} U \quad \text{(by (11.740))}
\]
\[
\cong CG \otimes_{CI} U \quad \text{(since \( \text{Ind}_{\rho \tau} U = CG \otimes_{CI} U \) as left \( CG \)-modules)}
\]
as left \( CG \)-modules. This solves Exercise 4.16(a).

596 Indeed, this is a particular case of the following fundamental fact from linear algebra: If \( A \) is a \( \mathbb{C} \)-algebra, and if \( M \) is a right \( A \)-module, then there is a unique \( \mathbb{C} \)-vector space isomorphism \( \Phi : M \to M \otimes_{A} A \) which satisfies
\[
(\Phi(m) = m \otimes_{A} 1_{A} \quad \text{for all } m \in M).
\]
(b) First solution to Exercise 4.16(b). Let \( f \in R_C(I) \). Every \( r \in H \) satisfies
\[
(\text{Ind}_r f)(r) = \frac{1}{|I|} \sum_{(h,k) \in I \times H; \ k \tau(h) k^{-1} = r} f(h) \quad \text{(by the definition of Ind}_r f)\]
\[
= \frac{1}{|I|} \sum_{(i,v) \in I \times H; \ v \tau(i) v^{-1} = r} f(i) \quad \text{(here, we renamed the summation index \((h,k)\) as \((i,v)\))}
\]
\[
= \sum_{i \in I} \sum_{v \in H; \ v \tau(i) v^{-1} = r} f(i).
\]

Now, let \( g \in G \). Then, the definition of \( \text{Ind}_g (\text{Ind}_r f) \) yields
\[
(\text{Ind}_g (\text{Ind}_r f))(g) = \frac{1}{|H|} \sum_{(h,k) \in H \times G; \ k \rho(h) k^{-1} = g} (\text{Ind}_r f)(h) = \frac{1}{|H|} \sum_{(r,k) \in H \times G; \ k \rho(r) k^{-1} = g} \sum_{i \in I} \sum_{v \in H; \ v \tau(i) v^{-1} = r} (\text{Ind}_r f)(r)
\]
\[
= \sum_{k \in G} \sum_{r \in H; \ k \rho(r) k^{-1} = g} f(i) \quad \text{(by (11.742))}
\]
\[
= \sum_{i \in I} \sum_{r \in H; \ k \rho(r) k^{-1} = g} f(i) \quad \text{(here, we renamed the summation index \((h,k)\) as \((r,k)\))}
\]
\[
= \sum_{k \in G} \sum_{i \in I} \sum_{r \in H; \ k \rho(r) k^{-1} = g} f(i) \quad \text{(by (11.743))}
\]
\[
= \sum_{i \in I} \sum_{e \in H; \ k \rho(e) k^{-1} = g} f(e) \quad \text{(by (11.744))}
\]

But every \( k \in G \) and \( i \in I \) satisfy
\[
(11.744) \quad \sum_{r \in H; \ k \rho(r) k^{-1} = g} \sum_{v \in H; \ v \tau(i) v^{-1} = r} f(i) = \sum_{v \in H; \ k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} = g} f(i)
\]

597 Proof of (11.744): Let \( k \in G \) and \( i \in I \). We must prove (11.744). We have
\[
\sum_{v \in H; \ k \rho(v) k^{-1} = g} \sum_{v \tau(i) v^{-1} = r} f(i)
\]
\[
= \sum_{v \in H; k \rho(v) k^{-1} = g} \sum_{r \in H; v \tau(i) v^{-1} = r} f(i)
\]
\[
= \sum_{v \in H; r \in H; k \rho(v) k^{-1} = g} f(i) \quad \text{(because every \( v \in H \) satisfies either \( k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} = g \) or \( k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} \neq g \), but never both)}.
\]
Thus, (11.743) becomes

\[
(\text{Ind}_r (\text{Ind}_\rho f)) (g)
\]

\[
= \frac{1}{|H|} \sum_{i \in I} \sum_{k \in G} f(i) = \frac{1}{|H|} \sum_{i \in I} \sum_{v \in H; \ k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} = g} f(i)
\]

(by (11.744))

\[
= \frac{1}{|H|} \sum_{i \in I} \sum_{v \in H; \ k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} = g} f(i)
\]

(11.748)

\[
= \frac{1}{|H|} \sum_{i \in I} \sum_{v \in H; \ k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} = g} f(i).
\]

Now, we shall show that

\[
\sum_{r \in H; \ k \rho(r) = g; \ v \tau(r) = v} f(i) = 0
\]

for every \(v \in H\) which satisfies \(k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} = g\).

Proof of (11.746): Let \(v \in H\) be such that \(k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} = g\). If some \(r \in H\) satisfies \(k \rho(r) = g\) and \(v \tau(r) = v\), then this \(r\) must satisfy

\[
k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} = k \rho(v) \cdot \rho(\tau(i)) \cdot (\rho(v))^{-1} k^{-1}
\]

\[
= \rho(v) \cdot \rho(\tau(i)) \cdot (\rho(v))^{-1} k^{-1}
\]

\[
= \rho(v) \cdot \rho(\tau(i)) \cdot (\rho(v))^{-1}
\]

\[
(\text{since } \rho \text{ is a group homomorphism})
\]

\[
k \rho(v \tau(i)) k^{-1} = k \rho(r) = g,
\]

which contradicts \(k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} = g\). Hence, we have obtained a contradiction for every \(r \in H\) which satisfies \(k \rho(r) = g\) and \(v \tau(r) = v\). Therefore, the sum \(\sum_{r \in H; \ k \rho(r) = g; \ v \tau(r) = v} f(i)\) is an empty sum, and thus its value is 0. This proves (11.746).

On the other hand, let us show that

\[
\sum_{r \in H; \ k \rho(r) = g; \ v \tau(r) = v \tau(i)} f(i) = f(i)
\]

for every \(v \in H\) which satisfies \(k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} = g\).

Proof of (11.747): Let \(v \in H\) be such that \(k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} = g\). If some \(r \in H\) satisfies \(k \rho(r) = g\) and \(v \tau(r) = v\), then \(r\) must equal \(v \tau(i)\) \(v\) (because \(v \tau(i) = v\)). Hence, there exists at most one \(r \in H\) which satisfies \(k \rho(r) = g\) and \(v \tau(r) = v\).

On the other hand, the element \(v \tau(i)\) \(v\) of \(H\) satisfies

\[
k \rho(v \tau(i)) k^{-1} = k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} k^{-1} = k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1} g
\]

\[
= k \rho(v) \cdot \rho(\tau(i)) \cdot (k \rho(v))^{-1}
\]

\[
(\text{since } \rho \text{ is a group homomorphism})
\]

and \(v \tau(i) = v \tau(i) \tau(i) = v \). In other words, \(v \tau(i)\) \(v\) is an element \(r \in H\) which satisfies \(k \rho(r) = g\) and \(v \tau(i) = v\).

Hence, there exists at least one \(r \in H\) which satisfies \(k \rho(r) = g\) and \(v \tau(i) = v\). Since we also have shown that there exists at most one \(r \in H\), we can thus conclude that there exists exactly one \(r \in H\) which satisfies \(k \rho(r) = g\) and \(v \tau(i) = v\).

The sum \(\sum_{r \in H; \ k \rho(r) = g; \ v \tau(r) = v \tau(i)} f(i)\) has precisely one addend. Hence, this

sum rewrites as:

\[
\sum_{r \in H; \ k \rho(r) = g; \ v \tau(r) = v \tau(i)} f(i) = f(i).
\]

This proves (11.747).
But for every $i \in I$ and $v \in H$, we have
\begin{equation}
(11.749) \quad \sum_{k \in G; \ k p(v) \rho(\tau(i)) (k p(v))^{-1} = g} f(i) = \sum_{k \in G; \ k p(\tau(i)) k^{-1} = g} f(i)
\end{equation}

Thus, \((11.748)\) becomes
\begin{equation}
(\text{Ind}_\rho (\text{Ind}_\tau f))(g) = \frac{1}{|I| |H|} \sum_{i \in I} \sum_{v \in H} f(i) = \frac{1}{|I| |H|} \sum_{i \in I} \sum_{v \in H} f(i) = \frac{1}{|I| |H|} \sum_{i \in I} \sum_{v \in H} f(i) = \frac{1}{|I| |H|} \sum_{i \in I} \sum_{v \in H} f(i)
\end{equation}

(11.749) as (by \((11.746)\))

Compared with
\begin{equation}
(\text{Ind}_{\rho \circ \tau} f)(g) = \frac{1}{|I|} \sum_{(h,k) \in I \times G; \ k (\rho(\tau)(h) k^{-1} = g} f(h)
\end{equation}

this yields $(\text{Ind}_\rho (\text{Ind}_\tau f))(g) = (\text{Ind}_{\rho \circ \tau} f)(g)$.

Let us now forget that we fixed $g$. We thus have shown that $(\text{Ind}_\rho (\text{Ind}_\tau f))(g) = (\text{Ind}_{\rho \circ \tau} f)(g)$ for every $g \in G$. In other words, $\text{Ind}_\rho \text{Ind}_\tau f = \text{Ind}_{\rho \circ \tau} f$. This solves Exercise 4.16(b).

Now, \((11.745)\) becomes
\begin{equation}
\sum_{r \in H; \ k p(\tau)(r) k^{-1} = g} \sum_{v \in H; \ v \tau(r) v^{-1} = r} f(i) = \sum_{r \in H; \ k p(v) \rho(\tau(i)) (k p(v))^{-1} = g} \sum_{v \in H; \ v \tau(i) v^{-1} = r} f(i) + \sum_{r \in H; \ k p(v) \rho(\tau(i)) (k p(v))^{-1} \neq g} \sum_{v \in H; \ v \tau(i) v^{-1} = r} f(i)
\end{equation}

(11.745) as (by \((11.747)\))

This proves \((11.744)\).

\textbf{Proof of (11.749):} Let $i \in I$ and $v \in H$. We have $\rho(v) \in G$. Therefore, the map $G \to G$, $k \mapsto k p(v)$ is a bijection (since $G$ is a group). Therefore, we can substitute $k p(v)$ for $k$ in the sum $\sum_{k \in G; \ k p(\tau(i)) k^{-1} = g} f(i)$. We thus obtain
\begin{equation}
\sum_{k \in G; \ k p(\tau(i)) k^{-1} = g} f(i) = \sum_{k \in G; \ k p(v) \rho(\tau(i)) (k p(v))^{-1} = g} f(i) = \sum_{k \in G; \ k p(\tau(i)) (k p(v))^{-1} = g} f(i)
\end{equation}

This proves \((11.749)\).
Second solution to Exercise 4.16(b). We shall now give an alternative solution of Exercise 4.16(b) which relies on Exercise 4.16(d) and a certain fact about class functions:

Lemma 11.20. Let $G$ be a finite group. Let $u \in R_C(G)$. Assume that every $v \in R_C(G)$ satisfies $\langle u, v \rangle_G = 0$. Then, $u = 0$.

Proof of Lemma 11.20. Let us use the Iverson bracket notation; that is, for any statement $A$, we define $[A]$ to be the integer $\begin{cases} 1, & \text{if } A \text{ is true;} \\ 0, & \text{if } A \text{ is false.} \end{cases}$

Fix an element $h \in G$. We define a map $\alpha_h : G \to \mathbb{C}$ by

$$\alpha_h(g) = \sum_{k \in G} [khk^{-1} = g] \quad \text{for every } g \in G.$$ (Notice that this map $\alpha_h$ is identical with the map $\alpha_{G,h}$ defined in Exercise 4.34, but we will not use this.) Then, $\alpha_h \in R_C(G)$ $^{599}$. Hence, $\langle u, \alpha_h \rangle_G$ is well-defined. But recall that every $v \in R_C(G)$ satisfies $\langle u, v \rangle_G = 0$. Applying this to $v = \alpha_h$, we obtain $\langle u, \alpha_h \rangle_G = 0$. Thus,

$$0 = \langle u, \alpha_h \rangle_G = \frac{1}{|G|} \sum_{g \in G} u(g) \alpha_h(g^{-1}) = \sum_{k \in G} [khk^{-1} = g^{-1}]$$ (by the definition of the bilinear form $\langle \cdot, \cdot \rangle_G$)

$$= \frac{1}{|G|} \sum_{g \in G} u(g) \sum_{k \in G} [khk^{-1} = g^{-1}] = \frac{1}{|G|} \sum_{g \in G} u(g) \left[ khhk^{-1} = g^{-1} \right]$$ (by the definition of $\alpha_h$)

$$= \frac{1}{|G|} \sum_{k \in G} \sum_{g \in G} u(g) \left[ g = (khk^{-1})^{-1} \right].$$ (11.750)

But every $k \in G$ satisfies

$$\sum_{g \in G} u(g) \left[ g = (khk^{-1})^{-1} \right] = u(h^{-1})$$ (11.751)

$^{599}$Proof. Let $p$ and $q$ be two conjugate elements of $G$. Then, there exists some $r \in G$ such that $p = qrq^{-1}$ (since $p$ and $q$ are conjugate). Consider this $r$. The definition of $\alpha_h$ yields

$$\alpha_h(p) = \sum_{k \in G} [khk^{-1} = p] = \sum_{k \in G} [khk^{-1} = qrq^{-1}] = \sum_{k \in G} [r^{-1}khk^{-1}r = q].$$

Now, recall that $r \in G$. Hence, the map $G \to G$, $k \mapsto rk$ is a bijection (since $G$ is a group). Hence, we can substitute $rk$ for $k$ in the sum $\sum_{k \in G} [r^{-1}khk^{-1}r = q]$. We thus obtain

$$\sum_{k \in G} [r^{-1}khk^{-1}r = q] = \sum_{k \in G} [r^{-1}khk^{-1}r = q] = \sum_{k \in G} [r^{-1}khk^{-1}r = q] = \sum_{k \in G} [khk^{-1} = q] = \alpha_h(q).$$

(since $\alpha_h(q) = \sum_{k \in G} [khk^{-1} = q]$ (by the definition of $\alpha_h$)). Hence, $\alpha_h(p) = \sum_{k \in G} [r^{-1}khk^{-1}r = q] = \alpha_h(q)$.

Let us now forget that we fixed $p$ and $q$. We thus have shown that $\alpha_h(p) = \alpha_h(q)$ whenever $p$ and $q$ are two conjugate elements of $G$. In other words, the function $\alpha_h$ is constant on $G$-conjugacy classes. In other words, $\alpha_h$ is a class function on $G$ (because the class functions on $G$ are defined to be the functions $G \to \mathbb{C}$ which are constant on $G$-conjugacy classes). In other words, $\alpha_h \in R_C(G)$ (since $R_C(G)$ is the set of all class functions on $G$). Qed.
Hence, \((11.750)\) becomes
\[
0 = \frac{1}{|G|} \sum_{k \in G} \sum_{g \in G} u(g) \left[ g = (khk^{-1})^{-1} \right] \quad \text{where} \quad \sum_{k \in G} u(h^{-1}) = \frac{1}{|G|} |G| u(h^{-1}) = u(h^{-1}).
\]

Thus, \(u(h^{-1}) = 0\).

Let us now forget that we fixed \(h\). We thus have shown that
\[(11.753) \quad u(h^{-1}) = 0 \quad \text{for every } h \in G.\]

Let us now fix \(h \in G\). Then, \((11.753)\) (applied to \(h^{-1}\) instead of \(h\)) yields \(u(h^{-1})^{-1} = 0\). In other words, \(u(h) = 0\) (since \((h^{-1})^{-1} = h\)).

Let us now forget that we fixed \(h\). We thus have shown that \(u(h) = 0\) for every \(h \in G\). In other words, \(u = 0\). This proves Lemma 11.20.

Now, let us return to solving Exercise 4.16(b). Let \(f \in R_G(I)\). Let \(v \in R_G(G)\). Then,
\[
\langle \Ind_{\rho} \Ind_{\tau} f, v \rangle_G = \langle \Ind_{\tau} f, \Res_{\rho} v \rangle_H \quad \text{(by (4.17), applied to } \alpha = \Ind_{\tau} f \text{ and } \beta = v) = \langle f, \Res_{\rho} \Res_{\tau} v \rangle_I \quad \text{by (4.17), applied to } H, I, \tau, f \text{ and } \Res_{\rho} v \text{ instead of } G, H, \rho, \alpha \text{ and } \beta)
\]

Compared with
\[
\langle \Ind_{\rho \circ \tau} f, v \rangle_G = \langle f, \Res_{\rho \circ \tau} v \rangle_I \quad \text{(by (4.17), applied to } H, \rho \circ \tau, f \text{ and } v \text{ instead of } H, \rho, \alpha \text{ and } \beta),
\]
this yields \(\langle \Ind_{\rho} \Ind_{\tau} f, v \rangle_G = \langle \Ind_{\rho \circ \tau} f, v \rangle_G\). Now, the form \(\langle \cdot, \cdot \rangle_G\) is \(\mathbb{C}\)-bilinear, and therefore we have
\[
\langle \Ind_{\rho} \Ind_{\tau} f - \Ind_{\rho \circ \tau} f, v \rangle_G = \langle \Ind_{\rho} \Ind_{\tau} f, v \rangle_G - \langle \Ind_{\rho \circ \tau} f, v \rangle_G = \langle \Ind_{\rho \circ \tau} f, v \rangle_G - \langle \Ind_{\rho \circ \tau} f, v \rangle_G = 0.
\]

Now, let us forget that we fixed \(v\). We thus have shown that every \(v \in R_G(G)\) satisfies \(\langle \Ind_{\rho} \Ind_{\tau} f - \Ind_{\rho \circ \tau} f, v \rangle_G = 0\). Hence, Lemma 11.20 (applied to \(u = \Ind_{\rho} \Ind_{\tau} f - \Ind_{\rho \circ \tau} f\)) yields \(\Ind_{\rho} \Ind_{\tau} f - \Ind_{\rho \circ \tau} f = 0\). Thus, \(\Ind_{\rho} \Ind_{\tau} f = \Ind_{\rho \circ \tau} f\). This solves Exercise 4.16(b) again.

[Remark: Recall that Exercise 4.14 (specifically, its parts (c), (d), (e) and (f)) shows that \(\rho\)-induction (of modules and of class functions) generalizes both the usual notion of induction and the fixed point construction. More precisely, \(\rho\)-induction becomes usual induction (at least up to isomorphism) when \(\rho\) is injective, and becomes fixed point construction when \(\rho\) is surjective. Exercise 4.16 (specifically, its parts (a)]

\[600\] Proof of \((11.751):\) Let us first recall that \(u\) belongs to the set \(R_G(G)\). In other words, \(u\) is a class function on \(G\) (since \(R_G(G)\) is the set of all class functions on \(G\)). In other words, the function \(u\) is constant on \(G\)-conjugacy classes (because the class functions on \(G\) are defined to be the functions \(G \to \mathbb{C}\) which are constant on \(G\)-conjugacy classes). In other words, \((11.752)\)
\[
u(p) = u(q)
\]
whenever \(p\) and \(q\) are two conjugate elements of \(G\).

Now, let \(k \in G\). Then, the elements \(kh^{-1}k^{-1}\) and \(h^{-1}\) of \(G\) are conjugate. Hence, \(u(kh^{-1}k^{-1}) = u(h^{-1})\) (by (11.752), applied to \(p = kh^{-1}k^{-1}\) and \(q = h^{-1}\)).

All addends of the sum \(\sum_{g \in G} u(g) \left[ g = (khk^{-1})^{-1} \right]\) are zero except for the addend for \(g = (khk^{-1})^{-1}\) (because the factor \[g = (khk^{-1})^{-1}\] is zero unless \(g = (khk^{-1})^{-1}\)). Hence, this sum simplifies as follows:
\[
\sum_{g \in G} u(g) \left[ g = (khk^{-1})^{-1} \right] = u\left( \frac{(khk^{-1})^{-1}}{h^{-1}k^{-1}} \right) \left[ (khk^{-1})^{-1} = (khk^{-1})^{-1} \right] = u(kh^{-1}k^{-1}) = u(h^{-1}).
\]

This proves \((11.751)\).
and (b)), on the other hand, shows that in the general case, \(\rho\)-induction can be reduced to a composition of usual induction and fixed point construction, because every group homomorphism \(\rho : H \to G\) can be factored as a composition \(\alpha \circ \beta\) of a surjective group homomorphism \(\beta : H \to \overline{G}\) with an injective group homomorphism \(\alpha : \overline{G} \to G\). This allows some alternative proofs of some parts of Exercise 4.14, which the interested reader can find.]

11.84. Solution to Exercise 4.20. Solution to Exercise 4.20. We need to prove that

\[
\text{(11.754)} \quad \text{Ind}^S_{\mathcal{G}+j+k} \left( \text{Ind}^S_{\mathcal{G}+j} \left( U \otimes V \right) \otimes W \right) \cong \text{Ind}^S_{\mathcal{G}+j+k} \left( U \otimes V \otimes W \right)
\]

and

\[
\text{(11.755)} \quad \text{Ind}^S_{\mathcal{G}+j+k} \left( U \otimes \text{Ind}^S_{\mathcal{G}+j+k} \left( V \otimes W \right) \right) \cong \text{Ind}^S_{\mathcal{G}+j+k} \left( U \otimes V \otimes W \right)
\]

as \(\mathbb{C}[\mathcal{G} \times \mathcal{G} \times \mathcal{G}]\)-modules. We will only prove (11.754), since (11.755) is analogous.

It is easy to see that every finite group \(G\) and every \(CG\)-module \(P\) satisfy \(\text{Ind}_G^P P \cong P\). Applied to \(G = \mathcal{G}_k\) and \(P = W\), this yields \(\text{Ind}^{\mathcal{G}_k}_W W \cong W\).

Now, Exercise 4.2 (applied to \(\mathcal{G}+j+k, \mathcal{G}+j \times \mathcal{G}_k, \mathcal{G} \times \mathcal{G}_j \times \mathcal{G}_k\) and \(U \otimes V \otimes W\) instead of \(G, H, I\) and \(U\)) yields

\[
\text{Ind}^S_{\mathcal{G}+j+k} \left( \text{Ind}^S_{\mathcal{G}+j} \left( U \otimes V \otimes W \right) \right) \cong \text{Ind}^S_{\mathcal{G}+j+k} \left( U \otimes V \otimes W \right).
\]

Hence,

\[
\text{Ind}^S_{\mathcal{G}+j+k} \left( U \otimes V \otimes W \right) \cong \text{Ind}^S_{\mathcal{G}+j+k} \left( \text{Ind}^S_{\mathcal{G}+j+k} \left( U \otimes V \otimes W \right) \right)
\]

\[
\cong \text{Ind}^S_{\mathcal{G}+j+k} \left( \text{Ind}^S_{\mathcal{G}+j} \left( U \otimes V \right) \otimes \text{Ind}^S_{\mathcal{G}_k} W \right)
\]

\[
\cong \text{Ind}^S_{\mathcal{G}+j+k} \left( \text{Ind}^S_{\mathcal{G}+j} \left( U \otimes V \right) \otimes W \right)
\]

This proves (11.754). Thus, the solution of Exercise 4.20 is complete.

11.85. Solution to Exercise 4.29. Solution to Exercise 4.29. (a) Let \(n \in \mathbb{N}\). Let \(B\) denote the subgroup of \(GL_n(\mathbb{F})\) consisting of all upper-triangular matrices. Then, \(GL_n(\mathbb{F}) = \bigcup_{w \in \mathcal{S}_n} BwB\). (Indeed, this can be proven in the same way as we have shown the equality \(GL_n = \bigcup_{w \in \mathcal{S}_n} BwB\) in our above proof of Proposition 4.27, with the only difference that we are now working over \(\mathbb{F}\) instead of \(\mathbb{F}_q\).)

Let \(w_0\) denote the permutation in \(\mathcal{S}_n\) which sends every \(i \in \{1, 2, \ldots, n\}\) to \(n+1-i\); then, \(w_0^2 = \text{id}\). (The permutation \(w_0\) is written as \((n, n-1, \ldots, 1)\) in one-line notation.)

Let \(A \in GL_n(\mathbb{F})\). We have \(w_0A \in GL_n(\mathbb{F}) = \bigcup_{w \in \mathcal{S}_n} BwB\). Thus, there exists some \(w \in \mathcal{S}_n\) such that \(w_0A \in BwB\). Consider this \(w\). There exist two upper-triangular invertible matrices \(L'\) and \(U\) such that \(w_0A = L'wU\) (since \(w_0A \in BwB\)). Consider these \(L'\) and \(U\). The matrix \(w_0L'w_0\) is lower-triangular (since \(L'\) is upper-triangular, and since \(w_0\) is what it is). Set \(L = w_0L'w_0\). Then, \(L = w_0L'w_0\). Since \(P = w_0w_0\). Since \(P\) is clearly a permutation matrix (being the product of the permutation matrices \(w_0\) and \(w\)), we thus have shown that \(A = LPU\) for
a lower-triangular matrix $L \in GL_n(\mathbb{F})$, an upper-triangular matrix $U \in GL_n(\mathbb{F})$ and a permutation matrix $P \in \mathfrak{S}_n \subset GL_n(\mathbb{F})$. This solves Exercise 4.29(a).

(b) In the proof that follows, we shall essentially mimic the arguments used to prove $GL_n = \bigcup_{w \in \mathfrak{S}_n} BwB$ in our proof of Proposition 4.27 (but we will add some details in a few places).

Let us first show that the disjoint union $\bigcup_{f \in F_{n,m}} B_n f B_m$ is well-defined. This means proving that the sets $B_n f B_m$ for $f \in F_{n,m}$ are disjoint.

For every $A \in \mathbb{F}^{n \times m}$, every $i \in \{1, 2, \ldots, n+1\}$ and every $j \in \{0, 1, \ldots, m\}$, let $r_{i,j}(A)$ denote the rank of the matrix obtained by restricting $A$ to the rows $i, i+1, \ldots, n$ and columns $1, 2, \ldots, j$. This rank $r_{i,j}(A)$ does not change when we replace $A$ by $XA$ for some $X \in B_n$ (because replacing $A$ by $XA$ for some $X \in B_n$ means that we add to each row of $A$ a linear combination of the rows further below; but this row transformation does not change $r_{i,j}(A)$); neither does it change when we replace $A$ by $AY$ for some $Y \in B_m$ (for a similar reason). Hence, for every $f \in F_{n,m}$, $X \in B_n$, $Y \in B_m$, $i \in \{1, 2, \ldots, n+1\}$ and $j \in \{0, 1, \ldots, m\}$, we have

$$r_{i,j}(f) = r_{i,j}(Xf) = r_{i,j}(Yf).$$

Now, we claim that if $f \in F_{n,m}$ and $A \in B_n f B_m$, then we can reconstruct $f$ from $A$. Indeed, let $f \in F_{n,m}$ and $A \in B_n f B_m$. Then, there exist $X \in B_n$ and $Y \in B_m$ such that $A = XfY$ (since $A \in B_n f B_m$). Consider these $X$ and $Y$. Then, (11.756) yields $r_{i,j}(f) = r_{i,j}(XfY) = r_{i,j}(A)$ for every $i \in \{1, 2, \ldots, n+1\}$ and $j \in \{0, 1, \ldots, m\}$. Thus, from $A$ we can reconstruct the ranks $r_{i,j}(f)$ for all $i \in \{1, 2, \ldots, n+1\}$ and $j \in \{0, 1, \ldots, m\}$.

Notice that $f \in F_{n,m}$. Thus, $f$ is a matrix in $\{0, 1\}^{n \times m}$ such that each row of $f$ contains at most one 1 and each column of $f$ contains at most one 1.

Now, fix $i \in \{1, 2, \ldots, n\}$ and $j \in \{1, 2, \ldots, m\}$. Let $f_{i,j}$ be the entry of $f$ in row $i$ and column $j$. We will now show that

$$f_{i,j} = r_{i,j}(f) - r_{i,j-1}(f) - r_{i+1,j}(f) + r_{i+1,j-1}(f).$$

Proof of (11.757). First of all, if the first $j$ entries of the $i$-th row of $f$ are all 0, then (11.757) holds for obvious reasons (in fact, in this case, it is clear that $f_{i,j} = 0$, $r_{i,j}(f) = r_{i+1,j}(f)$ and $r_{i,j-1}(f) = r_{i+1,j-1}(f)$). We thus WLOG assume that the first $j$ entries of the $i$-th row of $f$ are not all 0. Then, there must be a 1 among these entries. It must lie in a different column than the 1’s appearing in all other rows of $f$ (because each column of $f$ contains at most one 1). Therefore, the first row of the matrix obtained by restricting $f$ to the rows $i, i+1, \ldots, n$ and columns $1, 2, \ldots, j$ is linearly independent from its other rows. Thus,

$$r_{i,j}(f) = r_{i+1,j}(f) + 1.$$
Thus we have proven that if \( f \in F_{n,m} \) and \( A \in B_n f B_m \), then we can reconstruct \( f \) from \( A \). In other words, the sets \( B_n f B_m \) for \( f \in F_{n,m} \) are disjoint. Thus, the disjoint union \( \bigcup_{f \in F_{n,m}} B_n f B_m \) is well-defined.

It remains to prove that this disjoint union \( \bigcup_{f \in F_{n,m}} B_n f B_m \) is \( \mathbb{F}^{n \times m} \). In order to do so, it is clearly enough to show that every \( g \in \mathbb{F}^{n \times m} \) belongs to \( \bigcup_{f \in F_{n,m}} B_n f B_m \).

Let \( g \in \mathbb{F}^{n \times m} \). We need to show that \( g \in \bigcup_{f \in F_{n,m}} B_n f B_m \). In other words, we need to find some \( f \in F_{n,m} \) such that \( g \in B_n f B_m \). In order to do so, we shall find a matrix \( f \in F_{n,m} \) which lies in \( B_n g B_m \). Once this is done, it will follow that \( g \in B_n f B_m \) (since every element of \( B_n \) and every element of \( B_m \) are invertible (because \( B_n \subseteq GL_n(\mathbb{F}) \) and \( B_m \subseteq GL_m(\mathbb{F}) \)), and we will be finished.

We refer to \( g B_m \) and \( B_n g \) as cosets, and to \( B_n g B_m \) as a double cosets, despite the fact that they are not subsets of a group.

The freedom to alter \( g \) within the coset \( g B_m \) allows one to scale columns and add scalar multiples of earlier columns to later columns. We claim that using such column operations, one can always find a representative \( g' \) for coset \( g B_m \) in which

- the bottommost nonzero entry of each nonzero column is 1 (call this entry a pivot),
- the entries to right of each pivot within its row are all 0, and
- there is at most one pivot in each row and at most one pivot in each column (so that their positions are the positions of the 1's in some matrix \( f \in F_{n,m} \)).

In fact, we will see below that \( B_n g B_m = B_n f B_m \) in this case. The algorithm which produces \( g' \) from \( g \) is simple: starting with the leftmost nonzero column, find its bottommost nonzero entry, and scale the column to make this entry a 1, creating the pivot in this column. Now use this pivot to clear out all entries in its row to its right, using column operations that subtract multiples of this column from later columns. Having done this, move on to the next nonzero column to the right, and repeat, scaling to create a pivot, and using it to eliminate entries to its right.\(^{601}\)

Having found this \( g' \) in \( g B_m \), a similar algorithm using left multiplication by \( B_n \) shows that \( f \) lies in \( B_n g' B_m = B_n g B_m \). This time no scalings are required to create the pivot entries: starting with the bottommost nonzero row, one uses its pivot to eliminate all the entries above it in the same column by adding multiples of this row to higher rows. Then do the same using the pivot in the next-to-bottom nonzero row, etc.\(^{602}\) The result is the matrix \( f \).

\(^{601}\)To see that this works, we need to check three facts:

(a) We will find a nonzero entry in every nonzero column during our algorithm.
(b) Our column operations preserve the zeroes lying to the right of already existing pivots.
(c) Every row contains at most one pivot at the end of the algorithm.

But fact (a) is a tautology. Fact (b) holds because all our operations either scale columns (which clearly preserves zero entries) or subtract a multiple of the column \( c \) containing the current pivot from a later column \( d \) (which will preserve every zero lying to the right of an already existing pivot, because any already existing pivot must lie in a column \( b < c \) and therefore both columns \( c \) and \( d \) have zeroes in its row). Fact (c) follows from noticing that the entries to the right of a pivot in its row are 0.

\(^{602}\)One thing that requires verification is the fact that these row operations preserve the following three properties of \( g' \):

- the bottommost nonzero entry of each nonzero column is 1 (call this entry a pivot),
- the entries to right of each pivot within its row are all 0, and
- there is at least one pivot in each row and at most one pivot in each column (so that their positions are the positions of the 1's in some matrix \( f \in F_{n,m} \)).

Let us show this, and also show that the positions of the pivots are preserved. Indeed, it is clear that the positions of the pivots are preserved (because zero columns stay zero, nonzero columns stay nonzero, and the bottommost entries of nonzero columns do not move); therefore it is enough to prove the following fact:

(d) Consider an elimination step, by which we mean a step in which a pivot in some position \( a \) is used to eliminate all the entries above it in the same column. Assume that, before the elimination step, the entries to the right of \( a \) within its row were all 0. Let \( b \) be the position of another pivot that existed before this elimination step. Assume that, before the elimination step, the pivot at \( b \) equaled 1, and the entries to the right of \( b \) within its row were all 0. Then, after the elimination step, \( b \) is a position of a pivot and this pivot still equals 1, and the entries to the right of \( b \) within its row are still 0.

So let us prove fact (d). We prove it by contradiction: Assume that it is not the case. Then, the elimination step must have changed at least one of the entries weakly to the right of \( b \) within its row (because we already know that \( b \) is still a position of a pivot after the elimination step). In particular, the elimination step must have changed at least one entry in the row of \( b \). Thus, \( b \) must lie in a row strictly above \( a \) (since otherwise, the elimination step would have not changed any entry in the row of \( b \)). Let \( c \) be the position in the same row as \( b \) and in the same column as \( a \). Then, the entry at \( c \) before the elimination step must not have been 0 (since otherwise, the elimination step would not have changed any entry in the row of \( b \)). As a consequence,
11.86. **Solution to Exercise 4.31.** Let us begin this solution by stating some trivialities. It is easy to see that every finite group $G$ and every $CG$-module $P$ satisfy

\[(11.758) \quad \text{Ind}_G^P P \cong P\]

and

\[(11.759) \quad \text{Infl}_G^G P \cong P.\]

Furthermore, if $G_1$ and $G_2$ be two groups, if $K_1 \triangleleft G_1$ and $K_2 \triangleleft G_2$ are normal subgroups, if $\mathcal{U}_1$ is a $C[G_1/K_1]$-module, and if $\mathcal{U}_2$ is a $C[G_2/K_2]$-module, then

\[(11.760) \quad \text{Infl}_{G_1 \times G_2}^{G_1} \mathcal{U}_1 \otimes \mathcal{U}_2 \cong \left( \text{Ind}_{G_1/K_1}^{G_1} \mathcal{U}_1 \right) \otimes \left( \text{Ind}_{G_2/K_2}^{G_2} \mathcal{U}_2 \right)\]

as $C[G_1 \times G_2]$-modules. (This is an analogue of (4.6), but is trivial to prove.) Also, if $G$, $H$, and $\mathcal{I}$ are three groups such that $\mathcal{I} < H < G$, and if $\mathcal{U}$ is a $C\mathcal{I}$-module, then

\[(11.761) \quad \text{Infl}_H^G \text{Infl}_\mathcal{I}^H \mathcal{U} = \text{Infl}_\mathcal{I}^G \mathcal{U}.\]

(This is an analogue of Exercise 4.2, and is again trivial.)

Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ be an almost-composition of an $n \in \mathbb{N}$ satisfying $\ell \geq 1$. Let $V_i$ be a $CG_i$-module for every $i \in \{1, 2, \ldots, \ell\}$. We need to prove the two isomorphisms

\[(11.762) \quad \text{ind}_{\alpha}^n (V_1 \otimes V_2 \otimes \cdots \otimes V_\ell)\]

\[(11.763) \quad \cong \text{ind}_{\alpha}^n (V_1 \otimes V_2 \otimes \cdots \otimes V_\ell)\]

We will only prove (11.762), since the proof of (11.763) is analogous.

Let $m$ be the nonnegative integer $\alpha_1 + \alpha_2 + \cdots + \alpha_{\ell-1}$. Then, $(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})$ is an almost-composition of $m$, and we have $\alpha_1 + \alpha_2 + \cdots + \alpha_{\ell-1} = m$.

Let $W$ denote the $C[G_1 \times G_2 \times \cdots \times G_{\ell-1}]$-module $V_1 \otimes V_2 \otimes \cdots \otimes V_{\ell-1}$. Then, $V_1 \otimes V_2 \otimes \cdots \otimes V_{\ell-1} = W$ and $V_1 \otimes V_2 \otimes \cdots \otimes V_{\ell} = \underbrace{(V_1 \otimes V_2 \otimes \cdots \otimes V_{\ell-1}) \otimes V_\ell = W \otimes V_\ell}$ and $\alpha_1 + \alpha_2 + \cdots + \alpha_{\ell-1} = m$. Hence, the relation (11.762) (which we want to prove) rewrites as

\[(11.764) \quad \text{ind}_{m, \alpha}^n (W \otimes V_\ell) \cong \text{ind}_{\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1}}^n (W \otimes V_\ell).\]

It thus remains to prove (11.764).

We distinguish between two cases:

**Case 1:** We have $G_\ast = \mathfrak{S}_\ast$ or $G_\ast = \mathfrak{S}_\ast \{\Gamma\}$.

**Case 2:** We have $G_\ast = GL_\ast$.

\[c\text{ must not lie to the right of } b\text{ within its row (because before the elimination step, the entries to the right of } b\text{ within its row were all 0). Thus, } c\text{ lies weakly to the left of } b\text{ within its row. In other words, the column containing } b\text{ lies weakly right of the column containing } c.\]

In other words, the column containing $b$ lies weakly right of the column containing $a$. Thus, the entries to the right of $b$ within its row in a. Then, the column containing $b$ must lie strictly right of the column containing $a$. Hence, the column containing $b$ must lie strictly right of the column containing $a$. But all these corresponding entries were silent entries (because the column containing $b$ lies strictly right of the column containing $a$) and thus were 0. Hence, the elimination step did not change the entries weakly to the right of $b$ within its row. This contradicts the fact that the elimination step must have changed at least one of the entries weakly to the right of $b$ within its row. This contradiction finishes the proof of fact (d).
Let us consider Case 1 first. In this case, we have $G_\star = \mathfrak{S}_\star$ or $G_\star = \mathfrak{S}_\star [\Gamma]$. Thus, $\text{ind}_N^N = \text{Ind}_{G_\beta}^{G_{\beta_i}}$ for every $N \in \mathbb{N}$ and every almost-composition $\beta$ of $N$ (by the definition of $\text{ind}_N^N$), and $\text{ind}_{i,j}^N = \text{Ind}_{G_{i} \times G_j}^{G_{\beta_i \times \beta_j}}$ for every $N \in \mathbb{N}$ and every $i, j \in \mathbb{N}$ satisfying $i + j = N$ (by the definition of $\text{ind}_{i,j}^N$).

Since $\text{ind}_{i,j}^N = \text{Ind}_{G_{i} \times G_j}^{G_{\beta_i \times \beta_j}}$ for every $N \in \mathbb{N}$ and every almost-composition $\beta$ of $N$, we have $\text{ind}_{m,\alpha_{\ell}}^m = \text{Ind}_{G_m \times G_{\alpha_{\ell}}}^{G_{m,\alpha_{\ell}}}$.

Thus,

\[
\text{ind}_{m,\alpha_{\ell}}^m \left( \text{ind}_{m,\alpha_{\ell}}^m \left( \mathcal{W} \otimes V_\ell \right) \right) = \text{Ind}_{G_m \times G_{\alpha_{\ell}}}^{G_{m,\alpha_{\ell}}} \left( \text{Ind}_{G_m \times G_{\alpha_{\ell}}}^{G_{m,\alpha_{\ell}}} \left( \mathcal{W} \otimes V_\ell \right) \right) = \text{Ind}_{G_m \times G_{\alpha_{\ell}}}^{G_{m,\alpha_{\ell}}} \left( \text{Ind}_{G_m \times G_{\alpha_{\ell}}}^{G_{m,\alpha_{\ell}}} \left( \mathcal{W} \otimes V_\ell \right) \right)
\]

(11.765)

This proves (11.762). Hence, (11.762) is proven in Case 1.

Let us now consider Case 2. In this case, we have $G_\star = GL_\star$. Hence, $\text{ind}_N^N = \text{Ind}_{G_\beta}^{G_{P_{G_\beta}}} \text{Ind}_{G_{P_{G_\beta}}}^{P_{G_\beta}}$ for every $N \in \mathbb{N}$ and every almost-composition $\beta$ of $N$ (by the definition of $\text{ind}_N^N$), and $\text{ind}_{i,j}^N = \text{Ind}_{G_{P_{G_{\beta_i}}}}^{P_{G_{\beta_i}}} \text{Ind}_{G_{P_{G_{\beta_j}}}}^{P_{G_{\beta_j}}}$ for every $N \in \mathbb{N}$ and every $i, j \in \mathbb{N}$ satisfying $i + j = N$ (by the definition of $\text{ind}_{i,j}^N$).

Since $\text{ind}_{i,j}^N = \text{Ind}_{G_{P_{G_{\beta_i}}}}^{P_{G_{\beta_i}}} \text{Ind}_{G_{P_{G_{\beta_j}}}}^{P_{G_{\beta_j}}}$ for every $N \in \mathbb{N}$ and every almost-composition $\beta$ of $N$, we have

\[
\text{ind}_{m,\alpha_{\ell}}^m \left( \text{ind}_{m,\alpha_{\ell}}^m \left( \mathcal{W} \otimes V_\ell \right) \right) = \text{Ind}_{G_m \times G_{\alpha_{\ell}}}^{G_{m,\alpha_{\ell}}} \left( \text{Ind}_{G_m \times G_{\alpha_{\ell}}}^{G_{m,\alpha_{\ell}}} \left( \mathcal{W} \otimes V_\ell \right) \right) = \text{Ind}_{G_m \times G_{\alpha_{\ell}}}^{G_{m,\alpha_{\ell}}} \left( \text{Ind}_{G_m \times G_{\alpha_{\ell}}}^{G_{m,\alpha_{\ell}}} \left( \mathcal{W} \otimes V_\ell \right) \right)
\]

(11.765)
(since $G_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})} = G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}}$). Therefore,

$$\text{ind}_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}^m W \otimes V_\ell \equiv \text{Ind}_{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}}^{G_{m}} \text{Inf}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}}}^{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}} W \otimes V_\ell \equiv \text{Ind}_{G_{\alpha_{\ell}}} V_\ell$$

(by (11.758) (applied to $G=G_{\alpha_{\ell}}$ and $P=V_\ell$) yields $\text{Ind}_{G_{\alpha_{\ell}}} V_\ell \equiv V_\ell$)

$$\equiv \left( \text{Ind}_{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}}^{G_{m}} \text{Inf}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}}}^{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}} W \otimes V_\ell \right) \equiv \text{Ind}_{G_{\alpha_{\ell}}} V_\ell$$

since Exercise 4.3 (applied to $G_{m}$, $G_{\alpha_{\ell}}$, $P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}$, $G_{\alpha_{\ell}}$, $\text{Inf}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}}}^{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}} W$ and $V_\ell$ instead of $G_1$, $G_2$, $H_1$, $H_2$, $U_1$ and $U_2$) yields

$$\text{Ind}_{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}}^{G_{m} \times G_{\alpha_{\ell}}} \left( \text{Inf}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}}}^{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}} W \otimes V_\ell \right) \equiv \text{Ind}_{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}}^{G_{m} \times G_{\alpha_{\ell}}} \left( \text{Inf}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}}}^{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}} W \otimes V_\ell \right)$$

(by (11.760), applied to $G_1=P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}$, $G_2=G_{\alpha_{\ell}}$, $K_1=K_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}$, $K_2=K_{\alpha_{\ell}}$, $U_1=W$ and $U_2=V_\ell$)

(because $K_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}$ is a normal subgroup of $P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}$ and the quotient is $P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}/K_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})} = G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}}$)

$$\equiv \text{Ind}_{G_{\alpha_{\ell}}} V_\ell$$

(11.766) $\equiv \text{Ind}_{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}}^{G_{m} \times G_{\alpha_{\ell}}} \text{Inf}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}}}^{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}} W \otimes V_\ell$.

Since \( \text{ind}_{n_{i,j}} = \text{Ind}_{P_{i,j}}^{G_{i} \times G_{j}} \) for every $N \in \mathbb{N}$ and every $i, j \in \mathbb{N}$ satisfying $i + j = N$, we have $\text{ind}_{n_{m,\alpha_{\ell}}} = \text{Ind}_{P_{m,\alpha_{\ell}}}^{G_{m} \times G_{\alpha_{\ell}}}$, so that

$$\text{ind}_{m,\alpha_{\ell}}^n \left( \text{ind}_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}^m W \otimes V_\ell \right)$$

$$= \text{Ind}_{P_{m,\alpha_{\ell}}}^{G_{m} \times G_{\alpha_{\ell}}} \left( \text{ind}_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}^m W \otimes V_\ell \right)$$

(by (11.766))

$$\equiv \text{Ind}_{P_{m,\alpha_{\ell}}}^{G_{m} \times G_{\alpha_{\ell}}} \left( \text{Ind}_{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}}^{G_{m} \times G_{\alpha_{\ell}}} \text{Inf}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}}}^{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}} W \otimes V_\ell \right)$$

(11.767) $\equiv \text{Ind}_{P_{m,\alpha_{\ell}}}^{G_{m} \times G_{\alpha_{\ell}}} \text{Inf}_{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}}^{G_{m} \times G_{\alpha_{\ell}}} \text{Inf}_{G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_{\ell-1}}}^{P_{(\alpha_1, \alpha_2, \ldots, \alpha_{\ell-1})}} W \otimes V_\ell$.
Now, we are going to prove that

\[(11.768) \quad \text{Ind}_{P_m,\alpha}^{G_m} \text{Ind}_{P_{(a_1, a_2, \ldots, a_{\ell-1})} \times G_{\alpha}}^{G_m \times G_{\alpha}} Z \cong \text{Ind}_{P_n}^{P_{(a_1, a_2, \ldots, a_{\ell-1})} \times G_{\alpha}} Z \]

for every \(C \left[ P_{(a_1, a_2, \ldots, a_{\ell-1})} \times G_{\alpha} \right] \)-module \(Z\). In order to do so, we try to apply Exercise 4.11 to \(G = P_{m,\alpha}\), \(H = P_{\alpha}\), \(K = K_{m,\alpha}\), and \(V = Z\). This yields (11.768) if we can prove the following two statements:

1. We have \(K_{m,\alpha} < P_{\alpha} < P_{m,\alpha}\) and \(K_{m,\alpha} < P_{m,\alpha}\).
2. The quotient \(P_{m,\alpha}/K_{m,\alpha}\) is canonically identified with \(G_m \times G_{\alpha}\) in such a way that its subgroup \(P_{\alpha}/K_{m,\alpha}\) is precisely \(P_{(a_1, a_2, \ldots, a_{\ell-1})} \times G_{\alpha}\). But this is obvious. Hence, the second statement is proven, and so we are able to apply Exercise 4.11 and therefore obtain (11.768).

Now, (11.767) becomes

\[
\text{Ind}_{m,\alpha}^{n} \left( \text{Ind}_{(a_1, a_2, \ldots, a_{\ell-1})}^{m} W \otimes V \right) = \text{Ind}_{P_n}^{G_n} \left( \text{Ind}_{(a_1, a_2, \ldots, a_{\ell-1})}^{P_m} \text{Ind}_{G_m \times G_{\alpha}}^{G_m} \text{Ind}_{P_{(a_1, a_2, \ldots, a_{\ell-1})} \times G_{\alpha}}^{P_{(a_1, a_2, \ldots, a_{\ell-1})} \times G_{\alpha}} (W \otimes V) \right)
\]

\[
= \text{Ind}_{(a_1, a_2, \ldots, a_{\ell-1})}^{P_n} \left( \text{Ind}_{G_m \times G_{\alpha}}^{G_m} \text{Ind}_{P_{(a_1, a_2, \ldots, a_{\ell-1})} \times G_{\alpha}}^{P_{(a_1, a_2, \ldots, a_{\ell-1})} \times G_{\alpha}} (W \otimes V) \right)
\]

This proves (11.762). Hence, (11.762) is proven in Case 2.

(b) Alternative solution to Exercise 4.20: Let \(G_s\) be the tower \(S_s\) of groups. We can then apply Exercise 4.31(a) to \(\ell = 3, n = i + j + k, \alpha = (i, j, k)\), \(V_1 = U, V_2 = V\) and \(V_3 = W\). As the result, we obtain

\[
\text{Ind}_{i,j,k}^{i+j+k} \left( \text{Ind}_{i,j}^{i+j} (U \otimes V) \right) \cong \text{Ind}_{i,j,k}^{i+j+k} (U \otimes V \otimes W) \cong \text{Ind}_{i,j,k}^{i+j+k} \left( U \otimes \text{Ind}_{j,k}^{j+k} (V \otimes W) \right).
\]
Thus, Exercise 4.20 is solved again. Hence, we have solved Exercise 4.31(b).

(c) Let \( \Sigma = \bigsqcup_{n \geq 0} \text{Irr}(G_n) \). We have to prove that the map \( m \) is associative. In other words, we have to prove that \( m(\beta \otimes \gamma) = m(\alpha \otimes m(\beta \otimes \gamma)) \) for any three elements \( \alpha, \beta, \gamma \) of \( A \). In order to do so, it is clearly enough to show that \( m(\alpha \otimes m(\beta \otimes \gamma)) \) for any three elements \( \alpha, \beta, \gamma \) of \( \Sigma \) is a \( \mathbb{Z} \)-module basis of \( A \), and the equality \( m(\alpha \otimes m(\beta \otimes \gamma)) = m(\alpha \otimes m(\beta \otimes \gamma)) \) is \( \mathbb{Z} \)-linear in each of \( \alpha, \beta, \gamma \). So let \( \alpha, \beta, \gamma \) be three elements of \( \Sigma \). Then, there exists \( i, j, k \in \mathbb{N} \) and \( s, t \in \mathbb{N} \) satisfying \( \alpha \in \text{Irr}(G_i) \), \( \beta \in \text{Irr}(G_j) \) and \( \gamma \in \text{Irr}(G_k) \) (since \( \alpha, \beta, \gamma \) belong to \( \Sigma = \bigsqcup_{n \geq 0} \text{Irr}(G_n) \)). Consider these \( i, j, k \).

There exists an irreducible \( CG_i \)-module \( U \) satisfying \( \alpha = \chi_U \). Similarly, there exists an irreducible \( CG_j \)-module \( V \) satisfying \( \beta = \chi_V \), and an irreducible \( CG_k \)-module \( W \) satisfying \( \gamma = \chi_W \). Consider these \( U, V, W \).

We can apply Exercise 4.31(a) to \( \ell = 3 \), \( n = i + j + k \), \( \alpha = (i, j, k) \), \( U = V = W = V_3 = W \). As the result, we obtain

\[
\text{Ind}_{i+j+k}^{i+j+k} \left( \text{Ind}_{i+j}^{i+j+k} (U \otimes V) \otimes W \right) \cong \text{Ind}_{i+j+k}^{i+j+k} (U \otimes V \otimes W).
\]

Thus,

\[
\text{Ind}_{i+j,k}^{i+j+k} \left( \text{Ind}_{i+j}^{i+j+k} (U \otimes V) \otimes W \right) \cong \text{Ind}_{i+j+k}^{i+j+k} \left( U \otimes \text{Ind}_{i+j,k}^{i+j+k} (V \otimes W) \right).
\]

Since isomorphic representations have equal characters, this yields

\[
(11.769) \quad \chi_{\text{Ind}_{i+j,k}^{i+j+k} (U \otimes V) \otimes W} = \chi_{\text{Ind}_{i+j,k}^{i+j+k} (U \otimes V) \otimes W}.
\]

Since \( \alpha = \chi_U \) and \( \beta = \chi_V \), we have

\[
m(\alpha \otimes \beta) = m(\alpha \otimes \beta) = \chi_{\text{Ind}_{i+j}^{i+j+k} (U \otimes V) \otimes W} \quad \text{(since } U \text{ is a } CG_i \text{-module and } V \text{ is a } CG_j \text{-module)}
\]

\[
= \chi_{\text{Ind}_{i+j}^{i+j+k} (U \otimes V) \otimes W} \quad \text{(since } i+j = i+j \text{)}.
\]

Thus,

\[
m \left( \frac{m(\alpha \otimes \beta) \otimes \gamma}{\chi_{\text{Ind}_{i+j}^{i+j+k} (U \otimes V) \otimes W}} \right) = \chi_{\text{Ind}_{i+j,k}^{i+j+k} (U \otimes V) \otimes W} \quad \text{(by (11.769))}
\]

Similarly, \( (11.771) \) becomes

\[
m(\alpha \otimes m(\beta \otimes \gamma)) = \chi_{\text{Ind}_{i+j,k}^{i+j+k} (U \otimes V) \otimes W} \quad \text{(by (11.769))}
\]

Now, \( (11.770) \) becomes

\[
m(\alpha \otimes \beta) \otimes \gamma = \chi_{\text{Ind}_{i+j,k}^{i+j+k} (U \otimes V) \otimes W} \quad \text{(by (11.769))}
\]

This is what we wanted to prove. Thus, Exercise 4.31(c) is solved.
Let $\Sigma = \bigsqcup_{n \geq 0} \text{Irr } (G_n)$.

We will solve Exercise 4.31(d) by induction over $\ell$. The induction base (the case $\ell = 0$) is trivial, so we come to the induction step. Fix a positive integer $\ell$. We assume that Exercise 4.31(d) is already solved for $\ell - 1$ instead of $\ell$. We now need to solve Exercise 4.31(d) for our $\ell$.

Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ be an almost-composition of an $n \in \mathbb{N}$. Let $\chi_i \in R(G_{\alpha_i})$ for every $i \in \{1, 2, \ldots, \ell\}$. We will show that

$$\chi_1 \chi_2 \cdots \chi_{\ell - 1} = \text{ind}_n^m \left( \chi_1 \otimes \chi_2 \otimes \cdots \otimes \chi_{\ell - 1} \right).$$

Since this equality is $\mathbb{Z}$-linear in each of $\chi_1, \chi_2, \ldots, \chi_\ell$, we can WLOG assume that $\chi_1, \chi_2, \ldots, \chi_\ell$ all lie in $\Sigma$ (since $\Sigma$ is a basis of $A$). Assume this. Then, $\chi_i \in \Sigma \cap R(G_{\alpha_i}) = \text{Irr } (G_{\alpha_i})$ for every $i \in \{1, 2, \ldots, \ell\}$. Hence, for every $i \in \{1, 2, \ldots, \ell\}$, there exists an irreducible $CG_{\alpha_i}$-module $V_i$ such that $\chi_i = \chi_i V_i$. Consider this $V_i$.

Let $m = \alpha_1 + \alpha_2 + \cdots + \alpha_\ell$.

By the induction hypothesis, we can apply Exercise 4.31(d) to the almost-composition $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ of $m$. As a result, we obtain

$$\chi_1 \chi_2 \cdots \chi_{\ell - 1} = \text{ind}_n^m \left( \chi_1 \otimes \chi_2 \otimes \cdots \otimes \chi_{\ell - 1} \right) \left( \chi_{\ell - 1} \right).$$

$$= \chi \text{ind}_n^m \left( \chi_{\ell - 1} \right) \left( \chi_{\ell - 1} \right).$$

Now,

$$\chi_1 \chi_2 \cdots \chi_{\ell} \cdots \chi_{\ell - 1} \chi_{\ell} = \chi \text{ind}_n^m \left( \chi_{\ell - 1} \right) \left( \chi_{\ell - 1} \right).$$

We assume that Exercise 4.31(d) is already solved for $\ell - 1$. Thus, $\chi_{\ell - 1}$ is an irreducible $CG_{\alpha_\ell}$-module.

This completes the induction step, and so Exercise 4.31(d) is solved.

(c) Let $n \in \mathbb{N}$, $\ell \in \mathbb{N}$, and $\chi \in R(G_n)$. We need to prove that $\Delta^{(\ell - 1)} \chi = \sum \text{res}_n^m \chi$. Since this equality is $\mathbb{Z}$-linear in $\chi$, we can WLOG assume that $\chi \in \text{Irr } (G_n)$ (since $\text{Irr } (G_n)$ is a $\mathbb{Z}$-module basis of $R(G_n)$). Assume this. Then, $\chi = \chi_P$. Thus, Exercise 4.31(d) is solved for our $\ell$.

Similarly to how we showed (4.18), we can prove that

$$(11.772) \quad \text{Hom}_{CG_n} (\text{ind}_n^m U, V) \cong \text{Hom}_{CG_n} (U, \text{res}_n^m V)$$

for every almost-composition $\alpha$ of $n$, every $CG_{\alpha}$-module $U$ and every $CG_n$-module $V$. Thus,

$$(11.773) \quad (\text{ind}_n^m \varphi, \psi)_{R(G_n)} = (\varphi, \text{res}_n^m \psi)_{R(G_n)}$$
for every almost-composition \( \alpha \) of \( n \), every \( \varphi \in R(G_\alpha) \) and every \( \psi \in R(G_\alpha) \).

The bilinear form \((\cdot,\cdot)_A\) on \( A \) induces a bilinear form \((\cdot,\cdot)_{A^\otimes \ell}\) on \( A^\otimes \ell \). It is easy to see that the maps \( m_{\ell-1} \) and \( \Delta_{\ell-1} \) are adjoint with respect to the forms \((\cdot,\cdot)_A\) and \((\cdot,\cdot)_{A^\otimes \ell}\). Hence, any \( \varphi \in A \) and \( \rho \in A^\otimes \ell \) satisfy

\[
(11.774) \quad (\Delta_{\ell-1} \varphi, \rho)_{A^\otimes \ell} = (\varphi, m_{\ell-1} \rho)_A.
\]

Since \( A = \bigoplus_{n \geq 0} R(G_n) \), we have

\[
A^\otimes \ell = \bigoplus_{n \geq 0} R(G_n) \otimes \ell = \bigoplus_{n_1, n_2, \ldots, n_\ell \geq 0} R(G_{n_1}) \otimes R(G_{n_2}) \otimes \cdots \otimes R(G_{n_\ell}) = R(G_{n_1} \times G_{n_2} \times \cdots \times G_{n_\ell})
\]

\[
(11.775) \quad = \bigoplus_{\alpha=\langle a_1, a_2, \ldots, a_\ell \rangle} R \left( G_{a_1} \times G_{a_2} \times \cdots \times G_{a_\ell} \right) = \bigoplus_{\alpha \text{ is an almost-composition of length } \ell} R(G_\alpha).
\]

This direct sum decomposition of \( A^\otimes \ell \) is orthogonal with respect to the bilinear form \((\cdot,\cdot)_{A^\otimes \ell}\); that is, if \( \alpha \) and \( \beta \) are two distinct almost-compositions of length \( \ell \), then

\[
(11.776) \quad (R(G_\alpha), R(G_\beta))_{A^\otimes \ell} = 0.
\]

Moreover, for every almost-composition \( \alpha \) of length \( \ell \), we have

\[
(11.777) \quad (\varphi, \psi)_{A^\otimes \ell} = (\varphi, \psi)_{R(G_\alpha)} \text{ for every } \varphi \in R(G_\alpha) \text{ and } \psi \in R(G_\alpha).
\]

\[\text{Proof of (11.777): Let } \alpha \text{ be an almost-composition of } n. \text{ Let } \varphi \in R(G_\alpha) \text{ and } \psi \in R(G_\alpha). \text{ We need to prove the equality (11.777). Since this equality is } \mathbb{Z}\text{-linear in each of } \varphi \text{ and } \psi, \text{ we can WLOG assume that } \varphi \in \text{Irr}(G_\alpha) \text{ and } \psi \in \text{Irr}(G_\alpha) \text{ (since } \text{Irr}(G_\alpha) \text{ and } \text{Irr}(G_\alpha) \text{ are } \mathbb{Z}\text{-module bases of } R(G_\alpha) \text{ and } R(G_\alpha), \text{ respectively). Assume this. Thus, there exist an irreducible } \mathbb{C}G_\alpha\text{-module } U \text{ and an irreducible } \mathbb{C}G_\alpha\text{-module } V \text{ such that } \varphi = \chi_U \text{ and } \psi = \chi_V. \text{ Consider these } U \text{ and } V. \text{ We have}
\]

\[
\left( \text{ind}_n^\alpha \varphi, \psi \right)_{R(G_\alpha)} = \left( \text{ind}_n^\alpha \chi_U, \chi_V \right)_{R(G_\alpha)} = \left( \chi_{\text{ind}_n^\alpha U}, \chi_V \right)_{R(G_\alpha)} = \dim_{\mathbb{C}} \text{Hom}_{\mathbb{C}G_\alpha}(\text{ind}_n^\alpha U, V) = \dim_{\mathbb{C}} \text{Hom}_{\mathbb{C}G_\alpha}(U, \text{res}_n^\alpha V) \quad \text{by (11.772)}.
\]

Compared with

\[
\left( \varphi, \text{res}_n^\alpha \psi \right)_{R(G_\alpha)} = \left( \chi_U, \text{res}_n^\alpha \chi_V \right)_{R(G_\alpha)} = \left( \chi_U, \chi_{\text{res}_n^\alpha V} \right)_{R(G_\alpha)} = \dim_{\mathbb{C}} \text{Hom}_{\mathbb{C}G_\alpha}(U, \text{res}_n^\alpha V),
\]

this yields \( (\text{ind}_n^\alpha \varphi, \psi)_{R(G_\alpha)} = (\varphi, \text{res}_n^\alpha \psi)_{R(G_\alpha)} \). This proves (11.777).

\[\text{Indeed, this follows from the inductive definitions of } m_{\ell-1} \text{ and } \Delta_{\ell-1}, \text{ if one recalls that the maps } m \text{ and } \Delta \text{ are adjoint with respect to the forms } (\cdot,\cdot)_A \text{ and } (\cdot,\cdot)_{A^\otimes \ell}. \text{ This follows from the definition of } (\cdot,\cdot)_{A^\otimes \ell} \text{ and the fact that the direct sum decomposition } A = \bigoplus_{n \geq 0} R(G_n) \text{ is orthogonal with respect to the bilinear form } (\cdot,\cdot)_A.
\]

\[\text{Proof of (11.777): Let } \alpha \text{ be an almost-composition of length } \ell. \text{ Let } \varphi \in R(G_\alpha) \text{ and } \psi \in R(G_\alpha). \text{ We need to prove the equality (11.777). Since this equality is } \mathbb{Z}\text{-linear in each of } \varphi \text{ and } \psi, \text{ we can WLOG assume that } \varphi \in \text{Irr}(G_\alpha) \text{ and } \psi \in \text{Irr}(G_\alpha) \text{ (since } \text{Irr}(G_\alpha) \text{ is a } \mathbb{Z}\text{-module basis of } R(G_\alpha) \text{). Assume this. Thus, there exist an irreducible } \mathbb{C}G_\alpha\text{-module } V \text{ and an irreducible } \mathbb{C}G_\alpha\text{-module } W \text{ such that } \varphi = \chi_V \text{ and } \psi = \chi_W. \text{ Consider these } V \text{ and } W.
\]

Write the almost-composition \( \alpha \) in the form \( (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \). Since \( V \) is an irreducible representation of \( G_\alpha = G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_\ell} \), we can write \( V \) in the form \( V = V_1 \otimes V_2 \otimes \cdots \otimes V_\ell \), where each \( V_i \) is an irreducible representation of \( G_{\alpha_i} \). Similarly, we can write \( W \) in the form \( W = W_1 \otimes W_2 \otimes \cdots \otimes W_\ell \), where each \( W_i \) is an irreducible representation of \( G_{\alpha_i} \). Consider these \( V_i \) and \( W_i \).
We need to show that $\Delta^{(\ell - 1)} \chi = \sum \text{res}_n^\chi$ (where the sum ranges over all almost-compositions $\alpha$ of $n$ having length $\ell$). In order to do so, it is clearly enough to prove that

$$
(11.778) \quad \left( \Delta^{(\ell - 1)} \chi, \rho \right)_{A^\otimes \ell} = \left( \sum \text{res}_n^\chi, \rho \right)_{A^\otimes \ell}
$$

for every $\rho \in A^\otimes \ell$.

(since the bilinear form $(\cdot, \cdot)_{A^\otimes \ell}$ is nondegenerate). So, let $\rho \in A^\otimes \ell$. It remains to prove (11.778).

The equality (11.778) is $\mathbb{Z}$-linear in $\rho$. Since $\bigcup_\alpha$ is an almost-composition $\text{Irr}(G_\alpha)$ is a $\mathbb{Z}$-module basis of $A^\otimes \ell$ of length $\ell$ (this follows from (11.775) and the fact that each $R(G_\alpha)$ has $\mathbb{Z}$-module basis $\text{Irr}(G_\alpha)$), we can therefore WLOG assume that $\rho \in \bigcup_\alpha$ is an almost-composition $\text{Irr}(G_\alpha)$. Assume this. Then, there exists an almost-composition $\beta$ of length $\ell$ such that $\rho \in \text{Irr}(G_\beta)$. Consider this $\beta$, and notice that $\rho \in \text{Irr}(G_\beta) \subset R(G_\beta)$.

For every almost-composition $\alpha$ of $n$ having length $\ell$ satisfying $\alpha \neq \beta$, we have

$$
(11.779) \quad (\text{res}_n^\chi, \rho)_{A^\otimes \ell} = 0
$$

Now, there exists a $\mathbb{C}$-vector space isomorphism

$$
\text{Hom}_{CG_\alpha} (V_1, W_1) \otimes \text{Hom}_{CG_\alpha} (V_2, W_2) \otimes \cdots \otimes \text{Hom}_{CG_\alpha} (V_\ell, W_\ell) \\
\rightarrow \text{Hom}_{\mathbb{C}[G_{\alpha_1} \times G_{\alpha_2} \times \cdots \times G_{\alpha_\ell}]} (V_1 \otimes V_2 \otimes \cdots \otimes V_\ell, W_1 \otimes W_2 \otimes \cdots \otimes W_\ell).
$$

(In fact, when $\ell = 2$, the existence of such an isomorphism follows from Exercise 4.9(a); otherwise it follows by induction over $\ell$ using Exercise 4.9(a).) The existence of this isomorphism yields

$$
dim_{\mathbb{C}} \left( \text{Hom}_{CG_\alpha} (V_1, W_1) \otimes \text{Hom}_{CG_\alpha} (V_2, W_2) \otimes \cdots \otimes \text{Hom}_{CG_\alpha} (V_\ell, W_\ell) \right) \\
= \prod_{i=1}^\ell \dim_{\mathbb{C}} \text{Hom}_{CG_\alpha} (V_i, W_i) = \prod_{i=1}^\ell \left( (V_i, W_i)_{A} = (x_{V_i}, x_{W_i})_{R(G_\alpha)} \right)
$$

Thus,

$$
(\varphi, \psi)_{A^\otimes \ell} = \text{dim}_{\mathbb{C}} \left( \text{Hom}_{CG_\alpha} (V_1, W_1) \otimes \text{Hom}_{CG_\alpha} (V_2, W_2) \otimes \cdots \otimes \text{Hom}_{CG_\alpha} (V_\ell, W_\ell) \right)
$$

This proves (11.777).
Thus, (11.778) is proven. This completes our solution of Exercise 4.31(e).

(11.782)

But let us write the almost-composition \( G \) of \( G \). Consider these \( \psi \) yields \((\text{ind}_n^\beta \rho, \chi)_{R(G_n)} = (\rho, \text{res}_n^\beta \chi)_{R(G_{\beta})}\).

(11.780)

Thus, (11.778) is proven. This completes our solution of Exercise 4.31(e).

11.87. Solution to Exercise 4.34. Definition to Exercise 4.34. Define the Iverson bracket notation as in Exercise 4.34(a).

(a) Let \( G \) be a finite group. For every subset \( P \) of \( G \), the definition of \( \text{1}_P \) yields

(11.781)

\( \text{1}_P (g) = [g \in P] \) for every \( g \in G \).

Also, for every subset \( P \) of \( G \), we have

(11.782)

\( |P| = \sum_{k \in G} [k \in P] \).

Let \( h \in G \) and \( g \in G \). Then, (11.782) (applied to \( P = Z_G (h) \)) yields

\[
|Z_G (h)| = \sum_{k \in G} \begin{cases} k \in Z_G (h) \\ \text{this is equivalent to } khk^{-1} = h \\ \text{by the definition of } Z_G (h) \end{cases} = \sum_{k \in G} [khk^{-1} = h].
\]

608. Since

\[
\left( \sum_{\alpha \in \text{Irr}(G_n)} \text{res}_n^\alpha \chi, \rho \right)_{A^{\otimes \ell}} \in \left( R(G_n), R(G_{\beta}) \right)_{A^{\otimes \ell}} = 0 \quad \text{(by (11.776))}
\]
Applying both sides of the identity \( \alpha_{G,h} = |Z_G(h)| \frac{1}{|\text{Conj}_G(h)|} \) to \( g \), we obtain

\[
\alpha_{G,h}(g) = \frac{|Z_G(h)| \frac{1}{|\text{Conj}_G(h)|}}{\sum_{k \in G} [khk^{-1} = h]} = \left( \sum_{k \in G} [khk^{-1} = h] \right) [g \in \text{Conj}_G(h)].
\]

We need to prove that \( \alpha_{G,h}(g) = \sum_{k \in G} [khk^{-1} = g] \). We must be in one of the following two cases:

**Case 1:** We have \( g \in \text{Conj}_G(h) \).

**Case 2:** We don't have \( g \in \text{Conj}_G(h) \).

Let us first consider Case 1. In this case, we have \( g = pgh^{-1} \). Consider this \( p \). Since \( g = pgh^{-1} \), we have \( h = p^{-1}gp \). Hence, \((11.783)\) becomes

\[
\alpha_{G,h}(g) = \left( \sum_{k \in G} [khk^{-1} = h] \right) [g \in \text{Conj}_G(h)] = \sum_{k \in G} \left[ \frac{g}{p^{-1}gp} k^{-1} h = 1 \right] = \sum_{k \in G} \left[ (pk^{-1})^{-1} g (pk^{-1})^{-1} = h \right]
\]

\[
= \sum_{k \in G} \left[ k^{-1} g k = h \right]
\]

This is equivalent to \( k^{-1} g k = h \). Therefore, (here, we have substituted \( k \) for \( pk^{-1} \) in the sum)

\[
= \sum_{k \in G} [khk^{-1} = g].
\]

Thus, \( \alpha_{G,h}(g) = \sum_{k \in G} [khk^{-1} = g] \) is proven in Case 1.

Let us now consider Case 2. In this case, we don't have \( g \in \text{Conj}_G(h) \). In other words, \( g \) is not in the conjugacy class of \( h \). In other words, \( g \) is not conjugate to \( h \). In other words, there exists no \( k \in G \) satisfying \( khk^{-1} = g \). In other words, for every \( k \in G \), we do not have \( khk^{-1} = g \). Hence, for every \( k \in G \), we have \( [khk^{-1} = g] = 0 \). Thus, \( \sum_{k \in G} [khk^{-1} = g] = \sum_{k \in G} 0 = 0 \). Comparing this with

\[
\alpha_{G,h}(g) = \sum_{k \in G} [khk^{-1} = h] = \sum_{g \in \text{Conj}_G(h)} 0 = 0,
\]

we obtain \( \alpha_{G,h}(g) = \sum_{k \in G} [khk^{-1} = g] \). Thus, \( \alpha_{G,h}(g) = \sum_{k \in G} [khk^{-1} = g] \) is proven in Case 2.

Thus, \( \alpha_{G,h}(g) = \sum_{k \in G} [khk^{-1} = g] \) is proven in both Cases 1 and 2. Hence, Exercise 4.34(a) is solved.

(b) Let \( H \) be a subgroup of a finite group \( G \). Let \( h \in H \). Let \( g \in G \). Then, the definition of \( \text{Ind}_H^G \alpha_{H,h} \) given in Exercise 4.1 yields

\[
\left( \text{Ind}_H^G \alpha_{H,h} \right)(g) = \frac{1}{|H|} \sum_{k \in G} \alpha_{H,h}(kg^{-1}) = \frac{1}{|H|} \sum_{p \in G; pgp^{-1} \in H} \alpha_{H,h}(pgp^{-1}) = \sum_{k \in H} [khk^{-1} = pgp^{-1}]
\]

(here, we renamed the summation index \( k \) as \( p \))

\[
= \frac{1}{|H|} \sum_{p \in G; pgp^{-1} \in H} \sum_{k \in H} [khk^{-1} = pgp^{-1}] = \frac{1}{|H|} \sum_{p \in G; pgp^{-1} \in H} \sum_{k \in H} [khk^{-1} = pgp^{-1}] = \sum_{k \in H} \sum_{p \in G; pgp^{-1} \in H} [khk^{-1} = pgp^{-1}] - (by \text{Exercise } 4.34(a), \text{applied to } H \text{ and } pgp^{-1} \text{ instead of } G \text{ and } g)
\]

\[
(11.784)\]
But every $k \in H$ satisfies

$$\sum_{p \in G; \ pgp^{-1} \in H} [khk^{-1} = pgp^{-1}] = \sum_{p \in G} [khk^{-1} = pgp^{-1}]$$

(11.785)

Now, (11.784) becomes

$$\left( \text{Ind}_H^G \alpha_{H,h} \right)(g) = \frac{1}{|H|} \sum_{k \in H} \sum_{p \in G; \ pgp^{-1} \in H} [khk^{-1} = pgp^{-1}] = \frac{1}{|H|} \sum_{k \in H} \sum_{p \in G} [khk^{-1} = pgp^{-1}]$$

(= \sum_{p \in G} [khk^{-1} = pgp^{-1}] \text{ (by (11.785))})

$$= \frac{1}{|H|} \sum_{k \in H} \sum_{p \in G} [p^{-1}kh \overset{k^{-1}p}{\rightarrow} = g] = \frac{1}{|H|} \sum_{k \in H} \sum_{p \in G} [p^{-1}kh (p^{-1}k)^{-1} = g]$$

(= |H| \sum_{p \in G} [php^{-1} = g])

$$= \frac{1}{|H|} |H| \sum_{k \in G} [php^{-1} = g] = \sum_{p \in G} [php^{-1} = g] = \sum_{k \in G} [khk^{-1} = g]$$

(here, we have substituted $k$ for $p$ in the sum)

$$= \alpha_{G,h} \ (g) \quad \text{ (by Exercise 4.34(a)).}$$

Let us now forget that we fixed $g$. We thus have proven that \( \left( \text{Ind}_H^G \alpha_{H,h} \right)(g) = \alpha_{G,h} \ (g) \) for every $g \in G$. Hence, \( \text{Ind}_H^G \alpha_{H,h} = \alpha_{G,h} \). This solves Exercise 4.34(b).

(c) Let $G_1$ and $G_2$ be two finite groups. Let $h_1 \in G_1$ and $h_2 \in G_2$. Let $I$ denote the canonical isomorphism $R_C(G_1) \otimes R_C(G_2) \rightarrow R_C(G_1 \times G_2)$. We need to prove that

$$\text{Ind}^G_{G_1 \times G_2} \alpha_{G_1 \times G_2, (h_1, h_2)} = \alpha_{G_1 \times G_2, (h_1, h_2)}. \quad (11.786)$$

Let $g \in G_1 \times G_2$. Let us write $g$ in the form $g = (g_1, g_2)$. Then, Exercise 4.34(a) (applied to $G_1$, $h_1$ and $g_1$ instead of $G$, $h$ and $g$) yields

$$\alpha_{G_1, h_1} \ (g_1) = \sum_{k \in G_1} [kh_1k^{-1} = g_1] = \sum_{k \in G_1} [k_1h_1k_1^{-1} = g_1]$$

(here, we have renamed the summation index $k$ as $k_1$). Similarly,

$$\alpha_{G_2, h_2} \ (g_2) = \sum_{k_2 \in G_2} [k_2h_2k_2^{-1} = g_2].$$

Proof of (11.785): Let $k \in H$. Then,

$$\sum_{p \in G} [khk^{-1} = pgp^{-1}] = \sum_{p \in G; \ pgp^{-1} \in H} [khk^{-1} = pgp^{-1}] + \sum_{p \in G; \ \ pgp^{-1} \notin H} [khk^{-1} = pgp^{-1}]$$

(since we don’t have $khk^{-1} = pgp^{-1}$ (since $khk^{-1} \in H$ (since $k \in H$ and $h \in H$) and $pgp^{-1} \notin H$))

$$= \sum_{p \in G; \ \ pgp^{-1} \in H} [khk^{-1} = pgp^{-1}] + \sum_{p \in G; \ \ pgp^{-1} \notin H} [khk^{-1} = pgp^{-1}]$$

This proves (11.785).
Now,

\[
(I(\alpha_{G_1,h_1} \otimes \alpha_{G_2,h_2}))(g) = \sum_{(k_1,k_2) \in G_1 \times G_2} k_1 h_1 k_2^{-1} = g_1 \text{ and } k_2 h_2 k_2^{-1} = g_2
\]

\[
= \sum_{(k_1,k_2) \in G_1 \times G_2} \left[ k_1 h_1 k_1^{-1} = g_1 \text{ and } k_2 h_2 k_2^{-1} = g_2 \right] = \sum_{(k_1,k_2) \in G_1 \times G_2} \left[ k_1 h_1 k_1^{-1} = g_1 \right] \left[ k_2 h_2 k_2^{-1} = g_2 \right]
\]

\[
= \sum_{(k_1,k_2) \in G_1 \times G_2} \left( k_1 h_1 (h_1, h_2) (k_1, k_2)^{-1} = g \right) = \sum_{k \in G_1 \times G_2} (k (h_1, h_2) k^{-1} = g)
\]

(here, we have renamed the summation index \((k_1,k_2)\) as \(k\)). Compared with

\[
\alpha_{G_1,G_2,(h_1,h_2)}(g) = \sum_{k \in G_1 \times G_2} [k (h_1, h_2) k^{-1} = g]
\]

(by Exercise 4.34(a), applied to \(G_1 \times G_2\) and \((h_1,h_2)\) instead of \(G\) and \(h\)),

this yields \((I(\alpha_{G_1,h_1} \otimes \alpha_{G_2,h_2}))(g) = \alpha_{G_1 \times G_2,(h_1,h_2)}(g)\).

Now, let us forget that we fixed \(g\). We thus have proven that \((I(\alpha_{G_1,h_1} \otimes \alpha_{G_2,h_2}))(g) = \alpha_{G_1 \times G_2,(h_1,h_2)}(g)\)

for every \(g \in G_1 \times G_2\). In other words, \((I(\alpha_{G_1,h_1} \otimes \alpha_{G_2,h_2})) = \alpha_{G_1 \times G_2,(h_1,h_2)}\). This proves (11.786). Exercise

4.34(c) is thus solved.

(d) For every partition \(\lambda\), let us define \(\overline{z}_\lambda\) as the size of the centralizer of a permutation in \(S_\lambda\) having
cycle type \(\lambda\). Note that this does not depend on the choice of said permutation, since all permutations in \(S_\lambda\) having
cycle type \(\lambda\) are mutually conjugate (and thus their centralizers are of the same size). It is well-known that \(\overline{z}_\lambda = z_\lambda\) (see Remark 2.44), but we shall avoid using this fact, as we can obtain an alternative proof of
it from the following argument.

We shall use the same notations as in the proof of Theorem 4.32. In particular, \(A_\mathbb{C} = \bigoplus_{n \geq 0} R_\mathbb{C}(S_n)\); this \(\mathbb{C}\)-vector space \(A_\mathbb{C}\) becomes a \(\mathbb{C}\)-algebra as explained in the proof of Theorem 4.32. Its multiplication is

given by \(\text{ind}_{n+m}^{\mathbb{C}}\); more precisely, every \(n \in \mathbb{N}\) and \(m \in \mathbb{N}\), and every \(\beta \in R_\mathbb{C}(S_n)\) and \(\gamma \in R_\mathbb{C}(S_m)\) satisfy

\[
(11.787) \quad \beta \gamma = \text{ind}_{n,m}^{\mathbb{C}}(\beta \otimes \gamma) = \text{Ind}_{S_n \times S_m}^{\mathbb{C}}(\beta \otimes \gamma).
\]

We define a \(\mathbb{C}\)-linear map \(\Phi : \Lambda_\mathbb{C} \rightarrow A_\mathbb{C}\) by setting

\[
\Phi(p_\lambda) = \overline{z}_\lambda 1_\lambda \quad \text{for every } \lambda \in \text{Par}.
\]

(This is well-defined, since \((p_\lambda)_{\lambda \in \text{Par}}\) is a basis of the \(\mathbb{C}\)-vector space \(\Lambda_\mathbb{C}\).) We notice that if \(\lambda\) is a partition
of a nonnegative integer \(n\), and if \(g \in S_n\) is a permutation having cycle type \(\lambda\), then

\[
\Phi(p_\lambda) = \alpha_{S_n,g}.
\]
Also, if \( \lambda \) is a partition of a nonnegative integer \( n \), then
\[
(11.789) \quad |\{h \in S_n \mid h \text{ has cycle type } \lambda\}| = n! / \tilde{z}_\lambda.
\]

We shall now prove that \( \Phi \) is a \( C \)-algebra homomorphism. Since \( \Phi (1) = 1 \) is true, we only need to verify that \( \Phi (uv) = \Phi (u) \Phi (v) \) for any \( u \in \Lambda_C \) and \( v \in \Lambda_C \). Let us prove this now. Fix \( u \in \Lambda_C \) and \( v \in \Lambda_C \). Since the equality \( \Phi (uv) = \Phi (u) \Phi (v) \) is \( C \)-linear in each of \( u \) and \( v \), we can WLOG assume that \( u \) and \( v \) are elements of the basis \( \{ p_\lambda \}_{\lambda \in \mathcal{P}_n} \) of the \( C \)-vector space \( \Lambda_C \). Assume this, and set \( u = p_\mu \) and \( v = p_\nu \) for two partitions \( \mu \) and \( \nu \). Let \( n = |\mu| \) and \( m = |\nu| \). Let \( g \) be a permutation in \( S_n \) having cycle type \( \mu \). (Such a \( g \) clearly exists.) Let \( h \) be a permutation in \( S_n \) having cycle type \( \nu \). (Such an \( h \) clearly exists.) Applying \( (11.788) \) to \( \mu \) instead of \( \lambda \), we obtain \( \Phi (p_\mu) = \alpha_{S_n,g} \). Thus,
\[
(11.790) \quad \Phi \left( \frac{u}{p_\mu} \right) = \Phi (p_\mu) = \alpha_{S_n,g}.
\]

---

**610 Proof of (11.788):** Let \( \lambda \) be a partition of a nonnegative integer \( n \). Let \( g \in S_n \) be a permutation having cycle type \( \lambda \). Then, \( g \) is a permutation in \( S_{|\lambda|} \) having cycle type \( \lambda \). Thus, \( \tilde{z}_\lambda \) is the size of the centralizer of \( g \) (by the definition of \( \tilde{z}_\lambda \)). In other words, \( \tilde{z}_\lambda = |Z_{S_n}(g)| \).

But any two permutations in \( S_n \) having the same cycle type are mutually conjugate. Hence, if \( h \) is any permutation in \( S_n \) having cycle type \( \lambda \), then \( h \) and \( g \) are conjugate (since \( h \) and \( g \) are permutations in \( S_n \) having the same cycle type). Conversely, if \( h \) is a permutation in \( S_n \) such that \( h \) and \( g \) are conjugate, then \( h \) has cycle type \( \lambda \) (because \( h \) and \( g \) are conjugate permutations and thus have the same cycle type, but the cycle type of \( g \) is \( \lambda \)). Combining these two statements, we conclude that if \( h \) is a permutation in \( S_n \), then \( h \) has cycle type \( \lambda \) if and only if \( h \) and \( g \) are conjugate. Hence,
\[
\left\{ \begin{array}{l}
  h \in S_n \mid h \text{ has cycle type } \lambda \\
  \text{this is equivalent to } \\
  \langle h \text{ and } g \text{ are conjugate} \rangle
\end{array} \right\} = \left\{ \begin{array}{l}
  h \in S_n \mid h \text{ and } g \text{ are conjugate} \\
  \text{this is equivalent to } h \in Conj_{S_n}(g)
\end{array} \right\} = \{ h \in S_n \mid h \in Conj_{S_n}(g) \} = Conj_{S_n}(g).
\]

On the other hand, \( \frac{1}{\tilde{z}_\lambda} \) is defined as the characteristic function for the \( S_n \)-conjugacy class of permutations of cycle type \( \lambda \). In other words, \( \frac{1}{\tilde{z}_\lambda} \) is the indicator function of the subset \( \{ h \in S_n \mid h \text{ has cycle type } \lambda \} \) of \( S_n \). In other words, \( \frac{1}{\tilde{z}_\lambda} \) is the indicator function of the subset \( Conj_{S_n}(g) \) (since \( \{ h \in S_n : h \text{ has cycle type } \lambda \} = Conj_{S_n}(g) \)). In other words, \( \frac{1}{\tilde{z}_\lambda} = \frac{1}{|Z_{S_n}(g)|} \).

Now, \( \Phi (p_\lambda) = \frac{\tilde{z}_\lambda}{|Z_{S_n}(g)|} \frac{1}{\tilde{z}_\lambda} = |Z_{S_n}(g)| \frac{1}{|Conj_{S_n}(g)|} \). Compared with \( \alpha_{S_n,g} = |Z_{S_n}(g)| \frac{1}{|Conj_{S_n}(g)|} \) (by the definition of \( \alpha_{S_n,g} \)), this yields \( \Phi (p_\lambda) = \alpha_{S_n,g} \). This proves (11.788).

---

**611 Proof of (11.789):** Let \( \lambda \) be a partition of a nonnegative integer \( n \). Fix a permutation \( g \in S_n \) having cycle type \( \lambda \). (Such a \( g \) clearly exists.) In the proof of (11.788), we have seen that \( \tilde{z}_\lambda = |Z_{S_n}(g)| \) and that \( \{ h \in S_n : h \text{ has cycle type } \lambda \} = Conj_{S_n}(g) \).

Now, it is well-known that for every finite group \( G \) and every element \( f \in G \), we have \( |G/Z_G(f)| = |Conj_G(f)| \) (in fact, there is a canonical bijection from the \( G \)-set \( G/Z_G(f) \) to \( Conj_G(f) \), which sends the equivalence class \( [\gamma] \in G/Z_G(f) \) of every \( \gamma \in G \) to \( \gamma f \gamma^{-1} \in Conj_G(f) \)). Applying this to \( G = S_n \) and \( f = g \), we obtain \( |S_n/Z_{S_n}(g)| = |Conj_{S_n}(g)| \). Thus, \( |Conj_{S_n}(g)| = |S_n/Z_{S_n}(g)| = |S_n| / |Z_{S_n}(g)| = n! / \tilde{z}_\lambda \), so that
\[
n! / \tilde{z}_\lambda = |Conj_{S_n}(g)| = |\{ h \in S_n \mid h \text{ has cycle type } \lambda \}|.
\]

This proves (11.789).

---

**612 This is because**
\[
\Phi \left( \frac{1}{p_\mu} \right) = \Phi (p_\mu) = \frac{\tilde{z}_\lambda}{|Z_{S_n}(g)|} \frac{1}{\tilde{z}_\lambda} = 1 \quad \text{(by the definition of } \Phi (p_\mu)\text{)}
\]

\[= 1.
\]
Applying (11.788) to $\nu$, $m$ and $h$ instead of $\lambda$, $n$ and $g$, we obtain $\Phi (p_\nu) = \alpha_{\mathfrak{g},m,h}$. Thus,

$$
(11.791) \quad \Phi \left( \frac{\nu}{=} = p_\nu \right) = \Phi (p_\nu) = \alpha_{\mathfrak{g},m,h}.
$$

The canonical isomorphism $R_C (\mathfrak{g}_n) \otimes R_C (\mathfrak{g}_m) \rightarrow R_C (\mathfrak{g}_n \times \mathfrak{g}_m)$ sends $\alpha_{\mathfrak{g},n,g} \otimes \alpha_{\mathfrak{g},m,h}$ to $\alpha_{\mathfrak{g},n \times \mathfrak{g}_m, (g,h)}$ (according to Exercise 4.34(c), applied to $G_1 = \mathfrak{g}_n$, $G_2 = \mathfrak{g}_m$, $h_1 = g$ and $h_2 = h$). Let us identify $R_C (\mathfrak{g}_n) \otimes R_C (\mathfrak{g}_m)$ with $R_C (\mathfrak{g}_n \times \mathfrak{g}_m)$ along this isomorphism. Then, the statement we just made rewrites as follows:

$$
(11.792) \quad \alpha_{\mathfrak{g},n,g} \otimes \alpha_{\mathfrak{g},m,h} = \alpha_{\mathfrak{g},n \times \mathfrak{g}_m, (g,h)}.
$$

Now, multiplying the equalities (11.790) and (11.791), we obtain

$$
\Phi (u) \Phi (v) = \alpha_{\mathfrak{g},n,g} \alpha_{\mathfrak{g},m,h}
= \text{Ind}_{\mathfrak{g}_n \times \mathfrak{g}_m}^{\mathfrak{g}_{n+m}} \left( \alpha_{\mathfrak{g},n,g} \otimes \alpha_{\mathfrak{g},m,h} \right)
= \alpha_{\mathfrak{g}_{n+m}, (g,h)} \quad \text{(by (11.787), applied to $\beta = \alpha_{\mathfrak{g},n,g}$ and $\gamma = \alpha_{\mathfrak{g},m,h}$)}
= \text{Ind}_{\mathfrak{g}_n \times \mathfrak{g}_m}^{\mathfrak{g}_{n+m}} \alpha_{\mathfrak{g},n \times \mathfrak{g}_m, (g,h)} = \alpha_{\mathfrak{g}_{n+m}, (g,h)}
$$

(by Exercise 4.34(b), applied to $\mathfrak{g}_{n+m}$, $\mathfrak{g}_n \times \mathfrak{g}_m$ and $(g,h)$ instead of $G, H$ and $h$).

Let $\lambda$ be the partition whose parts are $\mu_1, \mu_2, \ldots, \mu_{t(\mu)}$, $\nu_1, \nu_2, \ldots, \nu_{t(\nu)}$. Then, $(\lambda_1, \lambda_2, \ldots, \lambda_{t(\lambda)})$ is a permutation of the list $(\mu_1, \mu_2, \ldots, \mu_{t(\mu)}, \nu_1, \nu_2, \ldots, \nu_{t(\nu)})$, and the partition $\lambda$ has size

$$
|\lambda| = \mu_1 + \mu_2 + \ldots + \mu_{t(\mu)} + \nu_1 + \nu_2 + \ldots + \nu_{t(\nu)} = n + m.
$$

We also have $p_\lambda = p_\mu p_\nu$. $^6$ Since $p_\mu = u$ and $p_\nu = v$, this rewrites as $p_\lambda = uv$.

Recall that the permutation $g$ has cycle type $\mu$. In other words, the cycles of $g$ have lengths $\mu_1, \mu_2, \ldots, \mu_{t(\mu)}$. Similarly, the cycles of $h$ have lengths $\nu_1, \nu_2, \ldots, \nu_{t(\nu)}$. Hence, the cycles of the permutation $(g,h) \in \mathfrak{g}_{n+m}$ (which, as we recall, sends every $i \in \{1, 2, \ldots, n + m\}$ to $g (i)$, if $i \leq n$; $n + h (i - n)$, if $i > n$) have lengths $\mu_1, \mu_2, \ldots, \mu_{t(\mu)}, \nu_1, \nu_2, \ldots, \nu_{t(\nu)}$ (in fact, on the subset $\{1, 2, \ldots, n\}$ of $\{1, 2, \ldots, n + m\}$, the permutation $(g,h)$ acts as $g$ and thus has cycles of lengths $\mu_1, \mu_2, \ldots, \mu_{t(\mu)}$, whereas on the complementary subset $\{n + 1, n + 2, \ldots, n + m\}$ of $\{1, 2, \ldots, n + m\}$, the permutation $(g,h)$ acts as a (shifted version of) $h$ and thus has cycles of lengths $\nu_1, \nu_2, \ldots, \nu_{t(\nu)}$). In other words, the cycles of the permutation $(g,h) \in \mathfrak{g}_{n+m}$ have lengths $\lambda_1, \lambda_2, \ldots, \lambda_{t(\lambda)}$ (since $(\lambda_1, \lambda_2, \ldots, \lambda_{t(\lambda)})$ is a permutation of the list $(\mu_1, \mu_2, \ldots, \mu_{t(\mu)}, \nu_1, \nu_2, \ldots, \nu_{t(\nu)})$). In other words, the permutation $(g,h) \in \mathfrak{g}_{n+m}$ has cycle type $\lambda$. Thus, (11.788) (applied to $n + m$ and $(g,h)$ instead of $n$ and $g$) yields $\Phi (p_\lambda) = \alpha_{\mathfrak{g}_{n+m}, (g,h)}$. Compared with (11.793), this yields $\Phi (u) \Phi (v) = \Phi (p_\lambda)$. Since $p_\lambda = uv$, this rewrites as $\Phi (u) \Phi (v) = \Phi (uv)$. Thus, $\Phi (uv) = \Phi (u) \Phi (v)$ is proven. As explained, this completes the proof of the fact that $\Phi$ is a $\mathbb{C}$-algebra homomorphism.

But we know that $\text{ch} : A \rightarrow \Lambda$ is a $\mathbb{Z}$-Hopf algebra isomorphism, and thus the extension of $\text{ch}$ to a $\mathbb{C}$-linear map $A_C \rightarrow \Lambda_C$ is a $\mathbb{C}$-Hopf algebra isomorphism. We shall denote this extension by $\text{ch}_C$.

$^6$Proof. By the definition of $p_\mu$, we have $p_\mu = p_{\mu_1} p_{\mu_2} \cdots p_{\mu_{t(\mu)}}$. Similarly, $p_\nu = p_{\nu_1} p_{\nu_2} \cdots p_{\nu_{t(\nu)}}$. Multiplying these two equalities, we obtain $p_\mu p_\nu = \left( p_{\mu_1} p_{\mu_2} \cdots p_{\mu_{t(\mu)}} \right) \left( p_{\nu_1} p_{\nu_2} \cdots p_{\nu_{t(\nu)}} \right)$.

But the definition of $p_\lambda$ yields $p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{t(\lambda)}}$. The product $p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{t(\lambda)}}$ has the same factors as the product $p_{\mu_1} p_{\mu_2} \cdots p_{\mu_{t(\mu)}} p_{\nu_1} p_{\nu_2} \cdots p_{\nu_{t(\nu)}}$, but possibly in a different order (since $(\lambda_1, \lambda_2, \ldots, \lambda_{t(\lambda)})$ is a permutation of the list $(\mu_1, \mu_2, \ldots, \mu_{t(\mu)}, \nu_1, \nu_2, \ldots, \nu_{t(\nu)})$). Hence, $p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{t(\lambda)}} = p_{\mu_1} p_{\mu_2} \cdots p_{\mu_{t(\mu)}} p_{\nu_1} p_{\nu_2} \cdots p_{\nu_{t(\nu)}},$ so that

$$
p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{t(\lambda)}} = p_{\mu_1} p_{\mu_2} \cdots p_{\mu_{t(\mu)}} p_{\nu_1} p_{\nu_2} \cdots p_{\nu_{t(\nu)}} = \left( p_{\mu_1} p_{\mu_2} \cdots p_{\mu_{t(\mu)}} \right) \left( p_{\nu_1} p_{\nu_2} \cdots p_{\nu_{t(\nu)}} \right) = p_\mu p_\nu,
$$

qed.
We now shall show that \( \text{ch}_C \circ \Phi = \text{id}_{\Lambda_C} \). Indeed, let \( n \) be a positive integer. Then, \( \tilde{z}_{(n)} = n \) \(^{614}\). Now, the definition of \( \Phi \left( p_{(n)} \right) \) yields \( \Phi \left( p_{(n)} \right) = \tilde{z}_{(n)} \frac{1}{n} = n \frac{1}{n} \), so that

\[
(ch_C \circ \Phi) (p_n) = ch_C \left( \Phi \left( \frac{p_n}{n} \right) \right) = \frac{1}{n} \left( \Phi \left( p_{n_1} \right) \right) = \frac{1}{n} \left( n \frac{1}{n} \right) = n \frac{1}{n} = p_n = \text{id}_{\Lambda_C} (p_n).
\]

Now, let us forget that we fixed \( n \). We thus have shown that \( (ch_C \circ \Phi) (p_n) = \text{id}_{\Lambda_C} (p_n) \) for every positive integer \( n \). Thus, the \( \mathbb{C} \)-algebra homomorphisms \( ch_C \circ \Phi \) and \( \text{id}_{\Lambda_C} \) are equal to each other on \( p_n \) for every positive integer \( n \). Since \( (p_n)_{n \geq 1} \) is a generating set of the \( \mathbb{C} \)-algebra \( \Lambda_C \), this yields that these homomorphisms are equal to each other on a generating set of the \( \mathbb{C} \)-algebra \( \Lambda_C \). Hence, these homomorphisms must be identical. That is, we have \( ch_C \circ \Phi = \text{id}_{\Lambda_C} \).

Since \( ch_C \) is an isomorphism, this yields that \( \Phi \) is the inverse of \( ch_C \). That is, \( \Phi = (ch_C)^{-1} \).

Corollary 2.45(b) (applied to \( k = \mathbb{C} \)) yields that \( \{ p_\lambda \} \) and \( \{ z_\lambda^{-1} p_\lambda \} \) are dual bases of \( \Lambda_C \) with respect to the Hall inner product on \( \Lambda \). Thus,

\[
(p_\lambda, z_\mu^{-1} p_\mu)_{\Lambda_C} = \delta_{\lambda, \mu} \quad \text{for any partitions } \lambda \text{ and } \mu.
\]

Hence, every partition \( \lambda \) satisfies

\[
(11.795) \quad (p_\lambda, z_\lambda^{-1} p_\lambda)_{\Lambda_C} = z_\lambda \delta_{\lambda, \lambda} = z_\lambda.
\]

Our goal is to prove that \( ch_C \left( \frac{1}{n} \right) = \frac{p_\lambda}{z_\lambda} \) for every partition \( \lambda \).

\(^{614}\text{Proof. Let } g \text{ denote the } n \text{-cycle } (1,2, \ldots, n) \text{ in } \mathfrak{S}_n. \text{ Then, } g \text{ is a permutation in } \mathfrak{S}_n \text{ having cycle type } (n). \text{ But } \tilde{z}_{(n)} \text{ is defined as the size of the centralizer of a permutation in } \mathfrak{S}_{(n)} \text{ having cycle type } (n). \text{ Hence, } \tilde{z}_{(n)} \text{ is the size of the centralizer of } g \text{ in } \mathfrak{S}_n \text{ (since } g \text{ is a permutation in } \mathfrak{S}_n = \mathfrak{S}_{(n)} \text{ having cycle type } (n)). \text{ In other words, } \tilde{z}_{(n)} = |Z_{\mathfrak{S}_n}(g)|. \text{ It is clear that the subgroup } (g) \text{ of } \mathfrak{S}_n \text{ generated by } g \text{ satisfies } (g) \subseteq Z_{\mathfrak{S}_n}(g) \text{ (since every power of } g \text{ centralizes } g). \text{ We shall now show that } (g) = Z_{\mathfrak{S}_n}(g). \text{ Indeed, let } z \in Z_{\mathfrak{S}_n}(g). \text{ Then, } z \text{ must centralize } g. \text{ That is, we have } z g z^{-1} = g. \text{ Hence, } z g = g z, \text{ so that the elements } g \text{ and } z \text{ of } \mathfrak{S}_n \text{ commute. Hence, these elements } z \text{ and } g \text{ generate a commutative subgroup of } \mathfrak{S}_n. \text{ Denote this subgroup by } T. \text{ Now, every } i \in \{1, 2, \ldots, n\} \text{ satisfies }
\]

\[
(11.794) \quad g^{i-1} (1) = i
\]

(since \( g \) is the \( n \)-cycle \( (1,2, \ldots, n) \)). Now, let \( j = z (1) \). Then, \( g^{j-1} (1) = j \) (by \( 11.794 \), applied to \( i = j \)). On the other hand, let \( i \in \{1, 2, \ldots, n\} \). Then, \( g^{j-1} \) commutes with \( z \) (since \( g \) commutes with \( z \)); in other words, \( g^{j-1} z = z g^{j-1} \). Now,

\[
= g^{j-1} (1) = (z g^{j-1} (1)) (1) = g^{j-1} (z (1)) (1) = g^{j-1} (z (1)) = (g^{j-1} g^{j-1} (1)) (1) = g^{j-1} (g^{j-1} (1)) = g^{j-1} (1).
\]

Let us now forget that we fixed \( i \). We thus have shown that \( z (i) = g^{j-1} (i) \) for every \( i \in \{1, 2, \ldots, n\} \). Hence, \( z = g^{j-1} \in (g) \).

Now, let us forget that we fixed \( z \). We thus have proven that \( z \in (g) \) for every \( z \in Z_{\mathfrak{S}_n}(g) \). Hence, \( Z_{\mathfrak{S}_n}(g) \subseteq (g) \). Combined with \( (g) \subseteq Z_{\mathfrak{S}_n}(g) \), this yields \( (g) = Z_{\mathfrak{S}_n}(g) \). Hence, \( |g| = |Z_{\mathfrak{S}_n}(g)| \). Compared with \( \tilde{z}_{(n)} = |Z_{\mathfrak{S}_n}(g)| \), this yields \( \tilde{z}_{(n)} = |g| = (\text{the order of } g \text{ in } \mathfrak{S}_n) = n \) (since \( g \) is an \( n \)-cycle), qed.)
Let \( \lambda \) be a partition. Let \( n \) be the size of \( \lambda \). Then, \( \lambda \in \text{Par}_n \subset \text{Par} \) and \( \mathbf{1}_\lambda \in R_C(\mathfrak{S}_n) \). We know that \( \bar{z}_\lambda \) is a positive integer (since \( \bar{z}_\lambda \) is defined as the size of a centralizer, and centralizers are subgroups). Hence, we can divide by \( \bar{z}_\lambda \), and we have
\[
\Phi \left( \bar{z}_\lambda^{-1} p_\lambda \right) = \bar{z}_\lambda^{-1} \Phi (p_\lambda) = \bar{z}_\lambda^{-1} \bar{z}_\lambda \mathbf{1}_\lambda = \mathbf{1}_\lambda.
\]
(by the definition of \( \Phi(p_\lambda) \))
Since \( \Phi = (\text{ch})^{-1} \), this rewrites as \( (\text{ch})^{-1} \left( \bar{z}_\lambda^{-1} p_\lambda \right) = \mathbf{1}_\lambda \). Hence, \( \text{ch}_C \left( \mathbf{1}_\lambda \right) = \bar{z}_\lambda^{-1} p_\lambda \).

Recall that we want to prove that \( \text{ch}_C \left( \mathbf{1}_\lambda \right) = \frac{p_\lambda}{z_\lambda} \). If we can show that \( \bar{z}_\lambda = z_\lambda \), then (11.797) becomes
\[
\text{ch}_C \left( \mathbf{1}_\lambda \right) = \frac{p_\lambda}{z_\lambda}.
\]
Since \( \Phi = (\text{ch})^{-1} \), this concludes the proof of \( \text{ch}_C \left( \mathbf{1}_\lambda \right) = \frac{p_\lambda}{z_\lambda} \), and thus Exercise 4.34(d) is solved.

Thus, \( \bar{z}_\lambda = z_\lambda \) is proven. As we have explained, this concludes the proof of \( \text{ch}_C \left( \mathbf{1}_\lambda \right) = \frac{p_\lambda}{z_\lambda} \), and thus Exercise 4.34(d) is solved.
(e) Let \( \lambda \) be a partition. Let us work with the notations of the solution of Exercise 4.34(d) above. In the latter solution, we have shown that \( \tilde{z}_\lambda = z_\lambda \). Thus,

\[
\tilde{z}_\lambda = z_\lambda = \text{(the size of the centralizer of a permutation in } \mathfrak{S}_{|\lambda|} \text{ having cycle type } \lambda) \\
\text{(by the definition of } \tilde{z}_\lambda) \\
= \text{(the size of the centralizer in } \mathfrak{S}_n \text{ of a permutation having cycle type } \lambda, \text{ where } n = |\lambda|). 
\]

This proves Remark 2.44. Thus, Exercise 4.34(e) is solved.

(f) Let \( G \) and \( H \) be two finite groups. Let \( \rho : H \to G \) be a group homomorphism. Let \( y \in H \). We shall show that \( \text{Ind}_\rho \alpha_{H,y} = \alpha_{G,\rho(y)} \).

Let \( u \in G \). Then, the definition of \( \text{Ind}_\rho \alpha_{H,y} \) yields

\[
(\text{Ind}_\rho \alpha_{H,y})(u) = \frac{1}{|H|} \sum_{(h,k) \in H \times G; \ k \rho(h) k^{-1} = u} \alpha_{H,y}(h) = \frac{1}{|H|} \sum_{(h,x) \in H \times G; \ x \rho(h)x^{-1} = u} \alpha_{H,y}(h) \\
= \sum_{k \in H} \left[ k y k^{-1} = h \right] \quad \text{(by Exercise 4.34(a), applied to } H, y \text{ and } h \text{ instead of } G, h \text{ and } g) \\
= \sum_{x \in G} \sum_{h \in H; \ x \rho(h)x^{-1} = u} \sum_{k \in H} \left[ k y k^{-1} = h \right] \\
= \sum_{k \in H} \sum_{x \in G} \sum_{h \in H; \ x \rho(h)x^{-1} = u} \left[ k y k^{-1} = h \right]. 
\]

But every \( x \in G \) and \( k \in H \) satisfy

\[
\sum_{h \in H; \ x \rho(h)x^{-1} = u} \left[ k y k^{-1} = h \right] = \left[ x \rho(k) \rho(y) (x \rho(k))^{-1} = u \right] 
\]
615. Hence, (11.798) becomes

\[
(\text{Ind}_\rho \circ H, y)^{(u)} = \frac{1}{|H|} \sum_{x \in G} \sum_{h \in H} \sum_{h \in H; x \rho (h) x^{-1} = u} [ky^{-1} = h]
= \frac{1}{|H|} \sum_{x \in G} \sum_{h \in H; x \rho (h) x^{-1} = u} [x \rho (k \rho (y) (x \rho (k))^{-1})^{-1} = u]
\]

(by (11.799))

\[
= \frac{1}{|H|} \sum_{x \in G} \sum_{h \in H; x \rho (h) x^{-1} = u} [x \rho (y) x^{-1} = u] = \frac{1}{|H|} \sum_{x \in G} \sum_{h \in H; x \rho (h) x^{-1} = u} [x \rho (y) x^{-1} = u]
\]

(there, we have divided \(x\) for \(x \rho (k)\) in the sum
(because the map \(G \to G\), \(x \to x \rho (k)\) is a bijection
(since \(G\) is a group, and since \(\rho (k) \in G\)))

\[
= \frac{1}{|H|} \sum_{x \in G} [x \rho (y) x^{-1} = u] = \sum_{x \in G} [x \rho (y) x^{-1} = u]
= \sum_{k \in G} [k \rho (y) k^{-1} = u]
\]

(here, we renamed the summation index \(x\) as \(k\)).

Compared with

\[
\alpha_{G, \rho (y)}^{(u)} = \sum_{k \in G} [k \rho (y) k^{-1} = u]
\]

(by Exercise 4.34(a), applied to \(\rho (y)\) and \(u\) instead of \(h\) and \(g\)).

\[\text{Proof of (11.799): Let } x \in G \text{ and } k \in H. \text{ We have}
\]

\[
\sum_{h \in H; x \rho (h) x^{-1} = u} [x \rho (h) x^{-1} = u] [k^{-1} = h]
= \sum_{h \in H; x \rho (h) x^{-1} = u} [x \rho (h) x^{-1} = u] [k^{-1} = h] + \sum_{h \in H; x \rho (h) x^{-1} = u} [x \rho (h) x^{-1} = u] [k^{-1} = h]
\]

(since \(x \rho (h) x^{-1} = u\))

\[
= \sum_{h \in H; x \rho (h) x^{-1} = u} [k^{-1} = h] + \sum_{h \in H; x \rho (h) x^{-1} = u} [0 [k^{-1} = h] = \sum_{h \in H; x \rho (h) x^{-1} = u} [k^{-1} = h],
\]

so that

\[
\sum_{h \in H; x \rho (h) x^{-1} = u} [k^{-1} = h] = \sum_{h \in H} [x \rho (h) x^{-1} = u] [k^{-1} = h]
\]

\[
= \sum_{h \in H; h = k^{-1}} [x \rho (h) x^{-1} = u] [k^{-1} = h] + \sum_{h \in H; h = k^{-1}} [0 [x \rho (h) x^{-1} = u] [k^{-1} = h]
\]

(since \(k^{-1} = h\) and \(h \neq k^{-1}\))

\[
= \sum_{h \in H; h = k^{-1}} [x \rho (h) x^{-1} = u] + \sum_{h \in H; h = k^{-1}} [0 [x \rho (h) x^{-1} = u] [k^{-1} = h]
\]

\[
= \sum_{h \in H; h = k^{-1}} [x \rho (h) x^{-1} = u] = \sum_{h \in H; h = k^{-1}} [x \rho (h) x^{-1} = u]
\]

(since \(k^{-1} \in H\) (since \(k \in H\) and \(y \in H\))

\[
= \sum_{h \in H; h = k^{-1}} [x \rho (h) y x^{-1} = u] = \sum_{h \in H; h = k^{-1}} [x \rho (h) y x^{-1} = u]
\]

(since \(k^{-1} \in H\) (since \(k \in H\) and \(y \in H\))

\[
= \sum_{h \in H; h = k^{-1}} [x \rho (h) (\rho (k))^{-1} x^{-1} = u] = \sum_{h \in H; h = k^{-1}} [x \rho (h) (\rho (k))^{-1} x^{-1} = u]
\]

This proves (11.799).
this yields \((\text{Ind}_p \alpha_{H,y})(u) = \alpha_{G,\rho(y)}(u)\).

Let us now forget that we fixed \(u\). We thus have shown that \((\text{Ind}_p \alpha_{H,y})(u) = \alpha_{G,\rho(y)}(u)\) for every \(u \in G\). In other words, \(\text{Ind}_p \alpha_{H,y} = \alpha_{G,\rho(y)}\).

Let us now forget that we fixed \(y\). We thus have shown that \(\text{Ind}_p \alpha_{H,y} = \alpha_{G,\rho(y)}\) for every \(y \in H\). Renaming \(y\) as \(h\) in this statement, we obtain that \(\text{Ind}_p \alpha_{H,h} = \alpha_{G,\rho(h)}\) for every \(h \in H\). This solves Exercise 4.34(f).

11.88. **Solution to Exercise 4.35. Solution to Exercise 4.35.**

**Step 1: Study of inner tensor products.**

The well-definedness of the inner tensor product is clear (since the inclusion map \(G \to G \times G, \ g \mapsto (g,g)\) is a group homomorphism). We notice that if \(G\) is a group and \(U_1\) and \(U_2\) are two \(CG\)-modules, then the character \(\chi_{U_1 \otimes U_2}\) of the inner tensor product \(U_1 \otimes U_2\) of \(U_1\) and \(U_2\) is given by

\[
(11.800) \quad \chi_{U_1 \otimes U_2}(g) = \chi_{U_1}(g) \chi_{U_2}(g) \quad \text{for all } g \in G.
\]

**Step 2: The involutions \(\overline{\omega}_n\).**

Now let \(n \geq 0\). For every \(f \in R_C(\mathfrak{S}_n)\), it is easy to see that the map \(\mathfrak{S}_n \to C\) which sends every \(g \in \mathfrak{S}_n\) to \(\text{sgn}(g) f(g)\) is a class function (because both \(\text{sgn}(g)\) and \(f(g)\) are uniquely determined by the conjugacy class of \(g\)). This class function \(\mathfrak{S}_n \to C\) is denoted by \(\text{sgn}_{\mathfrak{S}_n} f\) and belongs to \(R_C(\mathfrak{S}_n)\) (being a class function). We can thus define a map \(\overline{\omega}_n : R_C(\mathfrak{S}_n) \to R_C(\mathfrak{S}_n)\) as follows:

\[
\overline{\omega}_n (f) = \text{sgn}_{\mathfrak{S}_n} f \quad \text{for all } f \in R_C(\mathfrak{S}_n).
\]

Consider this map \(\overline{\omega}_n\). It is \(C\)-linear (obviously) and an involution. Hence, \(\overline{\omega}_n\) is precisely the involution on class functions \(f : \mathfrak{S}_n \to C\) sending \(f \mapsto \text{sgn}_{\mathfrak{S}_n} f\).

Now, let \(V\) be any finite-dimensional \(C\mathfrak{S}_n\)-module. Then, every \(g \in \mathfrak{S}_n\) satisfies

\[
\overline{\omega}_n(\chi_V)(g) = (\text{sgn}_{\mathfrak{S}_n} \chi_V)(g) = \text{sgn}(g) \chi_V(g)
\]

(by the definition of \(\overline{\omega}_n\))

and

\[
\chi_{\text{sgn}_{\mathfrak{S}_n} \otimes V}(g) = \chi_{\text{sgn}_{\mathfrak{S}_n}}(g) \chi_V(g) \quad \text{(by (11.800))}
\]

\[
= \text{sgn}(g) \chi_V(g).
\]

Hence, every \(g \in \mathfrak{S}_n\) satisfies \((\overline{\omega}_n(\chi_V))(g) = \text{sgn}(g) \chi_V(g) = \chi_{\text{sgn}_{\mathfrak{S}_n} \otimes V}(g)\). In other words, \(\overline{\omega}_n(\chi_V) = \chi_{\text{sgn}_{\mathfrak{S}_n} \otimes V}\).

Forget that we fixed \(V\). We thus have proven that \(\overline{\omega}_n(\chi_V) = \chi_{\text{sgn}_{\mathfrak{S}_n} \otimes V}\) for every finite-dimensional \(C\mathfrak{S}_n\)-module \(V\). In particular, for every irreducible \(C\mathfrak{S}_n\)-module \(V\), we have \(\overline{\omega}_n(\chi_V) = \chi_{\text{sgn}_{\mathfrak{S}_n} \otimes V} \in R(\mathfrak{S}_n)\). Since the \(\chi_V\) span \(R(\mathfrak{S}_n)\) as a \(Z\)-module as \(V\) ranges through (a set of representatives of the isomorphism classes of) the irreducible \(C\mathfrak{S}_n\)-modules \(V\), this entails that the involution \(\overline{\omega}_n\) preserves the \(Z\)-lattice \(R(\mathfrak{S}_n)\). Since \(\overline{\omega}_n\) is the involution on class functions \(f : \mathfrak{S}_n \to C\) sending \(f \mapsto \text{sgn}_{\mathfrak{S}_n} f\), this rewrites as follows:

---

**616 This is because every \(g \in G\) satisfies**

\[
\chi_{U_1 \otimes U_2}(g) = \text{trace}(U_1 \otimes U_2 \to U_1 \otimes U_2, (g : U_1 \otimes U_2 \to U_1 \otimes U_2))
\]

(by the definition of the inner tensor product)

\[
= \text{trace}(U_1 \otimes U_2 \to U_1 \otimes U_2, (g : U_1 \otimes U_2 \to U_1 \otimes U_2))
\]

\[
= \text{trace}(U_1 \to U_1) \cdot \text{trace}(U_2 \to U_2) = \chi_{U_1}(g) \cdot \chi_{U_2}(g).
\]

**617 This is because multiplying a scalar by \(\text{sgn}(g)\) twice (for fixed \(g \in \mathfrak{S}_n\)) does nothing (since \((\text{sgn}(g))^2 = 1\)).**
The involution on class functions \( f : \mathcal{S}_n \to \mathbb{C} \) sending \( f \mapsto \text{sgn}_{\mathcal{S}_n} * f \) preserves the \( \mathbb{Z} \)-lattice \( R(\mathcal{S}_n) \). This proves one claim of Theorem 4.32(b).

Recall that the involution \( \bar{\omega}_n \) preserves the \( \mathbb{Z} \)-lattice \( R(\mathcal{S}_n) \). Thus, \( \bar{\omega}_n \) restricts to a \( \mathbb{Z} \)-linear involution \( R(\mathcal{S}_n) \to R(\mathcal{S}_n) \). Denote this involution by \( \bar{\omega}_n' \). It clearly satisfies

\[
\bar{\omega}_n' (f) = \bar{\omega}_n (f) \quad \text{(by the definition of } \bar{\omega}_n')
\]
\[
= \text{sgn}_{\mathcal{S}_n} * f \quad \text{for all } f \in R(\mathcal{S}_n).
\]

**Step 3: The involution \( \bar{\omega}_Z \).**

Now, forget that we fixed \( n \). We thus have constructed a \( \mathbb{Z} \)-linear involution \( \bar{\omega}_n' : R(\mathcal{S}_n) \to R(\mathcal{S}_n) \) for every \( n \geq 0 \). The direct sum of these involutions over all \( n \geq 0 \) is a graded \( \mathbb{Z} \)-linear involution \( \bigoplus_{n \geq 0} \bar{\omega}_n' : \bigoplus_{n \geq 0} R(\mathcal{S}_n) \to \bigoplus_{n \geq 0} R(\mathcal{S}_n) \). Denote this involution \( \bigoplus_{n \geq 0} \bar{\omega}_n' \) by \( \bar{\omega}_Z \). Then, \( \bar{\omega}_Z \) is a graded \( \mathbb{Z} \)-linear involution \( A(\mathcal{S}) \to A(\mathcal{S}) \) (since \( \bigoplus_{n \geq 0} R(\mathcal{S}_n) = A(\mathcal{S}) \)), and is precisely the involution on \( A(\mathcal{S}) \) defined in Theorem 4.32(b).

We have \( \bar{\omega}_Z = \bigoplus_{n \geq 0} \bar{\omega}_n' \). Hence, for every \( n \in \mathbb{N} \) and \( f \in R_C(\mathcal{S}_n) \), we have

\[
\bar{\omega}_Z (f) = \bar{\omega}_n (f) \quad \text{(since } \bar{\omega}_n' \text{ is defined as a restriction of } \bar{\omega}_n)
\]
\[
= \text{sgn}_{\mathcal{S}_n} * f.
\]

For every \( n \in \mathbb{N} \) and every finite-dimensional \( \mathbb{C} \mathcal{S}_n \)-module \( V \), we have

\[
\bar{\omega}_Z (\chi_V) = \bar{\omega}_n (\chi_V) \quad \text{(by (11.803), applied to } f = \chi_V)}
\]
\[
= \chi_{\text{sgn}_{\mathcal{S}_n} \otimes V}.
\]

In other words, the involution \( \bar{\omega}_Z \) sends \( \chi_V \) to \( \chi_{\text{sgn}_{\mathcal{S}_n} \otimes V} \) for every \( n \in \mathbb{N} \) and every finite-dimensional \( \mathbb{C} \mathcal{S}_n \)-module \( V \). In other words, the involution on \( A(\mathcal{S}) \) defined in Theorem 4.32(b) sends \( \chi_V \) to \( \chi_{\text{sgn}_{\mathcal{S}_n} \otimes V} \) for every \( n \in \mathbb{N} \) and every finite-dimensional \( \mathbb{C} \mathcal{S}_n \)-module \( V \) (since \( \bar{\omega}_Z \) is precisely the involution on \( A(\mathcal{S}) \) defined in Theorem 4.32(b)). This proves one of the claims of Exercise 4.35.

**Step 4: Properties of \( \bar{\omega}_Z \).**

We are now going to prove that \( \chi \circ \bar{\omega}_Z = \chi \circ \chi \text{ as maps } A(\mathcal{S}) \to \Lambda. \)

Indeed, let \( \lambda \) be a partition. Let \( n = |\lambda| \). Then, \( \lambda \in \text{Par}_n. \) Theorem 4.32(a) yields \( \chi (\text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda}) = h_\lambda \) and \( \chi (\text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \text{sgn}_{\mathcal{S}_\lambda}) = e_\lambda \). Since \( \chi (\text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda}) = h_\lambda \), we have \( \chi^{-1} (h_\lambda) = \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda}. \) Since \( \chi (\text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \text{sgn}_{\mathcal{S}_\lambda}) = e_\lambda \), we have \( \chi^{-1} (e_\lambda) = \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \text{sgn}_{\mathcal{S}_\lambda}. \)

On the other hand, \( \omega (h_\lambda) = e_\lambda \)

But it is easy to see that \( \omega (\text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda}) = \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \text{sgn}_{\mathcal{S}_\lambda}. \)

Now, comparing

\[
\left( \bar{\omega}_Z \circ \chi^{-1} \right) (h_\lambda) = \bar{\omega}_Z \left( \chi^{-1} (h_\lambda) \right) = \bar{\omega}_Z \left( \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda} \right) = \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \text{sgn}_{\mathcal{S}_\lambda},
\]

619 This follows from the fact that \( \omega \) is an algebra homomorphism and that \( \omega (h_m) = e_m \) for every \( m \in \mathbb{N}. \)

619 Proof. Let \( g \in \mathcal{S}_n. \) We have \( \bar{\omega}_Z \left( \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda} \right) = \text{sgn}_{\mathcal{S}_n} * \left( \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda} \right) \) (by (11.803), applied to \( f = \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda} \)). Hence,

\[
\left( \bar{\omega}_Z \left( \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda} \right) \right) (g) = \left( \text{sgn}_{\mathcal{S}_n} * \left( \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda} \right) \right) (g) = \text{sgn} (g) \cdot \left( \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda} \right) (g)
\]

(by the definition of \( \text{sgn}_{\mathcal{S}_n} * \left( \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda} \right) \)). Since (by the definition of \( \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda} \)) we have

\[
\text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda} (g) = \frac{1}{|\mathcal{S}_\lambda|} \sum_{k \in \mathcal{S}_n; \ k g^{-1} \in \mathcal{S}_\lambda} 1 = \frac{1}{|\mathcal{S}_\lambda|} \sum_{k \in \mathcal{S}_n; \ k g^{-1} \in \mathcal{S}_\lambda} 1,
\]

this rewrites as

\[
\left( \bar{\omega}_Z \left( \text{Ind}_{\mathcal{S}_\lambda}^\mathcal{S} \mathbf{1}_{\mathcal{S}_\lambda} \right) \right) (g) = \text{sgn} (g) \cdot \frac{1}{|\mathcal{S}_\lambda|} \sum_{k \in \mathcal{S}_n; \ k g^{-1} \in \mathcal{S}_\lambda} 1.
\]
with
\[(\text{ch}^{-1} \omega) (h_\lambda) = \text{ch}^{-1} \left( \omega \left( h_\lambda \right) \right) = \text{ch}^{-1} \left( e_\lambda \right) = \text{Ind}_{\mathfrak{e}_\lambda}^{\mathfrak{e}_n} \text{sgn}_{\mathfrak{e}_\lambda}, \]
we obtain \((\tilde{\omega}_Z \circ \text{ch}^{-1}) (h_\lambda) = (\text{ch}^{-1} \omega) (h_\lambda)\).

So we have shown that \((\tilde{\omega}_Z \circ \text{ch}^{-1}) (h_\lambda) = (\text{ch}^{-1} \omega) (h_\lambda)\) for every partition \(\lambda\). Thus, \(\tilde{\omega}_Z \circ \text{ch}^{-1} = \text{ch}^{-1} \omega\) (since the \(h_\lambda\) form a \(\mathbb{Z}\)-module basis of \(\Lambda\)). Hence, \(\text{ch}^{-1} \omega = \tilde{\omega}_Z \circ \text{ch}^{-1}\) so that \(\text{ch}^{-1} \omega \circ \text{ch} = \tilde{\omega}_Z\) and \(\omega \circ \text{ch} = \text{ch} \circ \tilde{\omega}_Z\).

We thus have \(\text{ch} \circ \tilde{\omega}_Z = \omega \circ \text{ch}\) as maps \(A(\mathfrak{S}) \rightarrow \Lambda\). In other words, the map \(\tilde{\omega}_Z\) corresponds under \(\text{ch}\) to the involution \(\omega\) on \(\Lambda\). In other words, the involution on \(A(\mathfrak{S})\) defined in Theorem 4.32(b) corresponds under \(\text{ch}\) to the involution \(\omega\) on \(\Lambda\) (since \(\tilde{\omega}_Z\) is precisely the involution on \(A(\mathfrak{S})\) defined in Theorem 4.32(b)). This completes the proof of Theorem 4.32(b).

**Step 5: The structure-preserving properties of \(\tilde{\omega}_Z\).**

Now, it only remains to show that the involution on \(A(\mathfrak{S})\) defined in Theorem 4.32(b) is a nontrivial PSH-automorphism of \(A(\mathfrak{S})\).

Recall that \(\chi\) is a PSH-isomorphism. Hence, its inverse \(\text{ch}^{-1}\) also is a PSH-isomorphism.

Exercise 2.76(c) shows that \(\omega : \Lambda \rightarrow \Lambda\) is a Hopf automorphism. Moreover, the map \(\omega\) is clearly graded. Let \(\Sigma\) denote the PSH-basis \(\{\lambda \in \text{Par} \mid \lambda \in \Lambda\}\). Then, \(\omega\) restricts to a bijection \(\Sigma \rightarrow \Sigma\) \(\mathbb{Z}\)-linearly. As a consequence, \(\omega\) \((\mathbb{N} \Sigma) = \mathbb{N} \omega \) \((\mathbb{N} \Sigma) \subset \mathbb{N} \Sigma\). Hence, \(\omega\) is a PSH-automorphism \(\Lambda \rightarrow \Lambda\) (since \(\omega\) is a graded Hopf algebra morphism), and thus a PSH-isomorphism \(\Lambda \rightarrow \Lambda\) (since \(\omega\) is an isomorphism and restricts to a bijection \(\Sigma \rightarrow \Sigma\)).

Now, \(\text{ch}^{-1} \omega \circ \text{ch}\) is the composition of three PSH-isomorphisms (since \(\text{ch}^{-1}\), \(\omega\) and \(\text{ch}\) are PSH-isomorphisms), therefore a PSH-automorphism itself. In other words, \(\tilde{\omega}_Z\) is a PSH-isomorphism (because \(\tilde{\omega}_Z = \text{ch}^{-1} \omega \circ \text{ch}\)). Thus, \(\tilde{\omega}_Z\) is a PSH-automorphism.

Since \(\omega(h_2) = e_2 \neq h_2\), we have \(\omega \neq \text{id}\) and therefore \(\tilde{\omega}_Z \neq \text{id}\) (since \(\tilde{\omega}_Z = \text{ch}^{-1} \omega \circ \text{ch}\)). In other words, \(\tilde{\omega}_Z\) is nontrivial.

So we know that \(\tilde{\omega}_Z\) is a nontrivial PSH-automorphism of \(A(\mathfrak{S})\). In other words, the involution on \(A(\mathfrak{S})\) defined in Theorem 4.32(b) is a nontrivial PSH-automorphism of \(A(\mathfrak{S})\) (since \(\tilde{\omega}_Z\) is precisely the involution on \(A(\mathfrak{S})\) defined in Theorem 4.32(b)). This completes the solution of Exercise 4.35.

**Remark.** There are some alternative ways to solve parts of this exercise.

For example, in order to prove that \(\tilde{\omega}_Z \circ \text{ch}^{-1} = \text{ch}^{-1} \omega\), we showed that \((\tilde{\omega}_Z \circ \text{ch}^{-1}) (h_\lambda) = (\text{ch}^{-1} \omega) (h_\lambda)\) for every partition \(\lambda\). Instead of doing this, it is possible to prove that \(\tilde{\omega}_Z \circ \text{ch}^{-1} = \text{ch}^{-1} \omega\) by showing that \((\tilde{\omega}_Z \circ \text{ch}^{-1}) (p_\lambda) = (\text{ch}^{-1} \omega) (p_\lambda)\) for every partition \(\lambda\). This, however, requires a little technicality (the \(p_\lambda\) do not span \(\Lambda\) as a \(\mathbb{Z}\)-module, only as a \(\mathbb{Q}\)-module).

We showed that \(\tilde{\omega}_Z \circ \text{ch}^{-1} = \text{ch}^{-1} \omega\) and used this to conclude that \(\tilde{\omega}_Z\) is a PSH-automorphism of \(A(\mathfrak{S})\). An alternative argument proceeds the other way round: The map \(\tilde{\omega}_Z\) is \(\mathbb{Z}\)-linear and graded and is easily seen to be self-adjoint with respect to the inner product on \(A(\mathfrak{S})\). It also is a coalgebra morphism, as

\[
\text{Compared with:} \quad \left( \text{Ind}_{\mathfrak{e}_\lambda}^{\mathfrak{e}_n} \text{sgn}_{\mathfrak{e}_\lambda} \right) (g) = \frac{1}{|\mathfrak{S}_\lambda|} \sum_{gk^{-1} \in \mathfrak{e}_\lambda, k \in \mathfrak{e}_n, k \neq 0} \text{sgn}_{\mathfrak{e}_\lambda}(gk^{-1}) = \text{Ind}_{\mathfrak{e}_\lambda}^{\mathfrak{e}_n} \text{sgn}_{\mathfrak{e}_\lambda} (g).
\]

\[
\text{(by the definition of } \text{Ind}_{\mathfrak{e}_\lambda}^{\mathfrak{e}_n} \text{sgn}_{\mathfrak{e}_\lambda} \text{)}
\]

\[
\frac{1}{|\mathfrak{S}_\lambda|} \sum_{gk^{-1} \in \mathfrak{e}_\lambda, k \in \mathfrak{e}_n, k \neq 0} \text{sgn}(k) \cdot 1 = \text{sgn}(g) \cdot \frac{1}{|\mathfrak{S}_\lambda|} \sum_{gk^{-1} \in \mathfrak{e}_\lambda} \text{sgn}(g) \cdot 1,
\]

this yields \((\tilde{\omega}_Z \left( \text{Ind}_{\mathfrak{e}_\lambda}^{\mathfrak{e}_n} 1_{\mathfrak{e}_\lambda} \right) \right) (g) = \left( \text{Ind}_{\mathfrak{e}_\lambda}^{\mathfrak{e}_n} \text{sgn}_{\mathfrak{e}_\lambda} \right) (g).

Now, forget that we fixed \(g\). We thus have shown that \((\tilde{\omega}_Z \left( \text{Ind}_{\mathfrak{e}_\lambda}^{\mathfrak{e}_n} 1_{\mathfrak{e}_\lambda} \right) \right) (g) = \left( \text{Ind}_{\mathfrak{e}_\lambda}^{\mathfrak{e}_n} \text{sgn}_{\mathfrak{e}_\lambda} \right) (g)\) for every \(g \in \mathfrak{e}_n\). In other words, \(\tilde{\omega}_Z \left( \text{Ind}_{\mathfrak{e}_\lambda}^{\mathfrak{e}_n} 1_{\mathfrak{e}_\lambda} \right) \right) = \text{Ind}_{\mathfrak{e}_\lambda}^{\mathfrak{e}_n} \text{sgn}_{\mathfrak{e}_\lambda}, \text{ qed.}

\(620\) This follows from the fact that \(\omega(s_{\lambda'}) = s_{\lambda'}\) for every partition \(\lambda\) (and this fact, in turn, is a consequence of (2.21) for \(\mu = \varnothing\)).
\[
\text{Res}^{S_n}_{S_{i+j}} \omega_n(f) = (\omega_i \otimes \omega_j) \left( \text{Res}^{S_{i+j}}_{S_i \times S_j} f \right)
\text{for all } n = i + j \text{ and all } f \in R(S_n).\]
Hence, this map \( \tilde{\omega}_Z \) also is an algebra morphism (since it is self-adjoint, and \( A(S) \) is self-dual), hence a bialgebra morphism and thus a Hopf morphism (by Proposition 1.46(c)). It also sends irreducible characters to irreducible characters (by (11.805)), and thus restricts to a bijection \( \{ \chi^A \} \to \{ \chi^B \} \). Hence, \( \tilde{\omega}_Z \) is a PSH-automorphism of \( A(S) \).

Every \( n \geq 0 \) satisfies
\[
(\tilde{\omega}_Z \circ \text{ch}^{-1})(h_n) = \tilde{\omega}_Z \left( \text{ch}^{-1}(h_n) \right) = \tilde{\omega}_Z (1_{\mathbb{S}_n}) = \omega_n (1_{\mathbb{S}_n}) = \text{sgn}_{\mathbb{S}_n},
\]
\[
= \text{ch}^{-1} \left( \frac{e_n}{\omega(h_n)} \right) = (\text{ch}^{-1} \circ \omega)(h_n).
\]

Since \( \tilde{\omega}_Z \circ \text{ch}^{-1} \) and \( \text{ch}^{-1} \circ \omega \) are algebra morphisms whereas the \( h_n \) generate \( \Lambda \), this yields \( \tilde{\omega}_Z \circ \text{ch}^{-1} = \text{ch}^{-1} \circ \omega \).

11.89. Solution to Exercise 4.36. Solution to Exercise 4.36. We need an auxiliary observation first. If \( G \) is a finite group, then a class function \( \chi \in R_C(G) \) of \( G \) will be called integral if every \( g \in G \) satisfies \( \chi(g) \in \mathbb{Z} \). Then, if \( G \) and \( H \) are two groups and if \( \phi \) and \( \psi \) are two integral class functions of \( G \) and \( H \), respectively, then

\[
(11.806) \quad \phi \otimes \psi \text{ is an integral class function of } G \times H.
\]

(This is a consequence of the fact that \( (\phi \otimes \psi)(g, h) = \phi(g)\psi(h) \) for all \( (g, h) \in G \times H \).) Moreover, if \( H \) is a subgroup of a group \( G \) and if \( \chi \) is an integral class function of \( H \), then

\[
(11.807) \quad \text{Ind}_H^G \chi \text{ is an integral class function of } G.
\]

(Proof of (11.807): Let \( J \) be a system of coset representatives for \( H \setminus G \), so that \( G = \bigsqcup_{j \in J} Hj \). Then, Exercise 4.1(a) (applied to \( f = \chi \)) shows that \( \text{Ind}_H^G \chi \) is a class function on \( G \). Also, Exercise 4.1(b) (applied to \( f = \chi \)) yields
\[
\left( \text{Ind}_H^G \chi \right)(g) = \sum_{j \in J} \chi(jgj^{-1}) \quad \in \mathbb{Z} \quad \text{for all } g \in G.
\]

In other words, \( \text{Ind}_H^G \chi \) is an integral class function of \( G \). This proves (11.807).)

Now let \( \bar{A} = A(S) \). For every \( a \in A \), every \( n \geq 0 \) and every \( g \in S_n \), we define \( a(g) \) to mean the value at \( g \) of the \( n \)-th homogeneous component of \( a \). Let \( \bar{A} \) denote the subset of \( A \) formed by all \( a \in A \) such that:

\[
(11.808) \quad \text{for every } n \geq 0 \text{ and every } g \in S_n, \text{ we have } a(g) \in \mathbb{Z}.
\]

It is clear that \( \bar{A} \) is a graded \( \mathbb{Z} \)-submodule of \( A \) and contains \( 1 = 1_S \).

We are now going to show that \( \bar{A} \) is closed under multiplication. In order to do so, it is clearly enough to prove that if \( a \in \bar{A} \) is homogeneous of degree \( n \) and \( b \in \bar{A} \) is homogeneous of degree \( m \), then \( ab \in \bar{A} \). But this is now easy: Since \( a \in A \), we know that \( a \) is an integral class function on \( S_m \). Similarly, \( b \) is an integral class function on \( S_n \). Hence, (11.806) shows that \( a \otimes b \) is an integral class function on \( S_n \times S_m \). Thus, \( \text{Ind}_{S_n \times S_m}^{S_{n+m}} (a \otimes b) \) is an integral class function on \( S_{n+m} \) (by (11.807)). In other words, \( ab \) is an integral class function on \( S_{n+m} \) (since \( ab = \text{Ind}_{S_n \times S_m}^{S_{n+m}} (a \otimes b) = \text{Ind}_{S_n \times S_m}^{S_{n+m}} (a \otimes b) \)). In other words, \( ab \in \bar{A} \). This proves that \( \bar{A} \) is closed under multiplication. Hence, \( \bar{A} \) is a subring of \( A \) (since \( \bar{A} \) is a \( \mathbb{Z} \)-submodule of \( A \) and contains \( 1 = 1_S \)). Thus, \( \langle \bar{A} \rangle \) (where \( \langle \cdot \rangle \) is defined as in Theorem 4.32) is a subring of \( \Lambda \). For every \( n \geq 1 \), we have \( 1_{S_n} \in \bar{A} \) (by the definitions), so that \( \text{ch}(1_{S_n}) \in \text{ch}(\bar{A}) \). Since \( \text{ch}(1_{S_n}) = n_h \), this rewrites as \( h_n \in \text{ch}(\bar{A}) \). Thus, the subring \( \text{ch}(\bar{A}) \) of \( \Lambda \) contains \( h_n \) for all \( n \geq 1 \). Hence, \( \text{ch}(\bar{A}) = \Lambda \) (because the
\(h_n\) generate \(\Lambda\) as a ring). Therefore, \(A = A\) (since \(\text{ch}\) is an isomorphism). If we recall how \(A\) was defined, we thus see that every \(a \in A\) satisfies (11.808).

Now, fix \(n \geq 0\) and \(g \in \mathcal{S}_n\) and a finite-dimensional \(\mathbb{C}\mathcal{S}_n\)-module \(V\). Then, \(\chi_V \in R(\mathcal{S}_n) \subset A\). Thus, (11.808) (which holds for every \(a \in A\), as we now know) yields that \(\chi_V(g) \in \mathbb{Z}\), which solves part (a) of the problem.

Note that a different way to solve part (a) would be by showing that \(\Lambda\) is contained in the \(\mathbb{Z}\)-submodule of \(\Lambda\) spanned by the \(p\lambda\) for all partitions \(\lambda\). This yields (by taking preimages under \(\text{ch}\)) that \(A\) is contained in the \(\mathbb{Z}\)-submodule of \(R(\mathcal{S})\) generated by the \(1_\lambda\). This way, we wouldn’t have to show that \(A\) is closed under multiplication; instead, we would obtain \(A = A\) by noticing that all \(1_\lambda\) are integral class functions.

(b) This can be done using the Noether-Deuring theorem (in fact, it is easy to show that there are two \(\mathbb{C}\mathcal{S}_n\)-modules \(U\) and \(U'\) defined over \(\mathbb{Q}\) satisfying \(U' \oplus V \cong U\), and then the Noether-Deuring theorem allows to “pull back” \(V\) to a \(\mathbb{Q}\mathcal{S}_n\)-module as well, showing that \(V\) is also defined over \(\mathbb{Q}\)). We are omitting this argument because it is somewhat technical and not very enlightening. (The “right” approach, in my opinion, is to construct the required \(\mathbb{Q}\mathcal{S}_n\)-module \(W\) explicitly; this is done, e.g., in [54, §7, between Proposition 1 and Lemma 3]. Of course, this does not have much to do with what we are doing in our notes.)

11.90. Solution to Exercise 4.38. Solution to Exercise 4.38. (a) Let \(G\) be any group. If \(U_1\) and \(U_2\) are two \(\mathbb{C}\mathcal{G}\)-modules, then the character \(\chi_{U_1 \boxtimes U_2}\) of the inner tensor product \(U_1 \boxtimes U_2\) of \(U_1\) and \(U_2\) is given by

\[
\chi_{U_1 \boxtimes U_2}(g) = \chi_{U_1}(g) \chi_{U_2}(g) \quad \text{for all } g \in G.
\]

(This is just a restatement of (11.800) using our notation \(U\) and Lemma 3]. Of course, this does not have much to do with what we are doing in our notes.)

Define a map \(* : R(\mathcal{G}) \times R(\mathcal{G}) \to R(\mathcal{G})\), which will be written in infix notation (that is, we will write \(a * b\) instead of \(* (a, b)\)), by setting

\[
(a * b)(g) = a(g)b(g) \quad \text{for any } a \in R(\mathcal{G}), b \in R(\mathcal{G}) \text{ and } g \in G.
\]

(This notation \(a * b\) generalizes the notation \(\text{sgn}_{\mathcal{S}_n} * f\) used in Theorem 4.32.) The map \(*\) is clearly \(\mathbb{C}\)-bilinear. Notice that

\[
\chi_{U_1} * \chi_{U_2} = \chi_{U_1 \boxtimes U_2} \quad \text{for any two } \mathbb{C}\mathcal{G}\text{-modules } U_1 \text{ and } U_2\]

621

Let \(R_\mathbb{Q}(\mathcal{G})\) denote the set of class functions \(G \to \mathbb{Q}\). This is clearly a subset of \(R(\mathcal{G})\). In general, it is not true that \(R(\mathcal{G}) \subset R_\mathbb{Q}(\mathcal{G})\), but we will see that this holds for \(G = \mathcal{S}_n\) for any \(n \in \mathbb{N}\).

It is clear that \(a * b \in R_\mathbb{Q}(\mathcal{G})\) for any \(a \in R(\mathcal{G})\) and \(b \in R(\mathcal{G})\).

Now, forget that we fixed \(G\). We have thus introduced a map \(* : R(\mathcal{G}) \times R(\mathcal{G}) \to R(\mathcal{G})\) for every group \(G\), and we have proved some properties of this map.

Let now \(n \in \mathbb{N}\). Exercise 4.36(a) yields that \(\chi_V(g) \in \mathbb{Z} \subset \mathbb{Q}\) for every \(g \in \mathcal{S}_n\) and every finite-dimensional \(\mathbb{C}\mathcal{S}_n\)-module \(V\). Hence, \(\chi_V\) is a map from \(\mathcal{S}_n\) to \(\mathbb{Q}\) for every finite-dimensional \(\mathbb{C}\mathcal{S}_n\)-module \(V\). In other words, \(\chi_V \in R_\mathbb{Q}(\mathcal{S}_n)\) for every finite-dimensional \(\mathbb{C}\mathcal{S}_n\)-module \(V\). Thus, \(R(\mathcal{S}_n) \subset R_\mathbb{Q}(\mathcal{S}_n)\) (since \(R(\mathcal{S}_n)\) is the \(\mathbb{Z}\)-module generated by the \(\chi_V\) for \(V\) ranging over the irreducible \(\mathbb{C}\mathcal{S}_n\)-modules).

Every element of \(R_\mathbb{Q}(\mathcal{S}_n)\) is a class function \(\mathcal{S}_n \to \mathbb{Q}\), and thus a \(\mathbb{Q}\)-linear combination of the \(1_\lambda\) for \(\lambda \in \text{Par}_n\) (because these functions \(1_\lambda\) are the indicator functions for the conjugacy classes of \(\mathcal{S}_n\)). Thus, the \(\mathbb{Q}\)-module \(R_\mathbb{Q}(\mathcal{S}_n)\) is spanned by the \(1_\lambda\) for \(\lambda \in \text{Par}_n\).

As in Theorem 4.32, we extend the PSH-isomorphism \(\text{ch} : A \to \Lambda\) to a \(\mathbb{C}\)-Hopf algebra isomorphism \(A\mathcal{C} \to \Lambda\mathcal{C}\); we shall denote the latter isomorphism by \(\text{ch}\) as well. Now, let us notice that

\[
\text{ch} (R(\mathcal{S}_n)) \subset \Lambda_\mathcal{Q}.
\]

621Proof of (11.810): Let \(U_1\) and \(U_2\) be two \(\mathbb{C}\mathcal{G}\)-modules. Then, every \(g \in G\) satisfies

\[
(\chi_{U_1} * \chi_{U_2})(g) = \chi_{U_1}(g) \chi_{U_2}(g) \quad \text{(by the definition of \(*\))}
\]

\[
= \chi_{U_1 \boxtimes U_2}(g) \quad \text{(by (11.809))}.
\]

In other words, \(\chi_{U_1} * \chi_{U_2} = \chi_{U_1 \boxtimes U_2}\). This proves (11.810).
Furthermore,

\[(11.812) \quad \text{ch} (a \ast b) = (\text{ch} a) \ast (\text{ch} b) \quad \text{for any } a \in R_Q(\mathfrak{S}_n) \text{ and } b \in R_Q(\mathfrak{S}_n). \]

Now, let \(U_1\) and \(U_2\) be two \(C\mathfrak{S}_n\)-modules. Then, \(\chi_{U_1} \in R(\mathfrak{S}_n) \subset R_Q(\mathfrak{S}_n)\) and similarly \(\chi_{U_2} \in R_Q(\mathfrak{S}_n)\), so that \((11.812)\) (applied to \(a = \chi_{U_1}\) and \(b = \chi_{U_2}\)) yields \(\text{ch} (\chi_{U_1} \ast \chi_{U_2}) = \text{ch} (\chi_{U_1}) \ast \text{ch} (\chi_{U_2}).\) Due to \((11.810)\), this rewrites as \(\text{ch} (\chi_{U_1} \ast \chi_{U_2}) = \text{ch} (\chi_{U_1}) \ast \text{ch} (\chi_{U_2}).\) This solves Exercise 4.38(a).

(c) Let us first recall that \(p_\lambda \ast p_\mu = \delta_{\lambda,\mu} z_\lambda \delta_\lambda \delta_\mu\) for any two partitions \(\lambda\) and \(\mu\). Thus, \(p_\lambda \ast p_\mu = 0\) whenever \(\lambda \neq \mu\). This yields, in particular, that \(p_\lambda \ast p_\mu = 0\) whenever \(|\lambda| \neq |\mu|\). In other words, for any two distinct integers \(n\) and \(m\), we have \(p_\lambda \ast p_\mu = 0\) for every \(\lambda \in \text{Par}_n\) and every \(\mu \in \text{Par}_m\). This yields that, for any two distinct integers \(n\) and \(m\), we have \(a \ast b = 0\) for every \(a \in \Lambda_n\) and \(b \in \Lambda_m\) (because \(a\) is a \(Q\)-linear combination of the \(p_\lambda\) with \(\lambda \in \Lambda_n\), whereas \(b\) is a \(Q\)-linear combination of the \(p_\mu\) with \(\mu \in \Lambda_m\)). In particular, for any two distinct integers \(n\) and \(m\), we have

\[s_\mu \ast s_\nu = 0 \quad \text{for any } \nu \in \text{Par}_n \text{ and } \mu \in \text{Par}_m\]

(because \(s_\mu \in \Lambda_m\) and \(s_\nu \in \Lambda_n\)). In other words,

\[(11.813) \quad s_\mu \ast s_\nu = 0 \quad \text{for any partitions } \mu \text{ and } \nu \text{ satisfying } |\mu| \neq |\nu|.\]

Now, let \(\mu\) and \(\nu\) be two partitions. We need to prove that \(s_\mu \ast s_\nu \in \sum_{\lambda \in \text{Par}_n} N_{\lambda}^{\sigma_\lambda}.\) This is obvious (because of \((11.813)\)) in the case when \(|\mu| \neq |\nu|\), so we can WLOG assume that \(|\mu| = |\nu|\). Assume this, and let \(n = |\mu| = |\nu|\). Consider the two irreducible characters \(\chi^\mu\) and \(\chi^\nu\) of \(C\mathfrak{S}_n\) defined as in Theorem 4.32(a).

Then, \(\chi^\mu = \chi_{U_1}\) and \(\chi^\nu = \chi_{U_2}\) for two \(C\mathfrak{S}_n\)-modules \(U_1\) and \(U_2\). Consider these \(U_1\) and \(U_2\). We have

\[s_\mu \ast s_\nu = 0 \quad \text{for any } \nu \in \text{Par}_n \text{ and } \mu \in \text{Par}_m\]

(because \(s_\mu \in \Lambda_m\) and \(s_\nu \in \Lambda_n\)). In other words,

\[(11.813) \quad s_\mu \ast s_\nu = 0 \quad \text{for any partitions } \mu \text{ and } \nu \text{ satisfying } |\mu| \neq |\nu|.\]

\[\text{Now, let } \mu \text{ and } \nu \text{ be two partitions. We need to prove that } s_\mu \ast s_\nu \in \sum_{\lambda \in \text{Par}_n} N_{\lambda}^{\sigma_\lambda}.\] This is obvious (because of \((11.813)\)) in the case when \(|\mu| \neq |\nu|\), so we can WLOG assume that \(|\mu| = |\nu|\). Assume this, and let \(n = |\mu| = |\nu|\). Consider the two irreducible characters \(\chi^\mu\) and \(\chi^\nu\) of \(C\mathfrak{S}_n\) defined as in Theorem 4.32(a).

Then, \(\chi^\mu = \chi_{U_1}\) and \(\chi^\nu = \chi_{U_2}\) for two \(C\mathfrak{S}_n\)-modules \(U_1\) and \(U_2\). Consider these \(U_1\) and \(U_2\). We have

\[s_\mu \ast s_\nu = 0 \quad \text{for any } \nu \in \text{Par}_n \text{ and } \mu \in \text{Par}_m\]

(because \(s_\mu \in \Lambda_m\) and \(s_\nu \in \Lambda_n\)). In other words,

\[(11.813) \quad s_\mu \ast s_\nu = 0 \quad \text{for any partitions } \mu \text{ and } \nu \text{ satisfying } |\mu| \neq |\nu|.\]
Since the binary operation $\mu$ for any partitions $\lambda$ of $\mathfrak{S}_n$-module) must be a direct sum of finitely many irreducible $\mathfrak{S}_n$-modules, and thus the character $\chi_{U_1 \boxtimes U_2}$ is the sum of finitely many irreducible characters of $\mathfrak{S}_n$. Since the irreducible characters of $\mathfrak{S}_n$ are the $\chi^\lambda$ for $\lambda \in \text{Par}_n$, this shows that $\chi_{U_1 \boxtimes U_2}$ is the sum of finitely many $\chi^\lambda$. In other words, $\chi_{U_1 \boxtimes U_2} = \sum_{\lambda \in \text{Par}} \mathbb{N}\chi^\lambda$. Applying the map $\text{ch}$ to both sides of this relation, we obtain
\[
\text{ch} (\chi_{U_1 \boxtimes U_2}) \in \text{ch} \left( \sum_{\lambda \in \text{Par}} \mathbb{N}\chi^\lambda \right) = \sum_{\lambda \in \text{Par}} \mathbb{N} \text{ch} (\chi^\lambda) = \sum_{\lambda \in \text{Par}} \mathbb{N}s_\lambda,
\]
(by Theorem 4.32(a))

Since
\[
\text{ch} (\chi_{U_1 \boxtimes U_2}) = \text{ch} (\chi_{U_1}) \ast \text{ch} (\chi_{U_2}) = s_\mu \ast s_\nu = s_\mu \ast s_\nu,
\]
this rewrites as $s_\mu \ast s_\nu \in \sum_{\lambda \in \text{Par}} \mathbb{N}s_\lambda$. This solves Exercise 4.38(c).

(b) Alternative solution of Exercise 2.80(h). We need to prove that $f * g \in \Lambda$ for any $f \in \Lambda$ and $g \in \Lambda$. Since the binary operation $\ast$ is $\mathbb{Z}$-bilinear, it is clearly enough to prove that $s_\mu \ast s_\nu \in \Lambda$ for any partitions $\mu$ and $\nu$ (since $(s_\lambda)_{\lambda \in \text{Par}}$ is a $\mathbb{Z}$-basis of $\Lambda$). But this follows from the fact that
\[
s_\mu \ast s_\nu \in \sum_{\lambda \in \text{Par}} \mathbb{N}s_\lambda
\]
(by Exercise 4.38(c))

for any partitions $\mu$ and $\nu$. Thus, Exercise 2.80 (h) is solved again.

11.91. Solution to Exercise 4.39. Solution to Exercise 4.39. Let us introduce a fundamental construction. If $X$ is any set, then we let $\mathfrak{S}_X$ denote the symmetric group on $X$ (that is, the group of all permutations of $X$). For any two sets $X$ and $Y$ and any permutations $\sigma \in \mathfrak{S}_X$ and $\tau \in \mathfrak{S}_Y$, we define a permutation $\sigma \times \tau$ of $X \times Y$ by setting
\[
(\sigma \times \tau)(x, y) = (\sigma(x), \tau(y)) \quad \text{for every } x \in X \text{ and } y \in Y.
\]
Thus, for any two sets $X$ and $Y$, we can define a map
\[
\text{cross} : \mathfrak{S}_X \times \mathfrak{S}_Y \rightarrow \mathfrak{S}_{X \times Y},
\]
\[(\sigma, \tau) \mapsto \sigma \times \tau.
\]
We notice a few properties of this map (whose simple proof we leave the reader):

Lemma 11.21. Let $X$ and $Y$ be two sets.

(a) The map $\text{cross} : \mathfrak{S}_X \times \mathfrak{S}_Y \rightarrow \mathfrak{S}_{X \times Y}$ is a group homomorphism.

(b) If the sets $X$ and $Y$ are nonempty, then the map $\text{cross} : \mathfrak{S}_X \times \mathfrak{S}_Y \rightarrow \mathfrak{S}_{X \times Y}$ is injective.

The following deeper fact provides the first hint of a combinatorial interpretation of $\square$:

Lemma 11.22. Let $X$ and $Y$ be two finite sets. Let $\sigma \in \mathfrak{S}_X$ and $\tau \in \mathfrak{S}_Y$ be two permutations. Let $\lambda$, $\mu$ and $\kappa$ be the cycle types of the permutations $\sigma \in \mathfrak{S}_X$, $\tau \in \mathfrak{S}_Y$ and $\sigma \times \tau \in \mathfrak{S}_{X \times Y}$. Then,
\[
p_\lambda \square p_\mu = p_\kappa.
\]
**Proof of Lemma 11.22.** The partition $\lambda$ is the cycle type of the permutation $\sigma$. In other words, $\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)}$ are the lengths of the cycles of the permutation $\lambda$. Let $C_1, C_2, \ldots, C_{\ell(\lambda)}$ denote these cycles, labelled in such a way that each $C_i$ has length $\lambda_i$. Clearly, $(C_1, C_2, \ldots, C_{\ell(\lambda)})$ is a set partition of the set $X$; thus, $X = \bigsqcup_{i=1}^{\ell(\lambda)} C_i$.

The partition $\mu$ is the cycle type of the permutation $\sigma$. In other words, $\mu_1, \mu_2, \ldots, \mu_{\ell(\mu)}$ are the lengths of the cycles of the permutation $\mu$. Let $D_1, D_2, \ldots, D_{\ell(\mu)}$ denote these cycles, labelled in such a way that each $D_j$ has length $\mu_j$. Clearly, $(D_1, D_2, \ldots, D_{\ell(\mu)})$ is a set partition of the set $Y$; thus, $Y = \bigsqcup_{j=1}^{\ell(\mu)} D_j$.

Since $X = \bigsqcup_{i=1}^{\ell(\lambda)} C_i$ and $Y = \bigsqcup_{j=1}^{\ell(\mu)} D_j$, we have

$$X \times Y = \left( \bigcup_{i=1}^{\ell(\lambda)} C_i \right) \times \left( \bigcup_{j=1}^{\ell(\mu)} D_j \right) = \bigsqcup_{i=1}^{\ell(\lambda)} \bigsqcup_{j=1}^{\ell(\mu)} (C_i \times D_j).$$

In particular, the sets $C_i \times D_j$ for $(i, j)$ ranging over all $(i, j) \in \{1, 2, \ldots, \ell(\lambda)\} \times \{1, 2, \ldots, \ell(\mu)\}$ are disjoint.

We now make the following claim:

**Claim A:** Let $i \in \{1, 2, \ldots, \ell(\lambda)\}$ and $j \in \{1, 2, \ldots, \ell(\mu)\}$. Then, the subset $C_i \times D_j$ of $X \times Y$ is the union of gcd $(\lambda_i, \mu_j)$ disjoint cycles of the permutation $\sigma \times \tau$, and each of these cycles has length lcm $(\lambda_i, \mu_j)$.

**Proof of Claim A:** We know that $C_i$ is a cycle of $\sigma$; thus, $\sigma(C_i) \subset C_i$. Similarly, $\tau(D_j) \subset D_j$. Now, $(\sigma \times \tau)(C_i \times D_j) = \sigma(C_i) \times \tau(D_j) \subset C_i \times D_j$. Hence, the permutation $\sigma \times \tau$ restricts to a permutation of the subset $C_i \times D_j$ of $X \times Y$.

Let us first show that, for every $h \in C_i \times D_j$, the cycle of $\sigma \times \tau$ containing $h$ is a subset of $C_i \times D_j$ and has length lcm $(\lambda_i, \mu_j)$.

Indeed, fix $h \in C_i \times D_j$. Then, all of the elements $h, (\sigma \times \tau)(h), (\sigma \times \tau)^2(h), (\sigma \times \tau)^3(h), \ldots$ belong to $C_i \times D_j$. Hence, the cycle of $\sigma \times \tau$ containing $h$ is a subset of $C_i \times D_j$ (because this cycle consists of these very elements $h, (\sigma \times \tau)(h), (\sigma \times \tau)^2(h), (\sigma \times \tau)^3(h), \ldots$). The length of this cycle is the smallest positive integer $N$ such that $(\sigma \times \tau)^N(h) = h$. We shall now prove that this smallest positive integer is lcm $(\lambda_i, \mu_j)$.

Indeed, let us write $h$ in the form $h = (c, d)$ for some $c \in C_i$ and $d \in D_j$. Then, the element $c$ belongs to a cycle of $\sigma$ which has length $\lambda_i$ (namely, $C_i$). Hence, the sequence $c, \sigma(c), \sigma^2(c), \sigma^3(c), \ldots$ repeats every $\lambda_i$ elements (and not more frequently). Thus, for any $N \in \mathbb{N}$, we have the following equivalence of statements:

$$((\sigma^N(c) = c) \iff (\lambda_i | N)).$$

Similarly, for any $N \in \mathbb{N}$, we have the following equivalence of statements:

$$((\tau^N(d) = d) \iff (\mu_j | N)).$$

Now, for any $N \in \mathbb{N}$, we have the following equivalence of statements:

$$((\sigma \times \tau)^N(h) = h) \iff ((\sigma \times \tau)^N((c, d)) = (c, d)) \iff ((\sigma^N(c), \tau^N(d)) = (c, d)) \iff (\lambda_i | N) \text{ and } (\mu_j | N) \iff (\text{lcm}(\lambda_i, \mu_j) | N).$$

Thus, the smallest positive integer $N$ such that $(\sigma \times \tau)^N(h) = h$ is lcm $(\lambda_i, \mu_j)$. In other words, the length of the cycle of $\sigma \times \tau$ containing $h$ is lcm $(\lambda_i, \mu_j)$ (since the length of the cycle of $\sigma \times \tau$ containing $h$ is the smallest positive integer $N$ such that $(\sigma \times \tau)^N(h) = h$).

Now, let us forget that we fixed $h$. We thus have proven that, for every $h \in C_i \times D_j$,

$$((11.814) \quad \text{the cycle of } \sigma \times \tau \text{ containing } h \text{ is a subset of } C_i \times D_j \text{ and has length } \text{lcm}(\lambda_i, \mu_j)).$$

---

624: since the permutation $\sigma \times \tau$ restricts to a permutation of the subset $C_i \times D_j$ of $X \times Y$
These cycles (for \( h \) ranging over all \( C_i \times D_j \)) clearly cover the set \( C_i \times D_j \) (because each \( h \in C_i \times D_j \) is contained in its corresponding cycle). Thus, \( C_i \times D_j \) is the union of several cycles of the permutation \( \sigma \times \tau \), and each of these cycles has length \( \text{lcm}(\lambda_i, \mu_j) \). Since any two cycles of \( \sigma \times \tau \) are either disjoint or identical, we can get rid of redundant cycles in this union, and thus obtain the following conclusion: \( C_i \times D_j \) is the union of several disjoint cycles of the permutation \( \sigma \times \tau \), and each of these cycles has length \( \text{lcm}(\lambda_i, \mu_j) \).

In order to complete the proof of Claim A, it thus remains only to show that the number of these cycles is \( \gcd(\lambda_i, \mu_j) \). But this is easy: These cycles are all disjoint, and cover a set of size \( \lambda_i \times \mu_j \), ranging over all \( (\lambda_i, \mu_j) \) (in fact, they cover the set \( C_i \times D_j \), which has size \( |C_i \times D_j| = |C_i| \cdot |D_j| = \lambda_i \mu_j \); since they have length \( \text{lcm}(\lambda_i, \mu_j) \) each, their number must be \( \frac{\lambda_i \mu_j}{\text{lcm}(\lambda_i, \mu_j)} = \gcd(\lambda_i, \mu_j) \). Thus, the proof of Claim A is complete.

We shall now continue our proof of Lemma 11.22.

The partition \( \kappa \) is the cycle type of the permutation \( \sigma \times \tau \). In other words, \( \kappa_1, \kappa_2, \ldots, \kappa_{\ell(\kappa)} \) are the lengths of the cycles of the permutation \( \sigma \times \tau \). Hence,

\[
\prod_{u=1}^{\ell(\kappa)} p_{\kappa_u} = \prod_{E \text{ is a cycle of } \sigma \times \tau} p_{|E|},
\]

where \( |E| \) denotes the length of any cycle \( E \). But every cycle \( E \) of \( \sigma \times \tau \) must satisfy \( E \subset C_i \times D_j \) (where we regard \( E \) as a set) for precisely one \( (i, j) \in \{1, 2, \ldots, \ell(\lambda)\} \times \{1, 2, \ldots, \ell(\mu)\} \). Hence,

\[
\prod_{E \text{ is a cycle of } \sigma \times \tau} p_{|E|} = \prod_{i=1}^{\ell(\lambda)} \prod_{j=1}^{\ell(\mu)} \prod_{\substack{E \text{ is a cycle of } \sigma \times \tau; \ \ E \subset C_i \times D_j}} p_{|E|}.
\]

But every \( (i, j) \in \{1, 2, \ldots, \ell(\lambda)\} \times \{1, 2, \ldots, \ell(\mu)\} \) satisfies

\[
\prod_{E \subset C_i \times D_j} p_{|E|} = p_{\text{lcm}(\lambda_i, \mu_j)}^{\gcd(\lambda_i, \mu_j)}
\]

\footnote{Proof. Let \( E \) be a cycle of \( \sigma \times \tau \). Then, \( E \) is nonempty, so that there exists an element \( h \) of \( E \). Fix such an \( h \). We have \( h \in E \subset X \times Y = \bigcup_{i=1}^{\ell(\lambda)} \bigcup_{j=1}^{\ell(\mu)} (C_i \times D_j) \). Hence, there exists some \( (i, j) \in \{1, 2, \ldots, \ell(\lambda)\} \times \{1, 2, \ldots, \ell(\mu)\} \) such that \( h \in C_i \times D_j \). Consider this \( (i, j) \). Then, \( E \) is the cycle of \( \sigma \times \tau \) containing \( h \). Hence, (11.814) shows that \( E \) is a subset of \( C_i \times D_j \) and has length \( \text{lcm}(\lambda_i, \mu_j) \). So we know that \( E \subset C_i \times D_j \).

So we have found a pair \( (i, j) \in \{1, 2, \ldots, \ell(\lambda)\} \times \{1, 2, \ldots, \ell(\mu)\} \) such that \( E \subset C_i \times D_j \). There clearly cannot be two distinct such pairs \( (i, j) \) (since the sets \( C_i \times D_j \) for \( (i, j) \) ranging over all \( (i, j) \in \{1, 2, \ldots, \ell(\lambda)\} \times \{1, 2, \ldots, \ell(\mu)\} \), are disjoint), and so this pair \( (i, j) \) is unique.}
Now, the definition of \( p_\kappa \) yields

\[
p_\kappa = p_{\kappa_1} p_{\kappa_2} \cdots p_{\kappa_\ell(\kappa)} = \prod_{u=1}^{\ell(\kappa)} p_{\kappa_u} = \prod_{i=1}^{\ell(\lambda)} \prod_{j=1}^{\ell(\mu)} p_{|E|} = \prod_{E \subset C_i \times D_j} p_{|E|}
\]

\[
= \prod_{i=1}^{\ell(\lambda)} \prod_{j=1}^{\ell(\mu)} p_{\text{lcm}(\lambda_i, \mu_j)} = p_\lambda \Box p_\mu
\]

since \( p_\lambda \Box p_\mu \) is defined to be \( \prod_{i=1}^{\ell(\lambda)} \prod_{j=1}^{\ell(\mu)} p_{\text{lcm}(\lambda_i, \mu_j)} \).

This proves Lemma 11.22. \( \square \)

(a) We already know that the operation \( \Box \) is \( \mathbb{Q} \)-bilinear. Thus, we only need to prove that the binary operation \( \Box \) is commutative and associative and has unity \( p_1 \).

We shall only prove the associativity of \( \Box \) (since the other two properties can be proven similarly). In other words, we shall prove that \((u \Box v) \Box w = u \Box (v \Box w)\) for any three elements \( u, v \) and \( w \) of \( \Lambda_\mathbb{Q} \).

So let \( u, v \) and \( w \) be three elements of \( \Lambda_\mathbb{Q} \). We need to prove the equality \((u \Box v) \Box w = u \Box (v \Box w)\). Since this equality is \( \mathbb{Q} \)-linear in each of \( u, v \) and \( w \) (this is because the operation \( \Box \) is \( \mathbb{Q} \)-bilinear), we can WLOG assume that \( u, v \) and \( w \) belong to the basis \((p_\lambda)_{\lambda \in \Par} \) of the \( \mathbb{Q} \)-vector space \( \Lambda_\mathbb{Q} \). Assume this. Then, there exist three partitions \( \lambda, \mu \) and \( \nu \) such that \( u = p_\lambda, v = p_\mu \) and \( w = p_\nu \). Consider these \( \lambda, \mu \) and \( \nu \).

There exist a finite set \( A \) and a permutation \( \alpha \in \Sym_A \) such that \( \lambda \) is the cycle type of \( \alpha \). Consider these \( A \) and \( \alpha \).

There exist a finite set \( B \) and a permutation \( \beta \in \Sym_B \) such that \( \mu \) is the cycle type of \( \beta \). Consider these \( B \) and \( \beta \).

There exist a finite set \( C \) and a permutation \( \gamma \in \Sym_C \) such that \( \nu \) is the cycle type of \( \gamma \). Consider these \( C \) and \( \gamma \).

For every permutation \( \pi \) of any finite set \( X \), we let type\( \pi \) denote the cycle type of \( \pi \). Lemma 11.22 (applied to \( X = A, Y = B, \sigma = \alpha, \tau = \beta \) and \( \kappa = \text{type}(\alpha \times \beta) \)) yields \( p_\alpha \Box p_\tau = p_{\text{type}(\alpha \times \beta)} \). Lemma 11.22 (applied to \( A \times B, C, \alpha \times \beta, \gamma, \text{type}(\alpha \times \beta), \nu \) and type\( (\alpha \times \beta \times \gamma) \)) instead of \( X, Y, \sigma, \tau, \lambda, \mu, \) and \( \kappa \) yields \( p_{\text{type}(\alpha \times \beta)} \Box p_\nu = p_{\text{type}(\alpha \times \beta) \times \gamma} \). Thus,

\[
\left( \begin{array}{c}
\mu \\
\end{array} \right) \Box \left( \begin{array}{c}
\nu \\
\end{array} \right) = \left( \begin{array}{c}
p_\mu \\
\end{array} \right) \Box \left( \begin{array}{c}
p_\nu \\
\end{array} \right) = p_{\text{type}(\alpha \times \beta) \times \gamma}
\]

A similar argument shows that \( u \Box (v \Box w) = p_{\text{type}(\alpha \times (\beta \times \gamma)}) \).

Let us now say that if \( U \) and \( V \) are two finite sets, and if \( \sigma \in \Sym_U \) and \( \tau \in \Sym_V \) are two permutations, then the permutations \( \sigma \) and \( \tau \) are isomorphic if and only if there exists a bijection \( \varphi : U \to V \) such that \( \varphi \circ \sigma = \tau \circ \varphi \). The intuition behind this meaning of “isomorphism” is that two permutations are isomorphic if one of them becomes the other after a relabelling of its ground set. It is clear that two isomorphic permutations of finite sets must have the same cycle type.

But the permutations \( (\alpha \times \beta) \times \gamma \in \Sym_{(A \times B) \times C} \) and \( \alpha \times (\beta \times \gamma) \in \Sym_{A \times (B \times C)} \) are isomorphic (as witnessed by the bijection \( \varphi : (A \times B) \times C \to A \times (B \times C) \) sending every \( (a, b, c) \in (A \times B) \times C \) to \( (a, (b, c)) \)). Hence, type\( (\alpha \times (\beta \times \gamma)) \) is type\( (\alpha \times (\beta \times \gamma)) \) (since two isomorphic permutations of finite sets must have the same

---

626 Proof. Let \( \{i, j\} \in \{1, 2, \ldots, \ell(\lambda)\} \times \{1, 2, \ldots, \ell(\mu)\} \). Claim \( A \) yields that the subset \( C_i \times D_j \) of \( X \times Y \) is the union of gcd\( (\lambda_i, \mu_j) \) disjoint cycles of the permutation \( \sigma \times \tau \), and each of these cycles has length \( \text{lcm}(\lambda_i, \mu_j) \). Obviously, these gcd\( (\lambda_i, \mu_j) \) disjoint cycles are exactly all the cycles in the class of \( \sigma \times \tau \) which satisfy \( E \subset C_i \times D_j \). Thus, there are precisely gcd\( (\lambda_i, \mu_j) \) cycles \( E \subset \sigma \times \tau \) which satisfy \( E \subset C_i \times D_j \), and each of these cycles \( E \) has length \( \text{lcm}(\lambda_i, \mu_j) \). Hence,

\[
\prod_{E \subset C_i \times D_j} p_{\text{lcm}(\lambda_i, \mu_j)} = \prod_{E \subset C_i \times D_j} p_{\text{lcm}(\lambda_i, \mu_j)} \left( \text{since } E \text{ has length } \text{lcm}(\lambda_i, \mu_j) \right) \]

that is, we have \( |E| = \text{lcm}(\lambda_i, \mu_j) \). (since there are precisely gcd\( (\lambda_i, \mu_j) \) cycles \( E \subset \sigma \times \tau \) which satisfy \( E \subset C_i \times D_j \), qed.
cycle type). Thus,
\[(u \Box v) \Box w = p_{\text{type}((\alpha \times \beta) \times \gamma)} = p_{\text{type}(\alpha \times (\beta \times \gamma))} \quad \text{(since type} ((\alpha \times \beta) \times \gamma) = \text{type} (\alpha \times (\beta \times \gamma)))
= u \Box (v \Box w).\]
This finishes the proof of the equality \((u \Box v) \Box w = u \Box (v \Box w)\), and thus Exercise 4.39(a) is solved.

(b) Let \(f \in \Lambda_Q\). We need to prove the equality \(1 \Box f = \epsilon_1(f) 1\). Since this equality is \(\mathbb{Q}\)-linear in \(f\) (because the operation \(\Box\) is \(\mathbb{Q}\)-bilinear and the map \(\epsilon_1\) is \(\mathbb{Q}\)-linear), we can WLOG assume that \(f\) belongs to the basis \((p_\lambda)_{\lambda \in \mathcal{P}_n}\) of the \(\mathbb{Q}\)-vector space \(\Lambda_Q\). Assume this. Then, there exists a partition \(\lambda\) such that \(f = p_\lambda\). Consider this \(\lambda\). The definition of \(p_\lambda \Box \lambda\) yields
\[
p_{\emptyset} \Box \lambda = \prod_{i=1}^{\ell(2)} \prod_{j=1}^{\ell(\lambda)} p_{\gcd(\ell(i), \ell(\lambda))} = \prod_{i=1}^{\ell(2)} \prod_{j=1}^{\ell(\lambda)} p_{\gcd(\ell(i), \ell(\lambda))} \quad \text{(since } \ell(2) = 0)\]
= (empty product) = 1.

On the other hand, \(\epsilon_1(p_\lambda) = 1\) \(^{627}\). Thus, \(\epsilon_1\left(\frac{f}{p_\lambda}\right) = \epsilon_1(p_\lambda) = 1\). Hence, \(1 \Box f = 1 = \epsilon_1(f)\). This solves Exercise 4.39(b).

(c) For every finite group \(G\) and every \(h \in G\), we define a class function \(\alpha_{G,h} \in R_C(G)\) as in Exercise 4.34.

We next recall the notation \(U^\tau\) defined in Definition 4.23. We state a few properties of this notation:

**Lemma 11.23.** Let \(K\) and \(H\) be two finite groups. Let \(\tau : K \to H\) be a group homomorphism.

(a) For every \(f \in R_C(H)\), the map \(\tau \circ \tau : K \to H\) belongs to \(R_C(K)\). We can thus define a \(\mathbb{C}\)-linear map \(\tau^* : R_C(H) \to R_C(K)\) by
\[
(\tau^* f) = f \circ \tau \quad \text{for every } f \in R_C(H).
\]
(b) For every (finite-dimensional) \(\mathbb{C}H\)-module \(U\), we have \(\chi_{U^\tau} = \tau^*(\chi_U)\).
(c) We have \(\tau^*(R(H)) \subset R(K)\).
(d) Assume that \(\tau : K \to H\) is a group isomorphism. Then, \(\tau^*(\alpha_{H,\tau(g)}) = \alpha_{K,h}\) for every \(g \in K\).

The proof of Lemma 11.23 is straightforward and left to the reader. (We will not use its part (c).)

Finally, here come two more simple facts whose proofs we leave to the reader:

**Lemma 11.24.** Let \(G\) and \(H\) be two groups. Let \(\Omega : G \to H\) be an injective group homomorphism. Define a map \(\overline{\Omega} : G \to \Omega(G)\) by
\[
(\overline{\Omega}(g) = \Omega(g) \quad \text{for every } g \in G).
\]
Then, \(\overline{\Omega}\) is a well-defined group isomorphism.

**Lemma 11.25.** Let \(U\) and \(V\) be two finite sets. Let \(\varphi : U \to V\) be a bijection. Define a map \(\varphi^* : \mathcal{S}_U \to \mathcal{S}_V\) by
\[
(\varphi^*(\pi) = \varphi \circ \pi \circ \varphi^{-1} \quad \text{for every } \pi \in \mathcal{S}_U).
\]

\(^{627}\)Proof. We have \(\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{\ell(\lambda)})\). Thus, the definition of \(p_\lambda\) yields \(p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{\ell(\lambda)}} = \prod_{i=1}^{\ell(\lambda)} p_{\lambda_i}\). Applying the map \(\epsilon_1\) to both sides of this equality, we conclude
\[
\epsilon_1(p_\lambda) = \epsilon_1 \left( \prod_{i=1}^{\ell(\lambda)} p_{\lambda_i} \right) = \prod_{i=1}^{\ell(\lambda)} \epsilon_1(p_{\lambda_i}) \quad \text{(by the definition of } \epsilon_1)\]
\[= \prod_{i=1}^{\ell(\lambda)} 1 = 1.\]
qed.
(a) This map $\varphi^*$ is well-defined and a group isomorphism.
(b) Let $\pi \in \mathcal{S}_V$. Then, the cycle type of $\varphi^*(\pi)$ equals the cycle type of $\pi$.

Recall that $\boxtimes$ is a $\mathbb{Q}$-bilinear map $\Lambda_\mathbb{Q} \times \Lambda_\mathbb{Q} \to \Lambda_\mathbb{Q}$. Let us extend $\boxtimes$ to a $\mathbb{C}$-bilinear map $\Lambda_\mathbb{C} \times \Lambda_\mathbb{C} \to \Lambda_\mathbb{C}$; we will still denote this extended map by $\boxtimes$.

Recall that $\text{ch} : A \to \Lambda$ is a $\mathbb{Z}$-Hopf algebra isomorphism. Hence, the extension of $\text{ch}$ to a $\mathbb{C}$-linear map $A_\mathbb{C} \to \Lambda_\mathbb{C}$ is a $\mathbb{C}$-Hopf algebra isomorphism. We shall denote this extension by $\text{ch}_\mathbb{C}$.

We are going to repeatedly use the following fact (which is easy to obtain from the solution of Exercise 4.34): If $\lambda$ is a partition of a nonnegative integer $n$, and if $g \in \mathcal{S}_n$ is a permutation having cycle type $\lambda$, then

$$\text{ch}_\mathbb{C}(\alpha_{\mathbb{S}_n,g}) = p_\lambda.$$  

\textbf{Theorem 4.32(a)} yields $\text{ch}(\chi^\lambda) = s_\lambda$ for every partition $\lambda$ (where $\chi^\lambda$ is defined as in Theorem 4.32(a)). Since $\text{ch}_\mathbb{C}$ is the extension of $\text{ch}$ to a $\mathbb{C}$-linear map $A_\mathbb{C} \to \Lambda_\mathbb{C}$, we have

$$\text{ch}_\mathbb{C}(\chi^\lambda) = \text{ch}(\chi^\lambda) = s_\lambda$$

for every partition $\lambda$.

Let us now come back to solving Exercise 4.39(c). Let $\mu$ and $\nu$ be two partitions. We need to show that $s_\mu \boxtimes s_\nu \in \sum_{\lambda \in \text{Par}} \mathbb{N}s_\lambda$.

Let $m = |\mu|$ and $n = |\nu|$. Then, $\mu \in \text{Par}_m$ and $\nu \in \text{Par}_n$. If $\min\{m,n\} = 0$, then $s_\mu \boxtimes s_\nu \in \sum_{\lambda \in \text{Par}} \mathbb{N}s_\lambda$ is easily seen to hold.\textsuperscript{629} Hence, for the rest of this proof, we can WLOG assume that we don’t have $\min\{m,n\} = 0$. Assume this. Thus, $\min\{m,n\} \geq 1$. Hence, $m \geq 1$ and $n \geq 1$.

We notice that

$$\text{ch}_\mathbb{C}(\chi^\rho) \in \sum_{\lambda \in \text{Par}} \mathbb{N}s_\lambda$$

for every finite-dimensional $\mathbb{C}\mathcal{S}_{nm}$-module $P$.

\textbf{Proof of (11.815):} Let $\lambda$ be a partition of a nonnegative integer $n$. Let $g \in \mathcal{S}_n$ be a permutation having cycle type $\lambda$. Define a map $\Phi : \Lambda_\mathbb{C} \to A_\mathbb{C}$ as in the solution of Exercise 4.34(d). Then, (11.788) yields $\Phi(p_\lambda) = \alpha_{\mathbb{S}_n,g}$, so that $\alpha_{\mathbb{S}_n,g} = \Phi(p_\lambda)$.

But $\text{ch}_\mathbb{C} \circ \Phi = \text{id}_{A_\mathbb{C}}$ (this was shown in the solution of Exercise 4.34(d)). Now, \text{ch}_\mathbb{C}\left(\frac{\alpha_{\mathbb{S}_n,g}}{\Phi(p_\lambda)}\right) = \text{ch}_\mathbb{C}(\Phi(p_\lambda)) = \text{ch}_\mathbb{C}\circ\Phi(p_\lambda) = \text{id}_{A_\mathbb{C}}(p_\lambda) = p_\lambda$. This proves (11.815).

\textbf{Proof of (11.817):} Let $f$ be a finite-dimensional $\mathbb{C}\mathcal{S}_{nm}$-module. Then, $f$ must be a direct sum of finitely many simple $\mathbb{C}\mathcal{S}_{nm}$-modules. In other words, there exist some simple $\mathbb{C}\mathcal{S}_{nm}$-modules $V_1, V_2, \ldots, V_j$ such that $f \cong V_1 \oplus V_2 \oplus \cdots \oplus V_j$ as $\mathbb{C}\mathcal{S}_{nm}$-modules. Consider these $V_1, V_2, \ldots, V_j$. We shall now show that $\chi_{V_i} \in \sum_{\lambda \in \text{Par}} \mathbb{N}\chi^\lambda$ for every $i \in \{1, 2, \ldots, j\}$.

Indeed, let $i \in \{1, 2, \ldots, j\}$. Then, $V_i$ is a simple $\mathbb{C}\mathcal{S}_{nm}$-module. Hence, $\chi_{V_i}$ is an irreducible character of $\mathbb{C}\mathcal{S}_{nm}$. But (a part of) Theorem 4.32(a) (applied to $nm$ instead of $n$) says that all irreducible characters of $\mathcal{S}_{nm}$ have the form $\chi^\lambda$ with $\lambda \in \text{Par}_{nm}$.

Thus, $\chi_{V_i}$ has the form $\chi_{V_i} = \chi^\lambda$ for some $\lambda \in \text{Par}_{nm}$. In other words, $\chi_{V_i} \in \left\{\chi^\lambda \mid \lambda \in \text{Par}_{nm} \subset \text{Par}\right\} \subset \left\{\chi^\lambda \mid \lambda \in \text{Par}\right\} \subset \sum_{\lambda \in \text{Par}} \mathbb{N}\chi^\lambda$.

Now, let us forget that we fixed $i$. We thus have shown that $\chi_{V_i} \in \sum_{\lambda \in \text{Par}} \mathbb{N}\chi^\lambda$ for every $i \in \{1, 2, \ldots, j\}$. Thus,

$$\sum_{i=1}^j \chi_{V_i} \in \sum_{i=1}^j \sum_{\lambda \in \text{Par}} \mathbb{N}\chi^\lambda \subset \sum_{\lambda \in \text{Par}} \mathbb{N}\chi^\lambda$$

(since $\sum_{\lambda \in \text{Par}} \mathbb{N}\chi^\lambda$ is closed under addition).
for any two sets \( X \) and \( Y \). Now, set \( X = \{1, 2, \ldots, n\} \) and \( Y = \{1, 2, \ldots, m\} \), and consider the map 

\[
\text{cross} : \mathfrak{S}_X \times \mathfrak{S}_Y \to \mathfrak{S}_{X \times Y}.
\]

In the following, when we speak of cross, we will always mean this map. Define a map 

\[
\chi : \{1, 2, \ldots, n\} \times \{1, 2, \ldots, m\} \to \mathbb{N}.
\]

Applying the map \( \chi \), we obtain 

\[
\chi_{(x,y)} = \chi(x) \oplus \chi(y).
\]

In the following, when we speak of cross, we will always mean this map. Define a map 

\[
\varphi : \mathfrak{S}_X \to \mathfrak{S}_{X \times Y}.
\]

Hence, a \( \mathbb{C} \)-linear map \( \tau^* : R_\mathbb{C} (\mathfrak{S}_n \times \mathfrak{S}_m) \to R_\mathbb{C} (\Omega (\mathfrak{S}_n \times \mathfrak{S}_m)) \) is defined (according to Lemma 11.23(a), applied to \( \Omega (\mathfrak{S}_n \times \mathfrak{S}_m) \) and \( \mathfrak{S}_n \times \mathfrak{S}_m \) instead of \( K \) and \( H \)).
The following commutative diagram illustrates the group homomorphisms we have just introduced:

(11.818) \[ a \boxdot b = \text{ch}_C \left( \text{Ind}^{\text{sm}}_{\Omega(\mathfrak{g}_n \times \mathfrak{g}_m)} (\tau^* \left( (\text{ch}_C)^{-1} (a) \otimes (\text{ch}_C)^{-1} (b) \right) \right) \]

(where the cycle formed by the \( \overline{\Omega} \) and \( \tau \) arrows is not a mistake; these two maps are mutually inverse!).

In the following, for every \( k \in \mathbb{N} \), we let \((\Lambda_C)_k \) denote the \( k \)-th homogeneous component of the graded \( C \)-algebra \( \Lambda_C \). Note that \((p_\lambda)_{\lambda \in \text{Par}_k} \) is a basis of this \( C \)-vector space \((\Lambda_C)_k \).

Let us now claim that every \( a \in (\Lambda_C)_n \) and \( b \in (\Lambda_C)_m \) satisfy

(11.819) \[ \left( \begin{array}{c}
\text{ch}_C^{-1} (a) \\
\text{ch}_C^{-1} (b)
\end{array} \right) = (\alpha_{\mathfrak{g}_n,g} \otimes \alpha_{\mathfrak{g}_m,h}) \]

(Exercise 4.34(c) (applied to \( G_1 = \mathfrak{g}_n, G_2 = \mathfrak{g}_m, h_1 = g \) and \( h_2 = h \)) yields that the canonical isomorphism \( R_C (\mathfrak{g}_n) \otimes R_C (\mathfrak{g}_m) \to R_C (\mathfrak{g}_n \times \mathfrak{g}_m) \) sends \( \alpha_{\mathfrak{g}_n,g} \otimes \alpha_{\mathfrak{g}_m,h} \) to \( \alpha_{\mathfrak{g}_n \times \mathfrak{g}_m,(g,h)} \). Since we are regarding this isomorphism as an identity (because we have identified \( R_C (\mathfrak{g}_n) \otimes R_C (\mathfrak{g}_m) \) with \( R_C (\mathfrak{g}_n \times \mathfrak{g}_m) \) along this isomorphism), this yields \( \alpha_{\mathfrak{g}_n,g} \otimes \alpha_{\mathfrak{g}_m,h} = \alpha_{\mathfrak{g}_n \times \mathfrak{g}_m,(g,h)} \). Thus, (11.819) becomes

(11.819) \[ \left( \begin{array}{c}
\text{ch}_C^{-1} (a) \\
\text{ch}_C^{-1} (b)
\end{array} \right) = (\alpha_{\mathfrak{g}_n,g} \otimes \alpha_{\mathfrak{g}_m,h}) \]

\[ \text{where } \alpha_{\mathfrak{g}_n,g} \text{ and } \alpha_{\mathfrak{g}_m,h} \text{ are the } \alpha \text{ values given in Exercise 4.34(c).} \]

\[ \text{Note that the term } \tau^* \left( (\text{ch}_C)^{-1} (a) \otimes (\text{ch}_C)^{-1} (b) \right) \text{ on the right hand side of (11.818) is well-defined. (This is because} \]

(11.818) \[ \left( \begin{array}{c}
\alpha_{\mathfrak{g}_n,g} \\
\alpha_{\mathfrak{g}_m,h}
\end{array} \right) \]

(11.819) \[ \left( \begin{array}{c}
\alpha_{\mathfrak{g}_n,g} \\
\alpha_{\mathfrak{g}_m,h}
\end{array} \right) \]

(11.818) \[ \left( \begin{array}{c}
\alpha_{\mathfrak{g}_n,g} \\
\alpha_{\mathfrak{g}_m,h}
\end{array} \right) \]

(11.819) \[ \left( \begin{array}{c}
\alpha_{\mathfrak{g}_n,g} \\
\alpha_{\mathfrak{g}_m,h}
\end{array} \right) \]
Applying the map $\tau^*$ to both sides of this equality, we obtain

\[(11.820) \quad \tau^* \left( (\text{ch}_C)^{-1} (a) \otimes (\text{ch}_C)^{-1} (b) \right) = \tau^* \left( \alpha_{\mathfrak{S}_n \times \mathfrak{S}_m, (g,h)} \right).\]

But $\Omega ((g,h)) = \Omega ((g,h))$ (by the definition of $\Omega ((g,h))$), and so $(g,h) = \tau^{-1} \left( \Omega ((g,h)) \right) = \tau (\Omega ((g,h))).$

Thus,

\[\tau^* \left( \alpha_{\mathfrak{S}_n \times \mathfrak{S}_m, (g,h)} \right) = \tau^* \left( \alpha_{\mathfrak{S}_n \times \mathfrak{S}_m, \tau(\Omega ((g,h)))} \right) = \alpha_{\Omega(\mathfrak{S}_n \times \mathfrak{S}_m), \Omega((g,h))} \]

(by Lemma 11.23(d), applied to $\Omega (\mathfrak{S}_n \times \mathfrak{S}_m)$, $\mathfrak{S}_n \times \mathfrak{S}_m$, and $\Omega ((g,h))$ instead of $K$, $H$ and $g$). Thus, (11.820) becomes

\[\tau^* \left( (\text{ch}_C)^{-1} (a) \otimes (\text{ch}_C)^{-1} (b) \right) = \tau^* \left( \alpha_{\mathfrak{S}_n \times \mathfrak{S}_m, (g,h)} \right) = \alpha_{\Omega(\mathfrak{S}_n \times \mathfrak{S}_m), \Omega((g,h))}.\]

Applying the map $\text{Ind}_{\Omega(\mathfrak{S}_n \times \mathfrak{S}_m)}^{\mathfrak{S}_m \times \mathfrak{S}_m}$ to both sides of this equality, we obtain

\[
\begin{align*}
\text{Ind}_{\Omega(\mathfrak{S}_n \times \mathfrak{S}_m)}^{\mathfrak{S}_m \times \mathfrak{S}_m} \left( \tau^* \left( (\text{ch}_C)^{-1} (a) \otimes (\text{ch}_C)^{-1} (b) \right) \right) &= \text{Ind}_{\Omega(\mathfrak{S}_n \times \mathfrak{S}_m)}^{\mathfrak{S}_m \times \mathfrak{S}_m} \left( \tau^* \left( \alpha_{\mathfrak{S}_n \times \mathfrak{S}_m, \Omega((g,h))} \right) \right) \\
&= \tau^* \left( \alpha_{\mathfrak{S}_n \times \mathfrak{S}_m, \Omega((g,h))} \right) \text{ (by Exercise 4.34(b), applied to } \mathfrak{S}_n, \Omega \text{ and } \Omega \text{ instead of } n, \lambda \text{ and } g) \\
&= \alpha_{\mathfrak{S}_n \times \mathfrak{S}_m, \Omega((g,h))}. \tag*{(11.821)}
\end{align*}
\]

(by the definition of cross). In other words, there exists a simple $\mathfrak{S}_n \times \mathfrak{S}_m$-module $M$ such that $\chi^\mu = \chi_M$. Consider this $M$. Then, $(\text{ch}_C)^{-1} (s_\mu) = \chi^\mu = \chi_M$. We know that $\chi^\nu$ is an irreducible complex character of $\mathbb{C}\mathfrak{S}_m$. In other words, there exists a simple $\mathbb{C}\mathfrak{S}_m$-module $N$ such that $\chi^\nu = \chi_N$. Consider this $N$. Then, $(\text{ch}_C)^{-1} (s_\nu) = \chi^\nu = \chi_N$. This proves (11.818).
Recall that we are identifying $R_C(\mathcal{G}_n) \otimes R_C(\mathcal{G}_m)$ with $R_C(\mathcal{G}_n \times \mathcal{G}_m)$ along the canonical isomorphism $R_C(\mathcal{G}_n) \otimes R_C(\mathcal{G}_m) \rightarrow R_C(\mathcal{G}_n \times \mathcal{G}_m)$. Thus, for any finite-dimensional $\mathbb{C}\mathcal{G}_n$-module $U$ and any finite-dimensional $\mathbb{C}\mathcal{G}_m$-module $V$, we have $\chi_{U \otimes V} = \chi_U \otimes \chi_V$ (since this isomorphism sends $\chi_U \otimes \chi_V$ to $\chi_{U \otimes V}$).

Applying this to $U = N$ and $V = M$, we obtain $\chi_{N \otimes M} = \chi_N \otimes \chi_M$. Thus, $\chi_N \otimes \chi_M = \chi_{N \otimes M}$.

Lemma 11.23(b) (applied to $\Omega(\mathcal{G}_n \times \mathcal{G}_m)$, $\mathcal{G}_n \times \mathcal{G}_m$ and $N \otimes M$ instead of $K$, $H$ and $U$) yields $\chi_{(N \otimes M)^\tau} = \tau^* (\chi_{N \otimes M})$. Thus, $\tau^* (\chi_{N \otimes M}) = \chi_{(N \otimes M)^\tau}$. Hence,

$$(11.823) \quad \tau^* (\chi_{N \otimes \chi_M}) = \tau^* (\chi_{N \otimes M}) = (\tau^* (\chi_{N \otimes M})).$$

But (4.5) (applied to $\mathcal{G}_{nm}$, $\Omega(\mathcal{G}_n \times \mathcal{G}_m)$ and $(N \otimes M)^\tau$ instead of $G$, $H$ and $U$) yields $\chi_{\text{Ind}_{\Omega(\mathcal{G}_n \times \mathcal{G}_m)}(\chi_{(N \otimes M)^\tau})} = \text{Ind}_{\Omega(\mathcal{G}_n \times \mathcal{G}_m)}(\chi_{N \otimes M})$. Let $P$ denote the finite-dimensional $\mathbb{C}\mathcal{G}_{nm}$-module $\text{Ind}_{\Omega(\mathcal{G}_n \times \mathcal{G}_m)}((N \otimes M)^\tau)$. Then, $P = \text{Ind}_{\Omega(\mathcal{G}_n \times \mathcal{G}_m)}((N \otimes M)^\tau)$, so that

$$\chi_P = \chi_{\text{Ind}_{\Omega(\mathcal{G}_n \times \mathcal{G}_m)}((N \otimes M)^\tau)} = \text{Ind}_{\Omega(\mathcal{G}_n \times \mathcal{G}_m)}(\chi_{(N \otimes M)^\tau}) = \text{Ind}_{\Omega(\mathcal{G}_n \times \mathcal{G}_m)}(\tau^* (\chi_{N \otimes \chi_M})).$$

Hence,

$$(11.824) \quad \text{Ind}_{\Omega(\mathcal{G}_n \times \mathcal{G}_m)}(\tau^* (\chi_{N \otimes \chi_M})) = \chi_P.$$

We have $s_\mu \in (\Lambda_C)_m$ (since $\mu \in \text{Par}_m$) and $s_\nu \in (\Lambda_C)_n$ (since $\nu \in \text{Par}_n$). Thus, we can apply (11.818) to $a = s_\nu$ and $b = s_\mu$.

Exercise 4.39(a) yields that $\Lambda_Q$, equipped with the binary operation $\boxtimes$, becomes a commutative $\mathbb{Q}$-algebra with unity $p_1$. Thus, the operation $\boxtimes$ is commutative. Hence,

$$s_\mu \boxtimes s_\nu = s_\nu \boxtimes s_\mu = \chi_C \left( \text{Ind}_{\Omega(\mathcal{G}_n \times \mathcal{G}_m)}(\tau^* \left( \chi_{N \otimes \chi_M} \right)) \right) = \chi_C \left( \chi_P \right) \in \sum_{\lambda \in \text{Par}_n} \mathbb{N}s_\lambda \quad \text{(by (11.817)).}$$

This solves Exercise 4.39(c).

[Remark: The above solution can be simplified using the results of Exercise 4.14. We leave the details of this simplification to the reader, and only mention some highlights. Let us use the notations of Exercise 4.14. Then, the awkward equality (11.818) can be replaced by the simpler equality

$$(11.825) \quad a \boxtimes b = \chi_C \left( \text{Ind}_{\Omega(\text{ch}_C)^{-1}(a) \boxtimes (\text{ch}_C)^{-1}(b)} \right).$$

This makes the definition of the maps $\overline{\Omega}$ and $\tau$ unnecessary. The WLOG assumption that we don’t have $\min \{m,n\} = 0$ becomes unnecessary as well; the injectivity of the map cross can no longer be guaranteed without this assumption, but we do not need this map to be injective anymore. We no longer need Lemma 11.21(b), Lemma 11.23 and Lemma 11.24. However, we need to use Exercise 4.14(b) instead of 4.5, and we have to use Exercise 4.34(f) instead of Exercise 4.34(b).]

(d) We need to prove that $f \boxtimes g \in \Lambda$ for any $f \in \Lambda$ and $g \in \Lambda$. Since the binary operation $\boxtimes$ is $\mathbb{Z}$-bilinear (because it is $\mathbb{Q}$-bilinear), it is clearly enough to prove that $s_\mu \boxtimes s_\nu \in \Lambda$ for any partitions $\mu$ and $\nu$ (since...}
(s_\lambda)_{\lambda \in \mathfrak{Par}} is a \mathbb{Z}\text{-basis of } \Lambda). But this follows from the fact that

\[ s_\mu \boxplus s_\nu \in \sum_{\lambda \in \mathfrak{Par}} \mathbb{N}s_\lambda \]  
(by Exercise 4.39(c))

for any partitions \mu and \nu. Thus, Exercise 4.39(d) is solved.

11.92. Solution to Exercise 4.48. Solution to Exercise 4.48. (a) We shall prove the following fact:

**Lemma 11.26.** Let \( q \) be a prime power. For every positive integer \( n \), we have

\[
(\text{the number of all irreducible monic degree-} n \text{ polynomials in } \mathbb{F}_q[x]) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d.
\]

**Proof of Lemma 11.26.** We first recall that every positive integer \( n \) satisfies

\[
\sum_{d|n} \mu(d) = \delta_{n,1}. \tag{11.826}
\]

(This is precisely the equality (11.337), which was proven in the solution of Exercise 2.82.)

For every positive integer \( n \), define a nonnegative integer \( \text{irr} n \) by

\[
\text{irr} n = (\text{the number of all irreducible monic degree-} n \text{ polynomials in } \mathbb{F}_q[x]). \tag{11.827}
\]

Let \( \mathfrak{P} \) denote the set of all irreducible monic polynomials in \( \mathbb{F}_q[x] \). Then, every \( p \in \mathfrak{P} \) is an irreducible polynomial and thus satisfies \( \deg p \geq 1 \). Also, the irreducible monic polynomials in \( \mathbb{F}_q[x] \) are exactly the elements of \( \mathfrak{P} \) (since \( \mathfrak{P} \) is the set of all irreducible monic polynomials in \( \mathbb{F}_q[x] \)). Hence, for every positive integer \( n \), we have

\[
\text{irr} n = (\text{the number of all irreducible monic degree-} n \text{ polynomials in } \mathbb{F}_q[x].
\]

(11.828) (since the irreducible monic polynomials in \( \mathbb{F}_q[x] \) are exactly the elements of \( \mathfrak{P} \)).

Let \( \mathfrak{N} \) be the set of all families \( (k_p)_{p \in \mathfrak{P}} \in \mathbb{N}^{\mathfrak{P}} \) of nonnegative integers (indexed by the polynomials belonging to \( \mathfrak{P} \)) such that all but finitely many \( p \in \mathfrak{P} \) satisfy \( k_p = 0 \). Every monic polynomial \( P \in \mathbb{F}_q[x] \) has a unique factorization into irreducible monic polynomials. In other words, every monic polynomial \( P \in \mathbb{F}_q[x] \) can be written in the form \( P = \prod_{p \in \mathfrak{P}} p^{k_p} \) for a unique family \( (k_p)_{p \in \mathfrak{P}} \in \mathfrak{N} \). Thus, the map

\[
\mathfrak{N} \to \{ P \in \mathbb{F}_q[x] \mid P \text{ is monic} \},
\]

\[ (k_p)_{p \in \mathfrak{P}} \mapsto \prod_{p \in \mathfrak{P}} p^{k_p} \]

is bijective. Thus, in the ring \( \mathbb{Q}[[t]] \) of formal power series, we have

\[
\sum_{P \in \mathbb{F}_q[x]: P \text{ is monic}} t^{\deg P} = \sum_{(k_p)_{p \in \mathfrak{P}} \in \mathfrak{N}} t^{\deg (\prod_{p \in \mathfrak{P}} p^{k_p})}. \tag{11.829}
\]

But every \( (k_p)_{p \in \mathfrak{P}} \in \mathfrak{N} \) satisfies

\[
t^{\deg (\prod_{p \in \mathfrak{P}} p^{k_p})} = t^{\sum_{p \in \mathfrak{P}} k_p \deg p} \quad \text{(since } \deg \left( \prod_{p \in \mathfrak{P}} p^{k_p} \right) = \sum_{p \in \mathfrak{P}} k_p \deg p \text{)}
\]

\[ = \prod_{p \in \mathfrak{P}} t^{k_p \deg p} = \prod_{p \in \mathfrak{P}} (t^{\deg p})^{k_p}.
\]
Thus, (11.829) becomes

\[
\sum_{P \in \mathbb{F}_q[x]; \ P \text{ is monic}} t^{\deg P} = \sum_{n \in \mathbb{N}} \sum_{P \in \mathbb{F}_q[x]; \ P \text{ is monic; (since } \deg P=n)} t^{\deg P} = \sum_{n \in \mathbb{N}} \sum_{P \in \mathbb{F}_q[x]; \ P \text{ is monic; (since } \deg P=n)} t^n
\]

\[
= \sum_{n \in \mathbb{N}} \binom{t^n}{n} \cdot \sum_{n \in \mathbb{N}} q^n t^n = \sum_{n \in \mathbb{N}} \binom{t^n}{n} \cdot \frac{1}{1 - qt}
\]

\[
= \prod_{n \geq 1} \frac{1}{1 - t^n}
\]

Compared with

\[
\sum_{P \in \mathbb{F}_q[x]; \ P \text{ is monic}} t^{\deg P} = \sum_{n \in \mathbb{N}} \sum_{P \in \mathbb{F}_q[x]; \ P \text{ is monic; (since } \deg P=n)} t^{\deg P} = \sum_{n \in \mathbb{N}} \sum_{P \in \mathbb{F}_q[x]; \ P \text{ is monic; (since } \deg P=n)} t^n
\]

\[
= \sum_{n \in \mathbb{N}} \binom{t^n}{n} \cdot \sum_{n \in \mathbb{N}} q^n t^n = \sum_{n \in \mathbb{N}} \binom{t^n}{n} \cdot \frac{1}{1 - qt}
\]

\[
= \prod_{n \geq 1} \frac{1}{1 - t^n}
\]

\[
\frac{1}{1 - qt} = \prod_{n \geq 1} \left( \frac{1}{1 - t^n} \right)^{\text{irr } n}
\]
Taking the logarithm of both sides of this identity, we obtain
\[
\log \frac{1}{1-qt} = \log \left( \prod_{n \geq 1} \left( \frac{1}{1-t^n} \right)^{\text{irr} n} \right) = \sum_{n \geq 1} (\text{irr} n) \cdot \log \left( \frac{1}{1-t^n} \right)
\]
\[
= - \log(1-qt) = \sum_{n \geq 1} \frac{1}{n} (qt)^n
\]
(by the Mercator series for the logarithm)

\[
= \sum_{n \geq 1} \sum_{u \geq 1} \frac{(\text{irr} n)}{u} \left( \frac{qt}{n} \right)^u
\]
\[
= \sum_{n \geq 1} \sum_{v \geq 1} \frac{(\text{irr} n)}{v} \left( \frac{q}{n} \right)^v = \sum_{n \geq 1} \sum_{d|n} \frac{(\text{irr} d)}{d} \frac{q^n}{n}
\]
\[
\left( \text{here, we renamed the summation indices } v \text{ and } n \text{ as } n \text{ and } d \right).
\]

Since
\[
\log \frac{1}{1-qt} = - \log (1 - qt) = \sum_{n \geq 1} \frac{1}{n} (qt)^n
\]
(by the Mercator series for the logarithm)
\[
= \sum_{n \geq 1} \frac{1}{n} q^n t^n
\]
this rewrites as
\[
\sum_{n \geq 1} \frac{1}{n} q^n t^n = \sum_{n \geq 1} \sum_{d|n} \frac{(\text{irr} d)}{d} \frac{q^n}{n}
\]
Comparing coefficients, we conclude that every positive integer \( n \) satisfies
\[
\frac{1}{n} q^n = \sum_{d|n} \frac{(\text{irr} d)}{d}.
\]

Multiplying this with \( n \), we obtain
\[
(11.830) \quad q^n = \sum_{d|n} (\text{irr} d) \frac{d}{n}.
\]

Now, every positive integer \( n \) satisfies
\[
\sum_{d|n} \mu(d) q^{n/d} = \sum_{e|n} \mu(e) q^{n/e} = \sum_{d|n/e} \mu(d) \sum_{e|n/d} (\text{irr} d) d
\]
\[
= \sum_{e|n} \sum_{d|n/e} \mu(e) (\text{irr} d) d = \sum_{d|n} \sum_{e|n/d} \mu(d) (\text{irr} d) d
\]
\[
= \sum_{d|n} \delta_{n/d,1} (\text{irr} d) d = \sum_{d|n} \delta_{n,d} (\text{irr} d) d = (\text{irr} n) n.
\]
Dividing this by \( n \), we obtain \( \frac{1}{n} \sum_{d|n} \mu(d) q^{n/d} = \text{irr} n \). Now, \( (11.827) \) yields
\[
\text{the number of all irreducible monic degree-} n \text{ polynomials in } \mathbb{F}_q[x] = \text{irr} n = \frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}.
\]
This proves Lemma 11.26. 

Now, let us solve Exercise 4.48(a). Let \( n \geq 2 \) be an integer. We know that \( |\mathcal{F}_n| \) is the number of irreducible monic degree-\( n \) polynomials \( f(x) \neq x \) in \( \mathbb{F}_q[x] \) with nonzero constant term. Since the irreducible monic degree-\( n \) polynomials \( f(x) \neq x \) in \( \mathbb{F}_q[x] \) with nonzero constant term are precisely the irreducible monic degree-\( n \) polynomials in \( \mathbb{F}_q[x] \) \(^{632}\), this statement rewrites as follows: \( |\mathcal{F}_n| \) is the number of irreducible monic degree-\( n \) polynomials in \( \mathbb{F}_q[x] \). In other words,

\[
|\mathcal{F}_n| = \left( \text{the number of all irreducible monic degree-}n\text{-polynomials in } \mathbb{F}_q[x] \right) = \frac{1}{n} \sum_{d|n} \mu \left( \frac{n}{d} \right) q^{d} \quad \text{(by Lemma 11.26)}.
\]

This solves Exercise 4.48(a).

(b) We shall use the results of Exercise 6.34 (which provides a more systematic introduction to necklaces).

Let \( \mathfrak{A} \) be the set \( \mathbb{F}_q \). Consider the notion of an \( n \)-necklace defined in Exercise 6.34. The “\( n \)-necklaces” (as defined in Exercise 6.34) are precisely the “necklaces with \( n \) beads of \( q \) colors” (as defined in Exercise 4.48(b))\(^{633}\). Moreover, it is easy to see that the “aperiodic \( n \)-necklaces” (as defined in Exercise 6.34) are precisely the “primitive necklaces with \( n \) beads of \( q \) colors” (as defined in Exercise 4.48(b))\(^{634}\). Hence,

\[
\left( \text{the number of all aperiodic } n\text{-necklaces} \right) = \left( \text{the number of all primitive necklaces with } n \text{ beads of } q \text{ colors} \right),
\]

so that

\[
\left( \text{the number of all primitive necklaces with } n \text{ beads of } q \text{ colors} \right) = \left( \text{the number of all aperiodic } n\text{-necklaces} \right) = \frac{1}{n} \sum_{d|n} \mu \left( \frac{n}{d} \right) \left[ \mathfrak{A}^{n/d} \right] \quad \text{(by Exercise 6.34(f))}
\]

\[
= \frac{1}{n} \sum_{d|n} \mu \left( \frac{n}{d} \right) q^{n/d}.
\]

This solves Exercise 4.48(b).

\(^{632}\)Proof. Let \( f \) be an irreducible polynomial monic degree-\( n \) polynomial in \( \mathbb{F}_q[x] \). We will now show that \( f(x) \neq x \) and that \( f \) has nonzero constant term.

Since \( f \) is a degree-\( n \) polynomial, we have \( \deg f = n \geq 2 > 1 = \deg x \), thus \( \deg f \neq \deg x \) and therefore \( f \neq x \). In other words, \( f(x) \neq x \).

Now, let us assume (for the sake of contradiction) that the constant term of \( f \) is zero. Then, the polynomial \( x \) divides \( f \) in \( \mathbb{F}_q[x] \). But since \( f \) is irreducible, the polynomial \( f \) is a scalar multiple of every non-constant polynomial which divides \( f \). In particular, \( f \) is a scalar multiple of \( x \) (since \( x \) is a non-constant polynomial which divides \( f \)). Consequently, \( \deg f = \deg x \), which contradicts \( \deg f \neq \deg x \). This contradiction proves that our assumption (that the constant term of \( f \) is zero) was wrong. Hence, the constant term of \( f \) is nonzero. In other words, \( f \) has nonzero constant term.

Now, let us forget that we fixed \( f \). We thus have proven that every irreducible monic degree-\( n \) polynomial \( f \) in \( \mathbb{F}_q[x] \) satisfies \( f(x) \neq x \) and has nonzero constant term. Thus, all irreducible monic degree-\( n \) polynomials in \( \mathbb{F}_q[x] \) are irreducible monic degree-\( n \) polynomials \( f(x) \neq x \) in \( \mathbb{F}_q[x] \) with nonzero constant term. Combining this statement with the (obvious) converse statement (which states that all irreducible monic degree-\( n \) polynomials \( f(x) \neq x \) in \( \mathbb{F}_q[x] \) with nonzero constant term are irreducible monic degree-\( n \) polynomials in \( \mathbb{F}_q[x] \)), we conclude that the irreducible monic degree-\( n \) polynomials \( f(x) \neq x \) in \( \mathbb{F}_q[x] \) with nonzero constant term are precisely the irreducible monic degree-\( n \) polynomials in \( \mathbb{F}_q[x] \). Qed.

\(^{633}\)Proof. Consider the group \( C \), its generator \( c \) and the action of \( C \) on \( \mathfrak{A}^n \) which are defined in Exercise 6.34. Then, the “\( n \)-necklaces” (as defined in Exercise 6.34) are the orbits of the \( C \)-action on \( \mathfrak{A}^n \). In other words, the “\( n \)-necklaces” (as defined in Exercise 6.34) are the equivalence classes of \( n \)-tuples \((a_1, a_2, ..., a_n) \) in \( \mathfrak{A}^n \) with respect to cyclic rotation (because \( C \) acts on \( \mathfrak{A}^n \) by cyclic rotation). But the same can be said about the “necklaces with \( n \) beads of \( q \) colors” (as defined in Exercise 4.48(b)). Thus, the “\( n \)-necklaces” (as defined in Exercise 6.34) are precisely the “necklaces with \( n \) beads of \( q \) colors” (as defined in Exercise 4.48(b)), qed.

\(^{634}\)This follows from Exercise 6.34(b) (because for an \( n \)-tuple \((w_1, w_2, ..., w_n) \) in \( \mathfrak{A}^n \), the statement that no nontrivial rotation (in \( \mathbb{Z}/n\mathbb{Z} \)) fixes \( w \) is equivalent to the statement that every \( k \in \{1, 2, ..., n-1\} \) satisfies \((w_{k+1}, w_{k+2}, ..., w_n, w_1, w_2, ..., w_k) \neq w) \).
11.93. Solution to Exercise 4.60. Solution to Exercise 4.60. Let us first notice that any $k \in \mathbb{N}$ and any $k$ partitions $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}$ satisfy
\begin{equation}
\prod_{\lambda^{(1)}} \prod_{\lambda^{(2)}} \cdots \prod_{\lambda^{(k)}} \prod_{\lambda^{(k)}} = \sum_{\lambda \in \text{Par}} g_{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)})} (q) \prod_{\lambda \in \text{Par}} g_{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)})} (q) \prod_{\lambda \in \text{Par}} g_{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)})} (q)
\end{equation}
(here, we renamed the summation index $\lambda$ as $\mu$). Any two partitions $\mu$ and $\nu$ satisfy
\begin{equation}
\prod_{\mu \in \text{Par}} g_{(\mu)} (q) \prod_{\nu \in \text{Par}} g_{(\nu)} (q) = \prod_{\mu \in \text{Par}} g_{(\mu)} (q) \prod_{\tau \in \text{Par}} g_{(\tau)} (q) \prod_{\tau \in \text{Par}} g_{(\tau)} (q)
\end{equation}
(here, we renamed the summation index $\lambda$ as $\tau$).

(a) We shall prove the statement of Exercise 4.60(a) by induction over $k$. The base cases ($k = 0$ and $k = 1$) are left to the reader. We will now handle the induction step. So let us solve Exercise 4.60(a) for $\lambda$ (here, we renamed the summation index $k$ as $\lambda$).

From (11.831), we obtain
\begin{align*}
\sum_{\mu \in \text{Par}} g_{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)})} (q) \prod_{\lambda \in \text{Par}} g_{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)})} (q) 
&= \prod_{\lambda \in \text{Par}} g_{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)})} (q) \prod_{\lambda \in \text{Par}} g_{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)})} (q) \\
&= \sum_{\tau \in \text{Par}} g_{(\tau)} (q) \prod_{\tau \in \text{Par}} g_{(\tau)} (q) \prod_{\tau \in \text{Par}} g_{(\tau)} (q)
\end{align*}
Combining coefficients in front of $\prod_{\lambda \in \text{Par}}$ on both sides of this equality, we obtain
\begin{equation}
\sum_{\mu \in \text{Par}} g_{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)})} (q) = \sum_{\mu \in \text{Par}} g_{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)})} (q) g_{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k-1)})} (q) g_{(\lambda^{(1)}, \lambda^{(2), \ldots, \lambda^{(k)})} (q)}
\end{equation}

Our $\mathbb{F}_q$-vector space $V$ is $n$-dimensional. Hence, we can WLOG assume that $V = \mathbb{F}_q^n$ (because nothing changes if we map $V$ isomorphically to $\mathbb{F}_q^n$ and change $g$ accordingly). Assume this.

In the following, an ender will mean a $g$-stable $\mathbb{F}_q$-vector subspace $W \subset \mathbb{F}_q^n$ for which the induced map $\tilde{g}$ on the quotient space $\mathbb{F}_q^n/W$ has Jordan type $\lambda^{(k)}$. Notice that if $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_k = V$ is a $(\lambda^{(1)}, \lambda^{(2), \ldots, \lambda^{(k)})}$-compatible $g$-flag, then $V_{k-1}$ is an ender (because the definition of a $(\lambda^{(1)}, \lambda^{(2), \ldots, \lambda^{(k)})}$-compatible $g$-flag shows that the endomorphism of $\mathbb{F}_q^n/W_{k-1} = V_k/V_{k-1}$ induced by $g$ has Jordan type $\lambda^{(k)}$). (This is why we have chosen the name “ender” – an ender is the last proper subspace in a $(\lambda^{(1)}, \lambda^{(2), \ldots, \lambda^{(k)})}$-compatible $g$-flag.)

Whenever $h$ is a unipotent endomorphism of a finite-dimensional vector space, we let type $h$ denote the Jordan type of $h$.

Proposition 4.58 (applied to $\nu = \lambda^{(k)}$) yields that for every $\mu \in \text{Par}$, the number $g_{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)})} (q)$ counts the $g$-stable $\mathbb{F}_q$-subspaces $W \subset \mathbb{F}_q^n$ for which the restriction $g|W$ acts with Jordan type $\mu$, and the induced map $\tilde{g}$ on the quotient space $\mathbb{F}_q^n/W$ has Jordan type $\lambda^{(k)}$. In other words, for every $\mu \in \text{Par}$, the number $g_{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)})} (q)$ counts the enders $W$ for which the restriction $g|W$ acts with Jordan type $\mu$. In other words, for every $\mu \in \text{Par}$, the number $g_{(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)})} (q)$ counts the enders $W$ for which type $(g|W) = \mu$. In other words, for

\footnote{Note that the variable that I am calling $W$ here has been denoted by $V$ in Proposition 4.58.}
every $\mu \in \text{Par}$, we have

$$g_{\mu,\lambda(k)}(q) = \sum_{\text{W is an ender; } \text{type}(g|W) = \mu} 1.$$ 

Hence, (11.833) becomes

$$g_{\lambda(1),\lambda(2),\ldots,\lambda(k)}(q) = \sum_{\mu \in \text{Par}} g_{\lambda(1),\lambda(2),\ldots,\lambda(k)}^\mu(q) = \sum_{\text{W is an ender; } \text{type}(g|W) = \mu} g_{\lambda(1),\lambda(2),\ldots,\lambda(k)}(q) \sum_{\text{W is an ender; } \text{type}(g|W) = \mu} 1.$$ 

(11.834)

Now, let us fix an ender $W$. By the definition of an ender, this $W$ is a $g$-stable $\mathbb{F}_q$-subspace of $\mathbb{F}_q^n$ for which the induced map $\bar{g}$ on the quotient space $\mathbb{F}_q^n/W$ has Jordan type $\lambda(k)$. Let $m = \text{dim } W$.

By the induction hypothesis, we can apply Exercise 4.60(a) to type $(g|W)$, $m$, $W$, $g|W$, $k-1$ and $(\lambda(1), \lambda(2), \ldots, \lambda(k-1))$ instead of $\lambda$, $n$, $V$, $g$ and $(\lambda(1), \lambda(2), \ldots, \lambda(k))$. As a result, we conclude that

$$g_{\lambda(1),\lambda(2),\ldots,\lambda(k-1)}^\mu(q) = \text{the number of } (\lambda(1), \lambda(2), \ldots, \lambda(k-1)) \text{-compatible } g|W\text{-flags.}$$

Now, forget that we fixed $W$. Combining (11.834) with (11.835), we see that $g_{\lambda(1),\lambda(2),\ldots,\lambda(k)}^\mu(q)$ is the number of all pairs $(W, F)$, where $W$ is an ender and $F$ is a $(\lambda(1), \lambda(2), \ldots, \lambda(k-1))$-compatible $g|W$-flag. But (in order to complete the induction step) we have to prove that $g_{\lambda(1),\lambda(2),\ldots,\lambda(k)}^\mu(q)$ is the number of $(\lambda(1), \lambda(2), \ldots, \lambda(k))$-compatible $g$-flags. So an obvious way to prove this would be to find a bijection between $(\lambda(1), \lambda(2), \ldots, \lambda(k))$-compatible $g$-flags and pairs $(W, F)$, where $W$ is an ender and $F$ is a $(\lambda(1), \lambda(2), \ldots, \lambda(k-1))$-compatible $g|W$-flag. This bijection is very easy to define: It sends a $(\lambda(1), \lambda(2), \ldots, \lambda(k))$-compatible $g$-flag $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_k = V$ to the pair $(W_{k-1}, F)$, where $F$ is the $(\lambda(1), \lambda(2), \ldots, \lambda(k-1))$-compatible $g|V_{k-1}$-flag $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_{k-1} = V_{k-1}$. Showing that this bijection is well-defined and bijective is completely straightforward\(^{636}\); thus, the induction step is complete, and the solution of Exercise 4.60(a) is finished.

(b) We shall prove the statement of Exercise 4.60(b) by induction over $k$. The base cases ($k = 0$ and $k = 1$) are left to the reader. We will now handle the induction step. So let us solve Exercise 4.60(b) for some positive integer $k > 1$, assuming (as the induction hypothesis) that Exercise 4.60(b) is already solved for $k-1$ instead of $k$.

We need to prove that $g_{\lambda(1),\lambda(2),\ldots,\lambda(k)}^\mu(q) = 0$ unless $|\lambda(1)| + |\lambda(2)| + \cdots + |\lambda(k)| = |\lambda|$ and $\lambda(1) + \lambda(2) + \cdots + \lambda(k) \triangleright \lambda$. In other words, we need to prove that if $g_{\lambda(1),\lambda(2),\ldots,\lambda(k)}^\mu(q) \neq 0$, then

$$|\lambda(1)| + |\lambda(2)| + \cdots + |\lambda(k)| = |\lambda| \quad \text{and} \quad \lambda(1) + \lambda(2) + \cdots + \lambda(k) \triangleright \lambda.$$ 

So let us assume that $g_{\lambda(1),\lambda(2),\ldots,\lambda(k)}^\mu(q) \neq 0$. Due to (11.833), this rewrites as

$$\sum_{\mu \in \text{Par}} g_{\lambda(1),\lambda(2),\ldots,\lambda(k-1)}^\mu(q) g_{\mu,\lambda(k)}^\mu(q) \neq 0.$$ 

Hence, there exists a $\mu \in \text{Par}$ satisfying $g_{\lambda(1),\lambda(2),\ldots,\lambda(k-1)}^\mu(q) \neq 0$ and $g_{\mu,\lambda(k)}^\mu(q) \neq 0$. Consider this $\mu$.

By the induction hypothesis, we can apply Exercise 4.60(b) to $\mu$, $k-1$ and $(\lambda(1), \lambda(2), \ldots, \lambda(k-1))$ instead of $\lambda$, $k$ and $(\lambda(1), \lambda(2), \ldots, \lambda(k))$. Thus, we conclude that $g_{\lambda(1),\lambda(2),\ldots,\lambda(k-1)}^\mu(q) = 0$ unless $|\lambda(1)| + |\lambda(2)| + \cdots + |\lambda(k-1)| = |\mu|$ and $\lambda(1) + \lambda(2) + \cdots + \lambda(k-1) \triangleright \mu$. Since we have $g_{\lambda(1),\lambda(2),\ldots,\lambda(k-1)}^\mu(q) \neq 0$, we therefore must have $|\lambda(1)| + |\lambda(2)| + \cdots + |\lambda(k-1)| = |\mu|$ and $\lambda(1) + \lambda(2) + \cdots + \lambda(k-1) \triangleright \mu$.

But let $n = |\lambda|$. Fix a unipotent endomorphism $g$ of $\mathbb{F}_q^n$ having Jordan type $\lambda$ (such a $g$ clearly exists). The number $g_{\mu,\lambda(k)}^\mu(q)$ counts the $g$-stable $\mathbb{F}_q$-subspaces $W \subset \mathbb{F}_q^n$ for which the restriction $g|W$ acts with

---

\(^{636}\)The inverse map takes a pair $(W, F)$, where $W$ is an ender and $F$ is a $(\lambda(1), \lambda(2), \ldots, \lambda(k-1))$-compatible $g|W$-flag $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_{k-1} = W$, and maps it to the $(\lambda(1), \lambda(2), \ldots, \lambda(k))$-compatible $g$-flag $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_{k-1} \subset V = V$. The number $g_{\mu,\lambda(k)}^\mu(q)$ counts the $g$-stable $\mathbb{F}_q$-subspaces $W \subset \mathbb{F}_q^n$ for which the restriction $g|W$ acts with
Thus, we have proven (11.836), and so the induction step is complete. We thus have solved Exercise 4.60(b).

Now, let $\lambda \in \mathbb{F}_q^n$ have Jordan type $\lambda \mid \ell$. Also, for every $\lambda = (\lambda(1), \lambda(2), \ldots, \lambda(k))$ having Jordan type $\lambda \mid \ell$, we obtain

\[ |\lambda| = |\lambda(1)| + |\lambda(2)| + \cdots + |\lambda(k)| = |\mu| + |\lambda(k)| = |\lambda| \]

and

\[ |\lambda(1)| + |\lambda(2)| + \cdots + |\lambda(k)| = (\lambda(1) + \lambda(2) + \cdots + \lambda(k-1)) + \lambda(k) \succ \mu + \lambda(k) \succ \lambda. \]

Thus, we have proven (11.836), and so the induction step is complete. We thus have solved Exercise 4.60(b).

(c) Let $n = |\lambda|$. Fix a unipotent endomorphism $g$ of $\mathbb{F}_q^n$ having Jordan type $\lambda$ (such a $g$ clearly exists).

Exercise 4.60(a) (applied to $V = \mathbb{F}_q^n$, $\ell = \ell$ and $\ell(i) = \ell(1(\lambda^i), \ell(1(\lambda^i)))$) shows that $g(1(\lambda^i), \ell(1(\lambda^i)))$ is the number of $(\ell(1(\lambda^i)))$-compatible g-flags. Hence, in order to prove

$g(1(\lambda^i)) \neq 0$ (and thus, to solve Exercise 4.60(c)), it will be enough to prove that there exists at least one $(\ell(1(\lambda^i)))$-compatible g-flag.

Let $f = g - \text{id}_{\mathbb{F}_q^n}$. Then, $f$ is a nilpotent endomorphism of $\mathbb{F}_q^n$ having Jordan type $\lambda$ (since $g$ is a unipotent endomorphism of $\mathbb{F}_q^n$ having Jordan type $\lambda$). Also, $g = f + \text{id}_{\mathbb{F}_q^n}$ (since $f = g - \text{id}_{\mathbb{F}_q^n}$).

Let us define a sequence $V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_\ell$ of $\mathbb{F}_q$-vector subspaces of $\mathbb{F}_q^n$ by setting

$V_i = \ker(f^i)$ for every $i \in \{0, 1, \ldots, \ell\}$.

Then, $V_0 = \ker(f^\ell) = \ker \text{id} = 0$. Also, it follows readily from Exercise 2.98(a) that

$\dim(\ker(f^k)) = (\lambda^i)_1 + (\lambda^i)_2 + \cdots + (\lambda^i)_k$ for every $k \in \mathbb{N}.$

Applying this to $k = \ell$, we obtain

$\dim(\ker(f^\ell)) = (\lambda^i)_1 + (\lambda^i)_2 + \cdots + (\lambda^i)_\ell = |\lambda^i|$

since $\lambda^i = (\lambda(1), \ell(1(\lambda^i)))$ which yields $\ker(f^\ell) = \mathbb{F}_q^n$ (since $\ker(f^\ell) \subset \mathbb{F}_q^n$). Thus, $V_\ell = \ker(f^\ell) = \mathbb{F}_q^n$. Since $V_0 = 0$ and $V_\ell = \mathbb{F}_q^n$, our sequence $V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_\ell$ thus can be written as $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_\ell = \mathbb{F}_q^n$.

For every $i \in \{0, 1, \ldots, \ell\}$, the $\mathbb{F}_q$-vector subspace $V_i$ of $\mathbb{F}_q^n$ is $g$-invariant (since $V_i = \ker(f^i)$ is $f$-invariant and thus $g$-invariant). Also, for every $i \in \{1, 2, \ldots, \ell\}$, we have $V_i / V_{i-1} = \ker(f^i) / \ker(f^{i-1});$

Proof of (11.837): Let $k \in \mathbb{N}$. Let $N \in \mathbb{F}_q^{n \times n}$ be the matrix representing the endomorphism $f$ of $\mathbb{F}_q^n$. Then, $N$ is nilpotent (since $f$ is nilpotent) and has Jordan type $\lambda$ (since $f$ has Jordan type $\lambda$), and thus satisfies $\dim(\ker(N^k)) = (\lambda^i)_1 + (\lambda^i)_2 + \cdots + (\lambda^i)_k$ (by Exercise 2.98(a)). But since $N$ is a matrix representing the map $f$, we have $\dim(\ker(N^k)) = \dim(\ker(f^k))$, so that $\dim(\ker(f^k)) = \dim(\ker(N^k)) = (\lambda^i)_1 + (\lambda^i)_2 + \cdots + (\lambda^i)_k$. This proves (11.837).
thus, the endomorphism of $V_i/V_{i-1}$ induced by $f$ is the zero map, and therefore the endomorphism of $V_i/V_{i-1}$ induced by $g$ is the identity map (since $g = f + \text{id}_{F^n}$). Hence, for every $i \in \{1, 2, \ldots, \ell\}$, this latter endomorphism has Jordan type $(1^{\dim(V_i/V_{i-1})})$. But since every $i \in \{1, 2, \ldots, \ell\}$ satisfies

$$
\dim \left( \frac{V_i}{V_{i-1}} \right) = \dim (\ker(f^i) / \ker(f^{i-1})) = \dim (\ker(f^i)) - \dim (\ker(f^{i-1})) = (\lambda^i_1 + \lambda^i_2 + \cdots + \lambda^i_i) - (\lambda^{i-1}_1 + \lambda^{i-1}_2 + \cdots + \lambda^{i-1}_{i-1}) = (\lambda^i_1 + \lambda^i_2 + \cdots + \lambda^i_i)
$$

(by (11.837), applied to $k = i$)

this rewrites as follows: For every $i \in \{1, 2, \ldots, \ell\}$, the endomorphism of $V_i/V_{i-1}$ induced by $g$ has Jordan type $(1^{(\lambda_i)})$.

Altogether, we now know that $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_{\ell} = \mathbb{F}_q^n$ is a sequence of $g$-invariant $\mathbb{F}_q$-vector subspaces $V_i$ of $\mathbb{F}_q^n$ such that for every $i \in \{1, 2, \ldots, \ell\}$, the endomorphism of $V_i/V_{i-1}$ induced by $g$ has Jordan type $(1^{(\lambda_i)})$. In other words, $0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_{\ell} = \mathbb{F}_q^n$ is a $(1^{(\lambda_1)})$-$1^{(\lambda_2)}$-$\cdots$-$1^{(\lambda_\ell)}$-compatible $g$-flag (according the definition of the latter notion). Hence, there exists at least one $(1^{(\lambda_1)})$-$1^{(\lambda_2)}$-$\cdots$-$1^{(\lambda_\ell)}$-compatible $g$-flag. Exercise 4.60(c) is solved.

(d) Write the partition $\lambda$ as $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ with $\ell = \ell(\lambda)$. Then, $(\lambda^i)^{\ell} = \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$. Hence, Exercise 4.60(c) (applied to $\lambda^i$ instead of $\lambda$) yields

$$
(11.838) \quad g^\mu_{\lambda_1} (\lambda_2, \ldots, \lambda_\ell) (q) \neq 0.
$$

But since $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$, we have $e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_\ell}$ and thus

$$
\varphi(e_\lambda) = \varphi(e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_\ell}) = \varphi(e_{\lambda_1}) \varphi(e_{\lambda_2}) \cdots \varphi(e_{\lambda_\ell}) \quad \text{(since $\varphi$ is a $\mathbb{C}$-algebra homomorphism)}
$$

$$
= \prod_{i=1}^{\ell} \varphi(e_{\lambda_i}) = \prod_{i=1}^{\ell} q\left(\frac{\lambda_i}{2}\right) \frac{1}{J(\lambda_i)} = q\left(\frac{\lambda}{2}\right) \frac{1}{J(\lambda)} \quad \text{(since Theorem 4.59 yields)}
$$

$$
\varphi(e_p) = q^{\frac{p}{2}} \frac{1}{J(p)} \text{ for every } p \in \mathbb{N}
$$

$$
= \prod_{i=1}^{\ell} q\left(\frac{\lambda_i}{2}\right) \frac{1}{J(\lambda_i)}
$$

$$
= \prod_{i=1}^{\ell} q\left(\frac{\lambda_i}{2}\right) \frac{1}{J(\lambda_i)}
$$

$$
= \sum_{\mu \in \text{Par}} g^\mu_{\lambda_1} (\lambda_2, \ldots, \lambda_\ell) (q) \frac{1}{J_\mu}
$$

(by (11.831), applied to $k = \ell$ and $\lambda^{(i)} = (\lambda_i)$)

$$
(11.839) \quad \lambda^i = (1^{\lambda_1}) + (1^{\lambda_2}) + \cdots + (1^{\lambda_\ell})
$$

Now, we notice that

$$
(11.840) \quad \lambda^i = (1^{\lambda_1}) + (1^{\lambda_2}) + \cdots + (1^{\lambda_\ell})
$$
Now, for every partition $\mu$, we have $g^\mu_{(1^{\lambda_1}), (1^{\lambda_2}), \ldots, (1^{\lambda_\ell})} (q) = 0$ unless $\mu \in \text{Par}_n$ and $\lambda^t \triangleright \mu$. Hence, we can replace the summation sign $\sum_{\mu \in \text{Par}_n}$ on the right hand side of (11.839) by a more restricted summation $\sum_{\mu \in \text{Par}_n; \lambda^t \triangleright \mu}$ without changing the value of the sum (since all addends that we lose are 0). Thus, (11.839) rewrites as

$$\varphi(e_\lambda) = \left( \prod_{i=1}^{\ell} q^{(\lambda_i)_2} \right) \sum_{\mu \in \text{Par}_n; \lambda^t \triangleright \mu} g^\mu_{(1^{\lambda_1}), (1^{\lambda_2}), \ldots, (1^{\lambda_\ell})} (q) \updownarrow_\mu$$

Setting $\alpha_{\lambda, \mu} = \left( \prod_{i=1}^{\ell} q^{(\lambda_i)_2} \right) g^\mu_{(1^{\lambda_1}), (1^{\lambda_2}), \ldots, (1^{\lambda_\ell})} (q)$, we can rewrite this as $\varphi(e_\lambda) = \sum_{\mu \in \text{Par}_n; \lambda^t \triangleright \mu} \alpha_{\lambda, \mu} \updownarrow_\mu$. Thus, we will be done solving Exercise 4.60(d) as soon as we can prove the inequality

$$\left( \prod_{i=1}^{\ell} q^{(\lambda_i)_2} \right) g^\lambda_{(1^{\lambda_1}), (1^{\lambda_2}), \ldots, (1^{\lambda_\ell})} (q) \neq 0.$$

But the latter inequality follows from $q \neq 0$ and (11.838). Thus, Exercise 4.60(d) is solved.

(c) The map $\varphi : \Lambda_C \to \mathcal{H}$ is graded, and thus, in order to prove that $\varphi$ is injective, it is enough to show that the restriction $\varphi \mid_{(\Lambda_C)_n}$ of $\varphi$ to $(\Lambda_C)_n$ is injective for every $n \in \mathbb{N}$. So let us fix $n \in \mathbb{N}$.

---

\(^{638}\) Proof of (11.840): In the following, we use the so-called Iverson bracket notation: For every assertion $A$, we let $[A]$ denote the integer $\begin{cases} 1, & \text{if } A \text{ is true}; \\ 0, & \text{if } A \text{ is false} \end{cases}$. (This integer is called the truth value of $A$.)

For every $p \in \mathbb{N}$ and $i \in \{1, 2, 3, \ldots\}$, we have

$$\begin{equation}
(1^p)_i = [p \geq i].
\end{equation}$$

Now, for every $i \in \{1, 2, 3, \ldots\}$ satisfies

$$\begin{aligned}
\left( 1^{\lambda_1} + 1^{\lambda_2} + \ldots + 1^{\lambda_\ell} \right)_i &= \left( 1^{\lambda_1} \right)_i + \left( 1^{\lambda_2} \right)_i + \ldots + \left( 1^{\lambda_\ell} \right)_i \\
&= \sum_{k=1}^{\ell} \left( \lambda_k \right)_i \\
&= [\{j \in \{1, 2, \ldots, \ell\} \mid \lambda_j \geq i\}] (\lambda^t)_i \\
&= (1^i)_i. \\
\end{aligned}$$

(by the definition of $\mu + \nu$ for two partitions $\mu$ and $\nu$)

Hence, $(1^{\lambda_1}) + (1^{\lambda_2}) + \ldots + (1^{\lambda_\ell}) = \lambda^t$, qed.

\(^{639}\) Proof. Let $\mu$ be a partition. Exercise 4.60(b) (applied to $\ell, (1^{\lambda_1})$ and $\mu$ instead of $k, \lambda^{(0)}$ and $\lambda$) shows that we have $g^\mu_{(1^{\lambda_1}), (1^{\lambda_2}), \ldots, (1^{\lambda_\ell})} (q) = 0$ unless $|1^{\lambda_1}| + |1^{\lambda_2}| + \ldots + |1^{\lambda_\ell}| = |\mu|$ and $(1^{\lambda_1}) + (1^{\lambda_2}) + \ldots + (1^{\lambda_\ell}) \triangleright \mu$. Since

$$|1^{\lambda_1}| + |1^{\lambda_2}| + \ldots + |1^{\lambda_\ell}| = \left( \lambda^t \right)_i = |\lambda| = n$$

and $(1^{\lambda_1}) + (1^{\lambda_2}) + \ldots + (1^{\lambda_\ell}) = \lambda^t$ (by (11.840)), this rewrites as follows: We have $g^\mu_{(1^{\lambda_1}), (1^{\lambda_2}), \ldots, (1^{\lambda_\ell})} (q) = 0$ unless $n = |\mu|$ and $\lambda^t \triangleright \mu$. In other words, we have $g^\mu_{(1^{\lambda_1}), (1^{\lambda_2}), \ldots, (1^{\lambda_\ell})} (q) = 0$ unless $\mu \in \text{Par}_n$ and $\lambda^t \triangleright \mu$, qed.
We know that \((e_{\lambda})_{\lambda \in \text{Par}_n}\) is a basis of the \(\mathbb{C}\)-vector space \((A_{\mathbb{C}})_n\). Hence, \((e_{\lambda'})_{\lambda \in \text{Par}_n}\) also is a basis of the \(\mathbb{C}\)-vector space \((A_{\mathbb{C}})_n\). Every \(\lambda \in \text{Par}_n\) satisfies
\[
\varphi(e_{\lambda'}) = \sum_{\mu \in \text{Par}_n; (\lambda') \triangleright \mu} \alpha_{\lambda', \mu} \frac{1}{t_{\mu}}
\]
for some coefficients \(\alpha_{\lambda', \mu} \in \mathbb{C}\) satisfying \(\alpha_{\lambda', (\lambda')^t} \neq 0\) (according to Exercise 4.60(d)). In other words, every \(\lambda \in \text{Par}_n\) satisfies
\[
(11.842)
\]
for some coefficients \(\alpha_{\lambda', \mu} \in \mathbb{C}\) satisfying \(\alpha_{\lambda', \lambda} \neq 0\) (since \((\lambda')^t = \lambda\)).

The \(\mathbb{C}\)-vector space basis \((e_{\lambda'})_{\lambda \in \text{Par}_n}\) of \((A_{\mathbb{C}})_n\) and the \(\mathbb{C}\)-vector space basis \((\frac{1}{t_{\lambda}})_{\lambda \in \text{Par}_n}\) of \(\mathcal{H}_n\) are indexed by the same set, and the matrix that represents the map \(\varphi \mid (A_{\mathbb{C}})_n\) with respect to these bases is triangular with respect to the dominance order (according to (11.842)), with its diagonal entries being nonzero (due to \(\alpha_{\lambda', \lambda} \neq 0\)). Hence, this matrix is invertible, so that the map \(\varphi \mid (A_{\mathbb{C}})_n\) is invertible (as a linear map \((A_{\mathbb{C}})_n \to \mathcal{H}_n\)) and thus injective. This completes the solution to Exercise 4.60(e).

11.94. Solution to Exercise 5.21. Solution to Exercise 5.21.

Alternative proof of Theorem 5.19. Let \(P\) be a labelled poset.

First of all, let \(f\) be any map \(P \to \{1, 2, 3, \ldots\}\). We define a binary relation \(\prec_f\) on the set \(P\) by letting \(i \prec_f j\) hold if and only if
\[
(f(i) < f(j) \text{ or } (f(i) = f(j) \text{ and } i <_{\mathbb{Z}} j)).
\]
It is straightforward to see that this binary relation \(\prec_f\) is the smaller relation of a total order on \(P\). Let us define \(w(f)\) to be the set \(P\) endowed with this total order. Thus, \(w(f) = P\) as sets, but the smaller relation \(\prec_{w(f)}\) of the totally ordered set \(w(f)\) is the relation \(\prec_f\).

Let us now forget that we fixed \(f\). Thus, for every map \(f : P \to \{1, 2, 3, \ldots\}\), we have constructed a binary relation \(\prec_f\) on the set \(P\) and a totally ordered set \(w(f)\). It is easy to see that, for every \(f \in A(P)\), we have
\[
(11.843)
\]

Now, fix \(w \in \mathcal{L}(P)\). Thus, \(w\) is a linear extension of \(P\). That is, \(w\) is a totally ordered set with ground set \(P\), and extends the poset \(P\).

Let us first show that
\[
(11.844)
\]

\(640\) This has been proven in the proof of Proposition 2.17. (Alternatively, this can be easily concluded from Proposition 2.17.)

\(641\) Proof of (11.843): Let \(f \in A(P)\). We need to show that \(w(f) \in \mathcal{L}(P)\). In other words, we need to show that \(w(f)\) is a linear extension of \(P\). In order to show this, it clearly suffices to prove that \(i \prec_f j\) (for two elements \(i\) and \(j\) of \(P\)) implies \(i \prec_{w(f)} j\) (because we already know that \(w(f)\) is a totally ordered set). So, let us fix two elements \(i\) and \(j\) of \(P\) satisfying \(i \prec_f j\).

We need to prove that \(i \prec_{w(f)} j\). We have \(i \prec_f j\) and thus \(i \neq j\). Hence, either \(i \prec_{\leq} j\) or \(i \succ_{\leq} j\). In other words, we are in one of the following two Cases:

Case 1: We have \(i \leq j\).
Case 2: We have \(i \geq j\).

Let us first consider Case 1. In this case, \(i \leq j\). But \(f\) is a \(P\)-partition (since \(f \in A(P)\)), and thus, by the definition of a \(P\)-partition, we conclude that \(f(i) \leq f(j)\) (since \(i \prec_f j\) and \(i \leq j\)). In other words, either \(f(i) < f(j)\) or \(f(i) = f(j)\). Therefore, either \(f(i) < f(j)\) or \(f(i) = f(j)\) (because we have \(i \leq j\) by assumption). In other words, \(i \prec_j f\).

This rewrites as \(i \prec_{w(f)} j\) (since the relation \(\prec_{w(f)}\) is the relation \(\prec_f\)). Thus, \(i \prec_{w(f)} j\) is proven in Case 1.

Let us now consider Case 2. In this case, \(i \geq j\). But \(f\) is a \(P\)-partition (since \(f \in A(P)\)), and thus, by the definition of a \(P\)-partition, we conclude that \(f(i) < f(j)\) (since \(i \prec_f j\) and \(i \geq j\)). Therefore, either \(f(i) < f(j)\) or \(f(i) = f(j)\) (because we have \(i \leq j\) by assumption). In other words, \(i \prec_j f\).

This rewrites as \(i \prec_{w(f)} j\) (since the relation \(\prec_{w(f)}\) is the relation \(\prec_f\)). Thus, \(i \prec_{w(f)} j\) is proven in Case 2.

We have now proven \(i \prec_{w(f)} j\) in both Cases 1 and 2; thus, \(i \prec_{w(f)} j\) always holds. This completes the proof of (11.843).
Proof of (11.844): Let \( f \in \{ g \in \mathcal{A}(P) \mid w(g) = w \} \). Hence, \( f \) is an element of \( \mathcal{A}(P) \) and satisfies \( w(f) = w \). Therefore, \( w = w(f) = P \) as sets.

Our next goal is to prove \( f \in A(w) \). In other words, we want to prove that \( f \) is a \( w \)-partition.

Indeed, let \( i \) and \( j \) be two elements of \( w \). We claim that

\[
(i \prec w j \text{ and } i \preceq z j \text{ implies } f(i) \leq f(j))
\]

and

\[
(i \prec w j \text{ and } i > z j \text{ implies } f(i) < f(j)).
\]

Let us prove (11.845) first. So let \( i \) and \( j \) be two elements of \( w \). Assume that \( i \prec w j \) and \( i \preceq z j \). We have \( i \prec w j \), thus \( i \prec w(f) j \) (since \( w = w(f) \)), and thus \( i \prec f j \) (since the relation \( \prec w(f) \) is the relation \( \prec f \)). By the definition of \( \prec f \), this means that we have \( f(i) \prec f(j) \) or \( f(i) = f(j) \) and \( i \preceq z j \)). From this, we immediately obtain \( f(i) \leq f(j) \). Thus, (11.845) is proven.

The proof of (11.846) is similar to the proof that we just gave for (11.845), but with a minor twist: In order to derive \( f(i) \prec f(j) \) from \( (f(i) \prec f(j) \) or \( f(i) = f(j) \) and \( i \preceq z j \)) \), we need to recall the assumption \( i > z j \).

Thus, both (11.845) and (11.846) are proven. Now, let us forget that we fixed \( i \) and \( j \). We thus have proven that every two elements \( i \) and \( j \) of \( w \) satisfy (11.845) and (11.846). In other words, \( f \) is a \( w \)-partition. In yet other words, \( f \in A(w) \).

Let us now forget that we fixed \( f \). We thus have proven that every \( f \in \{ g \in \mathcal{A}(P) \mid w(g) = w \} \) satisfies \( f \in A(w) \). In other words, (11.844) is proven.

Let us next show that

\[
A(w) \subseteq \{ g \in \mathcal{A}(P) \mid w(g) = w \}.
\]

Proof of (11.847): Let \( f \in A(w) \). Thus, \( f \) is a \( w \)-partition. We shall next prove that \( f \in \{ g \in \mathcal{A}(P) \mid w(g) = w \} \).

Indeed, let us first make a general and trivial observation: If \( Q \) and \( R \) are two labelled posets such that \( Q = R \) as sets, and if every two elements \( i \) and \( j \) of \( Q \) satisfying \( i \prec Q j \) satisfy \( i \preceq R j \) (that is, the poset \( R \) is an extension of the poset \( Q \)), then every \( R \)-partition is a \( Q \)-partition. Applying this to \( Q = P = R \), we conclude that every \( w \)-partition is a \( P \)-partition. Thus, \( f \) is a \( P \)-partition (since \( f \) is a \( w \)-partition).

Next, we want to check that \( w(f) = w \).

It is easy to check that every two elements \( i \) and \( j \) of \( w \) satisfying \( i \prec w j \) satisfy

\[
i \prec w(f) j
\]

Now, let us first again state a triviality: If \( Q \) and \( R \) are two totally ordered sets such that \( Q = R \) as sets, and if every two elements \( i \) and \( j \) of \( Q \) satisfying \( i \prec Q j \) satisfy \( i \preceq R j \), then \( Q = R \) as totally ordered sets. Applying this to \( Q = w = \mathcal{A}(P) \) and \( R = w(f) \), we conclude that \( w = w(f) \) as totally ordered sets (since every two elements \( i \) and \( j \) of \( w \) satisfying \( i \prec w j \) satisfy \( i \prec w(f) j \)). In other words, \( w(f) = w \).

Note, we have that \( f \in \mathcal{A}(P) \) and \( w(f) = w \). In other words, \( f \in \{ g \in \mathcal{A}(P) \mid w(g) = w \} \).

Let us now forget that we fixed \( f \). We thus have shown that \( f \in \{ g \in \mathcal{A}(P) \mid w(g) = w \} \) for every \( f \in \mathcal{A}(w) \). This proves (11.847).

Combining (11.844) with (11.847), we obtain

\[
\{ g \in \mathcal{A}(P) \mid w(g) = w \} = A(w).
\]

Let us now forget that we fixed \( w \). We thus have proven (11.849) for every \( w \in \mathcal{L}(P) \).

---

642 Proof of (11.848): Let \( i \) and \( j \) be two elements of \( w \) satisfying \( i \prec w j \). We must prove that \( i \prec w(f) j \).

We have \( i \prec w j \) and thus \( i \neq j \). Hence, either \( i \preceq z j \) or \( i \succ z j \). In other words, we are in one of the following two Cases:

Case 1: We have \( i \preceq z j \).

Case 2: We have \( i \succ z j \).

Let us consider Case 1 first. In this case, we have \( i \preceq z j \). Since \( f \) is a \( w \)-partition, we have \( f(i) \leq f(j) \) (because \( i \prec w j \) and \( i \prec z j \)). In other words, \( (f(i) < f(j) \) or \( f(i) = f(j) \)). Hence, \( (f(i) < f(j) \) or \( f(i) = f(j) \) and \( i \preceq z j \)) (because we have assumed that \( i \preceq z j \)). Therefore, \( i \preceq f j \) (because of the definition of \( "i \preceq f j" \)). In other words, \( i \preceq w(f) j \) (since the relation \( \preceq_w(f) \) is the relation \( \preceq f \)). Thus, \( i \prec w(f) j \) is proven in Case 1.

We can similarly prove \( i \prec w(f) j \) in Case 2 (but now we obtain \( f(i) < f(j) \) instead of \( f(i) \leq f(j) \)).

We have now shown that \( i \prec w(f) j \) in both Cases 1 and 2. Thus, (11.848) is proven.
Now, the definition of $F_p(x)$ yields

$$F_p(x) = \sum_{f \in A(P)} \sum_{w \in \mathcal{L}(P)} x_f \sum_{f \in A(w) : w(f) = w} x_f$$

$$= \sum_{f \in A(w)} x_f = F_w(x).$$

(since $w(f) \in \mathcal{L}(P)$ for every $f \in A(P)$ (by (11.843)))

This completes our proof of Theorem 5.19. □

11.95. Solution to Exercise 5.37. Solution to Exercise 5.37. Consider the power series $\bar{H}(t)$ and $\xi(t)$ defined in (5.18). From (5.18), we know that

$$\sum_{n \geq 1} \xi_n t^n = \log \bar{H}(t) = \log \left( \sum_{n \geq 0} H_n t^n \right).$$

(a) The fact that $\xi_n$ is primitive was proven for $k = \mathbb{Q}$ in Remark 5.36, and can be proven in the same way for general $k$. It remains to show that $\xi_n$ is homogeneous of degree $n$ for each $n \geq 1$. This can be done as follows:

Let us say that a power series $f \in A[[t]]$ over a graded $k$-algebra $A$ is equigraded if, for every $n \in \mathbb{N}$, the coefficient of $t^n$ before $t^n$ is homogeneous of degree $n$. Then, the set of all equigraded power series in $A[[t]]$ is a subalgebra of $A[[t]]$ which is closed under the usual topology on $A[[t]]$. In particular, if $g$ is an equigraded power series in $A[[t]]$ having constant term 1, then $\log g$ is equigraded. Applied to $A = \operatorname{NSym}$ and $g = \sum_{n \geq 0} H_n t^n$, this yields that the power series $\log \left( \sum_{n \geq 0} H_n t^n \right)$ is equigraded. Due to (11.850), this rewrites as follows: The power series $\sum_{n \geq 1} \xi_n t^n$ is equigraded. In other words, $\xi_n$ is homogeneous of degree $n$ for each $n \geq 1$. This completes the solution of Exercise 5.37(a).

For an alternative proof of Exercise 5.37(a), one can simply notice that it follows immediately from Exercise 5.37(c).

(b) The ring homomorphism $\pi : \operatorname{NSym} \to \Lambda$ induces a ring homomorphism $\operatorname{NSym}[[t]] \to \Lambda[[t]]$ which is continuous with respect to the usual topology on power series. Applying this latter homomorphism to the equality (11.850), we obtain

$$\sum_{n \geq 1} \pi(\xi_n) t^n = \log \left( \sum_{n \geq 0} \pi(H_n) t^n \right) = \log \left( \sum_{n \geq 0} h_n t^n \right) = \sum_{m=1}^{\infty} \frac{1}{m} p_m t^m$$

(where in the last step, we have used the equality (2.29)). Comparing coefficients, we obtain $\pi(\xi_n) = \frac{1}{n} p_n$ for all $n \geq 1$. Multiplying this by $n$, we obtain $\pi(n\xi_n) = p_n$ for all $n \geq 1$. This solves part (b) of the exercise.

Remark: Another way to solve Exercise 5.37(b) proceeds as follows: We know that $\pi(n\xi_n)$ is a primitive homogeneous element of $\Lambda$ of degree $n$ (by Exercise 5.37(a) and since $\pi$ is a graded homomorphism of Hopf algebras). But Exercise 3.10 shows that all such elements are scalar multiples of $p_n$. Thus, $\pi(n\xi_n)$ is a scalar multiple of $p_n$. Finding the scalar is easy (e.g., it can be obtained by specializing at (1)).
(c) From (11.850), we have

\[
\sum_{n \geq 1} \xi_n t^n = \log \left( \sum_{n \geq 0} \frac{n}{1 + \sum_{n \geq 1} H_n t^n} \right) = \log \left( 1 + \sum_{n \geq 1} H_n t^n \right)
\]

\[
= \sum_{i \geq 1} \frac{(-1)^{i-1}}{i} \left( \sum_{n \geq 1} H_n t^n \right)^i
\]

\[
= \sum_{n_1, n_2, \ldots, n_i \geq 1} \left( H_{n_1} t^{n_1} \right) \left( H_{n_2} t^{n_2} \right) \ldots \left( H_{n_i} t^{n_i} \right)
\]

(by the Mercator series for the logarithm)

\[
= \sum_{(n_1, n_2, \ldots, n_i) \text{ is a composition of length } i} H_{(n_1, n_2, \ldots, n_i)} t^{\sum_{i=1}^{n_i} \xi_i}
\]

\[
= \sum_{(n_1, n_2, \ldots, n_i) \text{ is a composition of length } i} H_{(n_1, n_2, \ldots, n_i)} t^{\ell(n_1, n_2, \ldots, n_i)}
\]

(there, we renamed the summation index \((n_1, n_2, \ldots, n_i)\) as \(\alpha\))

\[
= \sum_{i \geq 1} \left( \sum_{\ell(\alpha) = i} \frac{(-1)^{\ell(\alpha)-1}}{\ell(\alpha)} H_\alpha t^{\ell(\alpha)} \right)
\]

\[
= \sum_{n \geq 1} \left( \sum_{\alpha \in \text{Comp}_n} \frac{(-1)^{\ell(\alpha)-1}}{\ell(\alpha)} H_\alpha \right) t^n.
\]

Comparing coefficients in this identity, we conclude that every \(n \geq 1\) satisfies

\[
\xi_n = \sum_{\alpha \in \text{Comp}_n} \frac{(-1)^{\ell(\alpha)-1}}{\ell(\alpha)} H_\alpha = \sum_{\alpha \in \text{Comp}_n} (-1)^{\ell(\alpha)-1} \frac{1}{\ell(\alpha)} H_\alpha.
\]

This solves part (c) of the exercise.
Comparing coefficients in this identity, we conclude that every element \(\xi\) (because the elements of \(\text{NSym}\) is generated by the elements \(A_\alpha\)) every \(\alpha\) (since the coefficient of \(\sum_{n \geq 1} \xi_n t^n\) family is a \(H\)-basis of \(\text{NSym}\))

\[
H(t) = \exp\left(\sum_{n \geq 1} \xi_n t^n\right) = \sum_{i \geq 0} \frac{1}{i!} \left(\sum_{n \geq 1} \xi_n t^n\right)^i
\]

\[
= \sum_{i \geq 0} \frac{1}{i!} \sum_{n_1, n_2, \ldots, n_i \geq 1} \frac{\left(\xi_{n_1} t^{n_1}\right) \left(\xi_{n_2} t^{n_2}\right) \cdots \left(\xi_{n_i} t^{n_i}\right)}{\left(n_1 + n_2 + \cdots + n_i\right)!} = \sum_{i \geq 0} \frac{1}{i!} \sum_{n \geq 1} \xi_n t^n
\]

\[
= \sum_{i \geq 0} \frac{1}{i!} \sum_{(n_1, n_2, \ldots, n_i) \text{ is a composition of length } i} \xi_{n_1} \xi_{n_2} \cdots \xi_{n_i} \frac{t^{n_1 + n_2 + \cdots + n_i}}{(n_1 + n_2 + \cdots + n_i)!} = \sum_{i \geq 0} \frac{1}{i!} \sum_{\ell(\alpha) \text{ is a composition of length } i} \xi_\alpha t^{\ell(\alpha)}\]

(here, we renamed the summation index \((n_1, n_2, \ldots, n_i)\) as \(\alpha\))

\[
= \sum_{i \geq 0} \frac{1}{i!} \sum_{\ell(\alpha) \text{ is a composition of length } i} \xi_\alpha t^{\ell(\alpha)} = \sum_{i \geq 0} \frac{1}{i!} \sum_{\ell(\alpha) \text{ is a composition of length } i} \xi_\alpha t^{\ell(\alpha)}
\]

\[
= \sum_{n \geq 0} \sum_{\alpha \in \text{Comp}_n} \frac{1}{\ell(\alpha)!} \xi_\alpha \frac{d^{\ell(\alpha)}}{dt^{\ell(\alpha)}} t^n = \sum_{n \geq 0} \sum_{\alpha \in \text{Comp}_n} \frac{1}{\ell(\alpha)!} \xi_\alpha t^n = \sum_{n \geq 0} \left(\sum_{\alpha \in \text{Comp}_n} \frac{1}{\ell(\alpha)!} \xi_\alpha \right) t^n.
\]

Comparing coefficients in this identity, we conclude that every \(n \geq 0\) satisfies

\[
H_n = \sum_{\alpha \in \text{Comp}_n} \frac{1}{\ell(\alpha)!} \xi_\alpha
\]

(since the coefficient of \(t^n\) in \(H(t)\) is \(H_n\)). This proves (5.20).

Notice that, for every \(n \geq 1\), the element \(\xi_n\) of \(\text{NSym}\) is homogeneous of degree \(n\) (by Exercise 5.37(a)). Hence, for every composition \(\alpha\), the element \(\xi_\alpha\) of \(\text{NSym}\) is homogeneous of degree \(|\alpha|\). In particular, for every \(n \in \mathbb{N}\), it is clear that \((\xi_\alpha)_{\alpha \in \text{Comp}_n}\) is a family of elements of \(\text{NSym}_n\). We now need to prove that this family is a \(k\)-basis of \(\text{NSym}_n\) for every \(n \in \mathbb{N}\).

Let \(\mathfrak{A}\) be the \(k\)-subalgebra of \(\text{NSym}\) generated by the elements \(\xi_1, \xi_2, \xi_3, \ldots\). Then, \(\mathfrak{A}\) contains \(\xi_\alpha\) for every composition \(\alpha\) (by the definition of \(\xi_\alpha\)). Therefore, \(\mathfrak{A}\) contains \(H_n\) for every \(n \geq 1\) (by (5.19)). Consequently, \(\mathfrak{A} = \text{NSym}\) (because \(\text{NSym}\) is generated as a \(k\)-algebra by \(H_1, H_2, H_3, \ldots\)). In other words, the \(k\)-algebra \(\text{NSym}\) is generated by the elements \(\xi_1, \xi_2, \xi_3, \ldots\) (since we defined \(\mathfrak{A}\) as the \(k\)-subalgebra of \(\text{NSym}\) generated by the elements \(\xi_1, \xi_2, \xi_3, \ldots\)). In other words, the \(k\)-module \(\text{NSym}\) is spanned by all possible products of the elements \(\xi_1, \xi_2, \xi_3, \ldots\). In other words, the \(k\)-module \(\text{NSym}\) is spanned by the elements \(\xi_\alpha\) with \(\alpha \in \text{Comp}\) (because the elements \(\xi_\alpha\) with \(\alpha \in \text{Comp}\) are precisely all possible products of the elements \(\xi_1, \xi_2, \xi_3, \ldots\)). In yet other words, the family \((\xi_\alpha)_{\alpha \in \text{Comp}}\) is a \(k\)-module \(\text{NSym}\).

Now, fix \(n \in \mathbb{N}\). Every element of \(\text{NSym}_n\) can be written as a \(k\)-linear combination of the elements \(\xi_\alpha\) with \(\alpha \in \text{Comp}\) (since the family \((\xi_\alpha)_{\alpha \in \text{Comp}}\) spans the \(k\)-module \(\text{NSym}\)). In this \(k\)-linear combination, we can remove all terms \(\xi_\alpha\) with \(\alpha \notin \text{Comp}_n\) without changing the result (by gradedness, because \(\xi_\alpha\) is homogeneous of degree \(|\alpha|\)), and so we conclude that every element of \(\text{NSym}_n\) can be written as a \(k\)-linear combination of the elements \(\xi_\alpha\) with \(\alpha \in \text{Comp}_n\). In other words, for every \(n \in \mathbb{N}\), the family \((\xi_\alpha)_{\alpha \in \text{Comp}_n}\) spans the \(k\)-module \(\text{NSym}_n\).
Now, we can apply Exercise 2.46(b) to \( A = \text{NSym}_n \), \( I = \text{Comp}_n \), \( (\gamma_i)_{i\in I} = (H_\alpha)_{\alpha\in\text{Comp}_n} \) and \( (\beta_i)_{i\in I} = (\xi_\alpha)_{\alpha\in\text{Comp}_n} \) (since we know that \((H_\alpha)_{\alpha\in\text{Comp}_n}\) is a \( k \)-basis of \( \text{NSym}_n \), whereas \((\xi_\alpha)_{\alpha\in\text{Comp}_n}\) spans the \( k \)-module \( \text{NSym}_n \)). We conclude that \((\xi_\alpha)_{\alpha\in\text{Comp}_n}\) is a \( k \)-basis of \( \text{NSym}_n \). Thus, Exercise 5.37(d) is solved.

11.96. **Solution to Exercise 5.38.** Solution to Exercise 5.38. We follow the hint.

Let \( f \) be the endomorphism \( \text{id} - u \cdot e \) of \( A \). Then, \( f = \sum_{n \geq 1} \pi_n \) (because the definition of the \( \pi_n \) yields that \( \text{id} = \sum_{n \geq 0} \pi_n = \pi_0 + \sum_{n \geq 1} \pi_n = u \cdot e + \sum_{n \geq 1} \pi_n \) and thus \( \text{id} - u \cdot e = \sum_{n \geq 1} \pi_n \)). Notice that \( f = \text{id} - u \cdot e \), so that \( \text{id} = f + u \cdot e \). Now, \( \epsilon = \log^* \left( \frac{\text{id}_A}{f + u \cdot e} \right) = \log^* (f + u \cdot e) = \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^{*n} \).

We let \( \text{End}_{gr} A \) be the \( k \)-submodule of \( \text{End} A \) consisting of all graded \( k \)-linear maps \( A \to A \). Then, it is easy to see that this \( k \)-submodule \( \text{End}_{gr} A \) is closed under convolution (i.e., if \( g_1 \) and \( g_2 \) are two graded \( k \)-linear maps \( A \to A \), then \( g_1 \star g_2 \) is also a graded \( k \)-linear map \( A \to A \)) and contains the unity \( u \cdot e \) of the algebra \( \text{End} A, \star \) (since \( u \cdot e : A \to A \) is a graded \( k \)-linear map). Hence, \( \text{End}_{gr} A, \ast \) is a \( k \)-subalgebra of \( (\text{End} A, \star) \). Moreover, \( \text{End}_{gr} A, \ast \) contains \( \pi_n \) for every \( n \geq 1 \). Hence, the \( k \)-algebra homomorphism \( \mathcal{W} : \text{NSym} \to (\text{End} A, \ast) \) maps the generators \( H_\alpha \) of \( \text{NSym} \) to elements of \( (\text{End}_{gr} A, \ast) \). Thus, the image of \( \mathcal{W} \) is contained in \( (\text{End}_{gr} A, \ast) \). In other words,

\begin{equation}
\mathcal{W} (x) \in \text{End}_{gr} A \quad \text{for every } x \in \text{NSym}.
\end{equation}

It is immediate to check that the \( k \)-subalgebra \( (\text{End}_{gr} A, \ast) \) of \( (\text{End} A, \star) \) is closed under the topology of pointwise convergence. Hence, it is closed under taking logarithms. In other words, \( \log^* g \in \text{End}_{gr} A \) for every \( g \in \text{End}_{gr} A \) for which \( \log^* g \) makes sense. Applied to \( g = \text{id} A \), this yields that \( \log^* (\text{id} A) \in \text{End}_{gr} A \) (since \( \text{id} A \in \text{End}_{gr} A \)). Since \( \log^* (\text{id} A) = \epsilon \), this yields that \( \epsilon \in \text{End}_{gr} A \). In other words, \( \epsilon \) is a graded \( k \)-linear map. The definition of \( \epsilon_n \) is legitimate because the gradedness of \( \epsilon \) yields \( \pi_n \circ \epsilon = \epsilon \circ \pi_n \). This solves Exercise 5.38(a).

We notice for future use the fact that \( \epsilon_0 = 0 \).

We also record the fact that

\begin{equation}
\Delta_A \circ \pi_n = \left( \sum_{k=0}^n \pi_k \otimes \pi_{n-k} \right) \circ \Delta_A \quad \text{for all } n \in \mathbb{N}.
\end{equation}

(11.852)

(This follows by checking that both sides of (11.852) are equal to \( \Delta_A \) on the \( n \)-th homogeneous component \( A_n \), while vanishing on all other components.)

\(\text{Proof.}\) The definition of \( \epsilon_0 \) yields \( \epsilon_0 = \pi_0 \circ \epsilon = \epsilon \circ \pi_0 \). Thus, \( \epsilon_0 (A) = (\epsilon \circ \pi_0) (A) = \epsilon \left( \begin{aligned} \pi_0 (A) \\
\end{aligned} \right) = \epsilon (k \cdot 1_A) = k \cdot \epsilon (1_A) \). But every \( n \geq 1 \) satisfies \( f^{*n} (1_A) = 0 \) (since \( \Delta^{(n-1)} (1_A) = \underbrace{1_A \otimes 1_A \otimes \ldots \otimes 1_A}_{n \text{ times}} \) and \( f (1_A) = 0 \)). Since \( \epsilon = \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^{*n} \), we have

\(\epsilon (1_A) \left( \begin{aligned} \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^{*n} \\
\end{aligned} \right) (1_A) = \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^{*n} (1_A) = \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} 0 = 0.\)

Now, \( \epsilon_0 (A) = k \cdot \epsilon (1_A) = 0 \), so that \( \epsilon_0 = 0 \), qed.

\(\text{And this is because } \Delta_A \text{ is graded.}\)
(b) From (11.850), we have

\[
\sum_{n \geq 1} \xi_n t^n = \log \left( \sum_{n \geq 0} H_n t^n \right) = \log \left( 1 + \sum_{n \geq 1} H_n t^n \right)
\]

\[
= \sum_{i \geq 1} (-1)^{i-1} \frac{1}{i} \left( \sum_{n \geq 1} H_n t^n \right)^i
\]

(by the Mercator series for the logarithm)

\[
= \sum_{i \geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1} H_{n_1} H_{n_2} \ldots H_{n_i} \sum_{n_1 + n_2 + \ldots + n_i = n} H_{n_1} H_{n_2} \ldots H_{n_i} t^n
\]

in the power series ring \( \text{NSym} [t] \). Comparing coefficients in this equality, we obtain

\[
(11.853) \quad \xi_n = \sum_{i \geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1; n_1 + n_2 + \ldots + n_i = n} H_{n_1} H_{n_2} \ldots H_{n_i}
\]

for every \( n \geq 1 \). Hence, every \( n \geq 1 \) satisfies

\[
\mathcal{M}(\xi_n) = \mathcal{M} \left( \sum_{i \geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1; n_1 + n_2 + \ldots + n_i = n} H_{n_1} H_{n_2} \ldots H_{n_i} \right)
\]

\[
= \sum_{i \geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1; n_1 + n_2 + \ldots + n_i = n} \mathcal{M}(H_{n_1}) \ast \mathcal{M}(H_{n_2}) \ast \ldots \ast \mathcal{M}(H_{n_i})
\]

(since \( \mathcal{M} \) is a \( k \)-algebra homomorphism)

\[
(11.854) \quad = \sum_{i \geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1; n_1 + n_2 + \ldots + n_i = n} \pi_{n_1} \ast \pi_{n_2} \ast \ldots \ast \pi_{n_i}
\]

(since \( \mathcal{M} \) maps every \( H_m \) to \( \pi_m \)).

On the other hand, \( f = \sum_{n \geq 1} \pi_n \). Thus,

\[
f^*i = \left( \sum_{n \geq 1} \pi_n \right)^*i = \sum_{n_1, n_2, \ldots, n_i \geq 1} \pi_{n_1} \ast \pi_{n_2} \ast \ldots \ast \pi_{n_i}
\]

for every \( i \in \mathbb{N} \). Now,

\[
\epsilon = \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^*n = \sum_{i \geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1} f^*i = \sum_{n_1, n_2, \ldots, n_i \geq 1} \pi_{n_1} \ast \pi_{n_2} \ast \ldots \ast \pi_{n_i}
\]

\[
(11.855) \quad = \sum_{i \geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1} \pi_{n_1} \ast \pi_{n_2} \ast \ldots \ast \pi_{n_i}.
\]
Now,
\[
\epsilon_n = \pi_n \circ \epsilon
\]
\[
= \pi_n \circ \left( \sum_{i \geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1} \pi_{n_1} \ast \pi_{n_2} \ast \ldots \ast \pi_{n_i} \right) \quad \text{(by (11.855))}
\]
\[
= \sum_{i \geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1} \pi_n \circ \left( \pi_{n_1} \ast \pi_{n_2} \ast \ldots \ast \pi_{n_i} \right)
\]
\[
= \sum_{i \geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1; n_1 + n_2 + \ldots + n_i = n} \pi_n \circ \left( \pi_{n_1} \ast \pi_{n_2} \ast \ldots \ast \pi_{n_i} \right)
\]
\[
= \mathcal{M}(\xi_n) + \sum_{i \geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1; n_1 + n_2 + \ldots + n_i \neq n} \pi_n \circ \left( \pi_{n_1} \ast \pi_{n_2} \ast \ldots \ast \pi_{n_i} \right)
\]
\[
= \mathcal{M}(\xi_n) \quad \text{(by (11.854))}
\]
\[
+ \sum_{i \geq 1} (-1)^{i-1} \frac{1}{i} \sum_{n_1, n_2, \ldots, n_i \geq 1; n_1 + n_2 + \ldots + n_i \neq n} 0 = \mathcal{M}(\xi_n).
\]

This solves Exercise 5.38(b).

(c) Exercise 1.59(a) (applied to \( C = A \)) yields that the comultiplication \( \Delta_A : A \to A \otimes A \) is a \( k \)-coalgebra homomorphism (since \( A \) is cocommutative). Also, the comultiplications \( \Delta_A \) and \( \Delta_{\text{NSym}} \) are \( k \)-algebra homomorphisms (as is the comultiplication of any \( k \)-bialgebra).

Let \( R \) be the \( k \)-linear map \((\text{End } A, \ast) \otimes (\text{End } A, \ast) \to (\text{End } A \otimes A, \ast) \) which sends every tensor \( f \otimes g \in (\text{End } A, \ast) \otimes (\text{End } A, \ast) \) to the map \( f \otimes g : A \otimes A \to A \otimes A \). This map \( R \) is a \( k \)-algebra homomorphism\(^{645}\).

Let \( \Omega_1 \) be the composition
\[
\text{NSym} \xrightarrow{\mathcal{M}} \text{End } A, \ast \xrightarrow{\text{post}(\Delta_A)} \text{Hom } (A, A \otimes A, \ast),
\]
where post(\( \Delta_A \)) denotes the \( k \)-linear map sending every \( \gamma \in (\text{End } A, \ast) \) to \( \Delta_A \circ \gamma \in (\text{Hom } (A, A \otimes A), \ast) \). Since the map \( \mathcal{M} \) is a \( k \)-algebra homomorphism, and since the map post(\( \Delta_A \)) is a \( k \)-algebra homomorphism\(^{646}\), their composition \( \Omega_1 \) also is a \( k \)-algebra homomorphism.

Let \( \Omega_2 \) be the composition
\[
\text{NSym} \xrightarrow{\Delta_{\text{NSym}}} \text{NSym} \otimes \text{NSym} \xrightarrow{\mathcal{M} \otimes \mathcal{M}} \text{End } (A, \ast) \otimes (A, \ast) \xrightarrow{R} \text{End } (A \otimes A, \ast),
\]
where pre(\( \Delta_A \)) denotes the \( k \)-linear map sending every \( \gamma \in (\text{End } A \otimes A, \ast) \) to \( \gamma \circ \Delta_A \in (\text{Hom } (A, A \otimes A), \ast) \). Since the maps \( \Delta_{\text{NSym}}, \mathcal{M} \otimes \mathcal{M} \) and \( R \) are \( k \)-algebra homomorphisms, and since the map pre(\( \Delta_A \)) is a \( k \)-algebra homomorphism\(^{647}\), their composition \( \Omega_2 \) also is a \( k \)-algebra homomorphism.

In order to solve Exercise 5.38(c), we need to show that every \( w \in \text{NSym} \) satisfies \( \Delta \circ (\mathcal{M}(w)) = \left( \sum_{(w)} \mathcal{M}(w_1) \otimes \mathcal{M}(w_2) \right) \circ \Delta \). Since
\[
\Delta \circ (\mathcal{M}(w)) = \Delta_A \circ (\mathcal{M}(w)) = (\text{post}(\Delta_A))(\mathcal{M}(w)) = \underbrace{(\text{post}(\Delta_A)) \circ \mathcal{M}}_{= \Omega_1}(w) = \Omega_1(w)
\]

\(^{645}\)In fact, this is a particular case of Exercise 1.29(b).

\(^{646}\)This is a particular case of the statement that \( \varphi \) is a convolution algebra morphism in the proof of Proposition 1.46 (because we know that \( \Delta_A \) is a \( k \)-algebra homomorphism).

\(^{647}\)This is a particular case of the statement that \( \varphi \) is a convolution algebra morphism in the proof of Proposition 1.46 (because we know that \( \Delta_A \) is a \( k \)-algebra homomorphism).
and

\[
\left( \sum_{(w)} (w_1) \otimes (w_2) \right) \circ \Delta_{\Delta} = R ( (\mathcal{W} \otimes \mathcal{W}) (\Delta_{\text{NSym}} (w))) \circ \Delta_A \\
= \text{(pre} (\Delta_A)) \left( R ( (\mathcal{W} \otimes \mathcal{W}) (\Delta_{\text{NSym}} (w))) \right) \\
= (\text{pre} (\Delta_A)) \circ R \circ (\mathcal{W} \otimes \mathcal{W}) \circ \Delta_{\text{NSym}} (w) = \Omega_2 (w),
\]

this is equivalent to showing that every \( w \in \text{NSym} \) satisfies \( \Omega_1 (w) = \Omega_2 (w) \). In other words, we need to prove that \( \Omega_1 = \Omega_2 \). Since \( \Omega_1 \) and \( \Omega_2 \) are \( k \)-algebra homomorphisms, it will be enough to verify this on the generators \( H_1, H_2, H_3, \ldots \) of the \( k \)-algebra \( \text{NSym} \). But on said generators, this is easily seen to hold, because every \( n \geq 1 \) satisfies

\[
\Omega_1 (H_n) = \text{(post} (\Delta_A) \circ \mathcal{W}) (H_n) = \text{(post} (\Delta_A) \left( \mathcal{W} (H_n) \right) \\
= \text{(post} (\Delta_A) (\pi_n) = \Delta_A \circ \pi_n = \left( \sum_{k=0}^{n} \pi_k \otimes \pi_{n-k} \right) \circ \Delta_A \quad \text{(by (11.852))}
\]

and

\[
\Omega_2 (H_n) = (\text{pre} (\Delta_A) \circ R \circ (\mathcal{W} \otimes \mathcal{W}) \circ \Delta_{\text{NSym}}) (H_n) \\
= (\text{pre} (\Delta_A) \circ R \circ (\mathcal{W} \otimes \mathcal{W}) \left( \Delta_{\text{NSym}} (H_n) \right) \\
= (\text{pre} (\Delta_A) \circ R \circ (\mathcal{W} \otimes \mathcal{W}) \left( \sum_{k=0}^{n} H_k \otimes H_{n-k} \right) \\
= (\text{pre} (\Delta_A) \circ R \left( \mathcal{W} \otimes \mathcal{W} \left( \sum_{k=0}^{n} H_k \otimes H_{n-k} \right) \right) \\
= (\text{pre} (\Delta_A) \circ R \left( \sum_{k=0}^{n} \mathcal{W} (H_k) \otimes \mathcal{W} (H_{n-k}) \right) \\
= (\text{pre} (\Delta_A) \circ R \left( \sum_{k=0}^{n} \pi_k \otimes \pi_{n-k} \right) \\
= (\text{pre} (\Delta_A)) \left( R \left( \sum_{k=0}^{n} \pi_k \otimes \pi_{n-k} \right) \right) = R \left( \sum_{k=0}^{n} \pi_k \otimes \pi_{n-k} \right) \circ \Delta_A \\
= \left( \sum_{k=0}^{n} \pi_k \otimes \pi_{n-k} \right) \circ \Delta_A.
\]

Exercise 5.38(c) is proven.

(d) Let \( n \geq 0 \). We need to prove that \( e_n (A) \subset p \). If \( n = 0 \), then this is clear because \( e_0 = 0 \). Hence, we assume WLOG that we don’t have \( n = 0 \). Thus, \( n \geq 1 \). Hence, Exercise 5.37(a) yields that \( \pi_n \) is primitive, so that \( \Delta (\pi_n) = \pi_n \otimes 1 + 1 \otimes \pi_n \).
Applying Exercise 5.38(c) to $w = \pi_n$, we obtain

$$
\Delta \circ (\mathcal{W} (\pi_n)) = \left( \sum_{(\pi_n)} \mathcal{W} ((\pi_n)_1) \otimes \mathcal{W} ((\pi_n)_2) \right) \circ \Delta \\
= (\mathcal{W} (\pi_n) \otimes \mathcal{W} (1) + \mathcal{W} (1) \otimes \mathcal{W} (\pi_n)) \circ \Delta \\
= \left( \text{since } \sum_{(\pi_n)} (\pi_n)_1 \otimes (\pi_n)_2 = \Delta (\pi_n) = \pi_n \otimes 1 + 1 \otimes \pi_n \right).
$$

Since $\mathcal{W} (1) = u\epsilon$ and $\mathcal{W} (\pi_n) = \epsilon_n$, this rewrites as

$$
\Delta \circ \epsilon_n = (\epsilon_n \otimes (u\epsilon) + (u\epsilon) \otimes \epsilon_n) \circ \Delta.
$$

Now, let $x \in A$. Then,

$$
\Delta (\epsilon_n (x)) = \left( \Delta \circ \epsilon_n \right) (x) = ((\epsilon_n \otimes (u\epsilon) + (u\epsilon) \otimes \epsilon_n) \circ \Delta) (x)
$$

Then,

$$
\Delta (\epsilon_n (x)) = \sum_{(x)} \epsilon_n (x_1) \otimes (u\epsilon) (x_2) + \sum_{(x)} (u\epsilon) (x_1) \otimes \epsilon_n (x_2)
$$

$$
= \sum_{(x)} \epsilon_n (x_1) \otimes 1_A (x_2) + \sum_{(x)} 1_A (x_1) \otimes \epsilon_n (x_2)
$$

$$
= \epsilon_n (x) \otimes 1_A + 1_A \otimes \epsilon_n (x).
$$

Hence, $\epsilon_n (x)$ is a primitive element of $A$; thus, $\epsilon_n (x) \in \mathfrak{p}$. Forget now that we fixed $x$. We thus have seen that every $x \in A$ satisfies $\epsilon_n (x) \in \mathfrak{p}$. In other words, $\epsilon_n (A) \subseteq \mathfrak{p}$. This solves Exercise 5.38(d).

(c) Every $n \geq 0$ satisfies $\epsilon_n (A) \subseteq \mathfrak{p}$ (by Exercise 5.38(d)).

We have $\text{id}_A = \sum_{n \geq 0} \pi_n$ (by the definition of $\pi_n$), and

$$
\epsilon = \frac{\text{id}_A}{\sum_{n \geq 0} \pi_n} \circ \epsilon = \sum_{n \geq 0} \frac{\pi_n \circ \epsilon}{\pi_n} = \sum_{n \geq 0} \epsilon_n.
$$

Hence, $\epsilon (A) = \left( \sum_{n \geq 0} \epsilon_n \right) (A) \subseteq \sum_{n \geq 0} \epsilon_n (A) \subseteq \sum_{n \geq 0} \mathfrak{p} \subseteq \mathfrak{p}$. This proves part (e) of the exercise.

(f) Let $x \in \mathfrak{p}$. Then, $\Delta (x) = x \otimes 1_A + 1_A \otimes x$.

Let $g (A, A)$ denote the $k$-submodule of $\text{End } A$ which consists of all $g \in \text{End } A$ satisfying $g (1_A) = 0$. Then, it is easy to see that $g (A, A)$ is an ideal of the algebra $(\text{End } A, \star)$ (indeed, any $g_1 \in \text{End } A$ and $g_2 \in g (A, A)$ satisfy $g_1 \star g_2 \in g (A, A)$ and $g_2 \star g_1 \in g (A, A)$). We have $f \in g (A, A)$. Since $g (A, A)$ is an ideal of $(\text{End } A, \star)$, this yields that every $n \geq 2$ satisfies $f^{*n} \in (g (A, A))^{*2}$ (where $I^{*2}$ denotes the square of the ideal $I$ for any ideal $I$ of the $k$-algebra $(\text{End } A, \star)$). But it is easy to see that

(11.856) $g (x) = 0$ for every $g \in (g (A, A))^{*2}$. 

In particular, this yields that
\[(11.857) \quad f^{*n}(x) = 0 \quad \text{for every } n \geq 2\]
(since \(f^{*n} \in (g(A,A))^2\) for every \(n \geq 2\)). Now, \(e = \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^{*n}\), so that
\[
\epsilon(x) = \left( \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^{*n} \right)(x) = \sum_{n \geq 1} (-1)^{n-1} \frac{1}{n} f^{*n}(x)
\]
\[
= (-1)^{i-1} \frac{1}{1} = f = \text{id}_A - u(e) (x) + \sum_{n \geq 2} (-1)^{n-1} \frac{1}{n} f^{*n}(x)
\]
(by \(11.857\))
\[
= (\text{id}_A - u(e))(x) + \sum_{n \geq 2} (-1)^{n-1} \frac{1}{n} = (\text{id}_A - u(e))(x) = x - u \epsilon(x) \quad (\text{by Proposition 1.37})
\]
\[
= x - u(0) = x.
\]

Now, forget that we fixed \(x\). We thus have shown that \(\epsilon(x) = x\) for every \(x \in p\). In other words, the map \(\epsilon\) fixes any element of \(p\). This solves Exercise 5.38(f).

Combining the results of parts (e) and (f) of Exercise 5.38, we conclude that \(\epsilon\) is a projection from \(A\) to the \(k\)-submodule \(p\). This completes the solution of the exercise.

11.97. Solution to Exercise 5.40. Solution to Exercise 5.40. Let us first show the following two lemmas, which have nothing to do with Hopf algebras:

Lemma 11.27. Let \(V\) be any torsionfree abelian group (written additively). Let \(N \in \mathbb{N}\). For every \(k \in \{0,1,\ldots,N\}\), let \(w_k\) be an element of \(V\). Assume that
\[
(11.858) \quad \sum_{k=0}^{N} w_k n^k = 0 \quad \text{for all } n \in \mathbb{N}.
\]
Then, \(w_k = 0\) for every \(k \in \{0,1,\ldots,N\}\).

Lemma 11.28. Let \(V\) be any torsionfree abelian group (written additively). Let \(N \in \mathbb{N}\). For every \((k,\ell) \in \{0,1,\ldots,N\}^2\), let \(v_{k,\ell}\) be an element of \(V\). Assume that
\[
(11.859) \quad \sum_{k=0}^{N} \sum_{\ell=0}^{N} v_{k,\ell} n^k m^\ell = 0 \quad \text{for all } n \in \mathbb{N} \text{ and } m \in \mathbb{N}.
\]
Then, \(v_{k,\ell} = 0\) for every \((k,\ell) \in \{0,1,\ldots,N\}^2\).

Proof of Lemma 11.27. The matrix \(\left( i^{N+1-j} \right)_{i,j=1,2,\ldots,N+1} \in \mathbb{Q}^{(N+1) \times (N+1)}\) is invertible (since its determinant is the Vandermonde determinant \(\prod_{1 \leq i < j \leq N+1} (i-j) \neq 0\)). Let \(\left( s_{i,j} \right)_{i,j=1,2,\ldots,N+1} \in \mathbb{Q}^{(N+1) \times (N+1)}\) be its inverse matrix. Then, \(\left( s_{i,j} \right)_{i,j=1,2,\ldots,N+1} \cdot \left( i^{N+1-j} \right)_{i,j=1,2,\ldots,N+1} = I_{N+1}\) (the \((N+1) \times (N+1)\) identity matrix) for all \(i,j \in \{1,2,\ldots,N+1\}\), so that
\[
\sum_{i=0}^{N} w_i n^i = 0 \quad \Rightarrow \quad \sum_{i=0}^{N} \sum_{j=0}^{N} s_{i,j} n^i m^j = 0
\]
for all \(n \in \mathbb{N}\).

This proves (11.856).

\footnote{Proof of (11.856): It is clearly enough to check that \((g_1 * g_2)(x) = 0\) for every \(g_1 \in g(A,A)\) and \(g_2 \in g(A,A)\). But this is easy: If \(g_1 \in g(A,A)\) and \(g_2 \in g(A,A)\), then \(g_1(1_A) = 0\) and \(g_2(1_A) = 0\), so that
\[
(g_1 * g_2)(x) = g_1(x) g_2(1_A) + g_1(1_A) g_2(x) = 0 + 0 = 0.
\]
This proves (11.856).}
matrix). Comparing the entries on both sides of this equality, we see: For every \((u, v) \in \{1, 2, ..., N + 1\}^2\), we have

\[
\sum_{j=1}^{N+1} s_{u,j} j^{N+1-v} = \delta_{u,v}
\]

(because \(\sum_{j=1}^{N+1} s_{u,j} j^{N+1-v}\) is the \((u, v)\)-th entry of the matrix \((s_{i,j})_{i,j=1,2,...,N+1} : (i^{N+1-j})_{i,j=1,2,...,N+1}\), while \(\delta_{u,v}\) is the \((u, v)\)-th entry of the matrix \(I_{N+1}\)).

The entries \(s_{i,j}\) of the matrix \((s_{i,j})_{i,j=1,2,...,N+1}\) are rational numbers, and therefore there exists a positive integer \(M\) such that every \((i, j) \in \{1, 2, ..., N + 1\}^2\) satisfies \(M s_{i,j} \in \mathbb{Z}\) (because finitely many rational numbers always have a common denominator). Consider this \(M\). For every \((i, j) \in \{1, 2, ..., N + 1\}^2\), define an element \(t_{i,j} \in \mathbb{Z}\) by \(t_{i,j} = M s_{i,j}\). Then,

\[
(11.860) \quad \sum_{j=1}^{N+1} t_{u,j} j^{N+1-v} = M \sum_{j=1}^{N+1} s_{u,j} j^{N+1-v} = M \delta_{u,v}
\]

for every \((u, v) \in \{1, 2, ..., N + 1\}^2\).

Now, let \(i \in \{0, 1, ..., N\}\). Then,

\[
\sum_{j=1}^{N+1} t_{N+1-i,j} \sum_{k=0}^{N} w_k j^k = \sum_{k=0}^{N} \left( \sum_{j=1}^{N+1} t_{N+1-i,j} j^k \right) w_k = M \sum_{k=0}^{N} \delta_{N+1-i,N+1-k} w_k = \delta_{i,k}
\]

(by \(11.860\), applied to \(u=N+1-i\) and \(v=N+1-k\))

\[
= M \sum_{k=0}^{N} \delta_{i,k} w_k = M w_i.
\]

Hence,

\[
M w_i = \sum_{j=1}^{N+1} t_{N+1-i,j} \sum_{k=0}^{N} w_k j^k = \sum_{j=1}^{N+1} t_{N+1-i,j} 0 = 0.
\]

(by \(11.858\), applied to \(n=j\))

Since \(M\) is a positive integer, we can cancel \(M\) from this equality (because \(V\) is torsionfree), and thus obtain \(w_i = 0\).

Now forget that we fixed \(i\). We thus have proven that \(w_i = 0\) for every \(i \in \{0, 1, ..., N\}\). In other words, \(w_k = 0\) for every \(k \in \{0, 1, ..., N\}\). Lemma 11.27 is thus proven.

\(\square\)

**Proof of Lemma 11.28.** Fix \(m \in \mathbb{N}\). Every \(n \in \mathbb{N}\) satisfies

\[
\sum_{k=0}^{N} \left( \sum_{\ell=0}^{N} v_{k,\ell} m^\ell \right) n^k = \sum_{k=0}^{N} \sum_{\ell=0}^{N} v_{k,\ell} m^\ell n^k = 0.
\]

Thus, Lemma 11.27 (applied to \(w_k = \sum_{\ell=0}^{N} v_{k,\ell} m^\ell\)) yields that \(\sum_{\ell=0}^{N} v_{k,\ell} m^\ell = 0\) for every \(k \in \{0, 1, ..., N\}\).

Now, forget that we fixed \(m\). We thus have proven that

\[
(11.861) \quad \sum_{\ell=0}^{N} v_{k,\ell} m^\ell = 0 \quad \text{for every } m \in \mathbb{N} \text{ and } k \in \{0, 1, ..., N\}.
\]
Now, fix $g \in \{0,1,\ldots,N\}$. For every $n \in \mathbb{N}$, we have
\[
\sum_{k=0}^{N} v_{g,k} k^k = \sum_{\ell=0}^{N} v_{g,\ell} \ell^\ell \quad \text{(here, we renamed the summation index $k$ as $\ell$)}
\]
\[
= 0 \quad \text{(by (11.861), applied to $k = g$ and $n = m$).}
\]
Hence, Lemma 11.27 (applied to $w_k = v_{g,k}$) yields that $v_{g,k} = 0$ for every $k \in \{0,1,\ldots,N\}$.

Now, forget that we fixed $g$. We thus have shown that $v_{g,k} = 0$ for every $g \in \{0,1,\ldots,N\}$ and $k \in \{0,1,\ldots,N\}$. Renaming the indices $g$ and $k$ as $k$ and $\ell$ in this statement, we obtain the following: We have $v_{k,\ell} = 0$ for every $k \in \{0,1,\ldots,N\}$ and $\ell \in \{0,1,\ldots,N\}$. In other words, $v_{k,\ell} = 0$ for every $(k,\ell) \in \{0,1,\ldots,N\}^2$. This proves Lemma 11.28. \(\square\)

Now, let us come to the solution of Exercise 5.40.

Define $\text{End}_{gr} A$ as in the solution of Exercise 5.38. We can prove (just as in the solution of Exercise 5.38) that $(\text{End}_{gr} A, \star)$ is a $k$-subalgebra of $(\text{End} A, \star)$. Since $\text{id}_A \in \text{End}_{gr} A$, we thus have $\text{id}_A^\ell \in \text{End}_{gr} A$ for every $\ell \in \mathbb{N}$. In other words,
\[
\text{id}_A^\ell \text{ is a graded $k$-linear map for every $\ell \in \mathbb{N}$.}
\]

We have $\varepsilon (1_A) = 0$ (as was proven in a footnote in the solution of Exercise 5.38). Hence, $\varepsilon \left( \begin{array}{c} A_0 \\ \vdots \\ \text{id}_A \\ \end{array} \right) = 0$.

Notice that $A$ is a $k$-module, hence a $\mathbb{Q}$-module (since $\mathbb{Q}$ is a subring of $k$), thus a torsionfree abelian group.

(a) For every $k$-linear map $f : A \to A$ which annihilates $A_0$, we can define an endomorphism $\exp^* f$ of $A$ by setting $\exp^* f = \sum_{\ell \geq 0} \frac{1}{\ell!} f^\ell$. (This follows from the same argument as the well-definedness of $\log^* (f + u\epsilon)$.) The usual rules for exponentials and logarithms apply: we have $\exp^* \log^* (f + u\epsilon) = f + u\epsilon$ and $\log^* \exp^* f = f$ for every $k$-linear map $f : A \to A$ which annihilates $A_0$. We also have $\exp^* (f + g) = \exp^* f \cdot \exp^* g$ for any two $k$-linear maps $f : A \to A$ and $g : A \to A$ which annihilate $A_0$ and satisfy $f \star g = g \star f$. Thus, $\exp^* (n f) = (\exp^* f)^{\ast n}$ for every $n \in \mathbb{N}$ and any $k$-linear map $f : A \to A$ which annihilates $A_0$.

The map $\varepsilon$ annihilates $A_0$ (since $\varepsilon (A_0) = 0$), and thus an endomorphism $\exp^* \varepsilon$ of $A$ is well-defined. We have $\varepsilon = \log^* (\text{id}_A)$, so that $\exp^* \varepsilon = \text{id}_A$.

Let us recall that any $f$ annihilating $A_0$ has the property that for each $n$ one has that $A_n$ is annihilated by $f^{\ast m}$ for every $m > n$ (we saw this in the proof of Proposition 1.44). Applying this to $f = \varepsilon$, we see that every $n \in \mathbb{N}$ satisfies
\[
\varepsilon^{\ast n} (A_N) = 0 \quad \text{for every $N \in \mathbb{N}$ and every $n \in \mathbb{N}$ satisfying $n > N$.}
\]

Recall that $\exp^* (n f) = (\exp^* f)^{\ast n}$ for every $n \in \mathbb{N}$ and any $k$-linear map $f : A \to A$ which annihilates $A_0$. Applying this to $f = \varepsilon$, we see that every $n \in \mathbb{N}$ satisfies
\[
\exp^* (n \varepsilon) = \left( \begin{array}{c} \exp^* \\ \varepsilon \\ \text{id}_A \\ \end{array} \right)^{\ast n} = \text{id}_A^{\ast n}.
\]

Now, let us fix $M \in \mathbb{N}$. Let also $N$ be any integer satisfying $N \geq M$. It is easy to see that
\[
\text{id}_A^{\ast n} (v) = \sum_{\ell=0}^{N} \frac{1}{\ell!} m^\ell \varepsilon^{\ast \ell} (v) \quad \text{for every $v \in A_M$ and every $m \in \mathbb{N}$.}
\]
Now, let \( v \in A_M \), \( n \in \mathbb{N} \) and \( m \in \mathbb{N} \) be arbitrary. Since \( \text{id}_A^m \) is a graded map (by (11.862), applied to \( \ell = m \)), we have \( \text{id}_A^m (v) \in A_M \) (since \( v \in A_M \)). Hence, (11.865) (applied to \( n \) and \( \text{id}_A^m (v) \) instead of \( m \) and \( v \)) yields

\[
\text{id}_A^n (\text{id}_A^m (v)) = \sum_{\ell=0}^{N} \frac{1}{\ell!} n^\ell e^{\ell} (\text{id}_A^m (v)) = \sum_{k=0}^{N} \frac{1}{k!} n^k e^k \left( \text{id}_A^m (v) \right) = \sum_{\ell=0}^{N} \frac{1}{\ell!} m^\ell e^{\ell} (v) \quad \text{(by (11.865)})
\]

(here, we have renamed the summation index \( \ell \) as \( k \))

\[
\sum_{k=0}^{N} \frac{1}{k!} n^k \left( \sum_{\ell=0}^{N} \frac{1}{\ell!} m^\ell e^{\ell} (v) \right) = \sum_{\ell=0}^{N} \frac{1}{\ell!} m^\ell e^k (e^k (v)) \quad \text{(since \( e^k \) is a \( k \)-linear map)}
\]

As a result,

\[
\sum_{\ell=0}^{N} \frac{1}{\ell!} n^\ell \left( e^k \circ e^\ell \right) (v) = \sum_{k=0}^{N} \sum_{\ell=0}^{N} \left( e^k \circ e^\ell \right) (v) \frac{n^k m^\ell}{k! \ell!}.
\]

But

\[
\text{id}_A^n (\text{id}_A^m (v)) = \frac{(\text{id}_A^n \circ \text{id}_A^m) (v)}{\text{id}_A^{nm} (v)} = \frac{(nm) \ell e^\ell (v)}{\text{id}_A^{nm} (v)} \quad \text{(by (11.865), applied to \( nm \) instead of \( m \))},
\]

649**Proof of (11.865):** Let \( v \in A_M \) and \( m \in \mathbb{N} \). Every integer \( \ell \geq M \) satisfies \( e^\ell \left( \begin{array}{c} v \\ \in A_M \end{array} \right) = 0 \) (by (11.863), applied to \( \ell \) and \( M \) instead of \( n \) and \( N \)). In other words,

\[(11.866) \quad \text{every integer } \ell \text{ satisfying } \ell > M \text{ satisfies } e^\ell (v) = 0.
\]

From (11.864) (applied to \( n \) instead of \( m \)), we obtain \( \exp^* (me) = \text{id}_A^m \), so that

\[
\text{id}_A^m = \exp^* (me) = \sum_{\ell=0}^{N} \frac{1}{\ell!} (me)^\ell = \sum_{\ell=0}^{N} \frac{1}{\ell!} m^\ell e^\ell
\]

and thus

\[
\text{id}_A^m (v) = \sum_{\ell=0}^{N} \frac{1}{\ell!} m^\ell e^\ell (v) = \sum_{\ell=0}^{N} \frac{1}{\ell!} m^\ell e^\ell (v) + \sum_{\ell=N+1}^{\infty} \frac{1}{\ell!} m^\ell e^\ell (v) \quad \text{(by (11.866), since } \ell > N \geq M)\]

\[
= \sum_{\ell=0}^{N} \frac{1}{\ell!} m^\ell e^\ell (v) + \sum_{\ell=N+1}^{\infty} \frac{1}{\ell!} m^\ell 0 = \sum_{\ell=0}^{N} \frac{1}{\ell!} m^\ell e^\ell (v).
\]

This proves (11.865).
Now,
\[
\sum_{k=0}^{N} \sum_{\ell=0}^{N} \left( \frac{(e^{*k} \circ e^{*\ell})(v)}{k! \ell!} - \frac{\delta_{k,\ell} e^{*k}(v)}{k!} \right) n^{k} m^{\ell} = \sum_{k=0}^{N} \sum_{\ell=0}^{N} \frac{(e^{*k} \circ e^{*\ell})(v)}{k! \ell!} n^{k} m^{\ell} - \sum_{k=0}^{N} \frac{\delta_{k,\ell} e^{*k}(v)}{k!} n^{k} m^{\ell} = \sum_{\ell=0}^{N} \frac{e^{*\ell}(v)}{\ell!} n^{\ell} m^{0} = \sum_{\ell=0}^{N} \frac{e^{*\ell}(v)}{\ell!} n^{0} m^{\ell} = \sum_{\ell=0}^{N} 1 \frac{(nm)^{\ell} e^{*\ell}(v)}{\ell!} = \text{id}_{A}^{*n}(\text{id}_{A}^{*m}(v)) \quad \text{(by (11.867))}
\]

= \text{id}_{A}^{*n}(\text{id}_{A}^{*m}(v)) - \text{id}_{A}^{*n}(\text{id}_{A}^{*m}(v)) = 0.

Now, let us forget that we fixed \( n \) and \( m \). We thus have proven that
\[
\sum_{k=0}^{N} \sum_{\ell=0}^{N} \left( \frac{(e^{*k} \circ e^{*\ell})(v)}{k! \ell!} - \frac{\delta_{k,\ell} e^{*k}(v)}{k!} \right) n^{k} m^{\ell} = 0
\]
for all \( n \in \mathbb{N} \) and \( m \in \mathbb{N} \). Hence, we can apply Lemma 11.28 to \( V = A \) and \( v_{k,\ell} = \frac{(e^{*k} \circ e^{*\ell})(v)}{k! \ell!} - \frac{\delta_{k,\ell} e^{*k}(v)}{k!} \) (because \( A \) is a torsionfree abelian group). As a result, we obtain \( \frac{(e^{*k} \circ e^{*\ell})(v)}{k! \ell!} - \frac{\delta_{k,\ell} e^{*k}(v)}{k!} = 0 \) for every \( (k, \ell) \in \{0, 1, ..., N\}^{2} \). In other words,
\[
(11.869) \quad \frac{(e^{*k} \circ e^{*\ell})(v)}{k! \ell!} = \frac{\delta_{k,\ell} e^{*k}(v)}{k!} \quad \text{for every } (k, \ell) \in \{0, 1, ..., N\}^{2}
\]

Now, forget that we fixed \( v \), \( M \) and \( N \). We thus have proven that every \( M \in \mathbb{N} \), every integer \( N \) satisfying \( N \geq M \), and every \( v \in A_{M} \) satisfy (11.869).

Now, let us fix two elements \( n \in \mathbb{N} \) and \( m \in \mathbb{N} \). In order to finish the solution of Exercise 5.40(a), it remains to show that \( e^{*n} \circ e^{*m} = n! \delta_{n,m} e^{*n} \). In order to show this, it is clearly enough to prove that \( (e^{*n} \circ e^{*m})(v) = n! \delta_{n,m} e^{*n}(v) \) for every \( v \in A \). So let us fix some \( v \in A \), and let us show that \( (e^{*n} \circ e^{*m})(v) = n! \delta_{n,m} e^{*n}(v) \).

Since both sides of the identity \( (e^{*n} \circ e^{*m})(v) = n! \delta_{n,m} e^{*n}(v) \) are \( k \)-linear in \( v \), we can WLOG assume that \( v \) is a homogeneous element of \( A \) (because every element of \( A \) is a \( k \)-linear combination of homogeneous elements). Assume this. Thus, \( v \in A_{M} \) for some \( M \in \mathbb{N} \). Consider this \( M \). Choose some \( N \in \mathbb{N} \) satisfying \( N \geq M \), \( N \geq n \) and \( N \geq m \). (Such an \( N \) clearly exists.) Then, \( N \geq M \) and \( (n, m) \in \{0, 1, ..., N\}^{2} \), and therefore we can apply (11.869) to \( k = n \) and \( \ell = m \). As a result, we obtain
\[
\frac{(e^{*n} \circ e^{*m})(v)}{n!} = \frac{n! \delta_{n,m} e^{*n}(v)}{n!}.
\]

Multiplying this identity with \( n!m! \), we obtain
\[
(e^{*n} \circ e^{*m})(v) = m! \delta_{n,m} e^{*n}(v).
\]

So we have proven \( (e^{*n} \circ e^{*m})(v) = n! \delta_{n,m} e^{*n}(v) \). As we have seen, this completes the solution of Exercise 5.40(a).

(b) Let \( n \in \mathbb{N} \) and \( m \in \mathbb{N} \). From (11.864) (applied to \( m \) instead of \( n \)), we obtain \( \text{exp}^{*}(me^{*}) = \text{id}_{A}^{*m} \), so that
\[
\text{id}_{A}^{*m} = \text{exp}^{*}(me^{*}) = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} (me^{*})^{\ell} = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} m^{\ell} e^{*\ell}
\]
Thus,
\[
e^{*n} \circ \text{id}_{A}^{*m} = e^{*n} \circ \left( \sum_{\ell=0}^{\infty} \frac{1}{\ell!} m^{\ell} e^{*\ell} \right) = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} m^{\ell} e^{*\ell} = \sum_{\ell=0}^{\infty} \frac{1}{\ell!} m^{\ell} n^{\ell} \delta_{n,m} e^{*n} = \frac{1}{n!} m^{n} n! e^{*n} = m^{n} e^{*n}
\]
(by Exercise 5.40(a), applied to \( \ell \) instead of \( m \))
and similarly \( \text{id}_A^* \circ e^* = m^* e^* \). This solves Exercise 5.40(b).

11.98. Solution to Exercise 5.44. Solution to Exercise 5.44. (a) This is proven by induction over \( i \). The induction step relies on the observation that \( R_{(1')n-i} + R_{(1'-1,n-i+1)} = R_{(1')} H_{n-i} \) (where \( R_{(-1,n+1)} \) is to be understood to mean 0 in the \( i = 0 \) case). Let us prove this observation. We assume WLOG that \( i > 0 \) (the proof in the \( i = 0 \) case is analogous but simpler). The equality (5.21) yields \( H_{n-i} = R_{(n-i)} \) (since only \((n-i)\) coarsens \((n-i)\)), so that

\[
R_{(1')} H_{n-i} = R_{(1')} R_{(n-i)} = R_{(1') \cdot (n-i)} + R_{(1') \odot (n-i)} \quad \text{(by (5.23))}
\]

qed.

(b) Part (a) yields

\[
(-1)^i R_{(1',n-i)} = (-1)^i \sum_{j=0}^i (-1)^{i-j} R_{(1')} H_{n-j} = \sum_{j=0}^i \left( (-1)^j R_{(1')} H_{n-j} \right) = S(R_{(1')}) H_{n-j} \quad \text{(by (5.24), since } \omega(j) = (1'))
\]

\[
= \sum_{j=0}^i S( H_{j}) H_{n-j}.
\]

This proves (b).

(c) We have

\[
\Psi_n = \sum_{i=0}^{n-1} (-1)^i R_{(1',n-i)} = \sum_{i=0}^{n-1} \sum_{j=0}^i S( H_{j}) H_{n-j} = \sum_{j=0}^n \sum_{i=j}^{n-1} S( H_{j}) H_{n-j} = H_n - \sum_{j=0}^{n-1} (n-j) S( H_{j}) H_{n-j} = \sum_{j=0}^{n-1} (\deg ((H_n)_2)) S( (H_n)_1 ) (H_n)_2
\]

\[
= (S \star E)(H_n),
\]

and thus \( \Psi_n \) is primitive (by Exercise 1.65 (a)). (That said, there are other ways to prove the primitivity of \( \Psi_n \).)
(d) Let \( n \in \mathbb{N} \). Then,
\[
\sum_{k=0}^{n-1} H_k \Psi_{n-k} = \sum_{k=0}^{n-1} H_k (S \ast E) (H_{n-k}) = \sum_{k=0}^{n} H_k (S \ast E) (H_{n-k})
\]
we have \( \alpha = (1^i) \) shows that \( \pi (R_{1^i,n-i}) = s_{\alpha} \), where \( \alpha \) is the ribbon diagram of the composition \( (1^i,n-i) \). But since the ribbon diagram of the composition \( (1^i,n-i) \) is the Ferrers diagram for the partition \( (n-i,1^i) \) (because its row lengths are \( 1,1,\ldots,1, n-i \) going from bottom to top, with an overlap of 1 between every two adjacent rows),
\[
\pi (R_{1^i,n-i}) = s_{\alpha} = s_{(n-i,1^i)}.
\]
As we have seen, this completes the solution of part (g).

(h) We will prove \((5.25)\) by strong induction over \( n \). So let \( N \) be an arbitrary positive integer, and let us assume that \((5.25)\) has been proven for every \( n < N \). We now need to prove \((5.25)\) for \( n = N \).
We notice that the family \((H_\alpha)_{\alpha \in \text{NSym}}\) is multiplicative, in the sense that any two compositions \(\beta\) and \(\gamma\) satisfy

\[
H_{\beta \cdot \gamma} = H_\beta \cdot H_\gamma
\]

(where, as we recall, \(\beta \cdot \gamma\) denotes the concatenation of the compositions \(\beta\) and \(\gamma\)). This follows from the definition of the \(H_\alpha\).

We have \(N > 0\). Hence, every \(\alpha \in \text{Comp}_N\) can be written uniquely in the form \(\alpha = (q) \cdot \beta\) for some \(q \in \{1, 2, \ldots, N\}\) and some \(\beta \in \text{Comp}_{N-q}\) (indeed, the \(q\) is just the first entry of the composition \(\alpha\), and \(\beta\) is the composition obtained by erasing this first entry). Hence,

\[
\sum_{\alpha \in \text{Comp}_N} (1)_{\ell(\alpha)-1} \ell_\beta(H_\alpha) = \sum_{q \in \{1, 2, \ldots, N\}; \beta \in \text{Comp}_{N-q}} (-1)_{\ell((q) \cdot \beta)-1} \ell_\beta(H_{(q) \cdot \beta}) (\text{since it is easy to see that } \ell((q) \cdot \beta)-1=\ell(\beta))
\]

\[
= \sum_{q=1}^{N} \beta \in \text{Comp}_{N-q} (-1)_{\ell(\beta)} \ell_\beta(H_{(q) \cdot \beta}) \cdot H_{\beta}
\]

\[
= \sum_{q=1}^{N-1} \beta \in \text{Comp}_{N-q} (-1)_{\ell(\beta)} \ell_\beta(H_{(q) \cdot \beta}) \cdot H_{\beta} + \sum_{\beta \in \text{Comp}_{N-N}} (-1)_{\ell(\beta)} \ell_\beta(H_{(N) \cdot \beta}) \cdot H_{\beta}
\]

\[
(11.870) = \sum_{q=1}^{N-1} H_{(q)} \cdot \left( \sum_{\beta \in \text{Comp}_{N-q}} (-1)_{\ell(\beta)} \ell_\beta(H_{\beta}) \right) + NH_N.
\]

But for every \(q \in \{1, 2, \ldots, N-1\}\), we can apply (5.25) to \(n = N - q\) (because \(N - q < N\), and because (5.25) has been proven for every \(n < N\)). Thus, for every \(q \in \{1, 2, \ldots, N-1\}\), we obtain

\[
\Psi_{N-q} = \sum_{\alpha \in \text{Comp}_{N-q}} (-1)_{\ell(\alpha)} \ell_\alpha(H_\alpha) = -\sum_{\alpha \in \text{Comp}_{N-q}} (-1)_{\ell(\alpha)} \ell_\alpha(H_\alpha) = -\sum_{\beta \in \text{Comp}_{N-q}} (-1)_{\ell(\beta)} \ell_\beta(H_{\beta})
\]

(here, we renamed the summation index \(\alpha\) as \(\beta\)), so that

\[
(11.871) -\Psi_{N-q} = \sum_{\beta \in \text{Comp}_{N-q}} (-1)_{\ell(\beta)} \ell_\beta(H_{\beta}).
\]

Thus, (11.870) becomes

\[
\sum_{\alpha \in \text{Comp}_N} (1)_{\ell(\alpha)-1} \ell_\beta(H_\alpha) = \sum_{q=1}^{N-1} H_{(q)} \cdot \left( \sum_{\beta \in \text{Comp}_{N-q}} (-1)_{\ell(\beta)} \ell_\beta(H_{\beta}) \right) + NH_N
\]

\[
= \sum_{q=1}^{N-1} H_{(q)} \cdot (-\Psi_{N-q}) + NH_N = -\sum_{q=1}^{N-1} H_{q}\Psi_{N-q} + NH_N.
\]
However, Exercise 5.44(d) (applied to \( n = N \)) yields \( \sum_{k=0}^{N-1} H_k \Psi_{N-k} = NH_N \), so that

\[
NH_N = \sum_{k=0}^{N-1} H_k \Psi_{N-k} = \sum_{q=0}^{N-1} H_q \Psi_{N-q} = H_0 \Psi_{N-0} + \sum_{q=1}^{N-1} H_q \Psi_{N-q} = \Psi_N + \sum_{q=1}^{N-1} H_q \Psi_{N-q}.
\]

Now, (11.872) becomes

\[
\sum_{\alpha \in \text{Comp}_N} (-1)^{\ell(\alpha)-1} \sum_{q=1}^{N-1} H_q \Psi_{N-q} + NH_N = -\sum_{q=1}^{N-1} H_q \Psi_{N-q} + \Psi_N + \sum_{q=1}^{N-1} H_q \Psi_{N-q} = \Psi_N.
\]

In other words, \( \Psi_N = \sum_{\alpha \in \text{Comp}_N} (-1)^{\ell(\alpha)-1} \sum_{q=1}^{N-1} H_q \Psi_{N-q} \). Thus, (5.25) is proven for \( n = N \). This completes the induction step, and therefore the proof of (5.25) is complete. That is, part (h) of the exercise is solved.

(i) First of all, for every positive integer \( n \), the element \( \Psi_n \) of NSym is homogeneous of degree \( n \) (this follows from the definition of \( \Psi_n \) or, alternatively, from Exercise 5.44(h)). Hence, for every composition \( \alpha \), the element \( \Psi_{\alpha} \) of NSym is homogeneous of degree \( |\alpha| \).

We notice that the family \( \{ \Psi_{\alpha} \}_{\alpha \in \text{NSym}} \) is multiplicative, in the sense that any two compositions \( \beta \) and \( \gamma \) satisfy \( \Psi_{\beta \cdot \gamma} = \Psi_{\beta} \cdot \Psi_{\gamma} \) (where, as we recall, \( \beta \cdot \gamma \) denotes the concatenation of the compositions \( \beta \) and \( \gamma \)). This follows from the definition of the \( \Psi_{\alpha} \).

Let us now prove (5.26). Indeed, we will show (5.26) by strong induction over \( n \). So let \( n \in \mathbb{N} \) be arbitrary, and let us assume that (5.26) has been proven for every \( n < N \). We now need to prove (5.26) for \( n = N \).

If \( N = 0 \), then (5.26) obviously holds for \( n = N \) (because both sides of (5.26) are 1 in NSym in this case). We thus WLOG assume that \( N \neq 0 \). Thus, every \( \alpha \in \text{Comp}_N \) can be written uniquely in the form \( \alpha = \beta \cdot (q) \) for some \( q \in \{1, 2, \ldots, N\} \) and some \( \beta \in \text{Comp}_{N-q} \) (indeed, the \( q \) is just the last entry of the composition \( \alpha \), and \( \beta \) is the composition obtained by erasing this last entry).

Hence,

\[
\sum_{\alpha \in \text{Comp}_N} \frac{1}{\pi_u(\alpha)} \Psi_{\alpha} = \sum_{\substack{q \in \{1, 2, \ldots, N\}; \\
\beta \in \text{Comp}_{N-q}}} \frac{1}{\pi_u(\beta \cdot (q))} \Psi_{\beta \cdot (q)} = \sum_{\substack{q \in \{1, 2, \ldots, N\}; \\
\beta \in \text{Comp}_{N-q}}} \frac{1}{\pi_u(\beta \cdot (q))} \Psi_{\beta} \cdot \Psi_{(q)}
\]

\[
= \frac{1}{N} \sum_{q=1}^{N} \sum_{\beta \in \text{Comp}_{N-q}} \frac{1}{\pi_u(\beta)} \Psi_{\beta} \cdot \Psi_{q} = \frac{1}{N} \sum_{q=1}^{N} \sum_{\beta \in \text{Comp}_{N-q}} \frac{1}{\pi_u(\beta)} \Psi_{\beta} \cdot \Psi_{q},
\]

(11.873)

But for every \( q \in \{1, 2, \ldots, N\} \), we can apply (5.26) to \( n = N - q \) (because \( N - q < N \), and because (5.26) has been proven for every \( n < N \)). Thus, for every \( q \in \{1, 2, \ldots, N\} \), we obtain

\[
H_{N-q} = \sum_{\alpha \in \text{Comp}_{N-q}} \frac{1}{\pi_u(\alpha)} \Psi_{\alpha} = \sum_{\beta \in \text{Comp}_{N-q}} \frac{1}{\pi_u(\beta)} \Psi_{\beta}
\]

(11.874)
(here, we renamed the summation index $\alpha$ as $\beta$). Thus, \((11.873)\) becomes

$$
\sum_{\alpha \in \text{Comp}_N} \frac{1}{\pi_u (\alpha)} \Psi_{\alpha} = \frac{1}{N} \sum_{q=1}^{N} \left( \sum_{\beta \in \text{Comp}_{N-q}} \frac{1}{\pi_u (\beta)} \Psi_{\beta} \right) \cdot \Psi_{q}
$$

(by \((11.874)\))

$$
= \frac{1}{N} \sum_{q=1}^{N} H_{N-q} \Psi_{q} = \frac{1}{N} \sum_{k=0}^{N-1} H_{k} \Psi_{N-k}
$$

(by Exercise 5.44(d), applied to $n=N$)

(here, we substituted $N - k$ for $q$ in the sum)

$$
= \frac{1}{N} N H_{N} = H_{N}.
$$

In other words, $H_{N} = \sum_{\alpha \in \text{Comp}_N} \frac{1}{\pi_u (\alpha)} \Psi_{\alpha}$. Thus, \((5.26)\) is proven for $n = N$. This completes the induction step, and therefore the proof of \((5.26)\) is complete.

We now need to prove that $(\Psi_{\alpha})_{\alpha \in \text{Comp}_n}$ is a $k$-basis of $\text{NSym}_n$ for every $n \in \mathbb{N}$. It is clear that $(\Psi_{\alpha})_{\alpha \in \text{Comp}_n}$ is a family of elements of $\text{NSym}_n$ (because for every composition $\alpha$, the element $\Psi_{\alpha}$ of $\text{NSym}$ is homogeneous of degree $|\alpha|$).

Let $\mathfrak{A}$ be the $k$-subalgebra of $\text{NSym}$ generated by the elements $\Psi_1, \Psi_2, \Psi_3, \ldots$. Then, $\mathfrak{A}$ contains $\Psi_{\alpha}$ for every composition $\alpha$ (by the definition of $\Psi_{\alpha}$). Therefore, $\mathfrak{A}$ contains $H_n$ for every $n \geq 1$ (by \((5.26)\)). Consequently, $\mathfrak{A} = \text{NSym}$ (because $\text{NSym}$ is generated as a $k$-algebra by $H_1, H_2, H_3, \ldots$). In other words, the $k$-algebra $\text{NSym}$ is generated by the elements $\Psi_1, \Psi_2, \Psi_3, \ldots$ (since we defined $\mathfrak{A}$ as the $k$-subalgebra of $\text{NSym}$ generated by the elements $\Psi_1, \Psi_2, \Psi_3, \ldots$). In other words, the $k$-module $\text{NSym}$ is spanned by all possible products of the elements $\Psi_1, \Psi_2, \Psi_3, \ldots$. In other words, the $k$-module $\text{NSym}$ is spanned by the elements $\Psi_{\alpha}$ with $\alpha \in \text{Comp}$ (because the elements $\Psi_{\alpha}$ with $\alpha \in \text{Comp}$ are precisely all possible products of the elements $\Psi_1, \Psi_2, \Psi_3, \ldots$). In yet other words, the family $(\Psi_{\alpha})_{\alpha \in \text{Comp}}$ spans the $k$-module $\text{NSym}$.

Now, fix $n \in \mathbb{N}$. Every element of $\text{NSym}_n$ can be written as a $k$-linear combination of the elements $\Psi_{\alpha}$ with $\alpha \in \text{Comp}$ (since the family $(\Psi_{\alpha})_{\alpha \in \text{Comp}}$ spans the $k$-module $\text{NSym}$). In this $k$-linear combination, we can remove all terms $\Psi_{\alpha}$ with $\alpha \notin \text{Comp}_n$ without changing its value (by gradedness, because $\Psi_{\alpha}$ is homogeneous of degree $|\alpha|$), and so we conclude that every element of $\text{NSym}_n$ can be written as a $k$-linear combination of the elements $\Psi_{\alpha}$ with $\alpha \in \text{Comp}_n$. In other words, for every $n \in \mathbb{N}$, the family $(\Psi_{\alpha})_{\alpha \in \text{Comp}_n}$ spans the $k$-module $\text{NSym}_n$.

Now, we can apply Exercise 2.46(b) to $A = \text{NSym}_n$, $I = \text{Comp}_n$, $(\gamma_i)_{i \in I} = (H_{\alpha})_{\alpha \in \text{Comp}_n}$ and $(\beta_i)_{i \in I} = (\Psi_{\alpha})_{\alpha \in \text{Comp}_n}$ (since we know that $(H_{\alpha})_{\alpha \in \text{Comp}_n}$ is a $k$-basis of $\text{NSym}_n$, whereas $(\Psi_{\alpha})_{\alpha \in \text{Comp}_n}$ spans the $k$-module $\text{NSym}_n$). We conclude that $(\Psi_{\alpha})_{\alpha \in \text{Comp}_n}$ is a $k$-basis of $\text{NSym}_n$. This completes the solution of part (i).

(j) For every composition $\alpha$, define an element $b_{\alpha}$ of $T(V)$ by $b_{\alpha} = b_{\alpha_1} b_{\alpha_2} \cdots b_{\alpha_{\ell}}$, where $\alpha$ is written in the form $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_{\ell})$ with $\ell = \ell(\alpha)$. Then, $(b_{\alpha})_{\alpha \in \text{Comp}}$ is a $k$-module basis of $T(V)$ (by the basic properties of tensor algebras, since $(b_{\alpha})_{\alpha \in \{1,2,3,\ldots\}}$ is a $k$-module basis of $V$). Notice that the family $(b_{\alpha})_{\alpha \in \{1,2,3,\ldots\}}$ generates the $k$-algebra $T(V)$ (since it is a basis of $V$).

For every composition $\alpha$, define an element $\Psi_{\alpha}$ of $\text{NSym}$ as in Exercise 5.44(i). We know from Exercise 5.44(i) that $(\Psi_{\alpha})_{\alpha \in \text{Comp}_n}$ is a $k$-basis of $\text{NSym}_n$ for every $n \in \mathbb{N}$. Hence, $(\Psi_{\alpha})_{\alpha \in \text{Comp}}$ is a $k$-basis of $\text{NSym}$. 
Every \( \alpha \in \text{Comp} \) satisfies \( F(b_n) = \Psi_\alpha \). The map \( F \) thus maps the basis \( (b_n)_{\alpha \in \text{Comp}} \) of the \( k \)-module \( T(V) \) to the basis \( (\Psi_\alpha)_{\alpha \in \text{Comp}} \) of the \( k \)-module \( \text{NSym} \). Hence, \( F \) is a \( k \)-module isomorphism (since any \( k \)-linear map mapping a basis to a basis is a \( k \)-module isomorphism).

Next, we are going to show the equality \( \Delta_{\text{NSym}} \circ F = (F \otimes F) \circ \Delta_{T(V)} \). Indeed, this is an equality between \( k \)-algebra homomorphisms, and thus needs only to be verified on a generating set of the \( k \)-algebra \( T(V) \). Picking \( (b_n)_{n \in \{1,2,3,\ldots\}} \) as this generating set, we thus only need to check that \( \Delta_{\text{NSym}} \circ F \) restricted to \( b_n \) for every positive integer \( n \). This is straightforward: If \( n \) is any positive integer, then comparing the equalities

\[
(\Delta_{\text{NSym}} \circ F) (b_n) = \Delta_{\text{NSym}} \left( \frac{F(b_n)}{f(b_n)} \right) = \Delta_{\text{NSym}} \left( \frac{f(b_n)}{f(b_n)} \right) = \Delta_{\text{NSym}} (\Psi_n)
\]

\[
= 1 \otimes \Psi_n + \Psi_n \otimes 1 \quad \text{(since Exercise 5.44(c) shows that } \Psi_n \text{ is primitive)}
\]

and

\[
((F \otimes F) \circ \Delta_{T(V)}) (b_n) = (F \otimes F) \left( \frac{\Delta_{T(V)} (b_n)}{1 \otimes b_n + b_n \otimes 1} \right) = (F \otimes F) (1 \otimes b_n + b_n \otimes 1)
\]

\[
= F(1) \otimes F(b_n) + F(b_n) \otimes F(1) \quad \text{(since } F \text{ is a } k \text{-algebra homomorphism)}
\]

\[
= 1 \otimes \frac{f(b_n)}{\Psi_n} + \frac{f(b_n)}{\Psi_n} \otimes 1 = 1 \otimes \Psi_n + \Psi_n \otimes 1 \quad \text{(by the definition of } f \text{)}
\]

yields \( (\Delta_{\text{NSym}} \circ F) (b_n) = ((F \otimes F) \circ \Delta_{T(V)}) (b_n) \). We thus have shown that \( \Delta_{\text{NSym}} \circ F = (F \otimes F) \circ \Delta_{T(V)} \). Combining this with the equality \( \text{en}_{\text{NSym}} \circ F = \epsilon_{T(V)} \) (whose proof is similar but yet simpler), we conclude that \( F \) is a \( k \)-coalgebra homomorphism, thus a \( k \)-bialgebra homomorphism, hence a \( k \)-Hopf algebra isomorphism (by Proposition 1.46(c)), therefore a \( k \)-Hopf algebra isomorphism (since it is a \( k \)-module isomorphism). This solves Exercise 5.44(j).

(k) Define the map \( F \) as in Exercise 5.44(j). Then, Exercise 5.44(j) shows that \( F \) is a Hopf algebra isomorphism, hence a \( k \)-module isomorphism.

We can endow the \( k \)-module \( V \) with a grading by assigning to each basis vector \( b_n \) the degree \( n \). Then, \( V \) is of finite type and satisfies \( V_0 = 0 \), and thus \( T(V) \) is a connected graded \( k \)-Hopf algebra.

The element \( \Psi_n \) of \( \text{NSym} \) is homogeneous of degree \( n \) for every positive integer \( n \). (This follows from the definition of \( \Psi_n \).)

The map \( f \) is graded (since it sends every basis vector \( b_n \) of \( V \) to the vector \( \Psi_n \in \text{NSym} \), which is homogeneous of the same degree \( n \) as \( b_n \)). Hence, the map \( F \) (being the \( k \)-algebra homomorphism \( T(V) \to \text{NSym} \) induced by \( f \)) is also graded. It is well-known that if a \( k \)-module isomorphism is graded, then it is an isomorphism of graded \( k \)-modules. Thus, \( F \) is an isomorphism of graded \( k \)-modules (since \( F \) is a \( k \)-module isomorphism and is graded), hence an isomorphism of graded Hopf algebras (since \( F \) is a Hopf algebra isomorphism).

\[\text{Proof. Let } \alpha \in \text{Comp}. \text{ Write } \alpha \in \text{Comp} \text{ in the form } \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \text{ with } \ell = \ell(\alpha). \text{ Then, } b_\alpha = b_{\alpha_1} b_{\alpha_2} \cdots b_{\alpha_\ell}, \text{ so that}
\]

\[
F(b_\alpha) = F(b_{\alpha_1} b_{\alpha_2} \cdots b_{\alpha_\ell}) = \frac{f(b_{\alpha_1})}{\Psi_{\alpha_1}} \frac{f(b_{\alpha_2})}{\Psi_{\alpha_2}} \cdots \frac{f(b_{\alpha_\ell})}{\Psi_{\alpha_\ell}} \quad \text{(by the definition of } f \text{)} \quad \text{(by the definition of } f \text{)} \quad \text{(by the definition of } f \text{)}
\]

\[
= \Psi_{\alpha_1} \Psi_{\alpha_2} \cdots \Psi_{\alpha_\ell} = \Psi_\alpha \quad \text{(since } \Psi_\alpha \text{ was defined to be } \Psi_{\alpha_1} \Psi_{\alpha_2} \cdots \Psi_{\alpha_\ell}, \text{ qed.)}
\]
Compared with $\alpha$ (This has been proven during the solution to Exercise 5.44(g).) Also, Theorem 5.42(b) (applied to (11.875) Exercise 5.44(k).

(i) We recall that

(11.875) \[ \pi(R(1^i,n+i)) = s(n+i,1^i) \]

for every positive integer $n$ and every $i \in \{0,1,...,n-1\}$.

(This has been proven during the solution to Exercise 5.44(g).) Also, Theorem 5.42(b) (applied to $\alpha = (1^i)$) yields that

(11.876) \[ \pi(R(1^i)) = s(1^i) = e_i \]

for every $i \in \mathbb{N}$.

In our solution to Exercise 5.44(a), we have shown that

(11.877) \[ R(1^i)H_{n-i} = R(1^i,n-i) + R(1^{i-1},n-i+1) \]

for every positive integer $n$ and every $i \in \{1,2,...,n-1\}$.

Now, we are ready to solve parts (a) and (b) of Exercise 2.90 anew.

Alternative solution to Exercise 2.90(a): Let $n$ and $m$ be positive integers. We need to prove that $e_n h_m = s_{(m+1,1^{n-1})} + s_{(m,1^n)}$.

Applying (11.877) to $n+m$ and $n$ instead of $n$ and $i$, we obtain $R(1^n)H_{(n+m)-n} = R(1^n,(n+m)-n) + R(1^{n-1},(n+m)-(n-1))$ (since $(n+m) - n = 1 = (n+m) - (n-1)$).

Applying the map $\pi$ to both sides of this equality, we obtain

\[
\pi(R(1^n)H_{(n+m)-n}) = \pi(R(1^n,(n+m)-n) + R(1^{n-1},(n+m)-(n-1)))
\]

\[
= \pi(R(1^n,(n+m)-n)) + \pi(R(1^{n-1},(n+m)-(n-1)))
\]

(by (11.875), applied to $n+m$ and $n$ instead of $n$ and $i$)

\[
= s_{((n+m)-n,1^n)} + s_{((n+m)-(n-1),1^{n-1})}
\]

(by (11.875), applied to $n+m$ and $n-1$ instead of $n$ and $i$)

\[
= s_{(m,1^n)} + s_{(m+1,1^{n-1})}
\]

(since $(n+m) - n = m$ and $(n+m) - (n-1) = m + 1$)

\[
= s_{(m+1,1^{n-1})} + s_{(m,1^n)}.
\]

Compared with

\[
\pi(R(1^n)H_{(n+m)-n}) = \frac{\pi(R(1^n))}{\pi(H_{(n+m)-n})} = \frac{\pi(H_{(n+m)-n})}{\pi(H_m)} = e_n H_m
\]

(by (11.876), applied to $n$ instead of $i$)

\[
= e_n \cdot \frac{\pi(H_m)}{\pi(H_m)} = e_n h_m,
\]

(by the definition of $\pi$)

this yields $e_n h_m = s_{(m+1,1^{n-1})} + s_{(m,1^n)}$. Thus, Exercise 2.90(a) is solved again.

Alternative solution to Exercise 2.90(b): Let $a \in \mathbb{N}$ and $b \in \mathbb{N}$. Applying Exercise 5.44(b) to $n = a + b + 1$ and $i = b$, we obtain

\[
(-1)^b R(1^b,(a+b+1)-b) = \sum_{j=0}^{b} S(H_j) H_{(a+b+1)-j} = \sum_{j=0}^{b} S(H_{b-j}) H_{(a+b+1)-(b-j)}
\]
(here, we substituted \( b - j \) for \( j \) in the sum). Multiplying both sides of this equality by \((-1)^b\), we obtain

\[
R_{(1^b,(a+b+1)-b)} = (-1)^b \sum_{j=0}^{b} S(H_{b-j}) \frac{H(a+b+1)-(b-j)}{H_{a+j+1}} = (-1)^b \sum_{j=0}^{b} S(H_{b-j}) \ H_{a+j+1} \\
\text{(since } (a+b+1)-(b-j)=a+j+1) \\
\]

\[
= (-1)^b \sum_{i=0}^{b} S(H_{b-i}) \ H_{a+i+1} \
\text{(here, we renamed the summation index } j \text{ as } i). 
\]

Applying the map \( \pi \) to both sides of this equality, we obtain

\[
\pi(R_{(1^b,(a+b+1)-b)}) = \pi\left((-1)^b \sum_{i=0}^{b} S(H_{b-i}) \ H_{a+i+1}\right) = (-1)^b \sum_{i=0}^{b} S\left(\frac{\pi(H_{b-i})}{\pi(H_{a+i+1})}\right) \\
\text{(by the definition of } \pi) \\
\]

\[
= (-1)^b \sum_{i=0}^{b} (-1)^{b-i} c_{b-i} h_{a+i+1} = \sum_{i=0}^{b} (-1)^i h_{a+i+1} c_{b-i}. 
\]

Thus,

\[
\sum_{i=0}^{b} (-1)^i h_{a+i+1} c_{b-i} = \pi(R_{(1^b,(a+b+1)-b)}) = s_{((a+b+1)-b,1^b)} \quad \text{(by (11.875), applied to } n = a + b + 1 \text{ and } i = b) \\
= s_{(a+1,1^b)}. 
\]

Thus, Exercise 2.90(b) is once again solved.

11.99. Solution to Exercise 5.45. Solution to Exercise 5.45.

(a) The solution of Exercise 5.45(a) can be obtained from the solution of Exercise 2.85(a) upon replacing \( f_n \) by \( F_n \).

(b) The solution of Exercise 5.45(b) can be obtained from the solution of Exercise 2.85(b) upon replacing \( f_n, f_m, f_{nm} \), and \( \Lambda \) by \( F_n, F_m, F_{nm} \) and QSym.

(c) The solution of Exercise 5.45(c) can be obtained from the solution of Exercise 2.85(c) upon replacing \( f_1 \) and \( \Lambda \) by \( F_1 \) and QSym.

(d) Let \( n \in \{1,2,3,\ldots\} \) and \((\beta_1, \beta_2, \ldots, \beta_s) \in \text{Comp.} \) The definition of \( M(\beta_1, \beta_2, \ldots, \beta_s) \) yields that \( M(\beta_1, \beta_2, \ldots, \beta_s) = \sum_{i_1 < i_2 < \cdots < i_s} x_{i_1}^{\beta_1} x_{i_2}^{\beta_2} \cdots x_{i_s}^{\beta_s} \) where the sum is over all \( s \)-tuples \((i_1, i_2, \ldots, i_s)\) of positive integers satisfying \( i_1 < i_2 < \cdots < i_s \). Applying the map \( F_n \) to both sides of this equality, we obtain

\[
F_n(M(\beta_1, \beta_2, \ldots, \beta_s)) = F_n\left(\sum_{i_1 < i_2 < \cdots < i_s} x_{i_1}^{\beta_1} x_{i_2}^{\beta_2} \cdots x_{i_s}^{\beta_s}\right) \\
= \left(\sum_{i_1 < i_2 < \cdots < i_s} x_{i_1}^{\beta_1} x_{i_2}^{\beta_2} \cdots x_{i_s}^{\beta_s}\right) \ (x_1^n, x_2^n, x_3^n, \ldots) \\
= \sum_{i_1 < i_2 < \cdots < i_s} (x_{i_1}^{n \beta_1} x_{i_2}^{n \beta_2} \cdots x_{i_s}^{n \beta_s})^\beta \\
= \sum_{i_1 < i_2 < \cdots < i_s} x_{i_1}^{n \beta_1} x_{i_2}^{n \beta_2} \cdots x_{i_s}^{n \beta_s}.
\]
Compared with

\[ M_{n\beta_1, n\beta_2, \ldots, n\beta_s} = \sum_{i_1 < i_2 < \cdots < i_s} x_{i_1}^{n\beta_1} x_{i_2}^{n\beta_2} \cdots x_{i_s}^{n\beta_s} \]  
(by the definition of \( M_{(n\beta_1, n\beta_2, \ldots, n\beta_s)} \)),

this yields \( F_n (M_{(\beta_1, \beta_2, \ldots, \beta_s)}) = M_{(n\beta_1, n\beta_2, \ldots, n\beta_s)} \). This solves Exercise 5.45(d).

(c) Fix \( n \in \{1, 2, 3, \ldots\} \). Let \( \alpha \in \text{Comp} \). Write \( \alpha \) in the form \( (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \). Then,

\[ \Delta M_\alpha = \sum_{k=0}^{\ell} M_{(\alpha_1, \alpha_2, \ldots, \alpha_k)} \otimes M_{(\alpha_{k+1}, \alpha_{k+2}, \ldots, \alpha_\ell)} \]  
(according to Proposition 5.7). Applying the map \( F \) to both sides of this equality, we obtain

\[ (F_n \otimes F_n) (\Delta M_\alpha) = (F_n \otimes F_n) \left( \sum_{k=0}^{\ell} M_{(\alpha_1, \alpha_2, \ldots, \alpha_k)} \otimes M_{(\alpha_{k+1}, \alpha_{k+2}, \ldots, \alpha_\ell)} \right) \]

\[ = \sum_{k=0}^{\ell} \left[ F_n \left( M_{(\alpha_1, \alpha_2, \ldots, \alpha_k)} \right) \right] \otimes \left[ F_n \left( M_{(\alpha_{k+1}, \alpha_{k+2}, \ldots, \alpha_\ell)} \right) \right] \]

\[ = \sum_{k=0}^{\ell} M_{(n\alpha_1, n\alpha_2, \ldots, n\alpha_k)} \otimes M_{(n\alpha_{k+1}, n\alpha_{k+2}, \ldots, n\alpha_\ell)}. \]

But

\[ (\Delta \circ F_n) (M_\alpha) = \Delta \left( F_n \left( M_\alpha \right) \right) \]

\[ = \left( F_n \left( M_\alpha \right) \right) = \Delta \left( F_n \left( M_{(\alpha_1, \alpha_2, \ldots, \alpha_\ell)} \right) \right) \]

\[ = \Delta M_{(n\alpha_1, n\alpha_2, \ldots, n\alpha_\ell)} = \sum_{k=0}^{\ell} M_{(n\alpha_1, n\alpha_2, \ldots, n\alpha_k)} \otimes M_{(n\alpha_{k+1}, n\alpha_{k+2}, \ldots, n\alpha_\ell)} \]

\[ = (F_n \otimes F_n) (\Delta M_\alpha) = (F_n \otimes F_n) (\Delta M_\alpha) \]  
(by (11.878), applied to \( (n\alpha_1, n\alpha_2, \ldots, n\alpha_\ell) \) and \( (n\alpha_1, n\alpha_2, \ldots, n\alpha_\ell) \) instead of \( \alpha \) and \( (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \))

\[ = (F_n \otimes F_n) (\Delta M_\alpha) \]

\[ = ((F_n \otimes F_n) \circ \Delta) (M_\alpha). \]

Let us now forget that we fixed \( \alpha \). We thus have shown that

\[ (\Delta \circ F_n) (M_\alpha) = (F_n \otimes F_n) (\Delta) (M_\alpha) \]  
for every \( \alpha \in \text{Comp} \).

Now, let us recall that the family \( (M_\alpha)_{\alpha \in \text{Comp}} \) is a basis of the \( k \)-module \( QSym \). The two \( k \)-linear maps \( \Delta \circ F_n \) and \( (F_n \otimes F_n) \circ \Delta \) are equal to each other on this basis (according to (11.880)), and therefore must be identical (because if two \( k \)-linear maps from the same domain are equal to each other on a basis of their domain, then these two maps must be identical). In other words, \( \Delta \circ F_n = (F_n \otimes F_n) \circ \Delta \). We can prove (using a similar but simpler argument) that \( \epsilon \circ F_n = \epsilon \). Thus, \( F_n \) is a \( k \)-coalgebra homomorphism (since \( \Delta \circ F_n = (F_n \otimes F_n) \circ \Delta \) and \( \epsilon \circ F_n = \epsilon \)).

We now know that \( F_n \) is a \( k \)-algebra homomorphism (due to Exercise 5.45(a)) and a \( k \)-coalgebra homomorphism. Hence, \( F_n \) is a \( k \)-bialgebra homomorphism, thus a Hopf algebra homomorphism (due to Proposition 1.46(c)). This solves Exercise 5.46(e).

(f) Let \( n \in \{1, 2, 3, \ldots\} \). For every \( \alpha \), we have

\[ (F_n \mid \Lambda) (a) = F_n(a) = a (x_1^n, x_2^n, x_3^n, \ldots) \]  
(by the definition of \( F_n \))

\[ = f_n(a) \]  
(since \( f_n(a) = a (x_1^n, x_2^n, x_3^n, \ldots) \) by the definition of \( f_n \)).

Thus, \( F_n \mid \Lambda = f_n \). This solves Exercise 5.45(f).
(g) The solution of Exercise 5.45(g) can be obtained from the solution of Exercise 2.85(f) upon replacing $f$ and $\Lambda$ by $F$ and $QSym$, and replacing the word “symmetric” by “quasisymmetric”.

(h) Alternative solution of Exercise 2.85(d). Fix $n \in \{1, 2, 3, \ldots\}$. Exercise 5.45(f) yields that $F_n|_{\Lambda} = f_n$. Hence, $f_n$ is the restriction of $F_n$ to the Hopf subalgebra $\Lambda$ of $QSym$. Thus, $f_n$ is the restriction of a Hopf algebra homomorphism to the Hopf subalgebra $\Lambda$ of $QSym$ (since $F_n$ is a Hopf algebra homomorphism (by Exercise 5.45(e))). Consequently, $f_n$ is a Hopf algebra homomorphism itself (since the restriction of a Hopf algebra homomorphism to a Hopf subalgebra must always be a Hopf algebra homomorphism). This gives a new solution to Exercise 2.85(d). Thus, Exercise 5.45(h) is solved.

11.100. Solution to Exercise 5.46. Solution to Exercise 5.46. Let us first notice that every positive integer $n$ satisfies

$$V_n(H_m) = \begin{cases} H_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{cases} \quad \text{for every } m \in \mathbb{N}$$

Recall that $NSym$ is (isomorphic to) the free associative algebra with generators $H_1, H_2, H_3, \ldots$ (according to (5.13)). Hence, the elements $H_1, H_2, H_3, \ldots$ generate the $k$-algebra $NSym$. In other words,

$$\text{(11.882) the family } (H_r)_{r \geq 1} \text{ generates the } k\text{-algebra } NSym.$$

(c) A solution to Exercise 5.46(c) can be obtained by copying the solution of Exercise 2.86(c) and making the following changes:

- Replace every appearance of $v_j$ (for some positive integer $j$) by $V_j$.
- Replace every appearance of $\Lambda$ by $NSym$.
- Replace every appearance of $h_j$ (for some $j \in \mathbb{N}$) by $H_j$.
- Replace the reference to Proposition 2.28 by a reference to (11.882).
- Replace the reference to Proposition 2.28(iii) by a reference to (5.14).

(d) A solution to Exercise 5.46(d) can be obtained by copying the solution of Exercise 2.86(d) and making the following changes:

- Replace every appearance of $v_1$ by $V_1$.
- Replace every appearance of $\Lambda$ by $NSym$.
- Replace every appearance of $h_j$ (for some $j \in \mathbb{N}$) by $H_j$.
- Replace the reference to Proposition 2.28 by a reference to (11.882).

(a) Here is one of several possible solutions of Exercise 5.46(a). Define a $k$-linear map $E : NSym \to NSym$ as in Exercise 1.65 (but with $NSym$ instead of $A$). Every positive integer $n$ satisfies

$$\Psi_n = (S \ast E)(H_n)$$

(according to Exercise 5.44(c)).

Fix a positive integer $n$. Let us make some auxiliary observations first:

---

651 Proof. To obtain a proof of (11.881), it is enough to repeat the proof of (11.386), making just the following changes:

- replacing every $h_i$ by $H_i$;
- replacing $v_n$ by $V_n$.

652 Another solution (which the reader can easily find) proceeds by making some relatively straightforward modifications to the solution of Exercise 2.86(a). (One needs to keep in mind that $NSym$, unlike $\Lambda$, is not commutative – but this does not prevent us from doing things such as substituting $t^n$ for $t$ in a power series over $NSym$ (since $t^n$ is a central element of $NSym[[t]]$).)
• Every composition \((\alpha_1, \alpha_2, \ldots, \alpha_\ell)\) satisfies
\[
H_{(\alpha_1, \alpha_2, \ldots, \alpha_\ell)} = H_{\alpha_1} H_{\alpha_2} \cdots H_{\alpha_\ell}.
\]
\[\text{(11.884)}\]

• If \((\alpha_1, \alpha_2, \ldots, \alpha_\ell)\) is a composition such that (not every \(i \in \{1, 2, \ldots, \ell\}\) satisfies \(n \mid \alpha_i\)), then
\[
\mathbf{V}_n \left( H_{(\alpha_1, \alpha_2, \ldots, \alpha_\ell)} \right) = 0.
\]
\[\text{(11.885)}\]

• If \((\alpha_1, \alpha_2, \ldots, \alpha_\ell)\) is a composition such that (every \(i \in \{1, 2, \ldots, \ell\}\) satisfies \(n \mid \alpha_i\)), then
\[
\mathbf{V}_n \left( H_{(\alpha_1, \alpha_2, \ldots, \alpha_\ell)} \right) = H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}.
\]
\[\text{(11.886)}\]

• We have
\[
\mathbf{V}_n \circ E = n \cdot (E \circ \mathbf{V}_n)
\]
(as endomorphisms of NSym).
\[\text{(11.887)}\]

\[\text{[653] This is precisely the equality (5.15), and has been proven before.}\]
\[\text{[654] Proof of (11.885): Let } (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \text{ be a composition such that (not every } i \in \{1, 2, \ldots, \ell\} \text{ satisfies } n \mid \alpha_i\), then}
\[\text{there exists an } i \in \{1, 2, \ldots, \ell\} \text{ satisfying } n \mid \alpha_i \text{ (since not every } i \in \{1, 2, \ldots, \ell\} \text{ satisfies } n \mid \alpha_i\). Consider this } i \text{. We have}
\[
\mathbf{V}_n (H_{\alpha_i}) = \begin{cases} H_{\alpha_i/n}, & \text{if } n \mid \alpha_i; \\ 0, & \text{if } n \nmid \alpha_i \text{. (by the definition of } \mathbf{V}_n (H_{\alpha_i}) \text{)} \end{cases}
\]
\[\text{But (11.884) yields } H_{(\alpha_1, \alpha_2, \ldots, \alpha_\ell)} = H_{\alpha_1} H_{\alpha_2} \cdots H_{\alpha_\ell}. \text{ Applying } \mathbf{V}_n \text{ to both sides of this equality, we obtain}
\]
\[
\mathbf{V}_n (H_{(\alpha_1, \alpha_2, \ldots, \alpha_\ell)}) = \mathbf{V}_n (H_{\alpha_1}) \mathbf{V}_n (H_{\alpha_2}) \cdots \mathbf{V}_n (H_{\alpha_\ell})
\]
\[\text{since } \mathbf{V}_n \text{ is a } k\text{-algebra homomorphism}
\]
\[
= 0 \quad \left( \text{since at least one factor of the product}
\right)
\]
\[
\mathbf{V}_n (H_{\alpha_i}) \mathbf{V}_n (H_{\alpha_2}) \cdots \mathbf{V}_n (H_{\alpha_\ell}) = 0 \quad \text{is 0 (namely, the factor } \mathbf{V}_n (H_{\alpha_i}) \text{)}
\]
\[\text{This proves (11.885).}
\]
\[\text{[655] Proof of (11.886): Let } (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \text{ be a composition such that (every } i \in \{1, 2, \ldots, \ell\} \text{ satisfies } n \mid \alpha_i\). Then}
\[\text{\((\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)\) is a composition. Therefore, (11.884) (applied to } (\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n) \text{ instead of } (\alpha_1, \alpha_2, \ldots, \alpha_\ell))
\]
\[\text{yields } H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)} = H_{\alpha_1/n} H_{\alpha_2/n} \cdots H_{\alpha_\ell/n}.
\]
\[\text{But every } i \in \{1, 2, \ldots, \ell\} \text{ satisfies}
\]
\[
\mathbf{V}_n (H_{\alpha_i}) = \begin{cases} H_{\alpha_i/n}, & \text{if } n \mid \alpha_i; \\ 0, & \text{if } n \nmid \alpha_i \text{. (by the definition of } \mathbf{V}_n (H_{\alpha_i}) \text{)} \end{cases}
\]
\[\text{Multiplying these equalities for all } i \in \{1, 2, \ldots, \ell\}, \text{ we obtain}
\]
\[
\mathbf{V}_n (H_{\alpha_1}) \mathbf{V}_n (H_{\alpha_2}) \cdots \mathbf{V}_n (H_{\alpha_\ell}) = H_{\alpha_1/n} H_{\alpha_2/n} \cdots H_{\alpha_\ell/n}.
\]
\[\text{But (11.884) yields } H_{(\alpha_1, \alpha_2, \ldots, \alpha_\ell)} = H_{\alpha_1} H_{\alpha_2} \cdots H_{\alpha_\ell}. \text{ Applying } \mathbf{V}_n \text{ to both sides of this equality, we obtain}
\]
\[
\mathbf{V}_n (H_{(\alpha_1, \alpha_2, \ldots, \alpha_\ell)}) = \mathbf{V}_n (H_{\alpha_1}) \mathbf{V}_n (H_{\alpha_2}) \cdots \mathbf{V}_n (H_{\alpha_\ell})
\]
\[\text{(since } \mathbf{V}_n \text{ is a } k\text{-algebra homomorphism)}
\]
\[
= H_{\alpha_1/n} H_{\alpha_2/n} \cdots H_{\alpha_\ell/n} = H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}.
\]
\[\text{This proves (11.886).}
\]
\[\text{[656] Proof of (11.887): Let us show that}
\]
\[\text{(11.888) } (\mathbf{V}_n \circ E) (H_{\alpha}) = (n \cdot (E \circ \mathbf{V}_n)) (H_{\alpha}) \quad \text{for every } \alpha \in \text{Comp}.
\]
\[\text{Proof of (11.888): Let } \alpha \in \text{Comp}. \text{ Then, } H_{\alpha} \text{ is a homogeneous element of NSym of degree } \deg (H_{\alpha}) = |\alpha|. \text{ Thus,}
\]
\[\text{the definition of } E (H_{\alpha}) \text{ yields } E (H_{\alpha}) = (\deg (H_{\alpha})) \cdot H_{\alpha} = |\alpha| \cdot H_{\alpha}. \text{ Now,}
\]
\[\begin{align*}
(\mathbf{V}_n \circ E) (H_{\alpha}) &= \mathbf{V}_n \left( E (H_{\alpha}) \right) \\
&= \mathbf{V}_n (|\alpha| \cdot H_{\alpha}) = |\alpha| \cdot \mathbf{V}_n (H_{\alpha}) \quad \text{(since the map } \mathbf{V}_n \text{ is } k\text{-linear)}.
\end{align*}
\]
\[\text{Now, let us write the composition } \alpha \text{ in the form } (\alpha_1, \alpha_2, \ldots, \alpha_\ell). \text{ We distinguish between two cases:}
\]
\[\text{Case 1: Not every } i \in \{1, 2, \ldots, \ell\} \text{ satisfies } n \mid \alpha_i.
\]
\[\text{Case 2: Every } i \in \{1, 2, \ldots, \ell\} \text{ satisfies } n \mid \alpha_i.
\]
• We have

\[(11.890) \quad V_n \circ (E \ast S) = n \cdot ((E \ast S) \circ V_n)\]

(where \(S\), as usual, denotes the antipode of \text{NSym}). \(\text{657}\)

Let us first consider Case 1. In this case, not every \(i \in \{1, 2, \ldots, \ell\}\) satisfies \(n \mid \alpha_i\). Thus, \((11.885)\) yields \(V_n \left(H_{(\alpha_1, \alpha_2, \ldots, \alpha_\ell)}\right) = 0\). Since \((\alpha_1, \alpha_2, \ldots, \alpha_\ell) = \alpha\), this rewrites as \(V_n \left(H_\alpha\right) = 0\). Thus, \((V_n \circ E) (H_\alpha) = [\alpha] \cdot V_n (H_\alpha) = 0\).

Compared with

\[
(n \cdot (E \circ V_n)) (H_\alpha) = n \cdot (E \circ V_n) (H_\alpha) = n \cdot E \left\{ \begin{array}{c}
V_n (H_\alpha) \\
= E(V_n(H_\alpha))
\end{array} \right. \quad (n \cdot E(0) = 0,\quad \text{(since \(E\) is \(k\)-linear})
\]

this yields \((V_n \circ E) (H_\alpha) = n \cdot (E \circ V_n) (H_\alpha)\). Hence, \((11.888)\) is proven in Case 1.

Let us now consider Case 2. In this case, every \(i \in \{1, 2, \ldots, \ell\}\) satisfies \(n \mid \alpha_i\). Thus, \((11.886)\) yields \(V_n \left(H_{(\alpha_1, \alpha_2, \ldots, \alpha_\ell)}\right) = H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}\). Since \((\alpha_1, \alpha_2, \ldots, \alpha_\ell) = \alpha\), this rewrites as \(V_n \left(H_{\alpha}\right) = H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}\). Thus,

\[(11.889) \quad (V_n \circ E) (H_\alpha) = [\alpha] \cdot \left\{ \begin{array}{c}
V_n (H_\alpha) \\
= H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}
\end{array} \right. = [\alpha] \cdot H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}.
\]

On the other hand, \(H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}\) is a homogeneous element of \text{NSym} of degree \(\text{deg} \left(H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}\right) = (\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)\) that satisfies \(n \mid \alpha_i\). Hence, the definition of \(E \left(H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}\right)\) yields

\[
E \left(H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}\right) = \left\{ \begin{array}{c}
\text{deg} \left(H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}\right) \\
= (\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)
\end{array} \right. \cdot H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)} = (\alpha_1/n) \cdot H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}.
\]

Now,

\[
(n \cdot (E \circ V_n)) (H_\alpha) = n \cdot (E \circ V_n) (H_\alpha) = n \cdot E \left\{ \begin{array}{c}
V_n (H_\alpha) \\
= E(V_n(H_\alpha))
\end{array} \right. = H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}
\]

\[
= (\alpha_1/n) \cdot H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}
\]

\[
= [\alpha] \cdot H_{(\alpha_1/n, \alpha_2/n, \ldots, \alpha_\ell/n)}.
\]

Compared with \((11.889)\), this yields \((V_n \circ E) (H_\alpha) = (n \cdot (E \circ V_n)) (H_\alpha)\). Hence, \((11.888)\) is proven in Case 2.

Thus, \((11.888)\) is proven in both Cases 1 and 2. Since these two Cases cover all possibilities, this shows that \((11.888)\) always holds.

Now, \((H_\alpha)_{\alpha \in \text{Comp}}\) is a basis of the \(k\)-module \text{NSym}. The equality \((11.888)\) shows that the two \(k\)-linear maps \(V_n \circ E\) and \(n \cdot (E \circ V_n)\) are equal to each other on every element of this basis. Hence, these two maps \(V_n \circ E\) and \(n \cdot (E \circ V_n)\) are identical (because if two \(k\)-linear maps having the same domain are equal to each other on every element of some basis of this domain, then these two maps must be identical). In other words, \(V_n \circ E = n \cdot (E \circ V_n)\). This proves \((11.887)\).

\(657\) Proof of \((11.890)\): We have \(E \ast S = m \circ (E \otimes S) \circ \Delta\) (by the definition of convolution). But we know (from Exercise 5.46(e)) that \(V_n\) is a Hopf algebra homomorphism; thus, \(V_n \circ S = S \circ V_n\). Also, \(V_n\) is a \(k\)-coalgebra homomorphism (since \(V_n\) is a Hopf algebra homomorphism); therefore, \((V_n \otimes V_n) \circ \Delta = \Delta \circ V_n\). Finally, \(V_n\) is a \(k\)-algebra homomorphism; thus,
But fix a positive integer $m$. We have $\Psi_m = (E \ast S)(H_m)$ (by (11.883)), applied to $m$ instead of $n$ and thus

\[
V_n\left(\frac{\Psi_m}{(E \ast S)(H_m)}\right) = V_n\left(\begin{array}{c}
((E \ast S)(H_m)) \\
\text{(by (11.880))}
\end{array}\right) = (n \cdot (E \ast S)(H_m)) = n \cdot (E \ast S)(H_m)
\]

This proves (11.890).

(b) Recall that a ring of formal power series $R[[t]]$ over a (not necessarily commutative) $k$-algebra $R$ has a canonical topology which makes it into a topological $k$-algebra. Thus, in particular, NSym $[[t]]$ becomes a topological $k$-algebra. We shall be considering this topology when we speak of continuity.

Assume that $\mathcal{Q}$ is a subring of $k$.

\[
\Psi_m = m \circ (V_n \otimes V_n). \quad \text{Now,}
\]

\[
\Psi_m = (E \ast S) \circ \Delta = m \circ ((E \otimes S) \circ (V_n \otimes V_n)) \circ \Delta
\]

This proves (11.890).

Proof. Assume that $m \mid n$. Then, $m/n$ is a positive integer. Hence, (11.883) (applied to $m/n$ instead of $n$) yields $\Psi_{m/n} = (E \ast S)(H_{m/n})$, so that $(E \ast S)(H_{m/n}) = \Psi_{m/n}$, qed.

This is proven just as in the case when $R$ is commutative.
Define two power series $\tilde{H}(t)$ and $\xi(t)$ in $\text{NSym}[[t]]$ by

$$\tilde{H}(t) = \sum_{n \geq 0} H_n t^n;$$
$$\xi(t) = \sum_{n \geq 1} \xi_n t^n = \log \left( \tilde{H}(t) \right).$$

(Here, the equality $\sum_{n \geq 1} \xi_n t^n = \log \left( \tilde{H}(t) \right)$ follows from (5.18).)

We have $\tilde{H}(t) = \sum_{n \geq 0} H_n t^n = \sum_{i \geq 0} H_i t^i$ (here, we renamed the summation index $n$ as $i$) and $\xi(t) = \sum_{n \geq 1} \xi_n t^n = \sum_{i \geq 1} \xi_i t^i$ (here, we renamed the summation index $n$ as $i$).

Now, let $n$ be a positive integer. The $k$-algebra homomorphism $V_n : \text{NSym} \to \text{NSym}$ induces a continuous $k$-algebra homomorphism $V_n[[t]] : \text{NSym}[[t]] \to \text{NSym}[[t]]$ given by

$$(V_n[[t]]) \left( \sum_{i \in \mathbb{N}} a_i t^i \right) = \sum_{i \in \mathbb{N}} V_n(a_i) t^i \quad \text{for all } (a_i)_{i \in \mathbb{N}} \in \text{NSym}^\mathbb{N}.$$ 

This $k$-algebra homomorphism $V_n[[t]]$ commutes with taking logarithms (since it is continuous and a $k$-algebra homomorphism) — i.e., we have

$$(11.891) \quad (V_n[[t]]) \left( \log Q \right) = \log ((V_n[[t]])(Q)) \quad \text{for every } Q \in \text{NSym}[[t]] \text{ having constant coefficient 1.}$$

We have $\xi(t) = \sum_{i \geq 1} \xi_i t^i$, thus

$$\sum_{i \geq 1} \xi_i t^i = \xi(t) = \log \left( \frac{\tilde{H}(t)}{= \sum_{i \geq 0} H_i t^i} \right) = \log \left( \sum_{i \geq 0} H_i t^i \right).$$

We can substitute $t^n$ for $t$ on both sides of this equality\(^{660}\). As a result, we obtain

$$(11.892) \quad \sum_{i \geq 1} \xi_i (t^n)^i = \log \left( \sum_{i \geq 0} H_i (t^n)^i \right).$$

\(^{660}\)This is allowed because the element $t^n$ of $\text{NSym}[[t]]$ is central. (Generally, if $R$ is a ring (possibly not commutative) and $Q \in R[[t]]$ is a power series, then every central element $z$ of $R[[t]]$ can be substituted for $t$ in $Q$ as long as the constant coefficient of $z$ is 0. The result of this substitution is denoted by $Q(z)$. The map $R[[t]] \to R[[t]]$ which sends every $Q$ to $Q(z)$ (for a fixed $z$) is a ring homomorphism $R[[t]] \to R[[t]]$.\)
But $\tilde{H}(t) = \sum_{i \geq 0} H_i t^i$. Applying the map $V_n[[t]]$ to both sides of this equality, we obtain

$$\left(V_n[[t]]\right) \left(\tilde{H}(t)\right) = \left(V_n[[t]]\right) \left(\sum_{i \geq 0} H_i t^i\right) = \sum_{i \geq 0} \left(\frac{V_n(H_i)}{H_i/n}, \begin{array}{ll} H_{i/n}, & \text{if } n \mid i; \\ 0, & \text{if } n \nmid i \end{array} \right) = \sum_{i \geq 0} H_{i/n} t^i$$

(by (11.881), applied to $m=n$)

$$= \sum_{i \geq 0; n \mid i} H_{i/n} = H_i (t^n)^i$$

(here, we have substituted $ni$ for $i$ in the sum)

$$\left(\sum_{i \geq 0; n \mid i} H_{i/n} \right) = \sum_{i \geq 0; n \mid i} H_i (t^n)^i$$

(11.893)

Now,

$$\left(V_n[[t]]\right) \left(\xi(t)\right) = \left(V_n[[t]]\right) \left(\log(\tilde{H}(t))\right) = \log\left(V_n[[t]]\right) \left(\tilde{H}(t)\right) = \log\left(\sum_{i \geq 0} H_i (t^n)^i\right)$$

(by (11.891), applied to $Q = \tilde{H}(t)$)

$$= \sum_{i \geq 1} \xi_i (t^n)^i$$

(by (11.892))

$$= \sum_{i \geq 1} \xi_i t^{ni}. \tag{11.894}$$

Comparing this with

$$\left(V_n[[t]]\right) \left(\xi(t)\right) = \left(V_n[[t]]\right) \left(\sum_{i \geq 1} \xi_i t^i\right) = \sum_{i \geq 1} V_n(\xi_i) t^i$$

(by the definition of $V_n[[t]]$),

we obtain

$$\sum_{i \geq 1} V_n(\xi_i) t^i = \sum_{i \geq 1} \xi_i t^{ni}. \tag{11.895}$$

Now, let $m$ be a positive integer. Comparing coefficients before $t^m$ in the equality (11.895), we obtain

$$V_n(\xi_m) = \left\{ \begin{array}{ll} \xi_{m/n}, & \text{if } n \mid m; \\ 0, & \text{if } n \nmid m \end{array} \right.$$
This proves (11.897).

(f) Recall first that \((H_\alpha)_{\alpha \in \Comp}\) and \((M_\alpha)_{\alpha \in \Comp}\) are mutually dual bases with respect to the dual pairing \(\NSym \otimes \QSym \xrightarrow{\cdot : \cdot} k\).  

Thus,

\[
(11.896) \quad (H_\alpha, M_\beta) = \delta_{\alpha, \beta} \quad \text{for any two compositions } \alpha \text{ and } \beta.
\]

Let us introduce a notation: For every composition \(\alpha\) and every positive integer \(s\), let \(a\{s\}\) denote the \(\ell\)-tuple \((s\alpha_1, s\alpha_2, \ldots, s\alpha_\ell)\), where \(\alpha\) is written in the form \((\alpha_1, \alpha_2, \ldots, \alpha_\ell)\). Notice that if \(\alpha\) is a composition and \(s\) is a positive integer, then \(a\{s\}\) is a composition again.

We have

\[
(11.897) \quad F_n a = M_\alpha\{n\} \quad \text{for every composition } \alpha.
\]

We need to prove that the maps \(F_n : \QSym \to \QSym\) and \(V_n : \NSym \to \NSym\) are adjoint with respect to the dual pairing \(\NSym \otimes \QSym \xrightarrow{\cdot : \cdot} k\). In other words, we need to prove that

\[
(11.899) \quad (b, F_n a) = (V_n b, a) \quad \text{for any } a \in \QSym \text{ and } b \in \NSym.
\]

**Proof of (11.899):** Fix \(a \in \QSym\) and \(b \in \NSym\).

Recall that \((M_\alpha)_{\alpha \in \Comp}\) is a basis of the \(k\)-module \(\NSym\). Hence, in proving (11.899), we can WLOG assume that \(a\) is an element of this basis \((M_\alpha)_{\alpha \in \Comp}\) (because the equality (11.899) is \(k\)-linear in \(a\)). Assume this. Thus, \(a\) is an element of the basis \((M_\alpha)_{\alpha \in \Comp}\). In other words, there exists a \(\beta \in \Comp\) such that \(a = M_\beta\). Consider this \(\beta\). We have \(F_n \underbrace{a}_{=M_\beta} = F_n M_\beta = M_\beta\{n\}\) (by (11.897), applied to \(\alpha = \beta\)).

Recall that \((H_\alpha)_{\alpha \in \Comp}\) is a basis of the \(k\)-module \(\Lambda\). Hence, in proving (11.899), we can WLOG assume that \(b\) is an element of this basis \((H_\alpha)_{\alpha \in \Comp}\) (because the equality (11.899) is \(k\)-linear in \(b\)). Assume this. Thus, \(b\) is an element of the basis \((H_\alpha)_{\alpha \in \Comp}\). In other words, there exists a \(\gamma \in \Comp\) such that \(b = H_\gamma\). Consider this \(\gamma\). We have

\[
(11.900) \quad \left(\underbrace{b}_{=H_\gamma} \cdot \underbrace{F_n a}_{=M_\beta}\right)(n) = (H_\gamma, M_\beta\{n\}) = \delta_{\gamma, \beta\{n\}}
\]

(by (11.896), applied to \(\gamma\) and \(\beta\{n\}\) instead of \(\alpha\) and \(\beta\)).

Let us write the composition \(\gamma\) in the form \((\gamma_1, \gamma_2, \ldots, \gamma_\ell)\). Then, \(\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_\ell)\), so that \(H_\gamma = H_{(\gamma_1, \gamma_2, \ldots, \gamma_\ell)} = H_{\gamma_1} H_{\gamma_2} \cdots H_{\gamma_\ell}\) (by (11.884), applied to \((\gamma_1, \gamma_2, \ldots, \gamma_\ell)\) instead of \((\alpha_1, \alpha_2, \ldots, \alpha_\ell)\)).

Let us first assume that

\[
(11.901) \quad \text{(not every } i \in \{1, 2, \ldots, \ell\} \text{ satisfies } n | \gamma_i)\]

---

\[\text{This follows from the definition of } (H_\alpha)_{\alpha \in \Comp}.\]

---

\[\text{Proof of (11.897): Let } \alpha \text{ be a composition. Write } \alpha \text{ in the form } (\alpha_1, \alpha_2, \ldots, \alpha_\ell). \text{ By the definition of } M_\alpha, \text{ we have } M_\alpha = \sum_{i_1 < i_2 < \cdots < i_\ell} x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_\ell}^{\alpha_\ell} \text{ (where the sum is over all } \ell\text{-tuples } (i_1, i_2, \ldots, i_\ell) \text{ of positive integers satisfying } i_1 < i_2 < \cdots < i_\ell)\).

On the other hand, the definition of \(a\{n\}\) yields \(a\{n\} = (n\alpha_1, n\alpha_2, \ldots, n\alpha_\ell)\) (since \(\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)\)). Hence, the definition of \(M_\alpha\{n\}\) yields

\[
(11.898) \quad M_\alpha\{n\} = \sum_{i_1 < i_2 < \cdots < i_\ell} x_{i_1}^{n\alpha_1} x_{i_2}^{n\alpha_2} \cdots x_{i_\ell}^{n\alpha_\ell}
\]

(where the sum is over all \(\ell\)-tuples \((i_1, i_2, \ldots, i_\ell)\) of positive integers satisfying \(i_1 < i_2 < \cdots < i_\ell\)).

Now, the definition of \(F_n\) yields

\[
F_n M_\alpha = \sum_{i_1 < i_2 < \cdots < i_\ell} \left(x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_\ell}^{\alpha_\ell}\right) \left(x_{i_1}^{n\alpha_1} x_{i_2}^{n\alpha_2} \cdots x_{i_\ell}^{n\alpha_\ell}\right) = M_\alpha\{n\} \quad \text{(by (11.898))}.\]

This proves (11.897).
Then, $V_n \left( \frac{H_{\gamma}}{H_{(\gamma_1, \gamma_2, \ldots, \gamma_\ell)}} \right) = V_n \left( H_{(\gamma_1, \gamma_2, \ldots, \gamma_\ell)} \right) = 0$ \((11.885)\), applied to \((\gamma_1, \gamma_2, \ldots, \gamma_\ell)\) instead of \((\alpha_1, \alpha_2, \ldots, \alpha_\ell)\). Also, $\beta \{n\} \neq \gamma$ \((11.901)\), so that $\delta_{\gamma, \beta \{n\}} = 0$. Thus, \((11.900)\) becomes $\langle b, F_n a \rangle = \delta_{\gamma, \beta \{n\}} = 0$.

Compared with $\left( V_n \frac{b}{H_{\gamma}}, a \right) = \left( V_n \left( \frac{H_{\gamma}}{a} \right) \right) = (0, a) = 0$ (since the dual pairing $\text{NSym} \otimes \text{QSym} \rightarrow k$ is $k$-bilinear), this yields $\langle b, F_n a \rangle = \langle V_n b, a \rangle$. Thus, \((11.899)\) holds.

Now, let us forget that we assumed that \((11.901)\) holds. We thus have proven \((11.899)\) under the assumption that \((11.901)\) holds. Hence, for the rest of our proof of \((11.899)\), we can WLOG assume that \((11.901)\) does not hold. Assume this.

We have assumed that \((11.901)\) does not hold. In other words,

\((11.902)\)

\[
every \ i \ in \ \{1, 2, \ldots, \ell\} \ satisfies \ n \mid \gamma_i.
\]

Thus, $\gamma_i/n$ is a positive integer for every $i \in \{1, 2, \ldots, \ell\}$ (since $\gamma_i$ is a positive integer for every $i \in \{1, 2, \ldots, \ell\}$). Thus, $(\gamma_1/n, \gamma_2/n, \ldots, \gamma_\ell/n)$ is a composition. Let us denote this composition by $\zeta$. We have $\zeta \{n\} = \gamma$ \(664\) and $V_n(H_{\gamma}) = H_{\zeta}$ \(665\) and thus

\[(11.903)\]

\[
\left( \frac{V_n b}{H_{\gamma}} = H_{\gamma}, a \right) = \left( V_n \left( \frac{H_{\gamma}}{a} \right) \right) = (H_{\zeta}, M_{\beta}) = \delta_{\zeta, \beta}
\]

(by \((11.896)\), applied to $\zeta$ instead of $\alpha$).

Now, let us write the composition $\beta$ in the form $(\beta_1, \beta_2, \ldots, \beta_\ell)$. Then, $\beta \{n\} = (n\beta_1, n\beta_2, \ldots, n\beta_\ell)$ (by the definition of $\beta \{n\}$).

Now, we have the following equivalence of assertions:

\[
(\gamma = \beta \{n\}) \iff ((\gamma_1, \gamma_2, \ldots, \gamma_\ell) = (n\beta_1, n\beta_2, \ldots, n\beta_\ell))
\]

(since $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_\ell)$ and $\beta \{n\} = (n\beta_1, n\beta_2, \ldots, n\beta_\ell)$)

\[
\iff (we \ have \ \ell = s, \ and \ every \ i \ in \ \{1, 2, \ldots, \ell\} \ satisfies \ \gamma_i = n\beta_i)
\]

(this is equivalent to $\gamma_i/n = \beta_i$)

\[
\iff (we \ have \ \ell = s, \ and \ every \ i \ in \ \{1, 2, \ldots, \ell\} \ satisfies \ \gamma_i/n = \beta_i)
\]

\[
\iff ((\gamma_1/n, \gamma_2/n, \ldots, \gamma_\ell/n) = (\beta_1, \beta_2, \ldots, \beta_\ell))
\]

\[
\iff (\zeta = \beta) \quad (since \ (\gamma_1/n, \gamma_2/n, \ldots, \gamma_\ell/n) = \zeta \ and \ (\beta_1, \beta_2, \ldots, \beta_\ell) = \beta).
\]

This equivalence shows that $\delta_{\gamma, \beta \{n\}} = \delta_{\zeta, \beta}$. But \((11.900)\) becomes

\[
(b, F_n a) = \delta_{\gamma, \beta \{n\}} = \delta_{\zeta, \beta} = \langle V_n b, a \rangle \quad \text{(by \(11.903)\)).}
\]

---

\(663\)Proof. Assume the contrary. Thus, $\beta \{n\} = \gamma$. Let us write the composition $\beta$ in the form $(\beta_1, \beta_2, \ldots, \beta_\ell)$. Then, $\beta \{n\} = (n\beta_1, n\beta_2, \ldots, n\beta_\ell)$ (by the definition of $\beta \{n\}$). Hence, $(\gamma_1, \gamma_2, \ldots, \gamma_\ell) = \gamma = \beta \{n\} = (n\beta_1, n\beta_2, \ldots, n\beta_\ell)$. As a consequence, $\ell = q$, so that $q = \ell$ and thus $(n\beta_1, n\beta_2, \ldots, n\beta_\ell) = (n\beta_1, n\beta_2, \ldots, n\beta_\ell)$. Hence, $(\gamma_1, \gamma_2, \ldots, \gamma_\ell) = (n\beta_1, n\beta_2, \ldots, n\beta_\ell) = (n\beta_1, n\beta_2, \ldots, n\beta_\ell)$. Consequently, every $i \ in \ \{1, 2, \ldots, \ell\}$ satisfies $\gamma_i = n\beta_i$. Thus, every $i \ in \ \{1, 2, \ldots, \ell\}$ satisfies $n \mid \gamma_i$. This contradicts \((11.901)\). This contradiction proves that our assumption was wrong, qed.

\(664\)Proof. We have $\zeta = (\gamma_1/n, \gamma_2/n, \ldots, \gamma_\ell/n)$. Hence, the definition of $\zeta \{n\}$ yields $\zeta \{n\} = (n\gamma_1/n, n\gamma_2/n, \ldots, n\gamma_\ell/n) = (\gamma_1, \gamma_2, \ldots, \gamma_\ell) = \gamma$, qed.

\(665\)Proof. We have $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_\ell)$, so that

\[
V_n(H_{\gamma}) = V_n \left( H_{(\gamma_1, \gamma_2, \ldots, \gamma_\ell)} \right) = H_{(\gamma_1/n, \gamma_2/n, \ldots, \gamma_\ell/n)} = H_{\zeta}
\]

(by \((11.886)\), applied to $(\gamma_1, \gamma_2, \ldots, \gamma_\ell)$ instead of $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$)

\[
= H_{\zeta} \quad (since \ (\gamma_1/n, \gamma_2/n, \ldots, \gamma_\ell/n) = \zeta).
\]

qed.
Thus, (11.899) is proven. As we know, this completes the solution of Exercise 5.46(f).

(g) We know that \( \pi : \text{NSym} \to \Lambda \) is a \( k \)-algebra homomorphism. Thus, \( \pi(0) = 0 \). Also, we know that

\[
\pi(H_n) = h_n \quad \text{for every positive integer } n.
\]

Now, let \( n \) be a positive integer. The maps \( v_n \circ \pi \) and \( \pi \circ V_n \) are \( k \)-algebra homomorphisms (since \( \pi \) and \( v_n \) are \( k \)-algebra homomorphisms).

Let \( r \) be a positive integer. Applying (11.905) to \( r \) instead of \( n \), we obtain \( \pi(H_r) = h_r \).

If \( n \mid r \), then \( r/n \) is a positive integer. Hence,

\[
\pi(H_r/n) = \begin{cases} 
    h_r/n, & \text{if } n \mid r; \\
    0, & \text{if } n \nmid r
\end{cases} \quad \text{(by the definition of } v_n) \text{.}
\]

Thus, (11.906) is proven. As we know, this completes the solution of Exercise 5.46(f).

Now, let \( \pi \circ V_n \) be defined as

\[
(\pi \circ V_n)(H_r) = \begin{cases}
    \pi(V_n(H_r)), & \text{if } n \mid r; \\
    V_n(h_r), & \text{if } n \nmid r
\end{cases}
\]

(11.907) \( \pi \circ V_n(h_r) = \begin{cases} 
    H_r/n, & \text{if } n \mid r; \\
    0, & \text{if } n \nmid r
\end{cases} \quad \text{(by the definition of } v_n) \text{.}
\]

This yields

\[
(\pi \circ V_n)(H_r) = \begin{cases} 
    0, & \text{if } n \nmid r
\end{cases} \quad \text{(since } \pi(0) = 0 \text{ in the case when } n \nmid r) 
\]

This completes the solution of Exercise 5.46(g).

(h) Alternative solution of Exercise 2.86(f). Let \( i \) denote the inclusion map \( \Lambda \to \text{QSym} \). Then, Corollary 5.35 yields that the map \( \pi \) is adjoint to the map \( i \) with respect to the dual pairing \( \text{NSym} \odot \text{QSym} \to k \). In other words,

\[
(\pi(b), a) = (b, i(a)) \quad \text{for every } b \in \text{NSym} \text{ and } a \in \Lambda.
\]

But Exercise 5.46(f) yields that the maps \( F_n : \text{QSym} \to \text{QSym} \) and \( V_n : \text{NSym} \to \text{NSym} \) are adjoint with respect to the dual pairing \( \text{NSym} \odot \text{QSym} \to k \). In other words,

\[
(b, F_n a) = (V_n b, a) \quad \text{for any } a \in \text{QSym} \text{ and } b \in \text{NSym}.
\]

Now, fix a positive integer \( n \). Let \( a \in \Lambda \) and \( b \in \Lambda \) be arbitrary. Then, \( a \in \Lambda \subset \text{QSym} \). Exercise 5.45(f) yields \( F_n |_\Lambda = f_n \). Thus, \( f_n = F_n |_\Lambda \), so that \( f_n a = (F_n |_\Lambda) a = F_n a \).
On the other hand, the projection \( \pi : \NSym \to \Lambda \) is surjective. Hence, there exists some \( b' \in \NSym \) satisfying \( b = \pi (b') \). Consider this \( b' \).

Since the Hall inner product on \( \Lambda \) is symmetric, we have

\[
(f_n a, b) = \left( b, f_n a \right) = \left( \pi (b'), f_n a \right) = \left( b', i(f_n a) \right) \quad \text{(since } i \text{ is an inclusion map)}
\]

(by (11.908), applied to \( b' \) and \( f_n a \) instead of \( b \) and \( a \))

\[
= \left( b', f_n a \right) = \left( b', F_n a \right) (\text{by (11.909), applied to } b' \text{ instead of } b).
\]

Comparing this with

\[
(a, v_n b) = \left( v_n \frac{b}{=\pi(b')} a \right) \quad \text{(since the Hall inner product on } \Lambda \text{ is symmetric)}
\]

\[
= \left( v_n \left( \pi \left( b' \right) \right), a \right) = \left( v_n \circ \pi \left( b' \right), a \right) = \left( \pi \circ v_n \left( b' \right), a \right) \quad \text{(by Exercise 5.46(g))}
\]

\[
= \left( \pi \left( v_n b' \right), a \right) = \left( v_n b', \underbrace{i(a)}_{=a} \right) \quad \text{(by (11.908), applied to } v_n b' \text{ instead of } b) \quad \text{(since } i \text{ is an inclusion map)}
\]

we obtain \((f_n a, b) = (a, v_n b)\).

Let us now forget that we fixed \( a \) and \( b \). We have thus shown that \((f_n a, b) = (a, v_n b)\) for every \( a \in \Lambda \) and \( b \in \Lambda \). In other words, the maps \( f_n : \Lambda \to \Lambda \) and \( v_n : \Lambda \to \Lambda \) are adjoint with respect to the Hall inner product on \( \Lambda \). Thus, Exercise 2.86(f) is solved once again. Hence, Exercise 5.46(h) is solved.


Proof of Proposition 6.2. (a) This can be easily verified by hand, but here is a slicker way to see it: Let \( \mathfrak{B} \) be the set \( \mathfrak{A} \cup \{-\infty\} \), where \(-\infty\) is a symbol. We define a total order on \( \mathfrak{B} \) by extending the given total order on \( \mathfrak{A} \) in such a way that every \( a \in \mathfrak{A} \) satisfies \(-\infty < a \). Consider the set \( \mathfrak{B}^\infty \) of all infinite sequences of elements of \( \mathfrak{B} \). Then, \( \mathfrak{A}^* \) embeds into \( \mathfrak{B}^\infty \) by identifying every word \((a_1, a_2, \ldots, a_n) \in \mathfrak{A}^* \) with the sequence \((a_1, a_2, \ldots, a_n, -\infty, -\infty, -\infty, \ldots) \in \mathfrak{B}^\infty \). It is easy to check that our order relation \( \leq \) on \( \mathfrak{A}^* \) thus becomes the restriction to \( \mathfrak{A}^* \) of the smaller-or-equal relation of the lexicographic order on \( \mathfrak{B}^\infty \). Consequently, it is a total order. Proposition 6.2(a) is thus proven.

(h) follows immediately from the definition of \( \leq \).

(i) follows from common sense.

(k) Let \( a \in \mathfrak{A}^* \) and \( b \in \mathfrak{A}^* \) be such that \( b \) is nonempty. Then, \( a \) is a prefix of \( ab \). Thus, \( a \leq ab \) (by Proposition 6.2(h), applied to \( ab \) instead of \( b \)). But \( a \neq ab \) \(666\). Combined with \( a \leq ab \), this yields \( a < ab \). Proposition 6.2(k) is thus proven.

(b) Proposition 6.2(b) is an almost immediate consequence of the definition of \( \leq \); its proof is thus left to the reader.

---

\(666\)Proof. Assume the contrary. Then, \( a = ab \). Hence, \( a\varnothing = a = ab \). Cancelling \( a \) from this equality, we obtain \( \varnothing = b \), so that the word \( b \) is empty. This contradicts the fact that \( b \) is nonempty. This contradiction proves that our assumption was wrong, qed.
(c) Let $a, c, d \in \mathfrak{A}^*$ satisfy $ac \leq ad$. We have $ac = (a_1, a_2, \ldots, a_{\ell(a)}, c_1, c_2, \ldots, c_{\ell(c)})$ and $ad = (a_1, a_2, \ldots, a_{\ell(a)}, d_1, d_2, \ldots, d_{\ell(d)})$, and we have $ac \leq ad$. Due to the definition of $\leq$, this means that we must be in one of the following two situations:

- There exists an $i \in \{1, 2, \ldots, \min \{\ell(a), \ell(c), \ell(a) + \ell(d)\}\}$ such that
  
  \[
  (ac)_i < (ad)_i, \quad \text{and every } j \in \{1, 2, \ldots, i - 1\} \text{ satisfies } (ac)_j = (ad)_j. 
  \]

  The word $ac$ is a prefix of $ad$.

  If we are in the first of these two situations, then we clearly must have $i > \ell(a)$ (since otherwise, both $(ac)_i$ and $(ad)_i$ would equal $a_i$, and then $(ac)_i < (ad)_i$ would contradict $(ac)_i = (ad)_i$), and therefore $i$ has the form $i = \ell(a) + i'$ for some $i' \in \{1, 2, \ldots, \min \{\ell(c), \ell(d)\}\}$. Then, (11.910) yields $(c_{i'}) < d_{i'}$, and every $j \in \{1, 2, \ldots, i' - 1\}$ satisfies $c_j = d_j$. But this shows that $c \leq d$, and so we are done in the first situation. In the second situation, we also have $c \leq d$ 667, and again we are done. Thus, Proposition 6.2(c) is proven.

(d) Let $a, b, c, d \in \mathfrak{A}^*$ satisfy $a \leq c$. We have $a = (a_1, a_2, \ldots, a_{\ell(a)})$ and $c = (c_1, c_2, \ldots, c_{\ell(c)})$, and we have $a \leq c$. Due to the definition of $\leq$, this means that we must be in one of the following two situations:

- There exists an $i \in \{1, 2, \ldots, \min \{\ell(a), \ell(c)\}\}$ such that
  
  \[
  (a_i < c_i, \text{ and every } j \in \{1, 2, \ldots, i - 1\} \text{ satisfies } a_j = c_j). 
  \]

  The word $a$ is a prefix of $c$.

  The first of these situations entails $ab \leq cd$ (because every $k \in \{1, 2, \ldots, \min \{\ell(a), \ell(c)\}\}$ satisfies $(ab)_k = a_k$ and $(cd)_k = c_k$). Hence, in both situations, Proposition 6.2(d) is proven.

(e) Let $a, b, c, d \in \mathfrak{A}^*$ satisfy $ab \leq cd$. We have $a = (a_1, a_2, \ldots, a_{\ell(a)})$, $b = (b_1, b_2, \ldots, b_{\ell(b)})$, $c = (c_1, c_2, \ldots, c_{\ell(c)})$, and $d = (d_1, d_2, \ldots, d_{\ell(d)})$, and we have $ab \leq cd$. Due to the definition of $\leq$, this means that we must be in one of the following two situations:

- There exists an $i \in \{1, 2, \ldots, \min \{\ell(a), \ell(b) + \ell(c), \ell(d)\}\}$ such that
  
  \[
  (ab)_i < (cd)_i, \quad \text{and every } j \in \{1, 2, \ldots, i - 1\} \text{ satisfies } (ab)_j = (cd)_j. 
  \]

  The word $ab$ is a prefix of $cd$.

  It is easy to see that if we are in the second situation, then either we have $a \leq c$ or the word $c$ is a prefix of a 668. We are therefore done in the second situation. We can thus WLOG assume that we are in the first situation. Assume this. We thus have an $i \in \{1, 2, \ldots, \min \{\ell(a), \ell(b) + \ell(c), \ell(d)\}\}$ satisfying (11.911). If $i \leq \min \{\ell(a), \ell(c)\}$, then this yields $a \leq c$ (because every $k \in \{1, 2, \ldots, \min \{\ell(a), \ell(c)\}\}$ satisfies $(ab)_k = a_k$ and $(cd)_k = c_k$), and we are done. If $i > \min \{\ell(a), \ell(c)\}$, then every $j \in \{1, 2, \ldots, \min \{\ell(a), \ell(c)\}\}$ satisfies $a_j = c_j$ (because every $j \in \{1, 2, \ldots, \min \{\ell(a), \ell(c)\}\}$ satisfies $j \in \{1, 2, \ldots, i - 1\}$ and thus

  \[
  a_j = (ab)_j = (cd)_j \quad \text{(by (11.911))}
  \]

  and thus we conclude that either $a$ is a prefix of $c$, or $c$ is a prefix of $a$. Again, this means that we are done. Thus, Proposition 6.2(e) is proven.

(f) follows from (e), because if $c$ is a prefix of $a$ satisfying $\ell(a) \leq \ell(c)$, then $c = a$.

(g) Let $a, b, c \in \mathfrak{A}^*$ satisfy $a \leq b \leq ac$. Then, $b \varnothing = b \leq ac$. Hence, Proposition 6.2(e) (applied to $b, \varnothing, a, c$ instead of $a, b, c, d$) yields that either we have $b \leq a$ or the word $a$ is a prefix of $b$. Since $b \leq a$ leads to $a = b$ (in view of $a \leq b$), we get in both of these cases that $a$ is a prefix of $b$. This proves Proposition 6.2(g).

(j) Let $a, b, c \in \mathfrak{A}^*$ satisfy $a \leq b$ and $\ell(a) \geq \ell(b)$. Proposition 6.2(d) (applied to $a, c, b$ and $c$) yields that either we have $ac \leq bc$ or the word $a$ is a prefix of $b$. In the first of these two cases, we are obviously done.

667 Proof. Let us consider the second situation. In this situation, the word $ac$ is a prefix of $ad$. In other words, there exists a word $z$ such that $ad = azc$. Consider this $z$. Cancelling $a$ from the equality $ad = azc$, we obtain $d = cz$. Hence, the word $c$ is a prefix of $d$. Thus, $c \leq d$, qed.

668 Proof. Assume that we are in the second situation. Then, the word $ab$ is a prefix of $cd$. In other words, there exists a word $z \in \mathfrak{A}^*$ such that $cd = abz$. Consider this $z$.

Now, $a$ is a prefix of the word $cd$ (since $cd = abz = a(bz)$). Also, $c$ is a prefix of the word $cd$. Hence, $a$ and $c$ are two prefixes of the word $cd$. Thus, Proposition 6.2(i) (applied to $c$ and $cd$ instead of $b$ and $c$) yields that either $a$ is a prefix of $c$, or $c$ is a prefix of $a$. Hence, either we have $a \leq c$ or $c$ is a prefix of $a$ (because if $a$ is a prefix of $c$, then $a \leq c$). Qed.
Hence, we WLOG assume that we are in the second of these two cases. Thus, the word $a$ is a prefix of $b$. Since the word $a$ is at least as long as $b$ (in fact, $\ell(a) \geq \ell(b)$), this yields that $a = b$, so that $a = b\cdot c = bc \leq bc$.

This proves Proposition 6.2(j). \qed

11.102. Solution to Exercise 6.7. Solution to Exercise 6.7.

Alternative proof of Corollary 6.6. We shall prove Corollary 6.6 by strong induction on $\ell(u) + \ell(v) + \ell(w)$. That is, we fix an $N \in \mathbb{N}$, and we assume (as the induction hypothesis) that Corollary 6.6 holds in the case when $\ell(u) + \ell(v) + \ell(w) < N$. We then need to prove that Corollary 6.6 holds in the case when $\ell(u) + \ell(v) + \ell(w) = N$.

So let $u, v$ and $w$ be words satisfying $w \geq vu$ and $vw \geq vw$ and $\ell(u) + \ell(v) + \ell(w) = N$. Assume that $v$ is nonempty. We want to prove that $uw \geq wu$.

If $u = \emptyset$, then $uw \geq wu$ holds obviously (because if $u = \emptyset$, then $uw = w = w \cdot \emptyset = wu$). Hence, for the rest of this proof, we can WLOG assume that $u \neq \emptyset$. Assume this. For a similar reason, we WLOG assume that $w \neq \emptyset$. Recall also that $v$ is nonempty; that is, $v \neq \emptyset$.

We have $\ell(u) > 0$ (since $u \neq \emptyset$) and $\ell(v) > 0$ (since $v \neq \emptyset$) and $\ell(w) > 0$ (since $w \neq \emptyset$).

Notice that

\begin{equation}
(11.912) \quad uvc \geq vuc \quad \text{for every } c \in \mathbb{A}^*. \tag{11.912}
\end{equation}

Also,

\begin{equation}
(11.913) \quad vwc \geq wvc \quad \text{for every } c \in \mathbb{A}^*. \tag{11.913}
\end{equation}

Let us first assume that $u$ is a prefix of $w$. Then, there exists a $w' \in \mathbb{A}^*$ satisfying $w = uw'$ (since $u$ is a prefix of $w$). Consider this $w'$. Applying (11.912) to $c = w'$, we obtain $uwv' \geq vuv' = vuv' \geq vuv'$. That is, $uwv' \leq uwv'$. Hence, Proposition 6.2(c) (applied to $u = u$, $w = wv$ and $d = vw'$) yields $wv' \leq vw'$, so that $wv' \geq wv'$. Furthermore, $\ell(u) + \ell(v) + \ell(w') > \ell(u) + \ell(v) + \ell(w) = N$. Hence, Corollary 6.6 holds for $w'$ instead of $w$ (by the induction hypothesis). Consequently, we obtain $uw' \geq w'u$ (since $vuw' \geq w'u$). Thus, $w'u \leq uw'$, so that $uw'u \leq uw'v$ (by Proposition 6.2(b), applied to $u = u$, $c = w'u$ and $d = uw'$). Thus, $uwv' \geq uwv' = uw'$.

Now, let us forget that we assumed that $u$ is a prefix of $w$. We thus have proven that $uw \geq wu$ under the assumption that $u$ is a prefix of $w$. Hence, for the rest of this proof of $uw \geq wu$, we can WLOG assume that $u$ is not a prefix of $w$.

Assume this.

Let us next assume that $v$ is a prefix of $w$. Then, there exists a $w' \in \mathbb{A}^*$ satisfying $w = vw'$ (since $v$ is a prefix of $w$). Consider this $w'$. We have $vuv' = vuv' \geq vvw', thus $vw'v \leq vw'$. Thus, Proposition 6.2(c) (applied to $a = v, c = wv'$ and $d = vw'$) yields $wv' \geq vw'$, so that $wv' \geq wv'$. Furthermore,

---

669 Proof of (11.912): Let $c \in \mathbb{A}^*$. We have $uw \geq vu$ and thus $vu \leq vu$. Also, $\ell(vu) = \ell(v) + \ell(u) = \ell(u) + \ell(v) = \ell(uv)$. Hence, Proposition 6.2(i) (applied to $a = vu$ and $b = uv$) yields $vu \geq uv$, so that $uv \geq uv$. This proves (11.912).

670 Proof of (11.913): Let $c \in \mathbb{A}^*$. We have $vw \geq vu$ and thus $vu \leq vu$. Also, $\ell(vw) = \ell(w) + \ell(v) = \ell(v) + \ell(w) = \ell(vw)$. Hence, Proposition 6.2(i) (applied to $a = vu$ and $b = uv$) yields $vu \geq uv$, so that $uv \geq uv$. This proves (11.913).
Proposition 6.2(d) (applied to a prefix of a word w yields that we have uw ≥ wv). Since the word u is a prefix of w, we can WLOG assume that v is not a prefix of w.

Assume this.

We have vw ≥ wv, thus wv ≤ vw. Hence, Proposition 6.2(e) (applied to a = w, b = v, c = v and d = w) yields that either we have w ≤ v or the word v is a prefix of w. Since the word v is not a prefix of w, this yields that we have w ≤ v. Thus,

\[ v \geq w. \]

Let us next assume that w is a prefix of u. Then, there exists a u' ∈ W* satisfying u = wu' (since w is a prefix of u). Consider this u'. We have uw ≥ wu' and thus wuv = uw ≥ v, so that uw' ≥ uwv (by (11.913), applied to c = u'). In other words, uw' ≤ uwv. Hence, Proposition 6.2(c) (applied to a = w, c = vu' and d = u'v) yields vu' ≤ u'v, so that u'v ≥ vu'. Furthermore, \( \ell \left( \frac{u}{w} \right) = \ell (w) + \ell (w') < \ell (u') \) yields that either we have w ≤ u or the word v is a prefix of w. Since the word v is not a prefix of w, this yields that we have w ≤ v. Thus,

\[ v \geq w. \]

Now, let us forget that we assumed that w is a prefix of u. We thus have proven that uw ≥ wu under the assumption that w is a prefix of u. Hence, for the rest of this proof of uw ≥ wu, we can WLOG assume that w is not a prefix of u.

Assume this.

Let us now assume that u is a prefix of v. Then, there exists a v' ∈ W* satisfying v = vu' (since u is a prefix of v). Consider this v'. We have uv' = v ≥ w = w∅. Hence, w∅ ≤ uv'. Thus, Proposition 6.2(e) (applied to a = w, b = ∅, c = u and d = v') yields that either we have w ≤ u or the word u is a prefix of w. Since the word u is not a prefix of w, we can conclude from this that w ≤ u. Thus, Proposition 6.2(d) (applied to a = w, b = u, c = u and d = v) shows that either we have wu ≤ uv or the word v is a prefix of u. Since the word w is not a prefix of u, this yields that wu ≤ uv, so that uw ≥ wu.

Now, let us forget that we assumed that u is a prefix of v. We thus have proven that uw ≥ wu under the assumption that u is a prefix of v. Hence, for the rest of this proof of uw ≥ wu, we can WLOG assume that u is not a prefix of v.

Assume this.

We have uw ≥ wu, thus vu ≤ uv. Hence, Proposition 6.2(e) (applied to a = v, b = u, c = u and d = v) yields that either we have v ≤ u or the word u is a prefix of v. Since the word u is not a prefix of v, this yields that we have u ≥ v. Combined with v ≥ w, this yields u ≥ w, so that w ≤ u. Thus, Proposition 6.2(d) (applied to a = w, b = u, c = u and d = w) shows that either we have wu ≤ uv or the word v is a prefix of u. Since the word v is not a prefix of u, this yields that wu ≤ uv, so that uw ≥ wu. Our proof of uw ≥ wu is thus complete.
Now, let us forget that we fixed $u$, $v$ and $w$. We have thus shown that if $u$, $v$ and $w$ are words satisfying $uv \geq vu$ and $uw \geq vu$ and $\ell(u) + \ell(v) + \ell(w) = N$, and if $v$ is nonempty, then $uw \geq vw$. In other words, Corollary 6.6 holds in the case when $\ell(u) + \ell(v) + \ell(w) = N$. This completes the induction step. We thus have proven Corollary 6.6 (again). □

11.103. Solution to Exercise 6.9. Solution to Exercise 6.9. If the word $u$ is empty, then Exercise 6.9 is easy to solve\(^\text{671}\). Hence, for the rest of this solution, we can WLOG assume that the word $u$ is nonempty. Assume this. The word $u^n$ is nonempty (since $u$ is nonempty and $n$ is a positive integer). In other words, the word $v^m$ is nonempty (since $u^n = v^m$). Thus, $v$ is nonempty.

We have $u \underbrace{v^n = u^n v^n = u^{n+1} = u^n u = v^m u}$ and $v^m v = v^{m+1} = vv^m$. Hence, Corollary 6.6 (applied to $u$, $v^m$ and $v$ instead of $u$, $v$ and $w$) yields $uw \geq vu$ (since $v^m$ is nonempty).

But we also have $v^m w = v^m v$ (since $v^m v = vv^m$) and $v^m u = uu^m$ (since $uu^m = v^m u$). Thus, Corollary 6.6 (applied to $v^m$, $u$ and $v$ instead of $u$, $v$ and $w$) yields $vu \geq uv$ (since $v^m$ is nonempty). Combined with $uv \geq vu$, this yields $uw = vu$. Hence, Proposition 6.4 yields that there exist a $t \in \mathbb{R}^*$ and two nonnegative integers $i$ and $j$ such that $u = t^i$ and $v = t^j$.\(^\text{672}\) Consider this $t$ and these $i$ and $j$. We have $i \neq 0$ (since $t^i = u$ is nonempty), so that $i$ is a positive integer. Also, $j \neq 0$ (since $t^j = v$ is nonempty), and therefore $j$ is a positive integer.

We have thus shown that there exists a word $t$ and positive integers $i$ and $j$ such that $u = t^i$ and $v = t^j$. Exercise 6.9 is thus solved.

11.104. Solution to Exercise 6.10. Solution to Exercise 6.10. We WLOG assume that $u$ is nonempty (because otherwise, both $uv \geq vu$ and $u^n v^m \geq v^m u^n$ hold for trivial reasons). Similarly, we WLOG assume that $v$ is nonempty.

The exercise asks us to prove the equivalence $(uv \geq vu) \iff (u^n v^m \geq v^m u^n)$. We shall verify its $\implies$ and $\iff$ parts separately:

$\implies$: Assume that $uv \geq vu$ holds. Then, Corollary 6.6 (applied to $w = v^m$) yields $u^m v^m \geq v^m u$ (since $v^m v = v^{m+1} = v^m v$). Thus, Corollary 6.6 (applied to $u^n$, $u$ and $v^m$ instead of $u$, $v$ and $w$) yields $u^n v^m \geq v^m u^n$ (since $u^n u = u^{n+1} = uu^n$). This proves the $\implies$ part of the equivalence $(uv \geq vu) \implies (u^n v^m \geq v^m u^n)$.

$\iff$: Assume that $u^n v^m \geq v^m u^n$ holds. Then, Corollary 6.6 (applied to $u$, $v^m$ and $v$ instead of $u$, $v$ and $w$) yields $w^m \geq vu$ (since $uw^m = u^{m+1} = uu^n$, and since the word $u^n$ is nonempty).\(^\text{673}\). Hence, Corollary 6.6 (applied to $u$, $v^m$ and $v$ instead of $u$, $v$ and $w$) yields $uw \geq vu$ (since $v^m v = v^{m+1} = vv^m$, and since the word $v^m$ is nonempty).\(^\text{674}\). This proves the $\iff$ part of the equivalence $(uv \geq vu) \iff (u^n v^m \geq v^m u^n)$.

Thus, both the $\implies$ and the $\iff$ part of the equivalence $(uv \geq vu) \iff (u^n v^m \geq v^m u^n)$ are proven, and Exercise 6.10 is solved.


Notice that $\ell(u^n) = n \ell(u) = m \ell(v)$ (since $\ell(v^m) = m \ell(v)$), thus $\ell(v^m) = \ell(u^n)$. Now, we have the following two logical implications:

\[
(u^n v^m \geq v^m u^n) \implies (u^n \geq v^m)
\]

\[
(u^n \geq v^m) \implies (u^n v^m \geq v^m u^n)
\]

\(^\text{671}\)Proof. Assume that the word $u$ is empty. Thus, $u = \varnothing$. Hence, $u^n = \varnothing^n = \varnothing$, so that $\varnothing = u^n = v^m$ and thus $v^m = \varnothing$. Consequently, $v = \varnothing$ (since $m$ is positive). Thus, there exists a word $t$ and positive integers $i$ and $j$ such that $u = t^i$ and $v = t^j$ (namely, we can take $t = \varnothing$ and $i = 1$ and $j = 1$). In other words, Exercise 6.9 is solved.

\(^\text{672}\)The $i$ and $j$ here are the variables called $n$ and $m$ in Proposition 6.4.

\(^\text{673}\)Because $u$ is nonempty and $n$ is positive

\(^\text{674}\)Because $v$ is nonempty and $m$ is positive

\(^\text{675}\)Proof. Assume that $u^n v^m \geq v^m u^n$. We need to prove that $u^n \geq v^m$. 

Combining these two implications, we obtain the equivalence \((u^n v^m \geq v^m u^n) \iff (u^n \geq v^m)\). But Exercise 6.10 shows that we have the equivalence \((uv \geq vu) \iff (u^n v^m \geq v^m u^n)\). Altogether, we thus obtain the following chain of equivalences:

\[
(uw \geq vu) \iff (u^n v^m \geq v^m u^n) \iff (u^n \geq v^m)
\]

This solves Exercise 6.11.

11.106. Solution to Exercise 6.12. Solution to Exercise 6.12. If \(w\) is a nonempty word, then let us denote by \(\text{rad}(w)\) the shortest word \(p\) such that \(w\) is a power of \(p\). (This is clearly well-defined because \(w\) is a power of itself, and because for every given integer \(\lambda\) there exists at most one word \(v\) of length \(\lambda\) such that \(w\) is a power of \(v\).)

Every nonempty word \(w\) and every positive integer \(n\) satisfy

\[
(11.914) \quad \text{rad}(w^n) = \text{rad}(w).
\]

Proof of (11.914): Let \(w\) be a nonempty word, and let \(n\) be a positive integer. Let \(q = \text{rad}(w^n)\). Thus, \(q\) is the shortest word \(p\) such that \(w^n\) is a power of \(p\). Consequently, \(w^n\) is a power of \(q\), so that there exists a positive integer \(N\) such that \(w^n = q^N\). Consider this \(N\).

The word \(w^n\) is nonempty (since \(w\) is nonempty and \(n\) is positive). That is, the word \(q^N\) is nonempty (since \(w^n = q^N\)). Hence, the word \(q\) is nonempty, so that \(\ell(q)\) is nonzero.

Exercise 6.9 (applied to \(w, q, n\) and \(N\) instead of \(u, v, n\) and \(m\)) yields that there exists a word \(t\) and positive integers \(i\) and \(j\) such that \(w = t^i\) and \(q = t^j\). Consider these \(t\), \(i\), and \(j\). We have \(q = t^j\) and thus \(q^N = (t^j)^N = t^{jn}\), so that \(t^{jn} = q^N = w^n\). Hence, \(w^n\) is a power of \(t\). The word \(t\) thus cannot be shorter than \(q\) (since \(q\) is the shortest word \(p\) such that \(w^n\) is a power of \(p\)). Consequently, \(j = 1\) (because otherwise, \(t\) would be shorter than \(t^j = q\)). Hence, \(q = t^j = t\) (since \(j = 1\)) and \(w = t^j = q^i\) (since \(t = q\)). Thus, \(w\) is a power of \(q\). Consequently, \(\ell(q) \geq \ell(\text{rad}(w))\) (since \(\text{rad}(w)\) is the shortest word \(p\) such that \(w\) is a power of \(p\)).

On the other hand, \(\text{rad}(w)\) is the shortest word \(p\) such that \(w\) is a power of \(p\). Thus, \(w\) is a power of \(\text{rad}(w)\). That is, there exists a \(P \in \mathbb{N}\) such that \(w = (\text{rad}(w))^P\). Consider this \(P\). We have \(P > 0\) (since \((\text{rad}(w))^P = w\) is nonempty), so that \(Pn > 0\) (since \(n > 0\)). Also, taking both sides of the equality \(w = (\text{rad}(w))^P\) to the \(n\)-th power, we obtain \(w^n = (\text{rad}(w))^P^n = (\text{rad}(w))^Pn\), so that \(w^n\) is a power of \(\text{rad}(w)\).

Recall that \(q\) is the shortest word \(p\) such that \(w^n\) is a power of \(p\). Since \(w^n\) is a power of \(\text{rad}(w)\), this yields that \(\ell(\text{rad}(w)) \geq \ell(q)\). Combined with \(\ell(q) \geq \ell(\text{rad}(w))\), this yields \(\ell(\text{rad}(w)) = \ell(q)\). Now,

\[
\ell \left( \frac{w^n}{(\text{rad}(w))^n} \right) = \ell \left( (\text{rad}(w))^Pn \right) = Pn \cdot \ell(\text{rad}(w)) = Pn \cdot \ell(q).
\]

We have \(v^m u^n \leq u^n v^m\) (since \(u^n v^m \geq v^m u^n\)) and \(\ell(v^m) \leq \ell(u^n)\) (since \(\ell(v^m) = \ell(u^n)\)). Hence, \(v^m \leq u^n\) (by Proposition 6.2(f), applied to \(a = v^m\), \(b = u^n\), \(c = u^n\) and \(d = v^m\)). Thus, \(u^n \geq v^m\), qed.

Proof. Assume \(u^n \geq v^m\). We need to prove that \(u^n v^m \geq v^m u^n\). Assume this.

If \(v^m u^n \leq u^n v^m\), then \(u^n v^m \geq v^m u^n\) is obviously true. Hence, for the rest of this proof of \(u^n v^m \geq v^m u^n\), we can WLOG assume that we don’t have \(v^m u^n \leq u^n v^m\). Assume this.

We have \(v^m \leq u^n\) (since \(u^n \geq v^m\)). Hence, Proposition 6.2(d) (applied to \(a = v^m\), \(b = u^n\), \(c = u^n\) and \(d = v^m\)) yields that either we have \(v^m u^n \leq u^n v^m\) or the word \(v^m\) is a prefix of \(u^n\). Since we don’t have \(v^m u^n \leq u^n v^m\), we therefore conclude that the word \(v^m\) is a prefix of \(u^n\). In other words, there exists a \(t \in \mathbb{N}^*\) such that \(u^n = v^m t\). Consider this \(t\).

We have \(\ell(v^m) = \ell \left( \frac{u^n}{v^m t} \right) = \ell \left( \frac{v^m}{u^n} \right) = \ell(v^m) + \ell(t)\), thus \(0 = \ell(t)\). Hence, the word \(t\) is empty, i.e., we have \(t = \emptyset\).

Thus, \(u^n = v^m \frac{v^m}{u^n} = v^m\), so that \(\frac{u^n}{v^m} = v^m u^n \geq v^m u^n\), qed.
Compared with $\ell \left( \frac{u^n}{q^n} \right) = \ell (q^N) = N \cdot \ell (q)$, this yields $Pn \cdot \ell (q) = N \cdot \ell (q)$. Division by $\ell (q)$ (which is nonzero) yields $Pn = N$. Thus, $q^{Pn} = q^N = w^n = (\text{rad} (w))^{Pn}$.

Since taking the $Pn$-th root of a word is unambiguous (when said root exists)\(^{677}\), this yields $q = \text{rad} (w)$. Because of $q = \text{rad} (w^n)$, this rewrites as $\text{rad} (w^n) = \text{rad} (w)$. This proves (11.914).

Next, we notice that (11.915) any two nonempty words $u$ and $v$ satisfying $uv = vu$ satisfy $\text{rad} (u) = \text{rad} (v)$.

Proof of (11.915): Let $u$ and $v$ be two nonempty words satisfying $uv = vu$. Proposition 6.4 yields that there exist a $t \in \mathcal{A}^*$ and two nonnegative integers $n$ and $m$ such that $u = t^n$ and $v = t^m$. Consider this $t$ and these $n$ and $m$. Since $u = t^n$, we have $\text{rad} (u) = \text{rad} (t^n) = \text{rad} (t)$ (by (11.914), applied to $t$ instead of $w$). Similarly, $\text{rad} (v) = \text{rad} (t)$. Thus, $\text{rad} (u) = \text{rad} (t) = \text{rad} (v)$, and this proves (11.915).

Now, we can finally solve the exercise. Let us show that (11.916) $u_{1i}u_i \geq u_iu_{1i}$ for every $i \in \{1, 2, \ldots, k\}$.

Indeed, (11.916) can be proven by induction over $i$: The base case ($i = 1$) is obvious, whereas the induction step (proving $u_{1i}u_{i+1} \geq u_{i+1}u_i$ using $u_{1i}u_i \geq u_iu_{1i}$) results from applying Corollary 6.6 to $u = u_1$, $v = u_i$ and $w = u_{i+1}$ (because we have $u_{1i}u_{i+1} \geq u_{i+1}u_i$ by assumption). We thus have shown (11.916).

Now, we can apply (11.916) to $i = k$, and obtain $u_1u_k \geq u_ku_1$. But on the other hand, we can apply $u_{1i}u_{i+1} \geq u_{i+1}u_i$ to $i = k$, and obtain $u_ku_{k+1} \geq u_{k+1}u_k$. Since $u_{k+1} = u_1$, this rewrites as $u_ku_1 \geq u_1u_k$. Contrasting this with $u_1u_k \geq u_ku_1$, we obtain $u_1u_k = u_ku_1$. Hence, (11.915) (applied to $u = u_1$ and $v = u_k$) yields $\text{rad} (u_1) = \text{rad} (u_k)$. Similarly, we can show that $\text{rad} (u_i) = \text{rad} (u_{i-1})$ for every $i \in \{1, 2, \ldots, k\}$, where $u_0$ means $u_k$ (in fact, our situation is invariant under cyclically shifting the $k$-tuple $(u_1, u_2, \ldots, u_k)$). Hence, $\text{rad} (u_k) = \text{rad} (u_{k-1}) = \cdots = \text{rad} (u_1)$. Denote this common value $\text{rad} (u_k) = \text{rad} (u_{k-1}) = \cdots = \text{rad} (u_1)$ by $t$. Then, for every $i \in \{1, 2, \ldots, k\}$, we have $t = \text{rad} (u_i)$, whence $u_i$ is a power of $t$ (because $\text{rad} (u_i)$ is defined as the shortest word $p$ such that $u_i$ is a power of $p$). This solves Exercise 6.12.

11.07. Solution to Exercise 6.21. Solution to Exercise 6.21. (a) We shall solve Exercise 6.21(a) by strong induction on $\ell (u) + \ell (v)$.

Induction step: Let $N \in \mathbb{N}$. Assume that Exercise 6.21(a) is already solved in the case when $\ell (u) + \ell (v) < N$. We now need to solve Exercise 6.21(a) in the case when $\ell (u) + \ell (v) = N$.

We have assumed that Exercise 6.21(a) is already solved in the case when $\ell (u) + \ell (v) < N$. In other words, we know that (11.917) if $u \in \mathcal{A}^*$ and $v \in \mathcal{A}^*$ are two words satisfying $uv < vu$ and $\ell (u) + \ell (v) < N$, then there exists a nonempty suffix $s$ of $u$ satisfying $sv < v$.

So let $u \in \mathcal{A}^*$ and $v \in \mathcal{A}^*$ be two words satisfying $uv < vu$ and $\ell (u) + \ell (v) = N$. We are going to prove that there exists a nonempty suffix $s$ of $u$ satisfying $sv < v$.

Indeed, let us assume the contrary. Thus, there exists no nonempty suffix $s$ of $u$ satisfying $sv < v$. Hence, (11.918) whenever $s$ is a nonempty suffix of $u$, we do not have $sv < v$.

We have $u \neq \varnothing$ (since otherwise, we would have $u = \varnothing$ and thus $u \circ v = v = \varnothing = vu$, which would contradict $uv < vu$). In other words, the word $u$ is nonempty. Similarly, the word $v$ is nonempty. We have $\ell (u) > 0$ (since $u$ is nonempty) and $\ell (v) > 0$ (since $v$ is nonempty).

We have $\ell (uv) = \ell (u) + \ell (v) > \ell (v)$, thus $\ell (uv) \neq \ell (v)$, hence $uv \neq v$.

\(^{677}\)because $Pn > 0$
Since $u$ is a nonempty suffix of $v$, we can apply (11.918) to $s = u$. As a result, we conclude that we do **not** have $uv < v$. Hence, we do **not** have $uv \leq v$\footnote{Proof. Assume the contrary. Then, $uv < v$. Since $uv \neq v$, this yields $uv < v$. This contradicts the fact that we do not have $uv < v$. This contradiction proves that our assumption was wrong, qed.}. In other words, we do **not** have $uv \leq v\emptyset$ (since $v\emptyset = v$).

But $uv\emptyset = uv < vu$. Thus, Proposition 6.2(e) (applied to $a = uv$, $b = \emptyset$, $c = v$ and $d = u$) yields that either we have $uv \leq v$ or the word $v$ is a prefix of $uv$. Since we do not have $uv \leq v$, we can therefore conclude that the word $v$ is a prefix of the word $uv$. In other words, there exists a word $g \in \mathfrak{A}^*$ such that $uv = vg$.

Consider this $g$.

Since $u$ is a prefix of $uv$, we have $u \leq uv < vu$. Thus, $u\emptyset = u < vu$. Hence, Proposition 6.2(e) (applied to $a = u$, $b = \emptyset$, $c = v$ and $d = u$) yields that either we have $u \leq v$ or the word $v$ is a prefix of $u$. In other words, we are in one of the following two cases:

Case 1: We have $u \leq v$.

Case 2: The word $v$ is a prefix of $u$.

Let us consider Case 1 first. In this case, we have $u \leq v$. Thus, Proposition 6.2(d) (applied to $a = u$, $b = v$, $c = v$ and $d = qu$) yields that either we have $uv \leq v\emptyset$ or the word $u$ is a prefix of $v$. Since we do not have $uv \leq v\emptyset$, this yields that the word $u$ is a prefix of $v$. In other words, there exists a word $q \in \mathfrak{A}^*$ such that $v = uq$. Consider this $q$. We have $\ell\left(\begin{array}{c} v \\ = uq \end{array}\right) = \ell\left(\begin{array}{c} uq \\ = uq \end{array}\right) = \ell(uq) = \ell(u) + \ell(q) > \ell(q)$, so that $\ell(q) < \ell(v)$. Also, $\ell(uq) = uq \leq vq$. Hence, $uq < qu$\footnote{Proof. If we had $uq = qu$, then we would have $u \hat{u}q = ugu$, which would contradict $uq < uqu$. Hence, we cannot have $uq = qu$. Thus, we have $uq \neq qu$. Combined with $qu \leq qu$, this yields $uq < qu$, qed.}.

Since $\ell(u) + \ell(q) < \ell(u) + \ell(v) = N$, we can therefore apply (11.917) to $q$ instead of $v$. As a result, we obtain that there exists a nonempty suffix $s$ of $u$ satisfying $sq < q$. Let us denote this $s$ by $t$. Thus, $t$ is a nonempty suffix of $u$ satisfying $tq < q$. Thus, Proposition 6.2(j) (applied to $a = tq$, $b = q$ and $c = g$) yields that $tgq \leq qg$ (because $\ell(tq) = \ell(t) + \ell(q) > \ell(q)$).

But $uq \leq v$ (since $uv = vg$). Cancelling $u$ from this equality, we obtain $qg = v$. Hence, $tqg \leq qg = v$. But we have $\ell(tv) = \ell(t) + \ell(v) > \ell(v)$, thus $\ell(tv) \neq \ell(v)$, hence $tv \neq v$. Combined with $tv \leq v$, this yields $tv < v$.

On the other hand, $t$ is a nonempty suffix of $u$. Hence, (11.918) (applied to $s = t$) yields that we do not have $tv < v$. This contradicts the fact that $tv < v$. Hence, we have obtained a contradiction in Case 1.

Let us now consider Case 2. In this case, the word $v$ is a prefix of $u$. Hence, there exists a word $r \in \mathfrak{A}^*$ satisfying $u = vr$. Consider this $r$. Clearly, $r$ is a suffix of $u$ (since $u = vr$).

We have $vr \leq uv < v \emptyset = vvr$. Thus, Proposition 6.2(c) (applied to $a = v$, $c = rv$ and $d = vr$) yields $rv \leq vr$. Hence, $rv < vr$\footnote{Proof. If we had $rv = vr$, then we would have $v \hat{r}v = vvr$, which would contradict $vrv < vvr$. Hence, we cannot have $rv = vr$. Thus, we have $rv \neq vr$. Combined with $rv \leq vr$, this yields $rv < vr$, qed.}.

Also, $\ell\left(\begin{array}{c} u \emptyset \\ = vr \end{array}\right) = \ell(vr) = \ell(v) + \ell(r) > \ell(r)$, so that $\ell(r) < \ell(u)$ and thus $\ell(r) + \ell(v) < \ell(u) + \ell(v) = N$. Therefore, (11.917) can be applied to $r$ instead of $u$. As a result,
we obtain that there exists a nonempty suffix s of r satisfying sv < v. Let us denote this by t. Thus, t is a nonempty suffix of r satisfying tv < v.

We know that t is a suffix of the word r, which (in turn) is a suffix of u. Hence, t is a suffix of u. Thus, (11.918) (applied to s = t) yields that we do not have tv < v. This contradicts the fact that tv < v. Hence, we have obtained a contradiction in Case 2.

We thus have obtained a contradiction in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that we always have a contradiction. Therefore, our assumption was wrong. So we have proven that there exists a nonempty suffix s of u satisfying sv < v.

Now, let us forget that we fixed u and v. We thus have proven that if u ∈ A* and v ∈ A* are two words satisfying uv < vu and ℓ(u) + ℓ(v) = N, then there exists a nonempty suffix s of u satisfying sv < v. In other words, we have solved Exercise 6.21(a) in the case when ℓ(u) + ℓ(v) = N. This completes the induction step. Hence, Exercise 6.21(a) is solved by strong induction.

(b)

Alternative proof of Theorem 6.20. Proof of the implication A → B: Assume that Assertion A holds. Thus, the word w is Lyndon. Hence, every nonempty proper suffix of w is > w (by the definition of a Lyndon word). Now, let u and v be any nonempty words satisfying w = uv. Then, v is a nonempty proper suffix of w, and therefore > w (since every nonempty proper suffix of w is > w). That is, we have v > w.

Now, let us forget that we fixed u and v. We thus have shown that any nonempty words u and v satisfying w = uv satisfy v > w. In other words, Assertion B holds. Thus, the implication A → B is proven.

Proof of the implication A → C: This implication follows from Proposition 6.14(b).

Proof of the implication A → D: This implication follows from Proposition 6.14(c).

Proof of the implication B → A: Assume that Assertion B holds. Let v be a nonempty proper suffix of w. Then, there exists a nonempty u ∈ A* satisfying w = uv (since v is a proper suffix of w). Consider this u. Assertion B yields v > w.

Now, let us forget that we fixed v. We thus have shown that every nonempty proper suffix v of w satisfies v > w. By the definition of a Lyndon word, this yields that w is Lyndon (since w is nonempty), so that Assertion A holds. Hence, the implication B → A is proven.

Proof of the implication C → B: Assume that Assertion C holds. Let us prove that Assertion B holds.

Indeed, let u and v be two nonempty words satisfying w = uv. We will show that v > w.

Let us first notice that Assertion C yields v > u. We have h = 0 belongs to this set R (since u^n = ∅ is a prefix of v), and therefore R is nonempty. Also, this set R is finite681. Hence, the set R has a maximum element (since it is nonempty and finite). Let m be this maximum element. Then, m ∈ R but m + 1 \∉ R.

We have h ∈ R = {k ∈ N | u^k is a prefix of v}. Thus, m is an element of N such that u^m is a prefix of v. In other words, there exists a word v' ∈ A* such that v = u^m v'. Consider this v'. Using m + 1 \∉ R, it is easy to see that u is not a prefix of v' 682.

681 Proof. Let i be an element of R. Thus, i ∈ R = {k ∈ N | u^k is a prefix of v}. In other words i is an element of N such that u^i is a prefix of v. The word u^i is not longer than v (since it is a prefix of v); thus \ell(u^i) ≤ \ell(v). But u is nonempty, and thus \ell(u^i) ≥ 1. Hence, \ell(u^i) = i \ell(u) ≥ i, so that i ≤ \ell(u^i) ≤ \ell(v).

Now, let us forget that we fixed i. We thus have proven that every element i of R satisfies i ≤ \ell(v). Thus, there are only finitely many elements of R (since there are only finitely many i ∈ N satisfying i ≤ \ell(v)). In other words, the set R is finite.

682 Proof. Assume the contrary. Thus, u is a prefix of v. In other words, there exists a word t ∈ A* such that v' = ut. Consider this t. We have v = u^m v' = u^m t = u^m+1 t, and thus the word u^m+1 is a prefix of v. Hence, m + 1 is an element of N such that u^m+1 is a prefix of v. In other words, m + 1 ∈ {k ∈ N | u^k is a prefix of v} = R. But this contradicts m + 1 \∉ R. This contradiction proves that our assumption was wrong, qed.
It is easy to see that the word $v'$ is nonempty. Also, the word $uu^m$ is nonempty (since $u$ is nonempty). We have $w = u \overset{v}{=} uu^m v'$. Therefore, Assertion $\mathcal{C}$ (applied to $uu^m$ and $v'$ instead of $u$ and $v$) yields $v' > uu^m \geq u$. Thus, $u \leq v'$. Hence, Proposition 6.2(d) (applied to $a = u$, $b = v'$, $c = v'$ and $d = \emptyset$) yields that either we have $uv' \leq v' \emptyset$ or the word $u$ is a prefix of $v'$. Since we know that the word $u$ is not a prefix of $v'$, we can thus conclude that $uv' \leq v' \emptyset = v'$. Hence, Proposition 6.2(b) (applied to $a = u^m$, $c = wv'$ and $d = v'$) yields $u^m uv' \leq u^m v'$. Thus, $u^m v' \geq u^m v' = v$. Since $u^m v' = v$, this rewrites as $v \geq w$. Hence, $v > w$.

Now, let us forget that we fixed $u$ and $v$. We thus have proven that any nonempty words $u$ and $v$ satisfying $w = uv$ satisfy $v > w$. In other words, Assertion $\mathcal{B}$ holds. Hence, the implication $\mathcal{C} \Rightarrow \mathcal{B}$ is proven.

Proof of the implication $\mathcal{D} \Rightarrow \mathcal{B}$: Assume that Assertion $\mathcal{D}$ holds. Let us prove that Assertion $\mathcal{B}$ holds. We shall prove Assertion $\mathcal{B}$ by strong induction over $\ell(u)$:

Induction step: Let $N \in \mathbb{N}$. Assume that Assertion $\mathcal{B}$ holds in the case when $\ell(u) < N$. We now need to prove Assertion $\mathcal{B}$ in the case when $\ell(u) = N$.

We have assumed that Assertion $\mathcal{B}$ holds in the case when $\ell(u) < N$. In other words, \[ (11.919) \quad \text{any nonempty words } u \text{ and } v \text{ satisfying } w = uv \text{ and } \ell(u) < N \text{ satisfy } v > w. \]

Now, let $u$ and $v$ be two nonempty words satisfying $w = uv$ and $\ell(u) = N$. We shall prove that $v > w$.

From Assertion $\mathcal{D}$, we obtain $uv > w$. Thus, $uv < vu$. Hence, Exercise 6.21(a) yields that there exists a nonempty suffix $s$ of $u$ satisfying $sv < v$. Consider this $s$. There exists a $p \in \mathbb{A}^*$ such that $u = ps$ (since $s$ is a suffix of $u$). Consider this $p$. Since $s$ is nonempty, we have $\ell(s) > 0$. Now, $\ell(\overset{w}{u}) = \ell(p) + \ell(s) > \ell(p)$, so that $\ell(p) < \ell(u) = N$. Also, $w = \overset{w}{u} v = psv$. Furthermore, if $p = \emptyset$, then $v > w$ is true. Hence, for the rest of our proof of $v > w$, we can WLOG assume that we don’t have $p = \emptyset$. Assume this. Thus, $p \neq \emptyset$, so that the word $p$ is nonempty. Also, the word $sv$ is nonempty (since $v$ is nonempty).

Now, the words $p$ and $sv$ are nonempty and satisfy $w = psv$ and $\ell(p) < N$. Hence, we can apply (11.919) to $p$ and $sv$ instead of $u$ and $v$. As a result, we obtain $sv > w$. But $sv < v$, so that $v > sv > w$. Hence, we have proven that $v > w$.

Now, let us forget that we fixed $u$ and $v$. We thus have proven that any nonempty words $u$ and $v$ satisfying $w = uv$ and $\ell(u) = N$ satisfy $v > w$. In other words, we have proven Assertion $\mathcal{B}$ in the case when $\ell(u) = N$. Hence, the induction step is complete. Assertion $\mathcal{B}$ is thus proven by strong induction. And so, we have established the implication $\mathcal{D} \Rightarrow \mathcal{B}$.

Combining the implications $\mathcal{A} \Rightarrow \mathcal{B}$, $\mathcal{A} \Rightarrow \mathcal{C}$, $\mathcal{A} \Rightarrow \mathcal{D}$, $\mathcal{B} \Rightarrow \mathcal{A}$, $\mathcal{C} \Rightarrow \mathcal{B}$ and $\mathcal{D} \Rightarrow \mathcal{B}$ that we have proven, we obtain the equivalence $\mathcal{A} \iff \mathcal{B} \iff \mathcal{C} \iff \mathcal{D}$. Hence, Theorem 6.20 is proven again. \[ \square \]

---

\[ ^{683} \text{Proof.} \] Assume the contrary. Thus, the word $v'$ is empty; that is, we have $v' = \emptyset$. Now, $v = u^m v' = u^m$ and $w = u \overset{v}{=} uu^m = u^m + 1 = u^m + 1 \equiv u = vu$. Hence, Assertion $\mathcal{C}$ (applied to $v$ and $u$ instead of $u$ and $v$) yields $u > v$. This contradicts $v > u$. This contradiction proves that our assumption was wrong, qed.

\[ ^{684} \text{Proof.} \] We have $\ell(\overset{w}{u}) = \ell(uv) = \ell(u) + \ell(v) > \ell(v)$, so that $\ell(u) \neq \ell(v)$ and thus $w \neq v$. Combined with $v \geq w$, this yields $v > w$, qed.

\[ ^{685} \text{Proof.} \] Assume that $p = \emptyset$. Then, $u = p$, so that $s = u$. Then, $w = uv = w$, so that $w = sv < v$ and thus $v > w$, qed.
11.108. \textbf{Solution to Exercise 6.22.} \textit{Solution to Exercise 6.22.} Let us first assume that \( w \) is Lyndon. We shall prove that
\begin{equation}
\text{(every nonempty word} \ t \text{and every positive integer} \ n \text{ satisfy (if} \ w \leq t^n, \text{then} \ w \leq t)\).
\end{equation}

Let \( t \) be a nonempty word, and let \( n \) be a positive integer. Assume that \( w \leq t^n \). We need to prove that \( w \leq t \).

Assume the contrary. Thus, we don’t have \( w \leq t \). In other words, we don’t have \( w \leq t^1 \).

Let \( m \) be the minimal \( i \in \{1, 2, \ldots, n\} \) satisfying \( w \leq t^i \). \(^{686}\) Then, \( w \leq t^m \). Hence, \( m \neq 1 \) (because \( w \leq t^m \), but we don’t have \( w \leq t^1 \)). Thus, \( m \geq 2 \), so that \( m - 1 \) is also an element of \( \{1, 2, \ldots, n\} \). If we had \( w \leq t^{m-1} \), then \( m - 1 \) would be an \( i \in \{1, 2, \ldots, n\} \) satisfying \( w \leq t^i \), which would contradict the fact that \( m \) is the minimal such \( i \). Thus, we cannot have \( w \leq t^{m-1} \).

But \( w\emptyset = w \leq t^{m-1}t \). Hence, Proposition 6.2(e) (applied to \( a = w \), \( b = \emptyset \), \( c = t^{m-1} \) and \( d = t \)) yields that either we have \( w \leq t^{m-1} \) or the word \( t^{m-1} \) is a prefix of \( w \). Since we cannot have \( w \leq t^{m-1} \), this shows that the word \( t^{m-1} \) is a prefix of \( w \). In other words, there exists \( v \in \mathfrak{A}^* \) such that \( w = t^{m-1}v \). Consider this \( v \). We have \( v \neq \emptyset \) (because otherwise, we would have \( v = \emptyset \) and thus \( w = t^{m-1} \emptyset = t^{m-1} \leq t^{m-1} \), contradicting the fact that we cannot have \( w \leq t^{m-1} \), so that \( v \) is nonempty. Hence, Proposition 6.14(b) (applied to \( u = t^{m-1} \)) now yields \( v > t^{m-1} \). Hence, \( t^{m-1} < v \). Thus, Proposition 6.2(b) (applied to \( a = t^{m-1} \), \( c = t^{m-1} \) and \( d = v \)) yields \( t^{m-1}t^{m-1} \leq t^{m-1}v = w \). Hence,
\begin{equation}
w \geq t^{m-1}t^{m-1} = t^{2(m-1)}.
\end{equation}

Combined with
\begin{equation}
w \leq t^m \leq t^mt^m
\end{equation}
(this makes sense since \( m \geq 2 \))
\begin{equation}
t^m = t^{m+1} = t^2(m-1),
\end{equation}
this yields \( w = t^{2(m-1)} = t^{m-1}t^{m-1} \), so that \( t^{m-1}t^{m-1} = w = t^{m-1}v \). Cancelling \( t^{m-1} \) in this, we obtain \( t^{m-1} = v \), which contradicts \( t^{m-1} < v \). This contradiction shows that our assumption (that we don’t have \( w \leq t \)) was false. Hence, \( w \leq t \).

Forget now that we assumed that \( w \leq t^n \). We thus have proven that if \( w \leq t^n \), then \( w \leq t \).

Now, forget that we fixed \( t \) and assumed that \( w \) is Lyndon. We thus have shown that
\begin{equation}
\text{(if} \ w \text{is Lyndon, then} \ (11.920) \text{ holds).}
\end{equation}

Now, conversely, assume that \( (11.920) \) holds. We will prove that \( w \) is Lyndon.

In fact, assume the contrary. Then, \( w \) is not Lyndon. Let \( v \) be the (lexicographically) smallest nonempty suffix of \( w \). Proposition 6.19(b) yields that there exists a nonempty \( u \in \mathfrak{A}^* \) such that \( w = uv \), \( u \geq v \) and \( uw \geq vu \). Consider this \( u \). We have \( u \geq v \), thus \( v \leq u \), and therefore Proposition 6.2(b) (applied to \( a = u \), \( c = v \) and \( d = u \)) yields \( uv \leq uu \). Now, \( w = w \leq uu = t^2 \). Thus, \( (11.920) \) (applied to \( t = u \) and \( n = 2 \)) yields \( w \leq u \). But this contradicts the fact that \( w = uv > u \) (since \( v \) is nonempty). As this contradiction shows, our assumption was wrong. Thus, we have shown that \( w \) is Lyndon.

Now, forget that we fixed \( w \). We thus have proven that
\begin{equation}
\text{(if} \ (11.920) \text{ holds, then} \ w \text{is Lyndon).}
\end{equation}

Combined with \( (11.921) \), this yields that \( w \) is Lyndon if and only if \( (11.920) \) holds. This solves the exercise.

11.109. \textbf{Solution to Exercise 6.23.} \textit{Solution to Exercise 6.23.} We will solve Exercise 6.23 by induction over \( n \):

The \textit{induction base} is the case \( n = 1 \); this case is vacuously true (since \( w_1 < w_n \) is impossible for \( n = 1 \)).

For the \textit{induction step}, we fix a positive integer \( N > 1 \), and we assume that Exercise 6.23 has been solved for \( n = N - 1 \). We now must solve Exercise 6.23 for \( n = N \).

So let \( w_1, w_2, \ldots, w_N \) be \( N \) Lyndon words. Assume that \( w_1 \leq w_2 \leq \cdots \leq w_N \) and \( w_1 < w_N \). We need to show that \( w_1w_2\cdots w_N \) is a Lyndon word.

\(^{686}\)Such an \( i \) exists, because \( n \in \{1, 2, \ldots, n\} \) satisfies \( w \leq t^n \).
Proposition 6.16(a) (applied to \( u = w_1 \) and \( v = w_N \)) yields that the word \( w_1w_N \) is Lyndon. If \( N = 2 \), then this yields that \( w_1w_2 \cdots w_N \) is Lyndon (because \( w_1w_2 \cdots w_N = w_1w_N \) if \( N = 2 \)). Thus, if \( N = 2 \), then we are done. We therefore WLOG assume that we don’t have \( N = 2 \). Hence, \( N \geq 3 \). Thus, \( w_3w_4 \cdots w_N \) is a nonempty product. More precisely, \( w_3w_4 \cdots w_N \) is a nonempty product of nonempty words (since the words \( w_3, w_4, \ldots, w_N \) are nonempty (because they are Lyndon)), and therefore a nonempty word itself. Hence, \( w_2 < w_2\left( w_3w_4 \cdots w_N \right) = w_2w_3 \cdots w_N \).

We have \( w_1 \leq w_2 \leq \cdots \leq w_N \). Hence, \( w_2 \leq w_3 \leq \cdots \leq w_N \) and, in particular, \( w_2 \leq w_N \). We distinguish between two cases:

Case 1: We have \( w_2 = w_N \).

Case 2: We have \( w_2 \neq w_N \).

Let us first consider Case 1. In this case, we have \( w_2 = w_N \). Hence, \( w_1 < w_N = w_2 \). Thus, Proposition 6.16(a) (applied to \( u = w_1 \) and \( v = w_2 \)) yields that the word \( w_1w_2 \) is Lyndon. Moreover, \( w_1w_2 < w_2 \) (by Proposition 6.16(b), applied to \( u = w_1 \) and \( v = w_2 \)). Hence, \( w_1w_2 \leq w_3 \leq w_4 \leq \cdots \leq w_N \). This follows by combining \( w_1w_2 < w_2 \) and \( w_2 \leq w_3 \leq \cdots \leq w_N \) and \( w_1w_2 < w = w_2 \). Thus, we can apply Exercise 6.23 to \( N - 1 \) and \( (w_1w_2, w_3, \ldots, w_N) \) instead of \( n \) and \( (w_1, w_2, \ldots, w_n) \) (because we have assumed that Exercise 6.23 has been solved for \( n = N - 1 \)). We thus obtain that \( w_1w_2w_3w_4 \cdots w_N \) is a Lyndon word. In other words, \( w_1w_2 \cdots w_N \) is a Lyndon word. So we have shown that \( w_1w_2 \cdots w_N \) is a Lyndon word in Case 1.

Let us now consider Case 2. In this case, we have \( w_2 \neq w_N \). Since \( w_2 \leq w_N \), this yields \( w_2 < w_N \). Thus, we can apply Exercise 6.23 to \( N - 1 \) and \( (w_2, w_3, \ldots, w_N) \) instead of \( n \) and \( (w_1, w_2, \ldots, w_n) \) (because we have assumed that Exercise 6.23 has been solved for \( n = N - 1 \)). We thus obtain that \( w_2w_3 \cdots w_N \) is a Lyndon word in Case 1. We thus obtain that \( w_1w_2w_3w_4 \cdots w_N \) is a Lyndon word. Therefore, Proposition 6.16(a) (applied to \( u = w_1 \) and \( v = w_2w_3 \cdots w_N \)) yields that the word \( w_1w_2w_3 \cdots w_N \) is Lyndon. In other words, the word \( w_1w_2 \cdots w_N \) is Lyndon. So we have shown that \( w_1w_2 \cdots w_N \) is a Lyndon word in both possible Cases 1 and 2. Hence, \( w_1w_2 \cdots w_N \) always is a Lyndon word. In other words, Exercise 6.23 is solved for \( n = N \). This completes the induction, and therefore Exercise 6.23 is solved.

11.110. Solution to Exercise 6.24. Solution to Exercise 6.24. We have assumed that

\[
(11.922) \quad w_1w_{i+1} \cdots w_n \geq w_1w_2 \cdots w_n \quad \text{for every } i \in \{1, 2, \ldots, n\}.
\]

As a consequence,

\[
(11.923) \quad w_1w_{i+1} \cdots w_n > w_1w_2 \cdots w_n \quad \text{for every } i \in \{2, 3, \ldots, n\}.
\]

Now, we claim that if \( j \) is any element of \( \{0, 1, \ldots, n\} \), then

\[
(11.924) \quad \text{every nonempty proper suffix } v \text{ of } w_{n-j+1}w_{n-j+2} \cdots w_n \text{ satisfies } v > w_1w_2 \cdots w_n.
\]

**Proof of (11.924):** We will prove (11.924) by induction over \( j \):

- **Induction base:** For \( j = 0 \), we have \( w_{n-j+1}w_{n-j+2} \cdots w_n = w_{n-0+1}w_{n-0+2} \cdots w_n = \text{(empty product)} = \emptyset \). Hence, for \( j = 0 \), the word \( w_{n-j+1}w_{n-j+2} \cdots w_n \) has no nonempty proper suffix. Thus, (11.924) is vacuously true for \( j = 0 \). The induction base is thus complete.

- **Induction step:** Let \( J \in \{0, 1, \ldots, n-1\} \). We assume that (11.924) holds for \( j = J \). We now need to prove that (11.924) holds for \( j = J + 1 \).

From \( J \in \{0, 1, \ldots, n-1\} \), we obtain \( n - J \in \{1, 2, \ldots, n\} \).

\[687\] **Proof of (11.923):** Let \( i \in \{2, 3, \ldots, n\} \). Then, \( i - 1 \geq 1 \). The words \( w_1, w_2, \ldots, w_{i-1} \) are Lyndon (since the words \( w_1, w_2, \ldots, w_n \) are Lyndon), and thus nonempty. Now, \( i - 1 \geq 1 > 0 \). Hence, the product \( w_1w_2 \cdots w_{i-1} \) is a nonempty product of nonempty words (since \( w_1, w_2, \ldots, w_{i-1} \) are nonempty words), and thus a nonempty word itself. Since \( w_1w_2 \cdots w_n = (w_1w_2 \cdots w_{i-1}) (w_i \cdots w_n) \), this yields that \( w_1w_2 \cdots w_n \) is a proper suffix of the word \( w_1w_2 \cdots w_n \). As a consequence, \( w_1w_{i+1} \cdots w_n \neq w_1w_2 \cdots w_n \). Combining this with \( w_1w_{i+1} \cdots w_n \geq w_1w_2 \cdots w_n \) (by (11.922)), we obtain \( w_1w_{i+1} \cdots w_n > w_1w_2 \cdots w_n \). This proves (11.923).
Let $g = w_{n-j+1}w_{n-j+2} \cdots w_n$. We have assumed that (11.924) holds for $j = J$. In other words, every nonempty proper suffix $v$ of $w_{n-J+1}w_{n-J+2} \cdots w_n$ satisfies $v > w_1w_2 \cdots w_n$. Since $w_{n-J+1}w_{n-J+2} \cdots w_n = g$, this rewrites as follows:

(11.925) 
Every nonempty proper suffix $v$ of $g$ satisfies $v > w_1w_2 \cdots w_n$.

Now, let us notice that

(11.926) 
$w_{n-(J+1)+1}w_{n-(J+1)+2} \cdots w_n = w_{n-J}w_{n-J+1} \cdots w_n = w_{n-J}(w_{n-J+1}w_{n-J+2} \cdots w_n) = w_{n-J}g.$

Let now $v$ be a nonempty proper suffix of $w_{n-J}g$. We are going to prove that $v > w_1w_2 \cdots w_n$.

We have $v \neq \emptyset$ (since $v$ is nonempty). Since $v$ is a nonempty suffix of $w_{n-J}g$, we must be in one of the following two cases (depending on whether this suffix begins before the suffix $g$ of $w_{n-J}g$ begins or afterwards):

Case 1: The word $v$ is a nonempty suffix of $g$. (Note that $v$ is allowed to be $g$.)

Case 2: The word $v$ has the form $hg$ where $h$ is a nonempty suffix of $w_{n-J}$.

Let us first consider Case 1. In this case, the word $v$ is a nonempty suffix of $g$. If $v$ is a proper suffix of $g$, then we immediately obtain $v > w_1w_2 \cdots w_n$ (from (11.925)). Thus, for the rest of the proof of $v > w_1w_2 \cdots w_n$ in Case 1, we can WLOG assume that $v$ is not a proper suffix of $g$. Assume this.

So we know that $v$ is a suffix of $g$, but not a proper suffix of $g$. Hence, $v$ must be $g$ itself. That is, we have $v = g$. Hence, $v = g = w_{n-J+1}w_{n-J+2} \cdots w_n$. Consequently, $J \neq 0$.

Combined with $J \in \{0, 1, \ldots, n-1\}$, this yields $J \in \{0, 1, \ldots, n-1\}$, so that $n-J+1 \in \{2, 3, \ldots, n\}$. Now,

$\begin{align*}
v &= g = w_{n-J+1}w_{n-J+2} \cdots w_n = w_{n-J+1}w_{n-(J+1)+1} \cdots w_n > w_1w_2 \cdots w_n
\end{align*}$

(by (11.923), applied to $i = n - J + 1$). Thus, $v > w_1w_2 \cdots w_n$ is proven in Case 1.

Let us now consider Case 2. In this case, the word $v$ has the form $hg$ where $h$ is a nonempty suffix of $w_{n-J}$. Consider this $h$. Since $h$ is a suffix of $w_{n-J}$, we have $\ell(h) \leq \ell(w_{n-J})$, so that $\ell(w_{n-J}) \geq \ell(h)$. But $w_{n-J}$ is a Lyndon word (since $w_1, w_2, \ldots, w_n$ are Lyndon words), and thus $h \geq w_{n-J}$ (by Corollary 6.15, applied to $w_{n-J}$ and $h$ instead of $w$ and $v$). Thus, $w_{n-J} \leq h$. Thus, Proposition 6.2(b) (applied to $a = w_{n-J}, b = h$ and $c = g$) yields $w_{n-J}g \leq hg = v$. Thus, $v \geq w_{n-J}g$.

But we also have $v \neq w_{n-J}g$ (since $v$ is a proper suffix of $w_{n-J}g$). Combined with $v \geq w_{n-J}g$, this yields $v > w_{n-J}g$.

Thus, $v > w_1w_2 \cdots w_n$ is proven in Case 2.

Now, $v > w_1w_2 \cdots w_n$ is proven in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that $v > w_1w_2 \cdots w_n$ always holds.

Now, let us forget that we fixed $v$. We thus have proven that every nonempty proper suffix $v$ of $w_{n-J}g$ satisfies $v > w_1w_2 \cdots w_n$. Since $w_{n-J}g = w_{n-(J+1)+1}w_{n-(J+1)+2} \cdots w_n$ (by (11.926)), this rewrites as follows: Every nonempty proper suffix $v$ of $w_{n-(J+1)+1}w_{n-(J+1)+2} \cdots w_n$ satisfies $v > w_1w_2 \cdots w_n$. In other words, (11.924) holds for $j = J+1$. This completes the induction step. Thus, (11.924) is proven by induction.

Now, we can apply (11.924) to $j = n$. As a result, we obtain that every nonempty proper suffix $v$ of $w_{n-n+1}w_{n-n+2} \cdots w_n$ satisfies $v > w_1w_2 \cdots w_n$. In other words, every nonempty proper suffix $v$ of $w_1w_2 \cdots w_n$ satisfies $v > w_1w_2 \cdots w_n$ (since $w_{n-n+1}w_{n-n+2} \cdots w_n = w_1w_2 \cdots w_n$). Since the word $w_1w_2 \cdots w_n$ is also nonempty, this shows that the word $w_1w_2 \cdots w_n$ is Lyndon (by the definition of a Lyndon word). This solves Exercise 6.24.

---

688Proof. Assume the contrary. Then, $J = 0$, so that

$v = w_{n-J+1}w_{n-J+2} \cdots w_n = w_{n-0+1}w_{n-0+2} \cdots w_n$ (since $J = 0$)

$= (\text{empty product}) = \emptyset,$

contradicting the fact that $v \neq \emptyset$. This contradiction shows that our assumption was wrong, qed.

689Proof. The product $w_1w_2 \cdots w_n$ is nonempty (since $n$ is a positive integer), and the words $w_1, w_2, \ldots, w_n$ are nonempty (since they are Lyndon words). Hence, $w_1w_2 \cdots w_n$ is a nonempty product of nonempty words. Thus, $w_1w_2 \cdots w_n$ is a nonempty word, qed.
11.11. Solution to Exercise 6.29. Solution to Exercise 6.29. For every positive integer \( n \), let \( \text{lyn} \) denote the number of Lyndon words of length \( n \). We need to prove that

\[
\text{lyn} \, n = \frac{1}{n} \sum_{d|n} \mu(d) q^{n/d}
\]

for every positive integer \( n \).

Let \( \mathfrak{M} \) denote the set of all finite multisets of Lyndon words. We can then define a map \( \mathbf{m} : \mathfrak{M} \rightarrow \mathfrak{A}^* \) as follows: Given an \( M \in \mathfrak{M} \), we set \( \mathbf{m}(M) = a_1 a_2 \cdots a_k \), where \( a_1, a_2, \ldots, a_k \) denote the elements of \( M \) listed in decreasing order and each as often as it appears in \( M \). This map \( \mathbf{m} \) is bijective (its inverse is given by sending every word \( w \) to the multiset \( \{a_1, a_2, \ldots, a_k\}_{\text{multiset}} \), where \( (a_1, a_2, \ldots, a_k) \) is the CFL factorization of \( w \).

On the other hand, let \( \mathcal{L} \) be the set of all Lyndon words. Thus, \( \mathfrak{M} \) is the set of all finite multisets of elements of \( \mathcal{L} \). Also, the definition of \( \text{lyn} \, n \) now rewrites as

\[
\text{lyn} \, n = |\{ w \in \mathcal{L} \mid \ell(w) = n \}| \quad \text{for every positive integer } n.
\]

Let \( \mathfrak{A} \) be the set of all families \( (k_w)_{w \in \mathcal{L}} \in \mathbb{N}^\mathcal{L} \) of nonnegative integers (indexed by the Lyndon words) such that all but finitely many \( w \in \mathcal{L} \) satisfy \( k_w = 0 \). We can now define a bijection \( \text{mult} : \mathfrak{M} \rightarrow \mathfrak{A} \) by sending every \( M \in \mathfrak{M} \) to the family \((\text{multiplicity of } w \text{ in the multiset } M)_{w \in \mathcal{L}} \in \mathfrak{A} \).

The composition \( \mathbf{m} \circ \text{mult}^{-1} : \mathfrak{M} \rightarrow \mathfrak{A}^* \) of the bijections \( \mathbf{m} \) and \( \text{mult}^{-1} \) is clearly a bijection. It can easily be seen to satisfy

\[
\ell \left( (\mathbf{m} \circ \text{mult}^{-1}) ((k_w)_{w \in \mathcal{L}}) \right) = \sum_{w \in \mathcal{L}} k_w \cdot \ell(w) \quad \text{for every } (k_w)_{w \in \mathcal{L}} \in \mathfrak{A}.
\]

Now, in the ring \( \mathbb{Q}[\llbracket t \rrbracket] \) of formal power series, we have

\[
\sum_{w \in \mathfrak{A}^*} t^{\ell(w)} = \sum_{n \in \mathbb{N}} \frac{|\{ w \in \mathcal{L} \mid \ell(w) = n \}| t^n}{(\text{since } |\mathfrak{A}| = q)} = \sum_{n \in \mathbb{N}} q^n t^n = \frac{1}{1 - qt}.
\]

Proof of (11.929): Let \( (k_w)_{w \in \mathcal{L}} \in \mathfrak{A} \). Let \( M = \text{mult}^{-1} ( (k_w)_{w \in \mathcal{L}} ) \). Then, \( M \) is a multiset of elements of \( \mathcal{L} \) and satisfies \( (k_w)_{w \in \mathcal{L}} = \text{mult} M = (\text{multiplicity of } w \text{ in the multiset } M)_{w \in \mathcal{L}} \). In other words, every \( w \in \mathcal{L} \) satisfies

\[
k_w = \text{(multiplicity of } w \text{ in the multiset } M) .
\]

Let \( a_1, a_2, \ldots, a_k \) denote the elements of this multiset \( M \) listed in decreasing order. Then, the definition of \( \mathbf{m} \) yields \( \mathbf{m}(M) = a_1 a_2 \cdots a_k \), so that

\[
\ell(\mathbf{m}(M)) = \ell(a_1 a_2 \cdots a_k) = \sum_{i \in \{1, 2, \ldots, k\}} \ell(a_i) = \sum_{w \in \mathcal{L} \mid \ell(w) = n} \ell \left( \sum_{a_i = w} \left( a_i \right) \right) = \sum_{w \in \mathcal{L}} \ell(w) \quad \text{(since every } a_i \text{ belongs to } \mathcal{L})
\]

\[
= \sum_{w \in \mathcal{L}} \ell(w) \quad \text{(number of } i \in \{1, 2, \ldots, k\} \text{ satisfying } a_i = w) \cdot \ell(w)
\]

Now,

\[
\ell \left( (\mathbf{m} \circ \text{mult}^{-1}) ((k_w)_{w \in \mathcal{L}}) \right) = \ell \left( \mathbf{m} \left( \text{mult}^{-1} ( (k_w)_{w \in \mathcal{L}} ) \right) \right) = \ell(\mathbf{m}(M)) = \sum_{w \in \mathcal{L}} k_w \cdot \ell(w) ,
\]

which proves (11.929).
Hence,
\[ \frac{1}{1 - qt} = \sum_{w \in \mathcal{A}^*} q^\ell(w) = \sum_{(k_w)_{w \in \mathcal{L}} \in \mathcal{A}} q^\ell((m \circ \text{mult}^{-1})(k_w)_{w \in \mathcal{L}}) \]
(here, we substituted \((m \circ \text{mult}^{-1})(k_w)_{w \in \mathcal{L}}\) for \(w\) in the sum, since the map \(m \circ \text{mult}^{-1} : \mathcal{L} \to \mathcal{A}^*\) is a bijection)
\[ = \sum_{(k_w)_{w \in \mathcal{L}} \in \mathcal{A}} \prod_{w \in \mathcal{L}} k_w^{-\ell(w)} = \prod_{n \geq 1} \sum_{w \in \mathcal{L}_n} \frac{1}{1 - t^{\ell(w)}} \quad \text{(by (11.929))} \]
\[ = \prod_{n \geq 1} \frac{1}{1 - t^n} = \prod_{n \geq 1} \prod_{w \in \mathcal{L}_n} \frac{1}{1 - t^{\ell(w)}} = \prod_{n \geq 1} \left( \frac{1}{1 - t^n} \right)^{\text{lyn}_n} \quad \text{(since } \ell(w) \geq 1 \text{ for every } w \in \mathcal{L}) \]

Taking the logarithm of both sides of this identity, we obtain
\[ \log \frac{1}{1 - qt} = \log \left( \prod_{n \geq 1} \left( \frac{1}{1 - t^n} \right)^{\text{lyn}_n} \right) = \sum_{n \geq 1} (\text{lyn}_n) \cdot \log \left( \frac{1}{1 - t^n} \right) = \sum_{n \geq 1} (\text{lyn}_n) \cdot \log \left( \frac{1}{1 - t^n} \right) \]
\[ = - \log(1 - t^n) = \sum_{n \geq 1} \frac{1}{u} (t^n)^u \quad \text{(by the Mercator series for the logarithm)} \]
\[ = \sum_{n \geq 1} (\text{lyn}_n) \cdot \sum_{u \geq 1} \frac{1}{u} (t^n)^u = \sum_{n \geq 1} (\text{lyn}_n) \frac{1}{u} (t^n)^u = \sum_{n \geq 1} \sum_{u \geq 1} (\text{lyn}_n) \frac{1}{u} t^{nu} \]
\[ = \sum_{n \geq 1} \sum_{v \geq 1} \frac{\text{lyn}_n}{v} \frac{1}{v} t^v = \sum_{n \geq 1} \sum_{v \geq 1} \frac{\text{lyn}_n}{v} \frac{n}{v} \frac{1}{v} t^v = \sum_{n \geq 1} \sum_{d | n} (\text{lyn}_{d}) \frac{d}{n} t^n \quad \text{(here, we substituted } v/n \text{ for } u \text{ in the second sum)} \]
\[ = \sum_{n \geq 1} \sum_{v \geq 1} \frac{\text{lyn}_n}{v} \frac{n}{v} \frac{t^v}{v} = \sum_{n \geq 1} \sum_{d | n} (\text{lyn}_{d}) \frac{d}{n} t^n = \sum_{n \geq 1} \sum_{d | n} (\text{lyn}_{d}) \frac{d}{n} t^n \]

(here, we renamed the summation indices \(v\) and \(n\) as \(n\) and \(d\)). Since
\[ \log \frac{1}{1 - qt} = - \log (1 - qt) = \sum_{n \geq 1} \frac{1}{n} (qt)^n \quad \text{(by the Mercator series for the logarithm)} \]
\[ = \sum_{n \geq 1} \frac{1}{n} q^n t^n, \]
this rewrites as
\[ \sum_{n \geq 1} \frac{1}{n} q^n t^n = \sum_{n \geq 1} \sum_{d | n} (\text{lyn}_{d}) \frac{d}{n} t^n. \]
Comparing coefficients, we conclude that every positive integer \( n \) satisfies
\[
\frac{1}{n} q^n = \sum_{d|n} (\text{lyn } d) \frac{d}{n}.
\]

Multiplying this with \( n \), we obtain
\[
q^n = \sum_{d|n} (\text{lyn } d) d.
\]

Now, recall that every positive integer \( N \) satisfies
\[
\sum_{d|N} \mu(d) = \delta_{N,1}.
\]

Now, every positive integer \( n \) satisfies
\[
\sum_{d|n} \mu(d) q^{n/d} = \sum_{e|n} \mu(e) \sum_{d|n/e} (\text{lyn } d) = \sum_{e|n} \mu(e) \left( \sum_{d|n} (\text{lyn } d) \right)
\]
\[
= \sum_{e|n} \mu(e) (\text{lyn } d) = \sum_{d|n} \mu(e) (\text{lyn } d)
\]
\[
= \sum_{d|n} \delta_{n/d,1} (\text{lyn } d) = \sum_{d|n} \delta_{n,d} (\text{lyn } d) d = (\text{lyn } n) n.
\]

Dividing this by \( n \), we obtain
\[
\frac{1}{n} \sum_{d|n} \mu(d) q^{n/d} = \text{lyn } n.
\]
This proves (11.927). Thus, Exercise 6.29 is solved.

\[\text{HOPF ALGEBRAS IN COMBINATORICS (VERSION CONTAINING SOLUTIONS)} \] 647

11.112. Solution to Exercise 6.31. Solution to Exercise 6.31. The word \( v \) is nonempty (since it is Lyndon). Thus, \( u \neq w \). Combined with \( u \leq w = w \), this yields \( u < w \).

(b) We have \( w = vw \); thus, \( v \) is a proper suffix of \( w \) (since \( u \) is nonempty). Also, the word \( w \) is Lyndon, and therefore
\[
(\text{11.933}) \quad \text{every nonempty proper suffix of } w \text{ is } > w
\]
(by the definition of a Lyndon word). Applying (11.933) to the nonempty proper suffix \( v \) of \( w \), we obtain \( v > w \). Thus, \( w < v \). Hence, \( u < w < v \), and this solves Exercise 6.31(b).

(a) The word \( v \) is Lyndon (by its definition). It remains to prove that the word \( u \) is Lyndon.

Recall that a word \( x \) is Lyndon if and only if it is nonempty and satisfies the property that every nonempty proper suffix \( y \) of \( x \) satisfies \( y > x \). \( \text{603} \) Applied to \( x = u \), this shows that the word \( u \) is Lyndon if and only if it is nonempty and satisfies the property that every nonempty proper suffix \( y \) of \( u \) satisfies \( y > u \). Let us now prove that
\[
(\text{11.934}) \quad \text{every nonempty proper suffix } y \text{ of } u \text{ satisfies } y > u.
\]

**Proof of (11.934):** Assume the contrary. Then, there exists a nonempty proper suffix \( y \) of \( u \) which does not satisfy \( y > u \). Let \( p \) be the shortest such suffix. Thus, \( p \) is a nonempty proper suffix of \( u \) which does not satisfy \( p > u \).

Notice that \( p \neq u \) (since \( p \) is a proper suffix of \( u \)).

\( \text{601} \) This is one of the most fundamental properties of the number-theoretic Möbius function. For a proof of (11.932), see the solution of Exercise 2.82. (More precisely, the equality (11.932) is obtained from (11.337) by renaming \( n \) as \( N \).)

\( \text{602} \) Proof. Assume the contrary. Then, \( u = w \). Hence, \( w \varnothing = w = uv \). Cancelling \( u \) from this equality, we obtain \( \varnothing = v \). Thus, the word \( v \) is empty; this contradicts the fact that \( v \) is nonempty. This contradiction proves that our assumption was wrong. QED.

\( \text{603} \) This is just a restatement of the definition of a Lyndon word.
There exists a nonempty \( t \in \mathbb{A}^* \) such that \( u = tp \) (since \( p \) is a proper suffix of \( u \)). Consider this \( t \). We have \( w = u \quad v = tpu = t(pv) \), and thus the word \( pv \) is a proper suffix of \( w \) (since \( t \) is nonempty). Also, the word \( pv \) is nonempty (since \( v \) is nonempty). Hence, \( pv \) is a nonempty proper suffix of \( w \). But applying (11.933) to the nonempty proper suffix \( pv \) of \( w \), we obtain \( pv > w = wv \). Thus, \( wv < pv \). Hence, Proposition 6.2(e) (applied to \( a = u, b = v, c = p \) and \( d = v \)) yields that either we have \( u \leq p \) or the word \( p \) is a prefix of \( u \). Since we don’t have \( u \leq p \) \(^{694}\), we therefore conclude that the word \( p \) is a prefix of \( u \). Hence, \( p \leq u \). Combined with \( p \neq u \), this yields \( p < u \). Combined with \( u < v \) (by Exercise 6.31(b)), this yields \( p < u < v \).

Let us now show that the word \( p \) is Lyndon. Indeed, let \( r \) be a nonempty proper suffix of \( p \). Then, \( r \) is shorter than \( p \) (being a proper suffix of \( p \)); in other words, \( \ell(r) < \ell(p) \). On the other hand, there exists some nonempty \( q \in \mathbb{A}^* \) satisfying \( p = qr \) (since \( r \) is a proper suffix of \( p \)). Consider this \( q \). We have \( u = t \quad p = tqr = (tq) \quad r \). Thus, \( r \) is a proper suffix of \( u \) (since \( tq \) is nonempty (because \( q \) is nonempty)). Hence, \( r > u \) \(^{695}\). Hence, \( r > u > p \) (since \( p < u \)). Now, let us forget that we fixed \( r \). We thus have shown that every nonempty proper suffix \( r \) of \( p \) satisfies \( r \geq p \).

But recall that the word \( p \) is Lyndon if and only if it is nonempty and satisfies the property that every nonempty proper suffix \( r \) of \( p \) satisfies \( r \geq p \). \(^{696}\) Hence, we conclude that the word \( p \) is Lyndon (since we have shown that \( p \) is nonempty and satisfies the property that every nonempty proper suffix \( r \) of \( p \) satisfies \( r \geq p \)).

Now, the words \( p \) and \( v \) are Lyndon and satisfy \( p < v \). Thus, Proposition 6.16(a) (applied to \( p \) instead of \( u \)) yields that the word \( pv \) is Lyndon. Notice also that \( w = tpu = t(pv) \), and thus \( pv \) is a proper suffix of \( w \) (since \( t \) is nonempty). Hence, \( pv \) is a proper suffix of \( w \) such that \( pv \) is Lyndon. In other words, \( pv \) is a proper suffix \( z \) of \( w \) such that \( z \) is Lyndon. Since \( v \) is the longest such suffix \(^{697}\), this yields that \( pv \) is not longer than \( v \). In other words, \( \ell(pv) \leq \ell(v) \). This contradicts \( \ell(pv) = \ell(p) + \ell(v) > \ell(v) \) (since \( p \) is nonempty).

This contradiction proves that our assumption was wrong. Hence, (11.934) is proven.

Now, recall that the word \( u \) is Lyndon if and only if it is nonempty and satisfies the property that every nonempty proper suffix \( y \) of \( u \) satisfies \( y > u \). Hence, the word \( u \) is Lyndon (because we have shown that \( u \) is nonempty and satisfies the property that every nonempty proper suffix \( y \) of \( u \) satisfies \( y > u \)). The solution of Exercise 6.31(a) is thus complete.

(c) Let us denote by \( u' \) and \( v' \) the words \( u \) and \( v \) constructed in Theorem 6.30 (to avoid confusing them with the words \( u \) and \( v \) defined in Exercise 6.31). We must then show that \( u = u' \) and \( v = v' \).

From Theorem 6.30, we see that \( v' \) is the (lexicographically) smallest nonempty proper suffix of \( w \), and that \( u' \) is a nonempty word such that \( w = u'v' \).

The word \( v \) is a nonempty proper suffix of \( w \). Since \( v' \) is the smallest such suffix, we thus conclude that \( v' \leq v \). We shall now prove that \( v' = v \).

Indeed, \( v \) is a proper suffix of \( w \). In other words, \( v \) is a proper suffix of \( u'v' \) (since \( w = u'v' \)). Hence, we must be in one of the following two cases (depending on whether this suffix begins before the suffix \( v' \) of \( u'v' \) begins or afterwards):

- Case 1: The word \( v \) is a proper suffix of \( v' \).
- Case 2: The word \( v \) has the form \( qv' \) where \( q \) is a proper suffix of \( u' \). (This suffix \( q \) may be empty.)

---

\(^{694}\)Proof. Assume the contrary. Then, \( u \leq p \). Thus, \( p \geq u \). Combined with \( p \neq u \), this yields \( p > u \). This contradicts the fact that \( p \) does not satisfy \( p > u \). This contradiction proves that our assumption was wrong, qed.

\(^{695}\)Proof. Assume the contrary. Then, we don’t have \( r > u \). Thus, \( r \) is a nonempty proper suffix of \( u \) which does not satisfy \( r > u \). In other words, \( r \) is a nonempty proper suffix \( y \) of \( u \) which does not satisfy \( y > u \). Since the shortest such suffix is \( p \) (by the definition of \( p \)), this yields that \( r \) is not shorter than \( p \). That is, \( \ell(r) \geq \ell(p) \). This contradicts \( \ell(r) < \ell(p) \). This contradiction proves that our assumption was wrong, qed.

\(^{697}\)This is just a restatement of the definition of a Lyndon word (applied to \( p \)).

\(^{696}\)Proof. We defined \( v \) as the longest proper suffix of \( w \) such that \( v \) is Lyndon. In other words, \( v \) is the longest proper suffix \( z \) of \( w \) such that \( z \) is Lyndon, qed.
Let us first consider Case 1. In this case, the word \( v \) is a proper suffix of \( v' \). Now, every nonempty proper suffix \( s \) of \( v' \) satisfies \( s > v' \) \( ^{698} \). Thus, the word \( v' \) is nonempty and satisfies the property that every nonempty proper suffix \( s \) of \( v' \) satisfies \( s > v' \). In other words, the word \( v' \) is Lyndon (according to the definition of a Lyndon word).

Now, \( v \) is the longest proper suffix of \( w \) such that \( v \) is Lyndon. In other words, \( v \) is the longest proper suffix \( z \) of \( w \) such that \( z \) is Lyndon. But \( v' \) also is a proper suffix \( z \) of \( w \) such that \( z \) is Lyndon (since \( v' \) is Lyndon). Since \( v \) is the longest such suffix, this yields that \( v' \) is not longer than \( v \). In other words, \( \ell(v') \leq \ell(v) \). But since \( v \) is a proper suffix of \( v' \), we have \( \ell(v) < \ell(v') \leq \ell(v) \). This is absurd. Hence, \( v' = v \) (since ex falso quodlibet). Thus, \( v' = v \) is proven in Case 1.

Let us now consider Case 2. In this case, the word \( v \) has the form \( qv' \) where \( q \) is a proper suffix of \( u' \). Consider this \( q \).

We need to prove that \( v' = v \). If \( q = \emptyset \), then this is obvious (because if \( q = \emptyset \), then \( v = \overbrace{v'}^{=v} \) and thus \( v' = v \)). Hence, for the rest of this proof, we can WLOG assume that we don’t have \( q = \emptyset \). Assume this. The word \( q \) is nonempty (since we don’t have \( q = \emptyset \)), and thus \( v' \) is a proper suffix of \( v \) (since \( v = qv' \)).

The word \( v \) is Lyndon. Hence, every nonempty proper suffix of \( v \) is \( > v \) (by the definition of a Lyndon word). Applying this to the nonempty proper suffix \( v' \) of \( v \), we conclude that \( v' > v \). This contradicts \( v' \leq v \). This contradiction shows that we have \( v' = v \) (since ex falso quodlibet). Thus, \( v' = v \) is proven in Case 2.

Now, \( v' = v \) is proven in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that \( v' = v \) always holds.

Now we have proven that \( v' = v \). Cancelling \( v \) from the equality \( uv = w = u' \overbrace{v'}^{=v} = u'v \), we obtain \( u = u' \). Thus, we have \( u = u' \) and \( v = v' \). This solves Exercise 6.31(c).

[Remark: We have not used any statement of Theorem 6.30 in our above solution. Thus, parts (a) and (b) of Exercise 6.31 provide an alternative proof of Theorem 6.30 (because Exercise 6.31(c) shows that the words \( u \) and \( v \) defined in Exercise 6.31 are precisely the words \( u \) and \( v \) constructed in Theorem 6.30).]

11.13. Solution to Exercise 6.32. Solution to Exercise 6.32. (a) Let us prove the implications \( \mathcal{A}' \implies \mathcal{D}' \) and \( \mathcal{D}' \implies \mathcal{A}' \).

Proof of the implication \( \mathcal{A}' \implies \mathcal{D}' \): Assume that Assertion \( \mathcal{A}' \) holds. Thus, \( w \) is a power of a Lyndon word. Let said Lyndon word be \( t \). Thus, \( w = t^n \) for some \( n \in \mathbb{N} \). Consider this \( n \). Since \( w \) is nonempty, we have \( n \geq 1 \). Hence, \( t^{n-1} \) is well-defined.

We will first show the following simple lemma:

Lemma A: Let \( u' \), \( v' \) and \( p \) be three words, and \( N \in \mathbb{N} \) be such that \( u'v' = p^N \). Assume that \( p \) is not a prefix of \( u' \), and that \( p \) is not a suffix of \( v' \). Then, \( N \leq 1 \).

Proof of Lemma A: Assume the contrary. Thus, \( N > 1 \), so that \( N \geq 2 \). The word \( v' \) is a suffix of \( p^{N-1}p \) (since \( u'v' = p^N = p^{N-1}p \)). Thus, we must be in one of the following two cases (depending on whether this suffix begins before the suffix \( p \) of \( p^{N-1}p \) begins, or afterwards):

Case 1: The word \( v' \) has the form \( rp \) where \( r \) is a suffix of \( p^{N-1} \).

Case 2: The word \( v' \) is a suffix of \( p \).

Let us first consider Case 1. In this case, the word \( v' \) has the form \( rp \) where \( r \) is a suffix of \( p^{N-1} \). Hence, \( p \) is a suffix of \( v' \), contradicting the fact that \( p \) is not a suffix of \( v' \). Hence, Case 1 leads to a contradiction.

Let us now consider Case 2. In this case, the word \( v' \) is a suffix of \( p \). In other words, there exists a \( q \in \mathbb{N}^* \) such that \( p = qv' \). Consider this \( q \). We have \( u'v' = p^N = p^{N-1}p = p^{N-1}qv' \). Cancelling \( v' \) from this equality, we obtain \( u' = p^{N-1}q \). Since \( p^{N-1} = pp^{N-2} \) (this is well-defined because \( N \geq 2 \)), this further becomes \( u' = p^{N-1} = q = pp^{N-2}q = p (p^{N-2}q) \). Hence, \( p \) is a prefix of \( u' \), contradicting the fact that \( p \) is not a prefix of \( u' \). We have thus obtained a contradiction in Case 2.

\(^{698} \text{Proof. Let } s \text{ be a nonempty proper suffix of } v'. \text{ The word } s \text{ is a proper suffix of } v', \text{ which (in turn) is a suffix of } w. \text{ Hence, } s \text{ is a proper suffix of } w. \text{ Thus, } s \text{ is a nonempty proper suffix of } w. \text{ Since the smallest such suffix is } v' \text{ (by the definition of } v' \text{), this yields } s \geq v'. \text{ Since } s \neq v' \text{ (because } s \text{ is a proper suffix of } v') \text{, this yields } s > v', \text{ qed.} \)
We have thus obtained a contradiction in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this shows that we always get a contradiction. This completes the proof of Lemma A.

Now, let us return to the proof of the implication $A' \Rightarrow D'$. Let $u$ and $v$ be nonempty words satisfying $w = uv$. We want to show that $uv 
subseteq w$.

The word $t$ is nonempty (since it is Lyndon); thus, $\ell (t) \geq 1$. Hence, for every sufficiently large $a \in \mathbb{N}$, we have that the word $t^a$ is not a prefix of $u$ \footnote{Proof. Let $a \in \mathbb{N}$ be such that $a > \ell (u)$. Then, $\ell (t^a) = a \ell (t) \geq a \ell (u) \geq 1$, so that the word $t^a$ is longer than $u$. Hence, the word $t^a$ is not a prefix of $u$.} \footnote{This is well-defined because of the following two facts:
  \begin{itemize}
    \item There exists an $a \in \mathbb{N}$ such that the word $t^a$ is a prefix of $u$ (namely, $a = 0$).
    \item For every sufficiently large $a \in \mathbb{N}$, we have that the word $t^a$ is not a prefix of $u$.}

Similarly, for every sufficiently large $b \in \mathbb{N}$, we have that the word $t^b$ is not a suffix of $v$.

Let $A$ be the largest nonnegative integer $a$ such that $t^a$ is a prefix of $u$ \footnote{This is well-defined because of the following two facts:
  \begin{itemize}
    \item There exists a $b \in \mathbb{N}$ such that the word $t^b$ is a suffix of $v$ (namely, $b = 0$).
    \item For every sufficiently large $b \in \mathbb{N}$, we have that the word $t^b$ is not a suffix of $v$.} \footnote{We have thus shown that for every $a \in \mathbb{N}$ satisfying $a > \ell (u)$, we have that the word $t^a$ is not a prefix of $u$. Consequently, for every sufficiently large $a \in \mathbb{N}$, we have that the word $t^a$ is not a prefix of $u$, qed.}
Thus, $t^A$ is a prefix of $u$, but $t^{A+1}$ is not a prefix of $u$. Let $B$ be the largest nonnegative integer $b$ such that $t^b$ is a suffix of $v$ \footnote{We have thus obtained a contradiction in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this shows that we always get a contradiction. This completes the proof of Lemma A.}

Thus, $t^B$ is a suffix of $v$, but $t^{B+1}$ is not a suffix of $v$.

There exists a $u' \in \mathcal{A}^*$ such that $u = t^A u'$ (since $t^A$ is a prefix of $u$). Consider this $u'$. Recall that $t^{A+1}$ is not a prefix of $u$. In other words, $t^A t$ is not a prefix of $t^A u'$ (since $t^{A+1} = t^A t$ and $u = t^A u'$).

There exists a $v' \in \mathcal{A}^*$ such that $v = v't^B$ (since $t^B$ is a suffix of $v$). Consider this $v'$. Recall that $t^{B+1}$ is not a suffix of $v$. In other words, $t^B t$ is not a prefix of $v't^B$ (since $t^{B+1} = t^B t$ and $v = v't^B$).

If the word $t$ was a prefix of $u'$, then $tA t$ would be a prefix of $t^A u'$, which would contradict the fact that $t^{A+1}$ is not a prefix of $t^A u'$. Thus, $t$ is not a prefix of $u'$. In particular, $t \neq u'$.

If the word $t$ was a suffix of $v'$, then $tB t$ would be a suffix of $v't^B$, which would contradict the fact that $t^{B+1}$ is not a suffix of $v't^B$. Thus, $t$ is not a suffix of $v'$. In particular, $t \neq v'$.

But we have $t^n = w = \underbrace{u \cdots u}_{=tA u'} v^t = t^A (u' v^t B) = t^A (u' v^t B)$. Hence, $t^A$ is a prefix of $t^n$; therefore, $A \leq n$ and thus $n - A \geq 0$. Hence, $t^n = t^A = t^A u' v^t B = t^A (u' v^t B)$, which shows that $t^B = t^n - A$. Thus, $B \leq n - A$ and therefore $n - A - B \geq 0$. Hence, $t^n = t^A t^B = t^A (u' v^t B)$. Cancelling $t^A$ in this yields $t^n = t^A (u' v^t B)$. Now, Lemma A (applied to $p = t$ and $N = n - A - B$) yields $n - A - B \leq 1$. Since $n - A - B \geq 0$, this shows that we have either $n - A - B = 0$ or $n - A - B = 1$. We thus must be in one of the following two cases:

Case 1: We have $n - A - B = 0$.

Case 2: We have $n - A - B = 1$.

Let us first consider Case 1. In this case, we have $n - A - B = 0$. Thus, $t^n = t^A u' v^t B$ rewrites as $t^n = t^A u' v^t B$, so that $u' v^t B = 0$. Consequently, $u' = \emptyset$ and $v^t B = \emptyset$. Now, $u = t^A$, $u' = t^A$ and $v = v^t B = t^B$, so that $t^n = t^A u' v^t B = t^A u' v^t B = t^A u' v^t B = \emptyset$. Hence, $n = n - A - B$ is proven in Case 1.

Let us now consider Case 2. In this case, we have $n - A - B = 1$. Thus, $t^n = t^A u' v^t B$ rewrites as $t^n = t^A u' v^t B$, so that $u' v^t B = t^A = t^B$. As a consequence, $u'$ is nonempty (because if $u'$ was empty, then we would have $u' v^t B = \emptyset$ and thus $t$ is not a prefix of $u'$). Also, $v'$ is nonempty (because if $v'$ was empty, then we would have $u' v^t B = \emptyset$ and thus $u' v^t B = t^A = t^B$, contradicting $t \neq u'$). But $t$ is Lyndon, and therefore Proposition 6.14(a) (applied to $t$, $u'$ and $v'$ instead of $u$, $w$ and $v$) yields $v' \geq t$. That is, $t \leq v'$. Thus, Proposition 6.2(d) (applied to $a = t$, $b = t^{n-1}$, $c = v'$ and $d = t^B u$) yields that either we have $tt^n \leq v t^B t u$ or the word $t$ is a prefix of $v'$. Since
that a prefix of \(v'\), this yields \(tt^{n-1} \leq v'tBu\). Since \(tt^{n-1} = t^n = uv\) and \(v'tBu = vu\), this rewrites as \(uv \leq vu\). Hence, \(vu \geq uv\) is proven in Case 2.

We thus have shown that \(vu \geq uv\) in each of the two cases 1 and 2. Since these two cases cover all possibilities, this yields that we always have \(vu \geq uv\). Thus, Assertion \(D'\) is satisfied, so we have proven the implication \(A' \implies D'\).

**Proof of the implication \(D' \implies A'\):** Assume that Assertion \(D'\) holds. Thus, if \(u\) and \(v\) are nonempty words satisfying \(w = uv\), then we have \(vu \geq uv\).

We need to prove that Assertion \(A'\) holds, i.e., that \(w\) is a power of a Lyndon word. Assume the contrary. Thus, \(w\) is not a power of a Lyndon word; hence, \(w\) is not a Lyndon word itself. Consequently, there exist nonempty words \(u\) and \(v\) such that \(w = uv\) and \(uv \leq w\) (702). Consider such a pair of nonempty words \(u\) and \(v\) with minimum \(\ell(u)\). The minimality of \(\ell(u)\) shows that

\[
\ell(u') \geq \ell(u).
\]

We have \(vu \leq w\). In combination with \(vu \geq w\) (which follows from Assertion \(D'\)), this yields \(vu = w\). Therefore, Proposition 6.4 yields that there exist a \(t \in \mathbb{N}^*\) and two nonnegative integers \(n\) and \(m\) such that \(u = t^n\) and \(v = t^m\). Consider this \(t\) and these \(n\) and \(m\). We have \(n \neq 0\) (since \(t^n = u\) is nonempty) and \(m \neq 0\) (since \(t^m = v\) is nonempty), and the word \(t\) is nonempty (since \(t^n = u\) is nonempty). Moreover, we have \(n = 1\) (703). Hence, \(t^n = t^1 = t\), so that \(u = t^n = t\) and \(w = u = t^m = t^{m+1}\). We shall now prove that the word \(t\) is Lyndon.

Assume the contrary. Then, \(t\) is not Lyndon. Let \(q\) be the (lexicographically) smallest nonempty suffix of \(t\). Then, Proposition 6.19(b) (applied to \(t\) and \(q\) instead of \(w\) and \(v\)) yields that there exists a nonempty \(p \in \mathbb{N}^*\) such that \(t = pq\), \(p \geq q\) and \(pq \geq qp\) (704). Consider this \(p\). Since \(q\) is nonempty, we have \(\ell(pq) > \ell(p)\), so that \(\ell\left(\frac{u}{t = pq}\right) = \ell(pq) > \ell(p)\). From \(pq \geq qp\), we obtain \(qp \leq pq\). Thus, it is easy to see that \(\left(\frac{qp}{p}\right)^m + 1 \leq \left(\frac{pq}{pq}\right)^m + 1\) (705).

We have \(v = t^m\). Since \(t = pq\), this rewrites as \(v = (pq)^m\). Now,

\[
\frac{q}{p} = (pq)^m p = (pq)^m q p = (qp)^m q = (pq)^m + 1 \leq (pq)^m + 1 \leq (pq)^m \quad \text{and} \quad \frac{p}{q} = (pq)^m (pq)^m = (pq)^m + 1 = (pq)^m \left(\frac{pq}{t}\right)^m = (pq)^m = (pq)^m = (pq)^m = (pq)^m = (pq)^m v = pqv.
\]

702**Proof.** Assume the contrary. Hence, any nonempty words \(u\) and \(v\) satisfying \(w = uv\) satisfy \(vu > uv\). In other words, \(w\) satisfies Assertion \(D\) of Theorem 6.20. Hence, \(w\) also satisfies Assertion \(A\) of Theorem 6.20 (since Theorem 6.20 yields the equivalence of these two assertions \(D\) and \(A\)). In other words, \(w\) is a Lyndon word. This contradicts the knowledge that \(w\) is not a Lyndon word, qed.

703**Proof.** Assume the contrary. Hence, \(n \neq 1\). Combined with \(n \neq 1\), this leads to \(n \geq 2\). As a consequence, \(u = t^n\) can be rewritten as \(u = tt^{n-1}\). The word \(tt^{n-1+m}\) is nonempty (since \(t\) is nonempty and since \(n-1 + m \geq 2 - 1 + 0 = 1\)). Now,

\[
\frac{w}{v} = \frac{u}{t = t^{n-1} = t^m} = \frac{v}{t = t^{n-1} = t^m} = \frac{tt^{n-1}}{tt^{n-1}} = tt^{n-1+m} = tt^{n-1} + 1.
\]

Also, \(t^{n-1+m} t = t^{n-1+m+1} = tt^{n-1} + 1\). Hence, (11.935) (applied to \(u^t = u\) and \(v^t = t^{n-1+m}\)) yields \(\ell(t) \geq \ell(u)\) (since \(t\) and \(t^{n-1+m}\) are nonempty). Since \(u = t^n\), this rewrites as \(\ell(t) \geq \ell(t^n) = n \ell(t) \geq 2 \ell(t)\), whence \(\ell(t) = 0\), which contradicts the fact that \(t\) is nonempty. This contradiction shows that our assumption was wrong, qed.

704**Notice that the variable \(p\) here is what has been called \(u\) in Proposition 6.19(b). (We had to rename it since the letter \(u\) is already in use.)**

705**Proof.** We have \(qp \leq pq\). Hence, Proposition 6.2(d) (applied to \(qp\), \((pq)^m\), \(pq\) and \((pq)^m\) instead of \(a, b, c\) and \(d\)) yields that either we have \((qp)(pq)^m \leq (pq)(pq)^m\) or the word \(qp\) is a prefix of \(pq\). In the first of these two cases, we are done (because in the first of these two cases, we have \((qp)(pq)^m \leq (pq)(pq)^m\) and thus \((pq)^{m+1} = (qp)(pq)^m \leq (pq)(pq)^m = (pq)^{m+1}\)). Hence, we can WLOG assume that we are in the second of these two cases. Assume this. Then, the word \(qp\) is a prefix of \(pq\). Since the word \(qp\) has the same length as \(pq\) (in fact, \(\ell(qp) = \ell(q) + \ell(p) = \ell(p) + \ell(q) = \ell(qp)\)), this yields that \(qp = pq\), so that \((pq)^{m+1} \leq (pq)^{m+1}\), qed.
Since \( w = u \cdot v = pqv \), we can thus apply (11.935) to \( u' = p \) and \( v' = qv \). As a result, we obtain 
\[
\ell(p) \geq \ell(u). 
\]
This contradiction shows that our assumption (that \( t \) is not Lyndon) was wrong. Hence, \( t \) is Lyndon. Thus, \( w \) is a power of a Lyndon word (since \( w = t^m+1 \) is a power of \( t \)). Thus, Assertion \( \mathcal{A}' \) is satisfied, so we have proven the implication \( \mathcal{D}' \Rightarrow \mathcal{A}' \).

Now we have proven both implications \( \mathcal{A}' \Rightarrow \mathcal{D}' \) and \( \mathcal{D}' \Rightarrow \mathcal{A}' \). Therefore, the equivalence \( \mathcal{A}' \iff \mathcal{D}' \) follows. Thus, Exercise 6.32(a) is solved.

(b) Consider the letter \( m \) and the alphabet \( \mathfrak{A} \cup \{m\} \) defined in Assertion \( \mathcal{F}' \). We notice that the lexicographic order on \( \mathfrak{A}^* \) is the restriction of the lexicographic order on \( (\mathfrak{A} \cup \{m\})^* \) to \( \mathfrak{A}^* \). Therefore, when we have two words \( p \) and \( q \) in \( \mathfrak{A}^* \), statements like “\( p < q \)” do not depend on whether we are regarding \( p \) and \( q \) as elements of \( \mathfrak{A}^* \) or as elements of \( (\mathfrak{A} \cup \{m\})^* \). It is easy to see that the one-letter word \( m \) satisfies

\[
(11.936) \quad m > p \quad \text{for every } p \in \mathfrak{A}^*. 
\]

We shall prove the implications \( \mathcal{B}' \Rightarrow \mathcal{E}' \), \( \mathcal{C}' \Rightarrow \mathcal{E}' \), \( \mathcal{G}' \Rightarrow \mathcal{H}' \), \( \mathcal{E}' \Rightarrow \mathcal{F}' \), \( \mathcal{F}' \Rightarrow \mathcal{C}' \), \( \mathcal{B}' \Rightarrow \mathcal{G}' \) and \( \mathcal{H}' \Rightarrow \mathcal{B}' \).

First of all, the implication \( \mathcal{B}' \Rightarrow \mathcal{E}' \) holds for obvious reasons (in fact, if two words \( u \) and \( v \) satisfying \( w = uv \) satisfy \( v \geq w \), then \( v \geq u \) (because \( v \geq w = uv \geq u \)). Also, the implication \( \mathcal{C}' \Rightarrow \mathcal{E}' \) holds for obvious reasons (in fact, if two words \( u \) and \( v \) satisfying \( w = uv \) satisfy \( v \) is a prefix of \( u \), then \( v \) is a prefix of \( w \) (because \( v \) is a prefix of \( u \), and \( u \) in turn is a prefix of \( uv = w \))). The implication \( \mathcal{G}' \Rightarrow \mathcal{H}' \) is also trivially true.

Proof of the implication \( \mathcal{E}' \Rightarrow \mathcal{F}' \): Assume that Assertion \( \mathcal{E}' \) holds.

Assume (for the sake of contradiction) that the word \( wm \in (\mathfrak{A} \cup \{m\})^* \) is not a Lyndon word. Clearly, this word \( wm \) is nonempty. Let \( v' \) be the the (lexicographically) smallest nonempty suffix of this word \( wm \). Since \( wm \) is not a Lyndon word, we can apply Proposition 6.19(b) to \( \mathfrak{A} \cup \{m\} \), \( wm \) and \( v' \) instead of \( \mathfrak{A} \), \( w \) and \( v \). As a result, we conclude that there exists a nonempty \( u \in (\mathfrak{A} \cup \{m\})^* \) such that \( wm = uv' \), \( u \geq v' \) and \( uv' \geq v'u \). Consider this \( u \).

We know that \( v' \) is a suffix of \( wm \). Thus, we must be in one of the following two cases (depending on whether this suffix begins before the suffix \( m \) of \( wm \) begins or afterwards):

Case 1: The word \( v' \) is a nonempty suffix of \( m \). (Note that \( v' = m \) is allowed.)

Case 2: The word \( v' \) has the form \( vm \) where \( v \) is a nonempty proper suffix of \( w \).

Let us consider Case 1 first. In this case, the word \( v' \) is a nonempty suffix of \( m \). Since the only nonempty suffix of \( m \) is \( m \) itself (because \( m \) is a one-letter word), this yields \( v' = m \). Now, \( wm = uv' = uvm \).

Cancelling \( m \) from this equality, we obtain \( w = u \), so that \( u = u \in \mathfrak{A}^* \). Hence, \( m > p \) (by (11.936), applied to \( p = u \)). This contradicts \( u \geq v' = m \). Thus, we have obtained a contradiction in Case 1.

Let us now consider Case 2. In this case, the word \( v' \) has the form \( vm \) where \( v \) is a nonempty proper suffix of \( w \). Consider this \( v \). We have \( wm = u \cdot v' = uvm \). By cancelling \( m \) from this equality, we obtain \( w = uv \). Thus, \( u \) and \( v \) are subwords of \( w \), and therefore belong to \( \mathfrak{A}^* \) (since \( w \in \mathfrak{A}^* \)). Moreover, \( u \) and \( v \) are nonempty. Hence, Assertion \( \mathcal{E}' \) yields that either we have \( v \geq u \) or the word \( v \) is a prefix of \( w \). Since we cannot have \( v \geq u \) (because if we had \( v \geq u \), then we would have \( v' = vm > v \geq u \geq v' \), which is absurd), we therefore must have that \( v \) is a prefix of \( w \). In other words, there exists a \( q \in \mathfrak{A}^* \) such that \( w = vq \). Consider this \( q \). We have \( m > q \) (by (11.936), applied to \( p = q \)). Thus, \( q \leq m \). Hence, Proposition 6.2(b) (applied to \( \mathfrak{A} \cup \{m\} \), \( v \), \( q \) and \( m \) instead of \( \mathfrak{A} \), \( a \), \( c \) and \( d \)) yields \( vq \leq vm \). Therefore, \( vq \geq w = wvq \), so that \( v' = vm \geq w = uv > u \) (since \( v \) is nonempty). This contradicts \( u \geq v' \). Thus, we have found a contradiction in Case 2.

\footnote{Proof of (11.936): Let \( p \in \mathfrak{A}^* \). If \( p \) is empty, then (11.936) is obviously satisfied. Hence, for the rest of the proof of (11.936), we WLOG assume that \( p \) is nonempty. Then, \( p \) has a first letter, and we have (the first letter of the word \( p \)) \in \mathfrak{A} \) (since \( p \in \mathfrak{A}^* \)). Thus,

\[
(\text{the first letter of the word } p) < m \quad \text{ (since } a < m \text{ for every } a \in \mathfrak{A})
\]

\[
= (\text{the first letter of the word } m). 
\]

Hence, \( p < m \) (by the definition of the lexicographic order), that is, \( m > p \). This proves (11.936).}
We have therefore obtained a contradiction in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this shows that we always get a contradiction. Hence, our assumption (that the word \(wm \in (\mathfrak{A} \cup \{m\})^*\) is not a Lyndon word) was false. Hence, the word \(wm \in (\mathfrak{A} \cup \{m\})^*\) is a Lyndon word.

**Proof of the implication \(\mathcal{F}'' \implies \mathcal{B}'\):** Assume that Assertion \(\mathcal{F}''\) holds. Thus, the word \(wm \in (\mathfrak{A} \cup \{m\})^*\) is a Lyndon word.

Let \(u\) and \(v\) be nonempty words satisfying \(w = uv\). We shall prove that either we have \(v \geq w\) or the word \(v\) is a prefix of \(w\).

Indeed, \(v\) is a suffix of \(w\) (since \(w = uv\)), so that \(vm\) is a suffix of \(wm\). Clearly, \(vm\) is nonempty. Thus, Corollary 6.15 (applied to \(\mathfrak{A} \cup \{m\}\), \(wm\) and \(vm\) instead of \(\mathfrak{A}, w\) and \(v\)\) yields \(vm \geq wm\). In other words, \(wm \leq vm\). Hence, Proposition 6.2(e) (applied to \(\mathfrak{A} \cup \{m\}\), \(w, m\) and \(v\) instead of \(\mathfrak{A}, a, b, c\) and \(d\)\) yields that either we have \(w \leq v\) or the word \(v\) is a prefix of \(w\). In other words, either we have \(v \geq w\) or the word \(v\) is a prefix of \(w\).

Now, let us forget that we fixed \(u\) and \(v\). We thus have proven that if \(u\) and \(v\) are nonempty words satisfying \(w = uv\), then either we have \(v \geq w\) or the word \(v\) is a prefix of \(w\). In other words, Assertion \(\mathcal{B}'\) holds. Hence, we have proven the implication \(\mathcal{F}'' \implies \mathcal{B}'\).

**Proof of the implication \(\mathcal{F}'' \implies \mathcal{C}'\):** Assume that Assertion \(\mathcal{F}''\) holds. Thus, the word \(wm \in (\mathfrak{A} \cup \{m\})^*\) is a Lyndon word.

Let \(u\) and \(v\) be nonempty words satisfying \(w = uv\). We shall prove that either we have \(v \geq u\) or the word \(v\) is a prefix of \(u\).

Indeed, Proposition 6.14(b) (applied to \(\mathfrak{A} \cup \{m\}\), \(wm\) and \(vm\) instead of \(\mathfrak{A}, w\) and \(v\)\) yields \(vm \geq u\) (since \(w = uvm\), and since \(vm\) is nonempty). Hence, \(vm \geq u = w\mathfrak{A}\), so that \(u\mathfrak{A} \leq w\). Thus, Proposition 6.2(c) (applied to \(\mathfrak{A} \cup \{m\}\), \(u, \mathfrak{A}, v\) and \(m\) instead of \(\mathfrak{A}, a, b, c\) and \(d\)\) yields that either we have \(u \leq v\) or the word \(v\) is a prefix of \(u\). In other words, either we have \(v \geq u\) or the word \(v\) is a prefix of \(u\).

Now, let us forget that we fixed \(u\) and \(v\). We thus have proven that if \(u\) and \(v\) are nonempty words satisfying \(w = uv\), then either we have \(v \geq u\) or the word \(v\) is a prefix of \(u\). In other words, Assertion \(\mathcal{C}'\) holds. Hence, we have proven the implication \(\mathcal{F}'' \implies \mathcal{C}'\).

**Proof of the implication \(\mathcal{B}' \implies \mathcal{G}'\):** Assume that Assertion \(\mathcal{B}'\) holds.

Let \(s\) be the longest suffix \(v\) of \(w\) satisfying \(v < w\). (This is well-defined, because there exists a suffix \(v\) of \(w\) satisfying \(v < w\) - namely, the empty word.) So we know that \(s\) is a suffix of \(w\) satisfying \(v < w\). In other words, \(s\) is a suffix of \(w\) and satisfies \(s < w\). As a consequence, \(s\) is a proper suffix of \(w\) (because otherwise, \(s\) would be \(w\), and this would contradict \(s < w\)). Hence, there exists a nonempty word \(h \in \mathfrak{A}^*\) satisfying \(w = hs\). Consider this \(h\). Using Assertion \(\mathcal{B}'\), it is easy to see that \(s\) is a prefix of \(w\) 707. In other words, there exists a \(g \in \mathfrak{A}\) such that \(w = sg\). Consider this \(g\).

We know that \(s\) is the longest suffix \(v\) of \(w\) satisfying \(v < w\). Hence, \((11.937)\)

\[
\text{if } v \text{ is a suffix of } w \text{ satisfying } v < w, \text{ then } \ell (v) \leq \ell (s) .
\]

There exists a nonnegative integer \(m\) such that \(h^m\) is a prefix of \(s\) (for example, the nonnegative integer \(m = 0\)). Consider the **maximal** such integer \(m\) 708. Then, \(h^m\) is a prefix of \(s\), but \(h^{m+1}\) is not a prefix of \(s\). Since \(h^m\) is a prefix of \(s\), there exists a word \(q \in \mathfrak{A}^*\) such that \(s = h^mq\). Consider this \(q\). Clearly, \(w = h s = h h^m q = h^{m+1} q = h^m h q\). Hence, \(h^m h q = w = s g = h^m q g\). Cancelling \(h^m\) from this

707 **Proof:** Assume the contrary. Thus, \(s\) is not a prefix of \(w\). Hence, \(s\) is nonempty. Therefore, Assertion \(\mathcal{B}'\) (applied to \(u = h\) and \(v = s\)) yields that either we have \(s \geq w\) or the word \(s\) is a prefix of \(w\). Since \(s\) is not a prefix of \(w\), we must thus have \(s \geq w\). But this contradicts \(s < w\). This contradiction shows that our assumption was wrong, qed.

708 **Proof:** This is well-defined, because of the following reason:

We have \(\ell (h) \geq 1\) (since the word \(h\) is nonempty). Thus, for every \(m \in \mathbb{N}\) satisfying \(m > \ell (s)\), we have \(\ell (h^m) = m \ell (h) \geq m \ell (s)\). In other words, for every \(m \in \mathbb{N}\) satisfying \(m > \ell (s)\), the word \(h^m\) is longer than \(s\). Hence, for every \(m \in \mathbb{N}\) satisfying \(m > \ell (s)\), the word \(h^m\) cannot be a prefix of \(s\). Thus, for every sufficiently high \(m \in \mathbb{N}\), the word \(h^m\) cannot be a prefix of \(s\).

We thus know the following:

- There exists a nonnegative integer \(m\) such that \(h^m\) is a prefix of \(s\).
- For every sufficiently high \(m \in \mathbb{N}\), the word \(h^m\) cannot be a prefix of \(s\).

Consequently, there exists a **maximal** nonnegative integer \(m\) such that \(h^m\) is a prefix of \(s\), qed.
equality, we obtain \( hq = qg \). It is now easy to see that \( h > q \).\(^{709}\) Hence, \( q \leq h \leq hq = qg \). Thus, Proposition 6.2(g) (applied to \( q, h \) and \( q \) instead of \( a, b \) and \( c \)) yields that \( q \) is a prefix of \( h \).

Next, we shall prove that the word \( h \) is Lyndon.

In fact, assume the contrary. Then, \( h \) is not Lyndon. Let \( v \) be the (lexicographically) smallest nonempty suffix of \( h \). Then, Proposition 6.19(b) (applied to \( h \) instead of \( w \)) yields that there exists a nonempty \( v \in \mathcal{A}^* \) such that \( h = uv \), \( u \geq v \) and \( uv \geq vu \). Consider this \( u \). Since \( w = \frac{h}{uv} s = uvs \), it is clear that the word \( vs \) is a suffix of \( w \). If we had \( vs < w \), then we could therefore obtain \( \ell(vs) \leq \ell(s) \) (by \((11.937)\)), applied to \( vs \) instead of \( v \), which would contradict \( \ell(vs) = \ell(v) + \ell(s) > \ell(s) \). Thus, we cannot have \( \ell(s) > \ell(v) \) (since \( v \) is nonempty).

Recall that \( s < v \). Hence, \( vs \leq vw \) (by Proposition 6.2(b), applied to \( v, s \) and \( w \) instead of \( a, c \) and \( d \)). Now, \( uv \geq vs \), so that \( uv \leq vs \leq v \). Hence, \( v = h \). Therefore, Proposition 6.2(g) (applied to \( vu, uv \) and \( vs \) instead of \( a, b \) and \( c \)) yields \( vu = v \). Thus, the elements \( u \) and \( v \) of the monoid \( \mathcal{A}^* \) commute. Since \( h = uv \), the element \( h \) lies in this submonoid, and therefore the element \( h^m \) lies in it as well. Thus, \( h^m \) commutes with \( v \) (since this submonoid is commutative, i.e.), we have \( vh^m = h^mv \). Thus, \( v = vh^m = vh^m \). Hence, \( h^mvq = vs \geq w = h \). Hence, \( \ell(vq) = \ell(v) + \ell(q) > \ell(v) \), so that \( \ell(vq) \leq \ell(hq) \). This contradicts the fact that \( hq \leq vq \) with \( hq \leq vq \) instead of \( a, b \) and \( c \) yields \( hq \leq vq \). Combined with \( hq \leq vq \), this yields \( vq = hq \). Hence, \( \ell(vq) = \ell(hq) \). This contradicts the fact that \( \ell(vq) > \ell(hq) \), which is absurd. This contradiction proves that our assumption is wrong. Thus, we have shown that the word \( h \) is Lyndon.

We now know that \( h \in \mathcal{A}^* \) is a Lyndon word, \( m + 1 \) is a positive integer, and \( q \) is a prefix of \( h \), and we have \( w = h^{m+1} q \). Hence, there exists a Lyndon word \( t \in \mathcal{A}^* \), a positive integer \( t \) and a prefix \( p \) of \( t \) (possibly empty) such that \( t = tp \) (namely, \( t = h \), \( t = m + 1 \) and \( p = q \)). In other words, Assertion \( \mathcal{G}' \) holds. This proves the implication \( \mathcal{B}' \implies \mathcal{G}' \).

We are now going to prove the implication \( \mathcal{F}' \implies \mathcal{B}' \); this implication will later be used in the proof of the implication \( \mathcal{H}' \implies \mathcal{B}' \).

Proof of the implication \( \mathcal{F}' \implies \mathcal{B}' \): Assume that Assertion \( \mathcal{F}' \) holds. In other words, the word \( w \) is a prefix of a Lyndon word in \( \mathcal{A}^* \). Let \( z \) be this Lyndon word. Thus, \( w \) is a prefix of \( z \). In other words, there exists a word \( q \in \mathcal{A}^* \) such that \( z = wq \). Consider this \( q \).

Let \( u \) and \( v \) be nonempty words satisfying \( w = uv \). We are going to prove that either we have \( v \geq w \) or the word \( v \) is a prefix of \( w \).

\(^{709}\)Proof: Assume the contrary. Then, \( h \leq q \). Hence, \( h \leq q \leq qq \) (since \( hq = qg \)). Therefore, Proposition 6.2(g) (applied to \( h, q \) and \( q \) instead of \( a, b \) and \( c \)) yields that \( h \) is a prefix of \( q \). In other words, there exists a word \( r \in \mathcal{A}^* \) such that \( q = hr \). Consider this \( r \). Now, \( s = h^m q = h^m q r = h^m + 1 r \), so that \( h^m + 1 \) is a prefix of \( s \). This contradicts the fact that \( h^m + 1 \) is not a prefix of \( s \). This contradiction proves that our assumption was wrong, qed.
We have $z = wq = uvq$. Thus, Proposition 6.14(a) (applied to $z$ and $vq$ instead of $w$ and $v$) yields $vq \geq z$ (since $vq$ is nonempty (since $v$ is nonempty)). Hence, $vq \geq z = vq$. In other words, $wq \leq vq$. Thus, Proposition 6.2(e) (applied to $w, q, v$ and $q$ instead of $a, b, c$ and $d$) yields that either we have $w \leq v$ or the word $v$ is a prefix of $w$.

Now, forget that we fixed $u$ and $v$. We thus have shown that if $u$ and $v$ are nonempty words satisfying $w = uv$, then either we have $v \geq w$ or the word $v$ is a prefix of $w$. In other words, Assertion $B'$ holds. Thus, the implication $F' \implies B'$ is proven.

Proof of the implication $H' \implies B'$: Assume that Assertion $H'$ holds. In other words, there exists a Lyndon word $t \in \mathfrak{A}^*$, a nonnegative integer $\ell$ and a prefix $p$ of $t$ (possibly empty) such that $w = t^\ell p$. Consider this $t$, this $\ell$ and this $p$.

We are going to prove that for every $m \in \mathbb{N}$,

\[(11.938) \quad \text{every suffix $s$ of $t^m p$ satisfies either } s \geq t^m p \text{ or (the word $s$ is a prefix of $t^m p$).}
\]

Proof of (11.938): We will prove (11.938) by induction over $m$:

Induction base: Using the implication $F' \implies B'$, it is easy to see that (11.938) holds for $m = 0$. This completes the induction base.

Induction step: Let $M$ be a positive integer. Assume that (11.938) is proven for $m = M - 1$. We will now show that (11.938) holds for $m = M$.

Let $r$ denote the word $t^M r^{-1} p$. It is easy to see that $r$ is a prefix of $t^M p$. In other words, there exists a word $g \in \mathfrak{A}^*$ such that $t^M p = rg$. Consider this $g$.

Let $s$ be a suffix of $t^M p$. We shall show that either $s \geq t^m p$ or (the word $s$ is a prefix of $t^M p$).

In order to prove this, let us assume the contrary (for the sake of contradiction). Then, neither $s \geq t^M p$ nor (the word $s$ is a prefix of $t^M p$). In other words, we have $s < t^M p$, and the word $s$ is not a prefix of $t^M p$. If the word $s$ was a prefix of $r$, then the word $s$ would be a prefix of $t^M p$ (since $r$ is a prefix of $t^M p$), which would contradict the fact that the word $s$ is not a prefix of $t^M p$. Hence, the word $s$ cannot be a prefix of $t^M p$.

In other words, the word $s$ cannot be a prefix of $t^M p$ (since $r = t^{M-1} p$).

The word $s$ is a suffix of $t^M p = t^{M-1} p = t r$. Therefore, we must be in one of the following two cases (depending on whether this suffix begins before the suffix $r$ of $tr$ begins or afterwards):

Case 1: The word $s$ is a suffix of $r$. (Note that $s = r$ is allowed.)

Case 2: The word $s$ has the form $s'r$ where $s'$ is a nonempty suffix of $t$.

Let us consider Case 1 first. In this case, the word $s$ is a suffix of $r$. In other words, the word $s$ is a suffix of $t^{M-1} p$ (since $r = t^{M-1} p$). Hence, (11.938) (applied to $m = M - 1$) yields that either $s \geq t^{M-1} p$.

Proof. Assume that $m = 0$. Then, $t^0 p = \emptyset = p = p$.

Let $s$ be a suffix of $t^m p$. Then, $s$ is a suffix of $t^m p = p$. In other words, there exists a word $g \in \mathfrak{A}^*$ satisfying $p = gs$. Consider this $g$.

We are going to prove that either $s \geq t^m p$ or (the word $s$ is a prefix of $t^m p$). If the word $s$ is empty, then this is obvious (because if the word $s$ is empty, then the word $s$ is a prefix of $t^m p$). Hence, we WLOG assume that the word $s$ is nonempty. If the word $g$ is empty, then it is also clear that either $s \geq t^m p$ or (the word $s$ is a prefix of $t^m p$) (because if the word $g$ is empty, then $g = \emptyset$ and thus $t^m p = p = g = s = s$, so that the word $s$ is a prefix of $t^m p$). Hence, we WLOG assume that the word $g$ is nonempty.

But the word $p$ is a prefix of a Lyndon word in $\mathfrak{A}^*$ (since $p$ is a prefix of $t$, and since $t$ is a Lyndon word in $\mathfrak{A}^*$). In other words, Assertion $F'$ with $w$ replaced by $p$ is satisfied. Hence, Assertion $B'$ with $w$ replaced by $p$ is satisfied as well (since we have already proven the implication $F' \implies B'$). In other words,

\[(11.939) \quad \left( \text{if } u \text{ and } v \text{ are nonempty words satisfying } p = uv, \text{ then } \text{either we have } v \geq p \text{ or the word } v \text{ is a prefix of } p \right).
\]

Since the words $g$ and $s$ are nonempty, we can apply (11.939) to $u = g$ and $v = s$. As a result, we obtain that either we have $s \geq p$ or the word $s$ is a prefix of $p$. In other words, either $s \geq p$ or (the word $s$ is a prefix of $p$). In other words, either $s \geq t^m p$ or (the word $s$ is a prefix of $t^m p$) (since $t^m p = p$). This proves (11.938).

Proof. There exists a word $q \in \mathfrak{A}^*$ such that $t = pq$ (since $p$ is a prefix of $t$). Consider this $q$. We have $t^{M-1} p = t^{M-1} p = t^{M-1} pq = rqp = r(qp)$. Hence, $r$ is a prefix of $t^M p$, qed.
or (the word $s$ is a prefix of $t^{M-1}p$) (since (11.938) holds for $m = M - 1$). Since the word $s$ cannot be a prefix of $t^{M-1}p$, we thus must have $s \geq t^{M-1}p$. Thus, $t^{M-1}p \leq s < t^M p = r g$. Hence, $r = t^{M-1}p \leq s < r g$. Therefore, Proposition 6.2(g) (applied to $r, s$ and $g$ instead of $a, b$ and $c$) yields that $r$ is a prefix of $s$. Since $\ell (r) \geq \ell (s)$ (because $s$ is a suffix of $r$), this can only hold if $r = s$. We thus have $r = s$. Thus, $s = r$, so that $s$ is a prefix of $s = r$. This contradicts the fact that the word $s$ cannot be a prefix of $r$. Thus, we have found a contradiction in Case 1.

Let us now consider Case 2. In this case, the word $s$ has the form $s' r$ where $s'$ is a nonempty suffix of $t$. Consider this $s'$. Corollary 6.15 (applied to $t$ and $s'$ instead of $w$ and $v$) yields $s' \geq t$. But $s'$ is a suffix of $t$, so that $\ell (s') \leq \ell (t)$. Also, $s = s' r$, so that $s' r = s < t^M p = r t$. Hence, Proposition 6.2(f) (applied to $s', r, t$ and $r$ instead of $a, b, c$ and $d$) yields that $s' \leq t$. Combined with $s' \geq t$, this yields $s' = t$. Hence, $s = s' r = t^M p$ (since $t^M p = r t$), which contradicts the fact that the word $s$ is not a prefix of $t^M p$. Thus, we have found a contradiction in Case 2.

We have thus obtained a contradiction in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this shows that we always get a contradiction. This completes the proof that either $s \geq t^M p$ or (the word $s$ is a prefix of $t^M p$).

Now, forget that we fixed $s$. We thus have shown that every suffix $s$ of $t^M p$ satisfies either $s \geq t^M p$ or (the word $s$ is a prefix of $t^M p$). In other words, (11.938) holds for $m = M$. This completes the induction step, and thus (11.938) is proven by induction.

Now, let $u$ and $v$ be nonempty words satisfying $w = w u$. Then, $v$ is a suffix of $w = t^\ell p$. Hence, (11.938) (applied to $m = \ell$ and $s = v$) yields that either $v \geq t^\ell p$ or (the word $v$ is a prefix of $t^\ell p$). In other words, either $v \geq w$ or (the word $v$ is a prefix of $w$) (since $w = t^\ell p$). In other words, either we have $v \geq w$ or the word $v$ is a prefix of $w$.

Now, forget that we fixed $u$ and $v$. We thus have shown that if $u$ and $v$ are nonempty words satisfying $w = w u$, then either we have $v \geq w$ or the word $v$ is a prefix of $w$. In other words, Assertion $B'$ holds. Thus, the implication $H' \implies B'$ is proven.

We have thus proven the implications $B' \implies E', C' \implies E', G' \implies H', E' \implies F''$, $F'' \implies B'$, $F' \implies C'$, $B' \implies G'$ and $H' \implies B'$. Combined, these yield the equivalence $B' \iff C' \iff E' \iff F'' \iff G' \iff H'$. This solves Exercise 6.32(b).

(c) The implication $F' \implies B'$ has already been proven in our solution of Exercise 6.32(b). Hence, Exercise 6.32(c) is solved.

(d) Assume that Assertion $D'$ holds. Then, Assertion $A'$ holds as well (because of the equivalence $A' \iff D'$). In other words, the word $w$ is a power of a Lyndon word. In other words, there exist a Lyndon word $z \in \mathfrak{A}^*$ and a nonnegative integer $m$ such that $w = z^m$. Consider these $z$ and $m$. There exists a Lyndon word $t \in \mathfrak{A}^*$, a nonnegative integer $\ell$ and a prefix $p$ of $t$ (possibly empty) such that $w = t^\ell p$ (namely, $t = z$, $\ell = m$ and $p = \varnothing$). In other words, Assertion $H'$ holds. Since Assertion $H'$ is equivalent to Assertion $B'$ (because of the equivalence $B' \iff C' \iff E' \iff F'' \iff G' \iff H'$), this yields that Assertion $B'$ holds. Thus, the implication $D' \implies B'$ is proven.

(e) Assume that there exists a letter $\mu \in \mathfrak{A}$ such that ($\mu > a$ for every letter $a$ of $w$). Consider this $\mu$. We need to prove that the equivalence $F' \iff F''$ holds.

Combining the implication $F' \implies B'$ (which has already been proven) and the implication $B' \implies F''$ (which follows from the equivalence $B' \iff C' \iff E' \iff F'' \iff G' \iff H'$ proven above), we obtain the implication $F' \implies F''$. Thus, in order to prove the equivalence $F' \iff F''$, it is enough to verify the implication $F'' \implies F'$. Let us do this now.

Assume that Assertion $F''$ holds. Let $\mathfrak{B}$ denote the alphabet consisting of all letters that appear in $w$. Clearly, $\mathfrak{B}$ is a subalphabet of $\mathfrak{A}$, and we have $w \in \mathfrak{B}^*$. Moreover, we have $\mu > a$ for every letter $a$ of $w$. Therefore, $\mu > a$ for every $a \in \mathfrak{B}$ (because the elements of $\mathfrak{B}$ are precisely the letters of $w$). In other words, $a < \mu$ for every $a \in \mathfrak{B}$. As a consequence, $\mu \notin \mathfrak{B}$; that is, $\mu$ is an object not in the alphabet $\mathfrak{B}$.

We know that Assertion $F''$ holds. Due to the implication $F'' \implies B'$ (which follows from the equivalence $B' \iff C' \iff E' \iff F'' \iff G' \iff H'$ proven above), this yields that Assertion $B'$ holds. Thus, Assertion $B'$ with $\mathfrak{A}$ replaced by $\mathfrak{B}$ holds as well (since the Assertion $B'$ does not change if we extend our alphabet). Due to the implication $B' \implies F''$ (which follows from the equivalence $B' \iff C' \iff E' \iff F'' \iff G' \iff H'$ proven above), this yields that Assertion $F''$ with $\mathfrak{A}$ replaced by $\mathfrak{B}$ holds as well. Thus, we can apply Assertion $F''$ with $\mathfrak{A}$ replaced by $\mathfrak{B}$ to $m = \mu$ (because $\mu$ is an object not in the alphabet $\mathfrak{B}$, and the total
order on \( \mathfrak{B} \) satisfies \((a < \mu \text{ for every } a \in \mathfrak{B})\). As a result, we conclude that the word \( w\mu \in (\mathfrak{B} \cup \{\mu\})^* \) is a Lyndon word. Since \( \mathfrak{B} \cup \{\mu\} \subset \mathfrak{A} \) (because \( \mathfrak{B} \subset \mathfrak{A} \) and \( \mu \in \mathfrak{A} \)), we have \( w\mu \in (\mathfrak{B} \cup \{\mu\}) \subset \mathfrak{A}^* \), and thus \( w\mu \in \mathfrak{A}^* \) is a Lyndon word. Of course, the word \( w \) is a prefix of \( w\mu \). As a consequence, the word \( w \) is a prefix of a Lyndon word in \( \mathfrak{A}^* \) (namely, of the word \( w\mu \)). In other words, Assertion \( F' \) holds. This proves the implication \( F'' \implies F' \). Thus, the solution of Exercise 6.32(e) is complete.

(f) Assume that there exists a letter \( \mu \in \mathfrak{A} \) such that \((\mu > a \text{ for some letter } a \text{ of } w)\). Consider this \( \mu \). We need to prove that the equivalence \( F' \iff F'' \) holds.

Just as in our solution of Exercise 6.32(e) above, we can see that it is enough to verify the implication \( F'' \implies F' \). Let us do this now.

Assume that Assertion \( F'' \) holds. Due to the implication \( F'' \implies B' \) (which follows from the equivalence \( B' \iff C' \iff E' \iff F'' \iff G' \iff H' \) proven above), this yields that Assertion \( B' \) holds.

Let \( \beta \) be the highest letter of the word \( w\mu \in \mathfrak{A}^* \) (the concatenation of the word \( w \) with the one-letter word \( \mu \)). The word \( w\beta\ell(w) \) is clearly nonempty (since \( w \) is nonempty). We shall now prove that the word \( w\beta\ell(w) \) is Lyndon.

Indeed, assume the contrary. Thus, \( w\beta\ell(w) \) is not Lyndon. Let \( v \) denote the (lexicographically) smallest nonempty suffix of \( w\beta\ell(w) \). Then, Proposition 6.19(b) (applied to \( w\beta\ell(w) \) instead of \( w \)) yields that there exists a nonempty \( u \in \mathfrak{A}^* \) such that \( w\beta\ell(w) = wu, u \geq v \) and \( uv \geq vu \).

Let \( f \) be the first letter of the word \( w \) (this is well-defined since \( w \) is nonempty). Then, we can write \( w \) in the form \( w = fs \) for some word \( s \in \mathfrak{A}^* \). Consider this \( s \).

We know that \( \mu > a \) for some letter \( a \) of \( w \). Consider this \( a \). We also know that \( \beta \) is the highest letter of the word \( w\mu \). Thus, \( \beta \geq \mu \) for every letter of the word \( w\mu \). In particular, this yields that \( \beta \geq \mu \) (since \( \mu \) is a letter of the word \( w\mu \)), so that \( \beta \geq \mu > a \). But it is fairly easy to see (using the fact that Assertion \( B' \) holds) that \( a \geq f \).

The word \( v \) is a proper suffix of \( w\beta\ell(w) \) (since \( w\beta\ell(w) = uv \) and since \( u \) is nonempty) and is nonempty. Hence, \( v \) is a nonempty proper suffix of \( w\beta\ell(w) \). Therefore, we must be in one of the following two cases (depending on whether this suffix begins before the suffix \( \beta\ell(w) \) of \( w\beta\ell(w) \) begins or afterwards):

Case 1: The word \( v \) is a nonempty suffix of \( \beta\ell(w) \). (Note that \( v = \beta\ell(w) \) is allowed.)

Case 2: The word \( v \) has the form \( q\beta\ell(w) \) where \( q \) is a nonempty proper suffix of \( w \).

Let us first consider Case 1. In this case, the word \( v \) is a nonempty suffix of \( \beta\ell(w) \). Thus, \( v \) has the form \( \beta^i \) for some \( i \in \{0,1,\ldots,\ell(w)\} \). Consider this \( i \). We have \( v = \beta^i \). Thus, \( i \neq 0 \) (since \( v \) is nonempty). Hence, (the first letter of \( v \) = \( \beta \). But since the word \( w \) is nonempty, we have

\[
\text{(the first letter of } w\beta\ell(w) \text{)} = (\text{the first letter of } w) = f
\]

(since \( f \) was defined to be the first letter of \( w \)). Now,

\[
\text{(the first letter of } v \text{)} = \beta > f = (\text{the first letter of } w\beta\ell(w) \rangle.
\]

By the definition of lexicographic order, this shows that \( v > w\beta\ell(w) \). But this contradicts \( w\beta\ell(w) = uv \geq vu \geq v \). Hence, we have found a contradiction in Case 1.

\[\text{Proof.} \quad \text{Assume the contrary. Thus, } a < f.\]

Since \( a \) is a letter of \( w \), the word \( w \) must have a suffix which begins with the letter \( a \). Let \( p \) be this suffix. Then, there exists a word \( q \in \mathfrak{A}^* \) such that \( w = qp \) (since \( p \) is a suffix of \( w \)). Consider this \( q \). The word \( p \) is nonempty (since it begins with \( a \)). Since the word \( p \) begins with the letter \( a \), we have

\[
(\text{the first letter of } p) = a < f = (\text{the first letter of } w).
\]

By the definition of the lexicographic order, this shows that \( p < w \). Hence, \( p \neq w \). Now, the word \( q \) is nonempty (since otherwise, we would have \( q = \varepsilon \) and thus \( w = q \cdot p = p \), contradicting \( p \neq w \)). Hence, applying Assertion \( B' \) to \( q \) and \( p \) instead of \( u \) and \( v \), we conclude that either we have \( p \geq w \) or the word \( p \) is a prefix of \( w \). Since \( p \geq w \) is impossible (because \( p < w \)), this yields that the word \( p \) is a prefix of \( w \). Since \( p \) is nonempty, this shows that \( p \) is a nonempty prefix of \( w \). But this yields that

\[
(\text{the first letter of } p) = (\text{the first letter of } w);
\]

this contradicts (11.940). This contradiction proves that our assumption was wrong, QED.
Let us now consider Case 2. In this case, the word \( v \) has the form \( q^β(w) \) where \( q \) is a nonempty proper suffix of \( w \). Consider this \( q \). Notice that \( \ell(q) < \ell(w) \) (since \( q \) is a proper suffix of \( w \)), so that \( \ell(w) > \ell(q) \).

Since \( q \) is a proper suffix of \( w \), there exists a nonempty word \( u \in \mathfrak{A}^* \) such that \( w = uq \). Consider this \( u \). Recall that Assertion \( B' \) holds. Applying this Assertion \( B' \) to \( q \) instead of \( v \), we conclude that either we have \( q \geq w \) or the word \( q \) is a prefix of \( w \). If the word \( q \) is not a prefix of \( w \), then it is very easy to derive a contradiction\(^{713}\). Hence, the word \( q \) must be a prefix of \( w \). In other words, there exists a word \( g \in \mathfrak{A}^* \) such that \( w = qg \). Consider this \( g \). Notice that \( \ell\left(\frac{w}{=qg}\right) = \ell(q) + \ell(g) \), so that

\[
\ell\left(\beta^\ell(w)\right) = \ell(w) > \ell(g).
\]

In other words, the word \( \beta^\ell(w) \) is longer than \( g \). It is now easy to see that \( \beta^\ell(w) \geq g \)\(^{714}\).

In other words, \( g \leq \beta^\ell(w) \). Hence, Proposition 6.2(b) (applied to \( q \), \( g \) and \( \beta^\ell(w) \) instead of \( a \), \( c \) and \( d \)) yields \( qg \leq \beta^\ell(w) \), so that \( w = qg \leq q\beta^\ell(w) = v \) (since \( v = q\beta^\ell(w) \)). Thus, \( v \geq g \), so that \( u \geq v \geq w \).

But \( v = q\beta^\ell(w) \), so that \( w\beta^\ell(w) = u \stackrel{\ell}{\longrightarrow} v = uq\beta^\ell(w) \). Cancelling \( \beta^\ell(w) \) from this equation, we obtain \( w = uq \). Thus, \( u \leq uq = w \). Since \( u \neq w \)\(^{715}\), this becomes \( u < w \). This contradicts \( u \geq w \). Hence, we have found a contradiction in Case 2.

We have now obtained a contradiction in each of our two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that we always obtain a contradiction. Thus, our assumption was wrong, and we conclude that the word \( w\beta^\ell(w) \) is Lyndon. Hence, \( w \) is a prefix of a Lyndon word in \( \mathfrak{A}^* \) (because \( w \) is a prefix of the word \( w\beta^\ell(w) \)). In other words, Assertion \( F' \) holds. This proves the implication \( \mathcal{F}'' \implies \mathcal{F}' \). This solves Exercise 6.32(f).

Remark: Of course, for a letter \( \mu \in \mathfrak{A} \), if we have \( \mu > a \) for every letter \( a \) of \( w \), then we also have \( \langle \mu > a \) for some letter \( a \) of \( w \) (since \( w \) is nonempty and thus has at least one letter). Hence, Exercise 6.32(e) is a particular case of Exercise 6.32(f).


(a) If we replace the word “total” by “partial” throughout the proof of Proposition 6.2, then the resulting argument can be used in the partial-order setting to prove Proposition 6.2 with “a total order” replaced by “a partial order”. Hence, Proposition 6.2 holds in the partial-order setting, as long as one replaces “a total order” by “a partial order” in part (a) of this Proposition. This solves Exercise 6.33(a).

---

\(^{713}\)Proof. Assume that the word \( q \) is not a prefix of \( w \). Then, we have \( q \geq w \) (since we know that either we have \( q \geq w \) or the word \( q \) is a prefix of \( w \)). In other words, \( w \leq q \). Hence, Proposition 6.2(d) (applied to \( w \), \( \beta^\ell(w) \), \( q \) and \( \beta^\ell(w) \) instead of \( a \), \( b \), \( c \) and \( d \)) yields that either we have \( w\beta^\ell(w) \leq q\beta^\ell(w) \) or the word \( w \) is a prefix of \( q \). Since the word \( w \) is not a prefix of \( q \) (because \( \ell(w) > \ell(q) \)), this yields that we have \( w\beta^\ell(w) \leq q\beta^\ell(w) \). Since \( \ell\left(w\beta^\ell(w)\right) = \ell(w) + \ell\left(\beta^\ell(w)\right) > \ell(q) + \ell\left(\beta^\ell(w)\right) = \ell\left(q\beta^\ell(w)\right) \), this shows that \( w\beta^\ell(w) < q\beta^\ell(w) = v \) (because \( v = q\beta^\ell(w) \)). Hence, \( v \geq w\beta^\ell(w) = uv \geq vu \geq v \), which is a contradiction, qed.

\(^{714}\)Proof. Assume the contrary. Then, \( \beta^\ell(w) < g \). By the definition of the lexicographic order, this yields that

\[
\text{either there exists an } i \in \{1,2,\ldots,\min\{\ell\left(\beta^\ell(w)\right),\ell(g)\}\} \\
\text{such that } \left(\beta^\ell(w)\right)_i < g_i, \text{ and every } j \in \{1,2,\ldots,i-1\} \text{ satisfies } \left(\beta^\ell(w)\right)_j = g_j,
\]

\[
\text{or the word } \beta^\ell(w) \text{ is a prefix of } g.
\]

Since the word \( \beta^\ell(w) \) cannot be a prefix of \( g \) (because the word \( \beta^\ell(w) \) is longer than \( g \)), this yields that there exists an \( i \in \{1,2,\ldots,\min\{\ell(\beta^\ell(w)),\ell(g)\}\} \) such that \( \left(\beta^\ell(w)\right)_i < g_i \). Consider this \( i \). We have \( \beta^\ell(w)_i = \beta_i \), so that \( \beta = \beta^\ell(w)_i < g_i \).

But \( g \) is a suffix of \( w \) (since \( w = qg \)), and thus every letter of \( g \) is a letter of \( w \). In particular, \( g_i \) is a letter of \( w \). Hence, \( g_i \) is a letter of \( w\beta^\ell(w) \) as well (since every letter of \( w \) is a letter of \( w\beta^\ell(w) \) because \( w \) is a prefix of \( w\beta^\ell(w) \)). Since \( \beta \) is the highest letter of \( w\beta^\ell(w) \), this yields \( \beta \geq g_i \). But this contradicts \( \beta < g_i \). This contradiction proves that our assumption was false, qed.

\(^{715}\)Proof. Assume the contrary. Then, \( u = w \). Hence, \( uv = u = w = uq \). Cancelling \( u \) from this equation, we obtain \( v = q \).

Hence, the word \( q \) is empty. This contradicts the fact that \( q \) is nonempty. This contradiction shows that our assumption was wrong, qed.
(b) Let us work in the partial-order setting. Let \( a, b, c, d \in \mathcal{A}^* \) be four words such that the words \( ab \) and \( cd \) are comparable. We need to prove that the words \( a \) and \( c \) are comparable.

The words \( ab \) and \( cd \) are comparable. In other words, we have either \( ab \leq cd \) or \( ab \geq cd \). In other words, we have either \( ab \leq cd \) or \( cd \leq ab \). We can WLOG assume that \( ab \leq cd \) (since otherwise, we can achieve \( ab \leq cd \) by switching \( a \) and \( b \) with \( c \) and \( d \), respectively). Assume this. We know (from Exercise 6.33(a)) that Proposition 6.2 holds in the partial-order setting, as long as one replaces “a total order” by “a partial order” in part (a) of this Proposition. In particular, Proposition 6.2(e) holds in the partial-order setting. Applying Proposition 6.2(e), we thus conclude that either we have \( a \leq c \) or the word \( c \) is a prefix of \( a \). Thus, either we have \( a \leq c \) or we have \( c \leq a \) (because if \( c \) is a prefix of \( a \), then \( c \leq a \)). In other words, either we have \( a \leq c \) or we have \( a \geq c \). In other words, the words \( a \) and \( c \) are comparable. This solves Exercise 6.33(b).

(c) Recall (from Exercise 6.33(a)) that Proposition 6.2 holds in the partial-order setting, as long as one replaces “a total order” by “a partial order” in part (a) of this Proposition.

Proposition 6.4 holds in the partial-order setting, because its proof applies verbatim in this setting.

Proposition 6.5 holds in the partial-order setting, because its proof applies verbatim in this setting.

The proof of Corollary 6.6 given above is no longer applicable in the partial-order setting. However, the alternative proof of Corollary 6.6 given in the solution to Exercise 6.7 does apply verbatim in the partial-order setting. Thus, Corollary 6.6 holds in the partial-order setting.

Corollary 6.8 holds in the partial-order setting, because its proof applies verbatim in this setting.

Exercise 6.9 and Exercise 6.10 hold in the partial-order setting, because their solutions apply verbatim in this setting.

Exercise 6.11 and Exercise 6.12 hold in the partial-order setting, because their solutions apply verbatim in this setting.

Proposition 6.14 holds in the partial-order setting, because its proof applies verbatim in this setting.

Corollary 6.15 holds in the partial-order setting, because its proof applies verbatim in this setting.

Proposition 6.16 holds in the partial-order setting, because its proof applies verbatim in this setting.

Corollary 6.17 holds in the partial-order setting, because its proof applies verbatim in this setting.

Our proof of Proposition 6.18 does not directly apply in the partial-order setting; however, it can be tweaked so that it does:

**Proof of Proposition 6.18 in the partial-order setting.** If the words \( u \) and \( v \) are incomparable, then Proposition 6.18 is easily seen to hold\(^{716}\). Hence, for the rest of this proof, we can WLOG assume that the words \( u \) and \( v \) are comparable. Assume this.

The words \( u \) and \( v \) are comparable. In other words, we have either \( u < v \) or \( u = v \) or \( u > v \). From here, we can proceed as in our proof of Proposition 6.18 in the total-order setting. Proposition 6.18 is thus proven in the partial-order setting.

Exercise 6.21(a) holds in the partial-order setting, because its solution applies verbatim in this setting.

The proof of Theorem 6.20 we gave (using Proposition 6.19) does not work in the partial-order setting\(^{717}\). However, the alternative proof of Theorem 6.20 given in Exercise 6.21(b) does apply verbatim in the partial-order setting. Thus, Theorem 6.20 holds in the partial-order setting.

Exercise 6.23 and Exercise 6.24 hold in the partial-order setting, because their solutions apply verbatim in this setting.

Exercise 6.31(a) and Exercise 6.31(b) hold in the partial-order setting, because their solutions apply verbatim in this setting.

Exercise 6.33(c) is thus solved.

\(^{716}\)Proof. Assume that \( u \) and \( v \) are incomparable. If the words \( uv \) and \( vu \) were comparable, then the words \( u \) and \( v \) would also be comparable (by Exercise 6.33(b), applied to \( a = u \), \( b = v \), \( c = v \) and \( d = u \)), which would contradict the fact that \( u \) and \( v \) are incomparable. Hence, the words \( uv \) and \( vu \) are incomparable. Thus, we cannot have \( u v \geq v u \). But we cannot have \( u \geq v \) either (since \( u \) and \( v \) are incomparable). Thus, neither \( u \geq v \) nor \( u v \geq v u \) holds. Hence, \( u \geq v \) if and only \( u v \geq v u \); therefore, Proposition 6.18 holds, qed.

\(^{717}\)One might try tweaking Proposition 6.19 for the partial-order setting by replacing “the (lexicographically) smallest nonempty suffix” by “a nonempty suffix which is (lexicographically) minimal among the nonempty suffixes”, and by replacing “\( u \geq v \) and \( u v \geq v u \)” by “\( u v \geq v u \)”. But this still would not hold. For example, if \( u \) is the word \( X X 12 \) over the partially ordered alphabet \( \{X, 1, 2\} \) with relation \( 1 < 2 \), then the nonempty suffix \( X X 12 \) is lexicographically minimal among such suffixes, but we do have \( X < X 12 \). (At least \( u v < v u \) does not hold indeed.)
Definition 11.29. Let \( P \) be a poset.

(a) A poset \( Q \) is said to be an extension of \( P \) if and only if the following two statements hold:

- We have \( Q = P \) as sets.
- Any two elements \( a \) and \( b \) of \( P \) satisfying \( a \leq b \) in \( P \) satisfy \( a \leq b \) in \( Q \).

(b) An extension \( Q \) of \( P \) is called a linear extension of \( P \) if and only if the poset \( Q \) is totally ordered.

Proposition 11.30. Let \( P \) be a finite poset.

(a) If \( a \) and \( b \) are two incomparable elements of \( P \), then there exists an extension \( Q \) of \( P \) such that we have \( a < b \) in \( Q \).

(b) If \( a \) and \( b \) are two elements of \( P \) which don’t satisfy \( a \geq b \), then there exists an extension \( Q \) of \( P \) such that we have \( a < b \) in \( Q \).

(c) There exists a linear extension of \( P \).

(d) If \( a \) and \( b \) are two elements of \( P \) which don’t satisfy \( a \geq b \), then there exists a linear extension \( Q \) of \( P \) such that we have \( a < b \) in \( Q \).

Proposition 11.31. Let \( \mathfrak{A} \) be a finite poset.

(a) If \( \mathfrak{B} \) is an extension of the poset \( \mathfrak{A} \), then \( \mathfrak{B}^* \) is an extension of the poset \( \mathfrak{A}^* \).

(b) Let \( u \in \mathfrak{A}^* \) and \( v \in \mathfrak{A}^* \) be two words. Then, \( u < v \) holds in \( \mathfrak{A}^* \) if and only if we have \( u < v \) in \( \mathfrak{B}^* \) for every linear extension \( \mathfrak{B} \) of \( \mathfrak{A} \).

Proof of Proposition 11.31. (a) Let \( \mathfrak{B} \) be an extension of the poset \( \mathfrak{A} \). Then, by the definition of an “extension”, we see that:

- We have \( \mathfrak{B} = \mathfrak{A} \) as sets.
- Any two elements \( a \) and \( b \) of \( \mathfrak{A} \) satisfying \( a \leq b \) in \( \mathfrak{A} \) satisfy \( a \leq b \) in \( \mathfrak{B} \).

Now, it is clear that \( \mathfrak{B}^* = \mathfrak{A}^* \) as sets (since \( \mathfrak{B} = \mathfrak{A} \) as sets). Also, any two elements \( a \) and \( b \) of \( \mathfrak{A}^* \) satisfying \( a \leq b \) in \( \mathfrak{A}^* \) satisfy \( a \leq b \) in \( \mathfrak{B}^* \). Consequently, \( \mathfrak{B}^* \) is an extension of the poset \( \mathfrak{A}^* \). This proves Proposition 11.31(a).
(b) It is easy to prove the logical implication

\[(u < v \text{ in } \mathfrak{A}^*) \implies (u < v \text{ in } \mathfrak{B}^* \text{ for every linear extension } \mathfrak{B} \text{ of } \mathfrak{A})\]

We shall now focus on proving the implication

\[(u < v \text{ in } \mathfrak{B}^* \text{ for every linear extension } \mathfrak{B} \text{ of } \mathfrak{A}) \implies (u < v \text{ in } \mathfrak{A}^*).\]

**Proof of (11.941):** Assume that

\[(u < v \text{ in } \mathfrak{A}^* \text{ for every linear extension } \mathfrak{A}).\]

We need to show that \(u < v \text{ in } \mathfrak{A}^*.\)

First of all, it is impossible that \(u \geq v \text{ in } \mathfrak{A}^*\). Consequently, we have \(u \neq v\) (because otherwise, we would have \(u = v\) and thus \(u \geq v \text{ in } \mathfrak{A}^*\), which would contradict the fact that it is impossible that \(u \geq v \text{ in } \mathfrak{A}^*\)).

Let \(m = \min \{\ell(u), \ell(v)\}\). Then, \(m \leq \ell(u)\) and \(m \leq \ell(v)\). Recall that \(u = (u_1, u_2, \ldots, u_{\ell(u)})\) and \(v = (v_1, v_2, \ldots, v_{\ell(v)})\). If every \(i \in \{1, 2, \ldots, m\}\) satisfies \(u_i = v_i\), then it is easy to see that \(u < v \text{ in } \mathfrak{A}^*\).

Hence, for the rest of the proof of (11.941), we can WLOG assume that not every \(i \in \{1, 2, \ldots, m\}\) satisfies \(u_i = v_i\). Assume this.

Not every \(i \in \{1, 2, \ldots, m\}\) satisfies \(u_i = v_i\). In other words, there exists an \(i \in \{1, 2, \ldots, m\}\) which does not satisfy \(u_i = v_i\). Let \(k\) be the smallest such \(i\). Thus, \(k \in \{1, 2, \ldots, m\}\) does not satisfy \(u_k = v_k\), whereas

\[(11.944)\]

\[
eq\]

\[
\text{such that } (u_i < v_i \text{ in } \mathfrak{B}, \text{ and every } j \in \{1, 2, \ldots, i \} \text{ satisfies } u_j = v_j) \text{ (namely, } i = i'). \text{ Thus, either}
\]

\[\text{either } \begin{cases} u \text{ is a prefix of } v, \text{ or } \end{cases}\]

\[\text{such that } (u_i < v_i \text{ in } \mathfrak{A}, \text{ and every } j \in \{1, 2, \ldots, i \} \text{ satisfies } u_j = v_j).\]

In other words, \(u \leq v \text{ in } \mathfrak{B}^*\). Thus, \(u \leq v \text{ in } \mathfrak{B}^*\) is proven in Case 1.

Let us now consider Case 2. In this case, the word \(u\) is a prefix of \(v\). Hence, \(u \leq v \text{ in } \mathfrak{B}^*\). Thus, \(u \leq v \text{ in } \mathfrak{B}^*\) is proven in Case 2.

Now, \(u \leq v \text{ in } \mathfrak{B}^*\) is proven in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that \(u \leq v \text{ in } \mathfrak{B}^*\) always holds.

Now, let us forget that we fixed \(u\) and \(v\). We have thus shown that any two elements \(u\) and \(v\) of \(\mathfrak{A}^*\) satisfying \(u \leq v \text{ in } \mathfrak{A}^*\) satisfy \(u \leq v \text{ in } \mathfrak{B}^*\). Renaming the variables \(a\) and \(b\) for \(u\) and \(v\) in this statement, we conclude the following: Any two elements \(a\) and \(b\) of \(\mathfrak{A}^*\) satisfying \(a \leq b \text{ in } \mathfrak{A}^*\) satisfy \(a \leq b \text{ in } \mathfrak{B}^*\). Qed.

**Proof of (11.941):** Assume that \(u \leq v \text{ in } \mathfrak{B}^*\). Thus, \(u \neq v\) and \(u \leq v \text{ in } \mathfrak{B}^*\). Let \(\mathfrak{B}\) be a linear extension of \(\mathfrak{A}\). Then, \(\mathfrak{B}^*\) is an extension of the poset \(\mathfrak{A}^*\) (by Proposition 11.31(a)). Hence, \(u \leq v \text{ in } \mathfrak{B}^*\) (since \(u \leq v \text{ in } \mathfrak{A}^*\)). Combined with \(u \neq v\), this yields \(u \leq v \text{ in } \mathfrak{B}^*\).

Now, let us forget that we fixed \(\mathfrak{B}\). We thus have proven that \(u \leq v \text{ in } \mathfrak{B}^*\) for every linear extension \(\mathfrak{B}\) of \(\mathfrak{A}\). This proves (11.941).

**Proof.** Assume the contrary. Thus, \(u \geq v \text{ in } \mathfrak{A}^*\). Hence, \(v \leq u \text{ in } \mathfrak{A}^*\).

Proposition 11.30(c) (applied to \(P = \mathfrak{A}^*\)) tells us that there exists a linear extension \(\mathfrak{A}\). Let \(\mathfrak{B}\) be such a linear extension. Then, \(u < v \text{ in } \mathfrak{B}^*\) (by (11.943)). But \(\mathfrak{B}^*\) is an extension of the poset \(\mathfrak{A}^*\) (by Proposition 11.30(a)). Thus, \(v \leq u \text{ in } \mathfrak{B}^*\) (since \(v \leq u \text{ in } \mathfrak{A}^*\)). This contradicts \(u < v \text{ in } \mathfrak{B}^*\). This contradiction proves that our assumption was wrong, qed.

**Proof.** Assume that every \(i \in \{1, 2, \ldots, m\}\) satisfies \(u_i = v_i\). Since \(u_i = v_i\) always holds.

Thus, \(u_i = \min \{\ell(u), \ell(v)\}\) and \(v_i = \min \{\ell(u), \ell(v)\}\) (since \(m \leq \ell(u)\)). In other words, the word \(v\) is a prefix of \(u\) (since \(v = (u_1, u_2, \ldots, u_m)\) and \(u = (u_1, u_2, \ldots, u_{\ell(u)})\)). Hence, \(v \leq u \text{ in } \mathfrak{A}^*\). In other words, \(u \geq v \text{ in } \mathfrak{A}^*\). This contradicts the fact that it is impossible that \(u \geq v \text{ in } \mathfrak{A}^*\).

This contradiction proves that our assumption was wrong. Hence, we cannot have \(\ell(u) \geq \ell(v)\). We thus have \(\ell(u) < \ell(v)\). Thus, \(m = \min \{\ell(u), \ell(v)\} = \ell(u)\) (since \(\ell(u) < \ell(v)\)). Now,

\[
u = (u_1, u_2, \ldots, u_{\ell(v)}) = (u_1, u_2, \ldots, u_m) \quad \text{(since } \ell(u) = m)\]

\[
\text{But the word } (v_1, v_2, \ldots, v_m) \text{ is clearly a prefix of } (v_1, v_2, \ldots, v_{\ell(u)}) \text{ (since } m \leq \ell(u)). \text{ In other words, the word } u \text{ is a prefix of } u \text{ (since } u = (u_1, u_2, \ldots, u_m) \text{ and } v = (v_1, v_2, \ldots, v_{\ell(v)})). \text{ Hence, } u \leq v \text{ in } \mathfrak{A}^*\). \text{ Combined with } u \neq v, \text{ this yields that } u < v \text{ in } \mathfrak{A}^*, \text{ qed.} \]
We have $u_k \neq v_k$ (since $k$ does not satisfy $u_k = v_k$). We must have $v_k \geq u_k$ in $\mathfrak{A}$. That is, $u_k \leq v_k$ in $\mathfrak{A}$. Combined with $u_k \neq v_k$, this yields that $u_k < v_k$ in $\mathfrak{A}$.

Every $j \in \{1, 2, \ldots, k-1\}$ satisfies $u_j = v_j$ (by (11.944), applied to $i = j$). Also, we have $k \in \{1, 2, \ldots, m\} = \{1, 2, \ldots, m\} \cap \{\ell(u), \ell(v)\}$ (since $m = \min \{\ell(u), \ell(v)\}$). Altogether, we thus have shown that $k \in \{1, 2, \ldots, m\} \cap \{\ell(u), \ell(v)\}$ satisfies $(u_k < v_k$ in $\mathfrak{A}$, and every $j \in \{1, 2, \ldots, k-1\}$ satisfies $u_j = v_j$). Thus, there exists an $i \in \{1, 2, \ldots, i-1\}$ such that $(u_i < v_i$ in $\mathfrak{A}$, and every $j \in \{1, 2, \ldots, i-1\}$ satisfies $u_j = v_j$ (namely, $i = k$). Hence,

- **either** there exists an $i \in \{1, 2, \ldots, m\} \cap \{\ell(u), \ell(v)\}$ such that $(u_i < v_i$ in $\mathfrak{A}$, and every $j \in \{1, 2, \ldots, i-1\}$ satisfies $u_j = v_j$),
- **or** the word $u$ is a prefix of $v$.

In other words, $u < v$ in $\mathfrak{A}^*$ (by the definition of the relation $<$ on $\mathfrak{A}^*$). This proves (11.942).

Combining the implications (11.941) and (11.942), we obtain the following equivalence of statements:

$$(u < v$ in $\mathfrak{A}^*) \iff (u < v$ in $\mathfrak{B}^*$ for every linear extension $\mathfrak{B}$ of $\mathfrak{A})$.

Thus, Proposition 11.31(b) is proven.

Let us now give a new proof of the fact that Corollary 6.6 holds in the partial-order setting:

**Alternative proof of Corollary 6.6 in the partial-order setting.** We need to prove that $uw \geq wu$. If $uw = wu$, then this is obvious. Hence, for the rest of this proof, we can WLOG assume that $uw \neq wu$. Assume this.

Let $\mathfrak{B}$ be any linear extension of the alphabet $\mathfrak{A}$. Then, $\mathfrak{B}^*$ is an extension of $\mathfrak{A}^*$ (according to Proposition 11.31(a)). Hence, $uw \geq wu$ in $\mathfrak{B}^*$ (since $uw \geq wu$ in $\mathfrak{A}^*$) and $wu \geq wu$ in $\mathfrak{B}^*$ (since $wu \geq wu$ in $\mathfrak{B}^*$). But the alphabet $\mathfrak{B}$ is totally ordered, and thus Corollary 6.6 (applied to $\mathfrak{B}$ instead of $\mathfrak{A}$) yields that $uw \geq wu$ in $\mathfrak{B}^*$ (since we know that Corollary 6.6 holds in the total-order setting). Thus, $uw > wu$ in $\mathfrak{B}^*$ (since $uw \neq wu$).

In other words, $wu < uw$ in $\mathfrak{B}^*$.

Now, let us forget that we fixed $\mathfrak{B}$. We thus have shown that $wu < uw$ in $\mathfrak{B}^*$ for every linear extension $\mathfrak{B}$ of $\mathfrak{A}$. But Proposition 11.31(b) (applied to $wu$ and $uw$ instead of $u$ and $v$) yields that $wu < uw$ holds in $\mathfrak{A}^*$ if and only if

$$(wu < uw$ in $\mathfrak{B}^*$ for every linear extension $\mathfrak{B}$ of $\mathfrak{A})$$

Thus, we conclude that $wu < uw$ holds in $\mathfrak{A}^*$ (since we already know that $wu < uw$ in $\mathfrak{B}^*$ for every linear extension $\mathfrak{B}$ of $\mathfrak{A}$). Thus, $uw > wu$ in $\mathfrak{A}^*$; hence, $uw \geq wu$ in $\mathfrak{A}^*$. This proves Corollary 6.6 in the partial-order setting.

Next, let us give a new proof of the fact that Theorem 6.20 holds in the partial-order setting:

**Alternative proof of Theorem 6.20 in the partial-order setting.** We need to prove the equivalence $\mathcal{A} \iff \mathcal{B} \iff \mathcal{C} \iff \mathcal{D}$. We can prove the implications $\mathcal{A} \implies \mathcal{B}$, $\mathcal{A} \implies \mathcal{C}$, $\mathcal{A} \implies \mathcal{D}$ and $\mathcal{B} \implies \mathcal{A}$ in the same way as we did in the total-order setting. Hence, in order to prove Theorem 6.20, it will be enough to prove the implications $\mathcal{C} \implies \mathcal{B}$ and $\mathcal{D} \implies \mathcal{A}$.

**Proof of the implication $\mathcal{C} \implies \mathcal{B}$:** Assume that Assertion $\mathcal{C}$ holds.

---

723 Proof. Assume the contrary. Thus, we don’t have $v_k \geq u_k$ in $\mathfrak{A}$. Hence, Proposition 11.30(d) (applied to $P = \mathfrak{A}^*$, $a = v_k$ and $b = u_k$) yields that there exists a linear extension $\mathfrak{Q}$ of $\mathfrak{A}$ such that we have $v_k < u_k$ in $\mathfrak{Q}$. Let $\mathfrak{B}$ be such a linear extension. Thus, $\mathfrak{B}$ is a linear extension of $\mathfrak{A}$ such that we have $v_k < u_k$ in $\mathfrak{B}$. We have $u < v$ in $\mathfrak{B}^*$ (by (11.943)).

Every $j \in \{1, 2, \ldots, k-1\}$ satisfies $u_j = v_j$ (by (11.944), applied to $i = j$). In other words, every $j \in \{1, 2, \ldots, k-1\}$ satisfies $v_j = u_j$. Also, we have $k \in \{1, 2, \ldots, m\} = \{1, 2, \ldots, m\} \cap \{\ell(v), \ell(u)\}$ (since $m = \min \{\ell(v), \ell(u)\}$). Altogether, we thus have shown that $k \in \{1, 2, \ldots, m\} \cap \{\ell(v), \ell(u)\}$ satisfies $(v_k < u_k$ in $\mathfrak{B}$, and every $j \in \{1, 2, \ldots, k-1\}$ satisfies $v_j = u_j$). Thus, there exists an $i \in \{1, 2, \ldots, \ell(v), \ell(u)\}$ such that $(v_i < u_i$ in $\mathfrak{B}$, and every $j \in \{1, 2, \ldots, i-1\}$ satisfies $v_j = u_j$ (namely, $i = k$). Hence,

- **either** there exists an $i \in \{1, 2, \ldots, \ell(v), \ell(u)\}$ such that $(v_i < u_i$ in $\mathfrak{B}$, and every $j \in \{1, 2, \ldots, i-1\}$ satisfies $v_j = u_j$),
- **or** the word $u$ is a prefix of $v$.

In other words, $v < u$ in $\mathfrak{B}^*$ (by the definition of the relation $<$ on $\mathfrak{B}^*$). This contradicts the fact that $u < v$ in $\mathfrak{B}^*$. This contradiction proves that our assumption was wrong, qed.
Let \( \mathcal{B} \) be any linear extension of the alphabet \( \mathcal{A} \). Then, \( \mathcal{B}^* \) is an extension of \( \mathcal{A}^* \) (according to Proposition 11.31(a)). It is easy to see that Assertion \( C \) (with \( \mathcal{A} \) replaced by \( \mathcal{B} \)) holds\(^{724}\). But we can apply Theorem 6.20 to \( \mathcal{B} \) instead of \( \mathcal{A} \) (since \( \mathcal{B} \) is totally ordered). As a consequence, we see that the four assertions \( A, B, C \) and \( D \) (all with \( \mathcal{A} \) replaced by \( \mathcal{B} \)) are equivalent. In particular, Assertion \( C \) (with \( \mathcal{A} \) replaced by \( \mathcal{B} \)) is equivalent to Assertion \( B \) (with \( \mathcal{A} \) replaced by \( \mathcal{B} \)). Since Assertion \( C \) (with \( \mathcal{A} \) replaced by \( \mathcal{B} \)) holds, we thus conclude that Assertion \( B \) (with \( \mathcal{A} \) replaced by \( \mathcal{B} \)) holds. In other words,

\[(11.945) \quad \text{any two nonempty words } u \in \mathcal{B}^* \text{ and } v \in \mathcal{B}^* \text{ satisfying } w = uv \text{ satisfy } v > w \text{ in } \mathcal{B}^*.\]

Now, let \( u \in \mathcal{A}^* \) and \( v \in \mathcal{A}^* \) be two nonempty words satisfying \( w = uv \). It is easy to see that

\[(11.946) \quad (w < v \text{ in } \mathcal{B}^* \text{ for every linear extension } \mathcal{B} \text{ of } \mathcal{A})\]

But Proposition 11.31(b) (applied to \( w \) instead of \( u \)) yields that \( w < v \) holds in \( \mathcal{A}^* \) if and only if we have

\[(w < v \text{ in } \mathcal{B}^* \text{ for every linear extension } \mathcal{B} \text{ of } \mathcal{A}).\]

Consequently, we conclude that \( w < v \) holds in \( \mathcal{A}^* \) (since we know that we have \( w < v \) in \( \mathcal{B}^* \) for every linear extension \( \mathcal{B} \) of \( \mathcal{A} \)). In other words, \( v > w \) in \( \mathcal{A}^* \).

Now, let us forget that we fixed \( u \) and \( v \). We have thus shown that any nonempty words \( u \in \mathcal{A}^* \) and \( v \in \mathcal{A}^* \) satisfying \( w = uv \) satisfy \( v > w \) in \( \mathcal{A}^* \). In other words, Assertion \( B \) holds. Thus, we have proven the implication \( C \Rightarrow B \).

Proof of the implication \( D \Rightarrow B \): The proof of the implication \( D \Rightarrow B \) is analogous to our above proof of the implication \( C \Rightarrow B \), and thus left to the reader.

The proof of Theorem 6.20 in the partial-order setting is complete. \( \square \)

The examples of Theorem 6.20 and Corollary 6.6 should have illustrated how Proposition 11.31 allows deriving certain facts about the partial-order setting from the corresponding facts about the total-order setting. This trick, however, has its limits. For example, in the total-order setting, the fact that any Lyndon word \( w \) of length \( > 1 \) can be written in the form \( w = uv \) for two Lyndon words \( u \) and \( v \) satisfying \( u < w < v \) is a consequence of Theorem 6.30. But in the partial-order setting, it is not clear how to derive it from the total-order setting, although it is true in the partial-order setting (and follows from Exercise 6.31).

(d) If \( \mathcal{A} \) is the partially ordered alphabet \( \{1, 2\} \) with no relations whatsoever (i.e., a 2-element antichain), and \( w \) is the word 12, then \( w \) does satisfy \( (w \leq t^n, \text{ then } w \leq t) \) for every nonempty word \( t \) and every positive integer \( n \), but \( w \) is not Lyndon.

[Remark: One direction of Exercise 6.22 does hold in the partial-order setting: Namely, if \( w \) is Lyndon, then every nonempty word \( t \) and every positive integer \( n \) satisfy \( (w \leq t^n, \text{ then } w \leq t) \). The proof of this is the same as in the total-order setting.]

(e) The following statement is clearly equivalent to Exercise 6.22 in the total-order setting, while still being valid in the partial-order setting:

Proposition 11.32. Let \( w \) be a nonempty word. Then, \( w \) is Lyndon if and only if every nonempty word \( t \) and every positive integer \( n \) satisfy (if \( w > t^n \), then \( w > t \)).

Proof of Proposition 11.32 in the partial-order setting. Let us first assume that \( w \) is Lyndon. We shall prove that

\[(11.947) \quad (\text{every nonempty word } t \text{ and every positive integer } n \text{ satisfy } (w > t, \text{ then } w > t^n))).\]

Let \( t \) be a nonempty word, and let \( n \) be a positive integer. Assume that \( w > t \). We need to prove that \( w > t^n \).

\(^{724}\)Proof. Let \( u \in \mathcal{B}^* \) and \( v \in \mathcal{B}^* \) be two nonempty words satisfying \( w = uv \). We have \( u \in \mathcal{B}^* = \mathcal{A}^* \) and \( v \in \mathcal{B}^* = \mathcal{A}^* \), and thus \( u \in \mathcal{A}^* \) and \( v \in \mathcal{A}^* \) are two nonempty words satisfying \( w = uv \). Hence, we have \( v > u \) in \( \mathcal{A}^* \) (since we assumed that Assertion \( C \) holds). Thus, \( u < v \) in \( \mathcal{A}^* \), so that \( u \neq v \) and \( u \leq v \) in \( \mathcal{A}^* \).

So we have \( u \leq v \) in \( \mathcal{A}^* \). Hence, \( u \leq v \) in \( \mathcal{B}^* \) (since \( \mathcal{B}^* \) is an extension of \( \mathcal{A}^* \)). Since \( u \neq v \), this becomes \( u < v \) in \( \mathcal{B}^* \). In other words, \( v > u \) in \( \mathcal{B}^* \).

Now, let us forget that we fixed \( u \) and \( v \). We have thus shown that any nonempty words \( u \in \mathcal{B}^* \) and \( v \in \mathcal{B}^* \) satisfying \( w = uv \) satisfy \( v > u \) in \( \mathcal{B}^* \).

\(^{725}\)Proof. Let \( \mathcal{B} \) be a linear extension of \( \mathcal{A} \). Then, \( \mathcal{B}^* \) is an extension of \( \mathcal{A}^* \) (according to Proposition 11.31(a)). Thus, \( \mathcal{B}^* = \mathcal{A}^* \) as sets, so that \( u \in \mathcal{B} = \mathcal{A}^* \) and \( v \in \mathcal{B} = \mathcal{A}^* \). Now, \( (11.945) \) yields \( v > w \) in \( \mathcal{B}^* \). In other words, \( w < v \) in \( \mathcal{B}^* \). This proves \( (11.946) \).
Assume the contrary. Thus, we don’t have \( w > t^m \). Hence, there exists an \( i \in \{1, \ldots, n\} \) such that we don’t have \( w > t^i \) (namely, \( i = n \)). Let \( m \) be the minimal such \( i \). Thus, \( m \in \{1, \ldots, n\} \), and we don’t have \( w > t^m \). Hence, \( m \neq 1 \) (since we don’t have \( w > t^m \), but we do have \( w > t = t^1 \)). Thus, \( m \geq 2 \), so that \( m - 1 \) is also an element of \( \{1, \ldots, n\} \). If we did not have \( w > t^{m-1} \), then \( m - 1 \) would therefore be an \( i \in \{1, \ldots, n\} \) such that we don’t have \( w > t^i \). But this would contradict the fact that \( m \) is the minimal such \( i \). Thus, we must have \( w > t^{m-1} \). In other words, \( t^{m-1} < w \).

Notice that \( m - 1 \geq 1 \) (since \( m \geq 2 \)). Hence, the word \( t^{m-1} \) is nonempty (since \( t \) is nonempty). In other words, \( t^{m-1} \neq \emptyset \).

Recall that we don’t have \( w > t^m \). It is now easy to see that we don’t have \( t^m \leq w \) \(^{726}\). In other words, we don’t have \( t^{m-1} \leq w \emptyset \) (since \( t^m = t^{m-1} t \) and \( w = w \emptyset \)).

But recall that \( t^{m-1} < w \). Hence, Proposition 6.2(d) (applied to \( a = t^{m-1} \), \( b = t \), \( c = w \) and \( d = \emptyset \)) yields that either we have \( t^{m-1} t \leq w \emptyset \) or the word \( t^{m-1} \) is a prefix of \( w \). Thus, the word \( t^{m-1} \) is a prefix of \( w \) (since we don’t have \( t^{m-1} t \leq w \emptyset \)). In other words, there exists a \( v \in \mathcal{A}^* \) such that \( w = t^{m-1} v \). Consider this \( v \). We have \( v \neq \emptyset \) (because otherwise, we would have \( v = \emptyset \) and thus \( w = t^{m-1} v = t^{m-1} \), contradicting the fact that \( w > t^{m-1} \), so that \( v \) is nonempty. Hence, Proposition 6.14(b) (applied to \( u = t^{m-1} \)) now yields \( v > t^{m-1} \). Hence, \( t^{m-1} < v \). Thus, Proposition 6.2(b) (applied to \( a = t^{m-1} \), \( c = t^{m-1} \) and \( d = v \)) yields \( t^{m-1} t^{m-1} \leq t^{m-1} v = w \). Hence,

\[
\begin{align*}
    w &\geq t^{m-1} t^{m-1} = t^2(m-1) = t^{m-m^2} = t^m t^{m-2} \quad \text{(this makes sense since} \ m \geq 2) \\
    &\geq t^m \quad \text{(since} \ t^m \text{is a prefix of} t^{m-m^2} \).
\end{align*}
\]

This contradicts the fact that we don’t have \( w > t^m \). This contradiction proves that our assumption (that we don’t have \( w > t^m \)) was false. Hence, \( w > t^m \).

Forget now that we assumed that \( w > t \). We thus have proven that if \( w > t \), then \( w > t^m \).

Now, forget that we fixed \( t \) and assumed that \( w \) is Lyndon. We thus have shown that

\[(11.948) \quad \text{(if} w \text{is Lyndon, then} (11.947) \text{holds).}\]

Now, conversely, assume that \( (11.947) \) holds. We will prove that \( w \) is Lyndon.

Let \( u \) and \( v \) be any nonempty words satisfying \( w = uv \). We have \( w \neq u \) \(^{727}\). Combined with \( w = uv \geq u \) (since \( u \) is a prefix of \( w \)), this yields \( w > u \). But \((11.947) \) (applied to \( t = u \) and \( n = 2 \)) yields that if \( w > u \), then \( w > u^2 \). Hence, \( w > u^2 \) (since we know that \( w > u \)). Thus, \( uv = w > u^2 = uu \), so that \( uu < uw \). Hence, Proposition 6.2(c) (applied to \( a = u \), \( c = u \) and \( d = v \)) yields \( u \leq v \). Since \( u \neq v \) \(^{728}\), this becomes \( u < v \), so that \( v > u \).

Now, let us forget that we fixed \( u \) and \( v \). We thus have proven that any nonempty words \( u \) and \( v \) satisfying \( w = uv \) satisfy \( v > u \). In other words, Assertion \( C \) of Theorem 6.20 holds. Hence, Assertion \( A \) of Theorem 6.20 holds as well (since Theorem 6.20 yields that these Assertions \( C \) and \( A \) are equivalent). In other words, the word \( w \) is Lyndon.

Now, forget that we fixed \( w \). We thus have proven that

\[(11.947) \quad \text{(if} (11.947) \text{holds, then} w \text{is Lyndon).}\]

Combined with \((11.948)\), this yields that \( w \) is Lyndon if and only if \((11.947)\) holds. This proves Proposition 11.32.

Thus, we have salvaged Exercise 6.22 in the partial-order setting. (And, as a side effect, we have obtained an alternative solution for Exercise 6.22 in the total-order setting.)

---

\(^{726}\) **Proof.** Assume the contrary. Thus, we have \( t^m \leq w \). Hence, \( \geq t^m \), so that \( w = t^m \) (since we don’t have \( w > t^m \)). Thus, \( w = t^m = t^{m-1} \). Since, \( t \) is a suffix of \( w \). Thus, Corollary 6.15 (applied to \( t \) instead of \( v \)) yields \( t \geq w \). Thus, \( t \geq w = t^m = t^{m-1} \). Combined with \( t^{m-1} \geq t \) (since \( t \) is a prefix of \( t^{m-1} \)), this yields \( t^{m-1} = t \). Thus, \( t^{m-1} = t = t \emptyset \). Cancelling \( t \) from this equality, we obtain \( t^{m-1} = \emptyset \). This contradicts \( t^{m-1} \neq \emptyset \). This contradiction proves that our assumption was wrong, qed.

\(^{727}\) **Proof.** Assume the contrary. Then, \( w = u \). Hence, \( w = w = uu = u \emptyset \). Cancelling \( u \) from this equality, we obtain \( v = \emptyset \), so that \( v \) is empty. This contradicts the fact that \( v \) is nonempty. This contradiction proves that our assumption was wrong, qed.

\(^{728}\) **Proof.** Assume the contrary. Then, \( u = v \). Hence, \( w = uw = v = uu = u^2 \). This contradicts \( w > u^2 \). This contradiction proves that our assumption was wrong, qed.
Remark: Speaking of salvaging, there is a little piece of Proposition 6.19 which can be salvaged for the partial-order case:

Proposition 11.33. Let $w$ be a nonempty word. Let $v$ be a (lexicographically) minimal nonempty suffix of $w$. Then, there exists a $u \in \mathcal{A}^*$ such that we don’t have $uw < vu$.

Proof of Proposition 11.33 in the partial-order case. We know that $v$ is a suffix of $w$. Hence, there exists a $u \in \mathcal{A}^*$ such that $w = uv$. Consider this $u$. It will clearly be enough to show that we don’t have $uw < vu$.

So let us prove that we don’t have $uw < vu$. Indeed, assume the contrary. Then, $uw < vu$. Thus, there exists at least one suffix $t$ of $u$ such that $tw < vt$ (namely, $t = u$). Let $p$ be the minimum-length such suffix. Then, $pv < vp$. Thus, $p$ is nonempty. In other words, $p \neq \emptyset$.

Since $p$ is a suffix of $u$, it is clear that $pv$ is a suffix of $uw = w$. So we know that $pv$ is a nonempty suffix of $w$. Since $v$ is the smallest such suffix, this yields that we don’t have $pv < v$. Hence, we don’t have $pv < v$.

But $pv \emptyset = pv < vp$. Hence, Proposition 6.2(e) (applied to $a = pv$, $b = \emptyset$, $c = v$ and $d = p$) yields that either we have $pv \leq v$ or the word $v$ is a prefix of $pv$. Thus, the word $v$ is a prefix of $pv$ (since we don’t have $pv \leq v$). In other words, there exists a $q \in \mathcal{A}^*$ such that $pv = vq$. Consider this $q$. This $q$ is nonempty (because otherwise we would have $pv = vq = v$, contradicting the fact that $p$ is nonempty).

From $vq = pv < vp$, we obtain $q \leq p$ (by Proposition 6.2(c), applied to $a = v$, $c = q$ and $d = p$).

We know that $q$ is a suffix of $pv$ (since $vq = pv$), whereas $pv$ is a suffix of $w$. Thus, $q$ is a suffix of $w$. So $q$ is a nonempty suffix of $w$. Since $v$ is the smallest such suffix, this yields that we don’t have $q < v$. But we have $pv \emptyset = p \leq pv < vp$. Hence, Proposition 6.2(e) (applied to $a = p$, $b = \emptyset$, $c = v$ and $d = p$) yields that either we have $p \leq v$ or the word $v$ is a prefix of $p$. From this, it is easy to obtain that $v$ is a prefix of $p$.

In other words, there exists an $r \in \mathcal{A}^*$ such that $p = vr$. Consider this $r$. Clearly, $r$ is a suffix of $p$, while $p$ is a suffix of $w$; therefore, $r$ is a suffix of $u$. Also, $pv < vp$ rewrites as $vrw < vvr$ (because $p = vr$). Thus, Proposition 6.2(c) (applied to $a = v$, $c = rc$ and $d = vr$) yields $rv \leq v$. Since $rv \neq vr$ (because otherwise, we would have $rv = vr$, thus $v_r = vrw$, contradicting $vrw < vvr$), this becomes $rv < vr$.

Now, $r$ is a suffix of $u$ such that $rv < vr$. Since $p$ is the minimum-length such suffix, this yields $\ell(r) \geq \ell(p)$. But this contradicts the fact that $\ell(p) = \ell(vr) = \ell(v) + \ell(r) = \ell(r) > \ell(r)$. This contradiction proves our assumption wrong. Thus, we have proven that $uw < vu$. This completes the proof of Proposition 11.33 in the partial-order case.

(f) Let us first establish a lemma which plays a role similar to that of Lemma 6.28 in the total-order setting:

Lemma 11.34. Let $(a_1,a_2,\ldots,a_k)$ be a Hazewinkel-CFL factorization of a nonempty word $w$ (in the partial-order setting). Let $p$ be a suffix of $w$ such that $p$ is Lyndon. Then, $p \geq a_k$.

Proof of Lemma 11.34. We will prove Lemma 11.34 by induction over the (obviously) positive integer $k$.

Induction base: Assume that $k = 1$. By the definition of a Hazewinkel-CFL factorization, $(a_1,a_2,\ldots,a_k)$ is a tuple of Lyndon words satisfying $w = a_1a_2\cdots a_k$ (since $(a_1,a_2,\ldots,a_k)$ is a Hazewinkel-CFL factorization

729By this, we mean that:

1. $v$ is a nonempty suffix of $w$;
2. no nonempty suffix $s$ of $w$ satisfies $s < v$.

Such a $v$ always exists (since $w$ is nonempty), but is not always unique.

730Proof. Assume the contrary. Thus, $pv \leq v$. This yields that $pv = v$ (since we don’t have $pv < v$). Thus, $pw = v = \emptyset v$. Cancelling $v$ from this equation, we obtain $p = \emptyset$. This contradicts $p \neq \emptyset$. This contradiction proves that our assumption was wrong. QED.

731Proof. Assume the contrary. Thus, $v$ is not a prefix of $p$. Hence, $p \leq v$ (since either we have $p \leq v$ or the word $v$ is a prefix of $p$). Now, $q \leq p \leq v$. Now, $q = v$ (since $q \leq v$ but not $q < v$). Hence, $v = q \leq p$. Combined with $p \leq v$, this yields $v = p$. Hence, $p, v = p, v = wp$, which contradicts $pv < vp$. This contradiction proves that our assumption was wrong. QED.
of $w$). Thus, $w = a_1a_2 \cdots a_k = a_1$ (since $k = 1$), so that $w$ is a Lyndon word (since $a_1$ is a Lyndon word). But $p$ is nonempty (since $p$ is Lyndon). Thus, Corollary 6.15 (applied to $v = p$) yields $p \geq w = a_1 = a_k$ (since $1 = k$). Thus, Lemma 11.34 is proven in the case $k = 1$. The induction base is complete.

**Induction step:** Let $K$ be a positive integer. Assume (as the induction hypothesis) that Lemma 11.34 is proven for $k = K$. We now need to show that Lemma 11.34 holds for $k = K + 1$.

So let $(a_1, a_2, \ldots, a_{K+1})$ be a Hazewinkel-CFL factorization of a nonempty word $w$. Let $p$ be a nonempty suffix of $w$ such that $p$ is Lyndon. We need to prove that $p \geq a_{K+1}$.

The tuple $(a_1, a_2, \ldots, a_{K+1})$ is a Hazewinkel-CFL factorization of $w$. By the definition of a Hazewinkel-CFL factorization, this yields that $(a_1, a_2, \ldots, a_{K+1})$ is a tuple of Lyndon words such that $w = a_1a_2 \cdots a_{K+1}$ and such that no $i \in \{1, 2, \ldots, K\}$ satisfies $a_i < a_{i+1}$.

Let $w' = a_2a_3 \cdots a_{K+1}$; then, $w = a_1a_2 \cdots a_{K+1} = a_1(a_2a_3 \cdots a_{K+1}) = a_1w'$. Hence, every nonempty suffix of $w$ is either a nonempty suffix of $w'$, or has the form $qw'$ for a nonempty suffix $q$ of $a_1$. Since $p$ is a nonempty suffix of $w$, we thus must be in one of the following two cases:

**Case 1:** The word $p$ is a nonempty suffix of $w'$.

**Case 2:** The word $p$ has the form $qw'$ for a nonempty suffix $q$ of $a_1$.

Let us first consider Case 1. In this case, $p$ is a nonempty suffix of $w'$. The $K$-tuple $(a_2, a_3, \ldots, a_{K+1})$ of Lyndon words satisfies $w' = a_2a_3 \cdots a_{K+1}$ and has the property that no $i \in \{1, 2, \ldots, K-1\}$ satisfies $a_{i+1} < a_{i+1+1}$ 732. Therefore, $(a_2, a_3, \ldots, a_{K+1})$ is a Hazewinkel-CFL factorization of $w'$. We can thus apply Lemma 11.34 to $K$, $w'$ and $(a_2, a_3, \ldots, a_{K+1})$ instead of $w$ and $(a_1, a_2, \ldots, a_k)$ (because we assumed that Lemma 11.34 is proven for $k = K$). As a result, we obtain that $p \geq a_{K+1}$. Thus, $p \geq a_{K+1}$ is proven in Case 1.

Let us now consider Case 2. In this case, $p$ has the form $qw'$ for a nonempty suffix $q$ of $a_1$. Consider this $q$. Since $a_1$ is a Lyndon word, we have $q \geq a_1$ (by Corollary 6.15, applied to $a_1$ and $q$ instead of $w$ and $v$).

Our goal, however, is to prove that $q \geq a_{K+1}$.

We will show that

$$q \geq a_i \quad \text{for every } i \in \{1, 2, \ldots, K+1\}. \quad (11.949)$$

**Proof of (11.949):** We will prove (11.949) by induction over $i$.

**Induction base:** We know that $q \geq a_1$. In other words, (11.949) holds for $i = 1$. This completes the induction base.

**Induction step:** Let $j \in \{1, 2, \ldots, K\}$. Assume that (11.949) holds for $i = j$. We must prove that (11.949) holds for $i = j+1$.

We have $j \leq K$. Hence, the product $a_{j+1}a_{j+2} \cdots a_{K+1}$ contains at least one factor. Also, the factors of this product are nonempty words (because the words $a_{j+1}, a_{j+2}, \ldots, a_{K+1}$ are nonempty (since these words are Lyndon)). Hence, $a_{j+1}a_{j+2} \cdots a_{K+1}$ is a nonempty product of nonempty words. Consequently, $a_{j+1}a_{j+2} \cdots a_{K+1}$ is a word.

The product $a_{j+2}a_{j+3} \cdots a_{K+1}$ is well-defined (because $j \leq K$). Denote this product by $g$. Thus, $g = a_{j+2}a_{j+3} \cdots a_{K+1}$.

We have $q \geq a_j$ (since (11.949) holds for $i = j$). But

$$p = q \cdot \frac{w'}{a_2a_3 \cdots a_{K+1}} = q(a_2a_3 \cdots a_j)(a_{j+1}a_{j+2} \cdots a_{K+1}) = (q(a_2a_3 \cdots a_j))(a_{j+1}a_{j+2} \cdots a_{K+1}).$$

Consequently, $a_{j+1}a_{j+2} \cdots a_{K+1}$ is a suffix of $p$. Thus, Corollary 6.15 (applied to $p$ and $a_{j+1}a_{j+2} \cdots a_{K+1}$ instead of $w$ and $v$) yields $a_{j+1}a_{j+2} \cdots a_{K+1} \geq p$ (since $p$ is Lyndon). Thus,

$$a_{j+1}a_{j+2} \cdots a_{K+1} \geq p = qw' \geq q \geq a_j.$$
In other words,

\[ a_j \leq a_{j+1}a_{j+2} \cdots a_{K+1} = a_{j+1}\left(a_{j+2}a_{j+3} \cdots a_{K+1}\right) = a_{j+1}g. \]

Hence, \( a_j \varnothing = a_j \leq a_{j+1}g \). Proposition 6.2(e) (applied to \( a_j, \varnothing, a_{j+1}\) and \( g \) instead of \( a, b, c \) and \( d \)) thus yields that either we have \( a_j \leq a_{j+1} \) or the word \( a_{j+1} \) is a prefix of \( a_j \). From this, it is easy to conclude that the word \( a_{j+1} \) is a prefix of \( a_j \).

\[ \text{Consequently, } a_{j+1} \leq a_j, \text{ so that } a_j \geq a_{j+1} \text{ and } g \geq a_j \geq a_{j+1}. \]

In other words, (11.949) holds for \( i = j + 1 \). This completes the induction step. Thus, (11.949) is proven by induction.

Now, (11.949) (applied to \( i = K + 1 \)) yields \( q \geq a_{K+1} \). Hence, \( p = qw' \geq q \geq a_{K+1} \). Thus, \( p \geq a_{K+1} \) is proven in Case 2.

We have now proven \( p \geq a_{K+1} \) in all cases. This proves that Lemma 11.34 holds for \( k = K + 1 \). The induction step is thus finished, and with it the proof of Lemma 11.34. \( \square \)

We can now conclude the solution of Exercise 6.33(f) by proving the following proposition:

**Proposition 11.35.** Let \( w \) be a word (in the partial-order setting). Then, there exists a unique Hazewinkel-CFL factorization of \( w \).

The proof of this proposition is almost a literal adaptation of the proof of Theorem 6.27:

**Proof of Proposition 11.35.** Let us first prove that there exists a Hazewinkel-CFL factorization of \( w \).

Indeed, there clearly exists a tuple \((a_1, a_2, \ldots, a_k)\) of Lyndon words satisfying \( w = a_1a_2 \cdots a_k \). Fix such a tuple with minimum \( k \). We claim that no \( i \in \{1, 2, \ldots, k, \ldots, K-1\} \) satisfies \( a_i < a_{i+1} \).

Indeed, if some \( i \in \{1, 2, \ldots, k-1\} \) would satisfy \( a_i < a_{i+1} \), then the word \( a_ia_{i+1} \) would be Lyndon (by Proposition 6.16(a), applied to \( u = a_i \) and \( v = a_{i+1} \)), whence \((a_1, a_2, \ldots, a_i, a_{i+1}, a_{i+2}, a_{i+3}, \ldots, a_k)\) would also be a tuple of Lyndon words satisfying \( w = a_1a_2 \cdots a_i(a_ia_{i+1})a_{i+2}a_{i+3} \cdots a_k \) but having length \( k-1 < k \), contradicting the fact that \( k \) is the minimum length of such a tuple. Hence, no \( i \in \{1, 2, \ldots, k-1\} \) can satisfy \( a_i < a_{i+1} \). Thus, \((a_1, a_2, \ldots, a_k)\) is a Hazewinkel-CFL factorization of \( w \), so we have shown that such a Hazewinkel-CFL factorization exists.

It remains to show that there exists at most one Hazewinkel-CFL factorization of \( w \). We shall prove this by induction over \( \ell(w) \). Thus, we fix a word \( w \) and assume that

\[ (11.950) \quad \text{for every word } v \text{ with } \ell(v) < \ell(w), \text{ there exists at most one Hazewinkel-CFL factorization of } v. \]

We now have to prove that there exists at most one Hazewinkel-CFL factorization of \( w \).

Indeed, let \((a_1, a_2, \ldots, a_k)\) and \((b_1, b_2, \ldots, b_m)\) be two Hazewinkel-CFL factorizations of \( w \). We need to prove that \((a_1, a_2, \ldots, a_k) = (b_1, b_2, \ldots, b_m)\). If \( w \) is empty, then this is obvious, so we WLOG assume that it is not; thus, \( k > 0 \) and \( m > 0 \).

The tuple \((a_1, a_2, \ldots, a_k)\) is a Hazewinkel-CFL factorization of \( w \). Thus, \((a_1, a_2, \ldots, a_k)\) is a tuple of Lyndon words satisfying \( w = a_1a_2 \cdots a_k \) and such that no \( i \in \{1, 2, \ldots, k-1\} \) satisfies \( a_i < a_{i+1} \).

The tuple \((b_1, b_2, \ldots, b_m)\) is a Hazewinkel-CFL factorization of \( w \). Thus, \((b_1, b_2, \ldots, b_m)\) is a tuple of Lyndon words satisfying \( w = b_1b_2 \cdots b_m \) and such that no \( i \in \{1, 2, \ldots, m-1\} \) satisfies \( b_i < b_{i+1} \). Now, \( b_m \) is a suffix of \( w \) (since \( w = b_1b_2 \cdots b_m \)). Also, \( b_m \) is Lyndon. Thus, Lemma 11.34 (applied to \( p = b_m \)) yields \( b_m \geq a_k \). The same argument (but with the roles of \((a_1, a_2, \ldots, a_k)\) and \((b_1, b_2, \ldots, b_m)\) switched) shows that \( a_k \geq b_m \). Combined with \( b_m \geq a_k \), this yields \( a_k = b_m \). Now let \( v = a_1a_2 \cdots a_k-1 \). Then, \((a_1, a_2, \ldots, a_k-1)\) is a tuple of Lyndon words satisfying \( v = a_1a_2 \cdots a_k-1 \) and such that no \( i \in \{1, 2, \ldots, (k-1)\} \) satisfies \( a_i < a_{i+1} \) (because no \( i \in \{1, 2, \ldots, K\} \) satisfies \( a_i < a_{i+1} \)). In other words, \((a_1, a_2, \ldots, a_k-1)\) is a Hazewinkel-CFL factorization of \( v \).

But \( w = a_1a_2 \cdots a_k = a_1a_2 \cdots a_{k-1}a_k = vb_m, \) so that

\[ vb_m = w = b_1b_2 \cdots b_m = b_1b_2 \cdots b_{m-1}b_m. \]

\[ \square \]

**Proof.** Assume the contrary. Thus, the word \( a_{j+1} \) is not a prefix of \( a_j \). Hence, we have \( a_j \leq a_{j+1} \) (because we know that either we have \( a_j \geq a_{j+1} \) or the word \( a_{j+1} \) is a prefix of \( a_j \)). Since we cannot have \( a_j < a_{j+1} \) (because no \( i \in \{1, 2, \ldots, K\} \) satisfies \( a_i < a_{i+1} \)), this yields that we have \( a_j = a_{j+1} \). Therefore, \( a_{j+1} \) is a prefix of \( a_j \); this contradicts our assumption that the word \( a_{j+1} \) is not a prefix of \( a_j \). This contradiction proves that our assumption was wrong. \( \square \)

**For instance,** the tuple \((w_1, w_2, \ldots, w_{\ell(w)})\) of one-letter words is a valid example (recall that one-letter words are always Lyndon).
Cancelling \( b_m \) from this equality yields \( v = b_1 b_2 \cdots b_{m-1} \). Thus, \((b_1, b_2, \ldots, b_{m-1})\) is a tuple of Lyndon words satisfying \( v = b_1 b_2 \cdots b_{m-1} \) and such that no \( i \in \{1, 2, \ldots, (m-1) - 1\} \) satisfies \( b_i < b_{i+1} \) (since no \( i \in \{1, 2, \ldots, m-1\} \) satisfies \( b_i < b_{i+1} \)). In other words, \((b_1, b_2, \ldots, b_{m-1})\) is a Hazewinkel-CFL factorization of \( v \). Since \( \ell(v) < \ell(w) \) (because \( v = a_1 a_2 \cdots a_{k-1} \) is shorter than \( w = a_1 a_2 \cdots a_k \)), we can apply (11.950) to obtain that there exists at most one Hazewinkel-CFL factorization of \( v \). But we already know two such Hazewinkel-CFL factorizations: \((a_1, a_2, \ldots, a_{k-1})\) and \((b_1, b_2, \ldots, b_{m-1})\). Thus, \((a_1, a_2, \ldots, a_k) = (b_1, b_2, \ldots, b_{m-1})\). Combining this with \( a_k = b_m \), we obtain \((a_1, a_2, \ldots, a_k) = (b_1, b_2, \ldots, b_m)\). This is exactly what we needed to prove. So we have shown (by induction) that there exists at most one Hazewinkel-CFL factorization of \( w \). This completes the proof of Proposition 11.35.

\[
\text{(g) Solution to Exercise 6.32 in the partial-order setting. Let us first observe the following fact (in the partial-order setting): If } h \text{ is a nonempty word which is not a Lyndon word, then}
\]

\[
(11.951) \quad \text{there exist nonempty words } u \text{ and } v \text{ such that } h = uv \text{ and not } vu > w.
\]

\[
\text{Renaming } u \text{ and } v \text{ as } p \text{ and } q \text{ in this result, we obtain the following: If } h \text{ is a nonempty word which is not a Lyndon word, then}
\]

\[
(11.952) \quad \text{there exist nonempty words } p \text{ and } q \text{ such that } h = pq \text{ and not } qp > pq.
\]

Furthermore, if \( h \) is a nonempty word which is not a Lyndon word, then

\[
(11.953) \quad \text{there exist nonempty words } u \text{ and } v \text{ such that } h = uv \text{ and not } v > u.
\]

Now, let us come to the solution of Exercise 6.32 in the partial-order setting.

\[
\text{Solution to Exercise 6.32(a) in the partial-order setting. The implication } A' \implies D' \text{ can be proven in the same way as it was proven in the total-order setting.}
\]

Let us now prove the implication \( D' \implies A' \):

\[
\text{Proof of the implication } D' \implies A': \text{ Assume that Assertion } D' \text{ holds. Thus, if } u \text{ and } v \text{ are nonempty words satisfying } w = uv, \text{ then we have } vu \geq wu.
\]

We need to prove that Assertion \( A' \) holds, i.e., that \( w \) is a power of a Lyndon word. Assume the contrary. Thus, \( w \) is not a power of a Lyndon word; hence, \( w \) is not a Lyndon word itself. Consequently, (11.951) (applied to \( h = w \)) shows that there exist nonempty words \( u \) and \( v \) such that \( w = uv \) and not \( vu > w \). Consider such a pair of nonempty words \( u \) and \( v \) with minimum \( \ell(u) \). The minimality of \( \ell(u) \) shows that

\[
(11.954) \quad \text{if } u' \text{ and } v' \text{ are nonempty words such that } w = u'v' \text{ and not } v'u' > u'v', \text{ then } \ell(u') \geq \ell(u).
\]

We have \( vu \geq uv \) (according to Assertion \( D' \)) but not \( vu > w \). Thus, \( vu = uv \). Therefore, Proposition 6.4 yields that there exist a \( t \in \mathbb{R^*} \) and two nonnegative integers \( n \) and \( m \) such that \( u = t^n \) and \( v = t^m \).

Consider this \( t \) and these \( n \) and \( m \). We have \( n \neq 0 \) (since \( t^n = u \) is nonempty) and \( m \neq 0 \) (since \( t^m = v \) is nonempty) what we needed to prove. So we have shown (by induction) that there exists at most one Hazewinkel-CFL factorization of \( w \). This completes the proof of Proposition 11.35.

\[\text{\underline{Proof of (11.951):}} \text{ Let } h \text{ be a nonempty word which is not a Lyndon word. We need to prove that (11.951) holds.}
\]

In fact, assume the contrary. Then, there exist no nonempty words \( u \) and \( v \) such that \( h = uv \) and not \( vu > w \). In other words, any nonempty words \( u \) and \( v \) satisfying \( h = uv \) must satisfy \( vu > w \). In other words, Assertion \( D \) of Theorem 6.20 (with \( h \) instead of \( w \)) holds.

But we know (from Exercise 6.33(c)) that Theorem 6.20 holds in the partial-order setting. Hence, we can apply Theorem 6.20 to \( h \) instead of \( w \). We thus conclude that Assertions \( A, B, C \) and \( D \) of Theorem 6.20 (with \( h \) instead of \( w \)) are equivalent. Hence, Assertion \( A \) of Theorem 6.20 (with \( h \) instead of \( w \)) holds (since Assertion \( D \) of Theorem 6.20 (with \( h \) instead of \( w \)) holds). In other words, the word \( h \) is Lyndon. This contradicts the fact that the word \( h \) is not Lyndon. This contradiction shows that our assumption was wrong. Hence, (11.951) is proven, qed.

\[\text{\underline{Proof of (11.953):}} \text{ Let } h \text{ be a nonempty word which is not a Lyndon word. We need to prove that (11.953) holds.}
\]

In fact, assume the contrary. Then, there exist no nonempty words \( u \) and \( v \) such that \( h = uv \) and not \( v > u \). In other words, any nonempty words \( u \) and \( v \) satisfying \( h = uv \) must satisfy \( v > u \). In other words, Assertion \( C \) of Theorem 6.20 (with \( h \) instead of \( w \)) holds.

But we know (from Exercise 6.33(c)) that Theorem 6.20 holds in the partial-order setting. Hence, we can apply Theorem 6.20 to \( h \) instead of \( w \). We thus conclude that Assertions \( A, B, C \) and \( D \) of Theorem 6.20 (with \( h \) instead of \( w \)) are equivalent. Hence, Assertion \( A \) of Theorem 6.20 (with \( h \) instead of \( w \)) holds (since Assertion \( C \) of Theorem 6.20 (with \( h \) instead of \( w \)) holds). In other words, the word \( h \) is Lyndon. This contradicts the fact that the word \( h \) is not Lyndon. This contradiction shows that our assumption was wrong. Hence, (11.953) is proven, qed.
nonempty), and the word \( t \) is nonempty (since \( t^n = u \) is nonempty). Moreover, we have \( n = 1 \) \(^{737} \). Hence, \( t^n = t^1 = t \), so that \( u = t^n = t \) and \( w = \frac{u}{v^0} = tt^m = t^{m+1} \). We shall now prove that the word \( t \) is Lyndon.

Assume the contrary. Then, \( t \) is not Lyndon. Hence, \((11.952) \) (applied to \( h = t \)) shows that there exist nonempty words \( p \) and \( q \) such that \( t = pq \) and not \( qp > pq \). Consider these \( p \) and \( q \). Since \( q \) is nonempty, we have \( \ell (q) > 0 \), so that \( \ell \left( \frac{u}{v^0 = pq} \right) = \ell (pq) = \ell (p) + \ell (q) > \ell (p) \).

We have \( w = \frac{u}{v^0 = pq} = p = pq = \frac{(pq)^m}{pq} \) and \( w = pq < (pq)^{m+1} \). This contradicts the fact that we do not have \( qp > pq \). This contradiction shows that our assumption (that \( t \) is not Lyndon) was wrong. Hence, \( t \) is Lyndon. Thus, \( w \) is a power of a Lyndon word (since \( w = t^{m+1} \) is a power of \( t \)). Thus, Assertion \( \mathcal{A}' \) is satisfied, so we have proven the implication \( \mathcal{D}' \implies \mathcal{A}' \).

\(^{738} \)Proof. Assume the contrary. Hence, \( n \neq 1 \). Combined with \( n \neq 0 \), this leads to \( n \geq 2 \). As a consequence, \( u = t^n \) can be rewritten as \( u = tt^{n-1} \). The word \( t^{n-1+m} \) is nonempty (since \( t \) is nonempty and since \( n-1+m \geq 2-1+0 = 1 \)). Now,

\[
\begin{align*}
\frac{u}{v^0 = pq} = tt^{n-1+m} = tt^{n-1+m}.
\end{align*}
\]

Hence, \((11.954) \) (applied to \( u' = t \) and \( v' = t^{n-1+m} \)) yields \( \ell (t) \geq \ell (u) \). Since \( t \) and \( t^{n-1+m} \) are nonempty. Since \( u = t^n \), this rewrites as \( \ell (t) = \ell (t^n) = \frac{n}{t} \ell (t) \geq 2 \ell (t) \), whence \( \ell (t) = 0 \), which contradicts the fact that \( t \) is nonempty. This contradiction shows that our assumption was wrong, qed.

\(^{737} \)Proof. Assume the contrary. Hence, \( n \neq 1 \). Combined with \( n \neq 0 \), this leads to \( n \geq 2 \). As a consequence, \( u = t^n \) can be rewritten as \( u = tt^{n-1} \). The word \( t^{n-1+m} \) is nonempty (since \( t \) is nonempty and since \( n-1+m \geq 2-1+0 = 1 \)). Now,

\[
\begin{align*}
\frac{u}{v^0 = pq} = tt^{n-1+m} = tt^{n-1+m}.
\end{align*}
\]

Hence, \((11.954) \) (applied to \( u' = t \) and \( v' = t^{n-1+m} \)) yields \( \ell (t) \geq \ell (u) \). Since \( t \) and \( t^{n-1+m} \) are nonempty. Since \( u = t^n \), this rewrites as \( \ell (t) = \ell (t^n) = \frac{n}{t} \ell (t) \geq 2 \ell (t) \), whence \( \ell (t) = 0 \), which contradicts the fact that \( t \) is nonempty. This contradiction shows that our assumption was wrong, qed.

\(^{740} \)Proof. Assume the contrary. Then, we do not have \( qr > pq \). We can thus apply \((11.954) \) to \( u' = p \) and \( v' = qv \). As a result, we obtain \( \ell (p) \geq \ell (u) \). This contradicts \( \ell (u) > \ell (p) \). This contradiction shows that our assumption was wrong, qed.

\(^{741} \)Proof. Assume the contrary. Thus, we don’t have \( pq < q \). But \( (pq)^m = (pq)^{m+1} < (pq)^{m+1} = (pq)^m \) and \( \ell (pq) = \ell (p) + \ell (q) = \ell (q) + \ell (p) = \ell (qp) \). Hence, Proposition \( 6.2(7) \) (applied to \( a = p \), \( b = (pq)^m \), \( c = qp \) and \( d = (qp)^m \)) yields \( pq \leq q \). Thus, \( pq = q \) (since we don’t have \( pq < q \)). Taking both sides of this equality to the \( (m+1) \)-th power, we obtain \( (pq)^{m+1} = (qp)^{m+1} \), which contradicts \( (pq)^{m+1} < (qp)^{m+1} \). This contradiction proves that our assumption was wrong, qed.
Now we have proven both implications \( A' \implies D' \) and \( D' \implies A' \). Therefore, the equivalence \( A' \iff D' \) follows. Thus, Exercise 6.32(a) is solved in the partial-order case.

Solution to Exercise 6.32(b) in the partial-order setting. Consider the letter \( m \) and the alphabet \( \mathfrak{A} \cup \{m\} \) defined in Assertion \( F'' \). We notice that the lexicographic order on \( \mathfrak{A}^* \) is the restriction of the lexicographic order on \( (\mathfrak{A} \cup \{m\})^* \) to \( \mathfrak{A}^* \). Therefore, when we have two words \( p \) and \( q \) in \( \mathfrak{A}^* \), statements like \( p < q \) do not depend on whether we are regarding \( p \) and \( q \) as elements of \( \mathfrak{A}^* \) or as elements of \( (\mathfrak{A} \cup \{m\})^* \). It is easy to see that the one-letter word \( m \) satisfies

\[(11.955) \quad m > p \quad \text{for every } p \in \mathfrak{A}^*.\]

The implications \( B' \implies E' \), \( C' \implies E' \), \( G' \implies H' \), \( F'' \implies B' \), \( F'' \implies C' \) and \( F' \implies B' \) can be proven in the same way as they were proven in the total-order setting. We shall now prove some further implications.

Proof of the implication \( E' \implies F'' \): Assume that Assertion \( E' \) holds.

Assume (for the sake of contradiction) that the word \( w m \in (\mathfrak{A} \cup \{m\})^* \) is not a Lyndon word. Clearly, this word \( w m \) is nonempty. Thus, (11.953) (applied to \( h = w m \)) shows that there exist nonempty words \( u \) and \( v \) in \( (\mathfrak{A} \cup \{m\})^* \) such that \( w m = w v \) and not \( v > u \). Denote these two nonempty words \( u \) and \( v \) by \( u' \) and \( v' \). Then, \( u \) and \( v' \) are nonempty words in \( (\mathfrak{A} \cup \{m\})^* \) such that \( w m = w v' \) and not \( v' > u \).

The word \( v' \) is a proper suffix of \( w m \) (since \( w m = w v' \) and since \( w \) is nonempty). Hence, \( v' \) is a nonempty suffix of \( w m \). Thus, we must be in one of the following two cases (depending on whether this suffix begins before the suffix \( m \) of \( w m \) begins or afterwards):

Case 1: The word \( v' \) is a nonempty suffix of \( m \). (Note that \( v' = m \) is allowed.)

Case 2: The word \( v' \) has the form \( v m \) where \( v \) is a nonempty proper suffix of \( w \).

Let us consider Case 1 first. In this case, the word \( v' \) is a nonempty suffix of \( m \). Since the only nonempty suffix of \( m \) is \( m \) itself (because \( m \) is a one-letter word), this yields \( v' = m \). Now, \( w m = u v' = u m \).

Cancelling \( m \) from this equality, we obtain \( w = u \). But \( v' = m > w \) (by (11.955)), applied to \( p = w \), thus \( v' > w = u \). This contradicts the fact that we don’t have \( v' > u \). Thus, we have obtained a contradiction in Case 1.

Let us now consider Case 2. In this case, the word \( v' \) has the form \( v m \) where \( v \) is a nonempty proper suffix of \( w \). Consider this \( v \). We have \( w m = u v' = u m \). By cancelling \( m \) from this equality, we obtain \( w = u v \).

Thus, \( u \) and \( v \) are subwords of \( w \), and therefore belong to \( \mathfrak{A}^* \) (since \( w \in \mathfrak{A}^* \)). Moreover, \( u \) and \( v \) are nonempty. Hence, Assertion \( E' \) yields that either we have \( v \geq u \) or the word \( v \) is a prefix of \( w \). Since we cannot have \( v \geq u \) (because if we had \( v \geq u \), then we would have \( v' = v m \geq v \geq u \), which would contradict the fact that we don’t have \( v' > u \)), we therefore must have that \( v \) is a prefix of \( w \). In other words, there exists \( q \in \mathfrak{A}^* \) such that \( w = v q \). Consider this \( q \). We have \( m > q \) (by (11.955)), applied to \( p = q \). Thus, \( q \leq m \). Hence, Proposition 6.2(b) (applied to \( \mathfrak{A} \cup \{m\} \), \( u \), \( q \) and \( m \) instead of \( \mathfrak{A} \), \( a \), \( c \) and \( d \)) yields \( v q \leq v m \).

Therefore, \( v m \geq v q = w \) (since \( w = v q \)), so that \( v' = v m \geq w = w v > u \) (since \( v \) is nonempty). This contradicts the fact that we don’t have \( v' > u \). Thus, we have found a contradiction in Case 2.

We have therefore obtained a contradiction in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this shows that we always get a contradiction. Hence, our assumption (that the word \( u m \in (\mathfrak{A} \cup \{m\})^* \) is not a Lyndon word) was false. Hence, the word \( w m \in (\mathfrak{A} \cup \{m\})^* \) is a Lyndon word. That is, Assertion \( F'' \) holds. Hence, we have proven the implication \( E' \implies F'' \).

Proof of the implication \( B' \implies G' \): Assume that Assertion \( B' \) holds.

Let \( s \) be the longest suffix \( v \) of \( w \) which does not satisfy \( v \geq w \). (This is well-defined, because there exists a suffix \( v \) of \( w \) which does not satisfy \( v \geq w \) — namely, the empty word.) So we know that \( s \) is a suffix of \( w \) which does not satisfy \( v \geq w \). In other words, \( s \) is a suffix of \( w \) and does not satisfy \( s \geq v \). Hence, \( s \neq w \) (because otherwise, we would have \( s = w \) and thus \( s \geq v \); but this would contradict the fact that \( s \) does not satisfy \( s \geq w \)). Thus, \( s \) is a proper suffix of \( w \) (because \( s \) is a suffix of \( w \) and satisfies \( s \neq w \)). Hence, there exists a nonempty word \( h \in \mathfrak{A}^* \) satisfying \( w = h s \). Consider this \( h \). Using Assertion \( B' \), it is easy to see that \( s \) is a prefix of \( w \). In other words, there exists a \( g \in \mathfrak{A}^* \) such that \( w = s g \). Consider this \( g \). We have

\[742\] This can be proven in the same way as it was proven in the total-order setting.

\[743\] Proof. Assume the contrary. Thus, \( s \) is not a prefix of \( w \). Hence, \( s \) is nonempty. Therefore, Assertion \( B' \) (applied to \( u = h \) and \( v = s \)) yields that either we have \( s \geq w \) or the word \( s \) is a prefix of \( w \). Since \( s \) is not a prefix of \( w \), we must thus have
it is clear that the word $q$ is nonempty. Thus, $s < sg = w$.

We know that $s$ is the longest suffix of $w$ which does not satisfy $v \geq w$. Hence, 

\[(11.956) \quad \text{(if $v$ is a suffix of $w$ which does not satisfy $v \geq w$, then $\ell(v) \leq \ell(s)$).}
\]

There exists a nonnegative integer $m$ such that $h^m$ is a prefix of $s$ (for example, the nonnegative integer $m = 0$). Consider the maximal such integer $m$. Then, $h^m$ is a prefix of $s$, but $h^{m+1}$ is not a prefix of $s$. Since $h^m$ is a prefix of $s$, there exists a word $q \in \mathfrak{A}^*$ such that $s = h^mq$. Consider this $q$. Clearly, $w = h^m \underbrace{q}_{=h^m} = \underbrace{hh^m}_{=h^{m+1}} = h^{m+1}q = h^m hq$. Hence, $h^m hq = w = s$. Cancelling $h^m$ from this equality, we obtain $hq = qg$. It is now easy to see that we don’t have $h \leq q$.

Thus, Proposition 6.2(e) (applied to $h$, $\varnothing$, $q$ and $g$ instead of $a$, $b$, $c$ and $d$) yields that either we have $h \leq q$ or the word $q$ is a prefix of $h$. Thus, the word $q$ is a prefix of $h$ (since we don’t have $h \leq q$).

Next, we shall prove that the word $h$ is Lyndon.

In fact, assume the contrary. Then, $h$ is not Lyndon. Hence, (11.951) shows that there exist nonempty words $u$ and $v$ such that $h = uv$ and not $vu > uv$. Consider these $u$ and $v$. Since $w = h^m \underbrace{s}_{=uv} = uv = s$, it is clear that the word $vs$ is a suffix of $w$. If this suffix $vs$ would not satisfy $vs \geq w$, then we could therefore obtain $\ell(vs) \leq \ell(s)$ (by (11.956), applied to $vs$ instead of $v$), which would contradict $\ell(v) = \ell(s)$. Thus, the suffix $vs$ must satisfy $vs \geq w$. We thus have $vs \geq w = hs \geq h = uv$ (since $v$ is nonempty). We thus have $uv$, so that $wv \leq vs$.

Recall that $s < w$. Hence, $vs \leq wv$ (by Proposition 6.2(b), applied to $v$, $s$ and $w$ instead of $a$, $c$ and $d$). Now, $wv\varnothing = wv \leq wv < v w = \underbrace{v}_{=hs} \underbrace{s}_{=uv} = uv = s = vs = vu s$. Since $wv \leq vs$, we therefore must have $vu = wv$. Thus, the elements $u$ and $v$ of the monoid $\mathfrak{A}^*$ commute. Thus, the submonoid of $\mathfrak{A}^*$ generated by $u$ and $v$ is commutative. Since $h = uv$, the element $h$ lies in this submonoid, and therefore the element $h^m$ lies in it as well. Thus, $h^m$ commutes with $v$ (since this submonoid is commutative), i.e., we have $vh^m = h^m v$. Thus, $u = h^m \underbrace{q}_{=h^m v} = \underbrace{vh^m}_{=h^m q} = q = h^m hq$, so that $h^m hq \leq h^m vq$. Hence, Proposition 6.2(c) (applied to $h^m$, $hq$ and $vq$ instead of $a$, $c$ and $d$) yields $hq \leq vq$. But since $q$ is a prefix of $h$, there exists a precursor of $q$ which does not satisfy $s$; this contradicts the fact that $s$ does not satisfy $s \geq w$. This contradiction shows that our assumption was wrong, qed.

---

\[\text{This is well-defined, because of the following reason:}
\]

We have $\ell(h) \geq 1$ (since the word $h$ is nonempty). Thus, for every $m \in \mathbb{N}$ satisfying $m > \ell(s)$, we have $\ell(h^m) = m \ell(h) \geq 1$.

$m > \ell(s)$. In other words, for every $m \in \mathbb{N}$ satisfying $m > \ell(s)$, the word $h^m$ is longer than $s$. Hence, for every $m \in \mathbb{N}$ satisfying $m > \ell(s)$, the word $h^m$ cannot be a prefix of $s$. Thus, for every sufficiently high $m \in \mathbb{N}$, the word $h^m$ cannot be a prefix of $s$.

We thus know the following:

- There exists a nonnegative integer $m$ such that $h^m$ is a prefix of $s$.
- For every sufficiently high $m \in \mathbb{N}$, the word $h^m$ cannot be a prefix of $s$.

Consequently, there exists a maximal nonnegative integer $m$ such that $h^m$ is a prefix of $s$, qed.

---

\[\text{Proof. Assume the contrary. Then, $h \leq q$. Hence, $h \leq q \leq hq = hq$ (since $hq = qg$). Therefore, Proposition 6.2(g) (applied to $h, q$ and $q$ instead of $a, b$ and $c$) yields that $h$ is a prefix of $q$. In other words, there exists a word $r \in \mathfrak{A}^*$ such that $q = hr$. Consider this $r$. Now, $s = h^m \underbrace{q}_{=hr} = \underbrace{h^m}_{=h^m q} r = h^m hq$, so that $h^m + 1$ is a prefix of $s$. This contradicts the fact that $h^m + 1$ is not a prefix of $s$. This contradiction proves that our assumption was wrong, qed.}\]
word \( z \in \mathfrak{A}^* \) such that \( h = qz \). Consider this \( z \). We have
\[
\ell \left( \frac{h}{u} \right) = \ell(uvq) = \ell(u + \ell(vq)) = \ell(u) + \ell(vq) > \ell(vq), \quad \text{so that} \quad \ell(vq) \leq \ell(hq).
\]
Hence, Proposition 6.2(f) (applied to \( vq \), \( z \), \( hq \) and \( z \) instead of \( a \), \( b \), \( c \) and \( d \)) yields \( vq \leq hq \). Combined with \( hq \leq vq \), this yields \( vq = hq \). Hence, \( \ell \left( \frac{vq}{hq} \right) = \ell(hq) > \ell(vq) \), which is absurd. This contradiction proves that our assumption is wrong. Thus, we have shown that the word \( h \) is Lyndon.

We now know that \( h \in \mathfrak{A}^* \) is a Lyndon word, \( m + 1 \) is a positive integer, and \( q \) is a prefix of \( h \), and we have \( w = h^{m+1}q \). Hence, there exists a Lyndon word \( t \in \mathfrak{A}^* \), a positive integer \( \ell \) and a prefix \( p \) of \( t \) (possibly empty) such that \( w = t^\ell p \) (namely, \( t = h, \ell = m + 1 \) and \( p = q \)). In other words, Assertion \( G' \) holds. This proves the implication \( B' \Rightarrow G' \).

We are now going to prove the implication \( F' \Rightarrow B' \); this implication will later be used in the proof of the implication \( H' \Rightarrow B' \).

Proof of the implication \( H' \Rightarrow B' \): Assume that Assertion \( H' \) holds. In other words, there exists a Lyndon word \( t \in \mathfrak{A}^* \), a nonnegative integer \( \ell \) and a prefix \( p \) of \( t \) (possibly empty) such that \( w = t^\ell p \). Consider this \( t \), this \( \ell \) and this \( p \).

We are going to prove that for every \( m \in \mathbb{N} \)
\[
\begin{align*}
(11.957) & \quad \text{(every suffix } s \text{ of } t^mp \text{ satisfies either } s \geq t^mp \text{ or } (\text{the word } s \text{ is a prefix of } t^mp)).
\end{align*}
\]

Proof of (11.957): We will prove (11.957) by induction over \( m \):

Induction base: Using the implication \( F' \Rightarrow B' \), it is easy to see that (11.957) holds for \( m = 0 \).

This completes the induction base.

Induction step: Let \( M \) be a positive integer. Assume that (11.957) is proven for \( m = M - 1 \). We will now show that (11.957) holds for \( m = M \).

Let \( r \) denote the word \( t^{M-1}p \). It is easy to see that \( r \) is a prefix of \( t^M p \).

In other words, there exists a word \( g \in \mathfrak{A}^* \) such that \( t^M p = rg \). Consider this \( g \).

Let \( s \) be a suffix of \( t^M p \). We shall show that either \( s \geq t^M p \) or (the word \( s \) is a prefix of \( t^M p \)).

\[\text{Proof. Assume that } m = 0. \text{ Then, } t^M p = t^M \frac{p}{g} = p = p. \]

Let \( s \) be a suffix of \( t^M p \). Then, \( s \) is a suffix of \( t^M p = p \). In other words, there exists a word \( g \in \mathfrak{A}^* \) satisfying \( p = gs \). Consider this \( g \).

We are going to prove that either \( s \geq t^M p \) or (the word \( s \) is a prefix of \( t^M p \)). If the word \( s \) is empty, then this is obvious (because if the word \( s \) is empty, then the word \( s \) is a prefix of \( t^M p \)). Hence, we WLOG assume that the word \( s \) is nonempty. If the word \( g \) is empty, then it is also clear that either \( s \geq t^M p \) or (the word \( s \) is a prefix of \( t^M p \)) (because if the word \( g \) is empty, then \( g = \varepsilon \) and thus \( t^M p = p = g = s \), so that the word \( s \) is a prefix of \( t^M p \)). Hence, we WLOG assume that the word \( g \) is nonempty.

But the word \( p \) is a prefix of a Lyndon word in \( \mathfrak{A}^* \) (since \( p \) is a prefix of \( t \), and since \( t \) is a Lyndon word in \( \mathfrak{A}^* \)). In other words, Assertion \( F' \) with \( w \) replaced by \( p \) is satisfied. Hence, Assertion \( B' \) with \( w \) replaced by \( p \) is satisfied as well (since we have already proven the implication \( F' \Rightarrow B' \)). In other words,
\[
(11.958) \quad \left( \begin{array}{c}
\text{if } u \text{ and } v \text{ are nonempty words satisfying } p = uv, \text{ then} \\
\text{either } u \text{ we have } v \geq p \text{ or the word } v \text{ is a prefix of } p
\end{array} \right).
\]

Since the words \( g \) and \( s \) are nonempty, we can apply (11.958) to \( u = g \) and \( v = s \). As a result, we obtain that we have \( s \geq p \) or the word \( s \) is a prefix of \( p \). In other words, either \( s \geq p \) or (the word \( s \) is a prefix of \( p \)). In other words, either \( s \geq t^M p \) or (the word \( s \) is a prefix of \( t^M p \)) (since \( t^M p = p \)). This proves (11.957).

\[\text{Proof. There exists a word } q \in \mathfrak{A}^* \text{ such that } t = pq \text{ (since } p \text{ is a prefix of } t \). \text{ Consider this } q. \text{ We have } t^M = t^M \frac{p}{pq} = t^M \frac{p}{q} = rqp = r(qp). \text{ Hence, } r \text{ is a prefix of } t^M p, \text{ qed.}\]
In order to prove this, let us assume the contrary (for the sake of contradiction). Then, neither $s \geq t^M p$ nor (the word $s$ is a prefix of $t^M p$). In other words, we don’t have $s \geq t^M p$, and the word $s$ is not a prefix of $t^M p$. If the word $s$ was a prefix of $r$, then the word $s$ would be a prefix of $t^M p$ (since $r$ is a prefix of $t^M p$), which would contradict the fact that the word $s$ is not a prefix of $t^M p$. Hence, the word $s$ cannot be a prefix of $r$. In other words, the word $s$ cannot be a prefix of $t^{M-1} p$ (since $r = t^{M-1} p$).

We don’t have $s \geq t^M p$. Since $s = s \emptyset$ and $t^M p = r g$, this rewrites as follows: We don’t have $s \emptyset \geq r g$. In other words, we don’t have $rg \leq s \emptyset$.

The word $s$ is a suffix of $\underbrace{t^M p}_{=t^{M-1}} p = t^{M-1} p = tr$. Therefore, we must be in one of the following two cases (depending on whether this suffix begins before the suffix $r$ of $tr$ begins or afterwards):

**Case 1:** The word $s$ is a suffix of $r$. (Note that $s = r$ is allowed.)

**Case 2:** The word $s$ has the form $s’r$ where $s’$ is a nonempty suffix of $t$.

Let us consider Case 1 first. In this case, the word $s$ is a suffix of $r$. In other words, the word $s$ is a suffix of $t^{M-1} p$ (since $r = t^{M-1} p$). Hence, (11.957) (applied to $m = M - 1$) yields that either $s \geq t^{M-1} p$ or (the word $s$ is a prefix of $t^{M-1} p$) (since (11.957) holds for $m = M - 1$). Since the word $s$ cannot be a prefix of $t^{M-1} p$, we thus must have $s \geq t^{M-1} p$. Thus, $t^{M-1} p \leq s$, so that $r = t^{M-1} p \leq s$. Therefore, Proposition 6.2(d) (applied to $r$, $g$, $s$ and $\emptyset$ instead of $a$, $b$, $c$ and $d$) yields that either we have $rg \leq s \emptyset$ or the word $r$ is a prefix of $s$. Thus, the word $r$ is a prefix of $s$ (since we don’t have $r g \leq s \emptyset$). But $\ell (r) \geq \ell (s)$ (since $s$ is a suffix of $r$). Hence, $r$ is a prefix of $s$ which is at least as long as $s$ itself. Consequently, $r = s$. Hence, $s = r$, so that $s$ is a prefix of $s = r$. This contradicts the fact that the word $s$ cannot be a prefix of $r$. Thus, we have found a contradiction in Case 1.

Let us now consider Case 2. In this case, the word $s$ has the form $s’r$ where $s’$ is a nonempty suffix of $t$. Consider this $s’$. Corollary 6.15 (applied to $t$ and $s’$ instead of $w$ and $v$) yields $s’ \geq t$. That is, $t \leq s’$. But $s’$ is a suffix of $t$, so that $\ell (s’) \leq \ell (t)$. Hence, $\ell (t) \geq \ell (s’)$. Thus, Proposition 6.2(j) (applied to $t$, $s’$ and $r$ instead of $a$, $b$ and $c$) yields $tr \leq s’r$. Hence, $s’r \geq tr$, so that $s = s’r \geq tr = t^M p$. This contradicts the fact that we don’t have $s \geq t^M p$. Thus, we have found a contradiction in Case 2.

We have thus obtained a contradiction in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this shows that we always get a contradiction. This completes the proof that either $s \geq t^M p$ or (the word $s$ is a prefix of $t^M p$).

Now, forget that we fixed $s$. We thus have shown that every suffix $s$ of $t^M p$ satisfies either $s \geq t^M p$ or (the word $s$ is a prefix of $t^M p$). In other words, (11.957) holds for $m = M$. This completes the induction step, and thus (11.957) is proven by induction.

Now, let $u$ and $v$ be nonempty words satisfying $w = uv$. Then, $v$ is a suffix of $w = t^M p$. Hence, (11.957) (applied to $m = \ell$ and $s = v$) yields that either $v \geq t^M p$ or (the word $v$ is a prefix of $t^M p$). In other words, either $v \geq w$ or (the word $v$ is a prefix of $w$) (since $w = t^M p$). In other words, either we have $v \geq w$ or the word $v$ is a prefix of $w$.

Now, forget that we fixed $u$ and $v$. We thus have shown that if $u$ and $v$ are nonempty words satisfying $w = uv$, then either we have $v \geq w$ or the word $v$ is a prefix of $w$. In other words, Assertion $B’$ holds. Thus, the implication $H’ \implies B’$ is proven.

We have thus proven the implications $B’ \implies E’, C’ \implies E’, G’ \implies H’, E’ \implies F’’, F’’ \implies B’, F’ \implies C’, B’ \implies G’$ and $H’ \implies B’$. Combined, these yield the equivalence $B’ \iff C’ \iff E’ \iff F’’ \iff G’ \iff H’$. This solves Exercise 6.32(b) in the partial-order setting.

**Solution to Exercise 6.32(c) in the partial-order setting.** The implication $F’ \implies B’$ has already been proven in our solution of Exercise 6.32(b). Hence, Exercise 6.32(c) is solved in the partial-order setting.

**Solution to Exercise 6.32(d) in the partial-order setting.** The solution of Exercise 6.32(d) in the partial-order setting proceeds precisely as in the total-order setting.

**Solution to Exercise 6.32(e) in the partial-order setting.** The solution of Exercise 6.32(e) in the partial-order setting proceeds precisely as in the total-order setting.

**Solution to Exercise 6.32(f) in the partial-order setting.** Assume that there exists a letter $\mu \in \mathbb{A}$ such that ($\mu > a$ for some letter $a$ of $w$). Consider this $\mu$. We need to prove that the equivalence $F’ \iff F’’$ holds.

Combining the implication $F’ \implies B’$ (which has already been proven) and the implication $B’ \implies F’’$ (which follows from the equivalence $B’ \iff C’ \iff E’ \iff F’’ \iff G’ \iff H’$ proven above), we obtain the implication $F’ \implies F’’$. Thus, in order to prove the equivalence $F’ \iff F’’$, it is enough to verify the implication $F’’ \implies F’$. Let us do this now.
Assume that Assertion $F''$ holds. Due to the implication $F'' \implies G'$ (which follows from the equivalence $B' \iff C' \iff E' \iff F'' \iff G' \iff H'$ proven above), this yields that Assertion $G'$ holds. In other words, there exists a Lyndon word $t \in \mathbb{A}^*$, a positive integer $\ell$ and a prefix $p$ of $t$ (possibly empty) such that $w = t^\ell p$. Consider these $t$, $\ell$ and $p$.

We shall first prove that

\[(11.959) \quad \exists z \in \mathbb{A} \text{ such that } z > t.
\]

**Proof of (11.959):** The word $t$ is Lyndon and thus nonempty. Hence, $\ell(t) \geq 1$. We thus are in one of the following two cases:

**Case 1:** We have $\ell(t) = 1$.

**Case 2:** We have $\ell(t) > 1$.

Let us consider Case 1 first. In this case, we have $\ell(t) = 1$. Thus, $t$ is a one-letter word. In other words, $t = b$ for some letter $b$. Let us consider this $b$.

Recall that $\mu > a$ for some letter $a$ of $w$. Consider this letter $a$. Both $a$ and $b$ are letters of $t$. Since the word $t$ has only one letter (because $\ell(t) = 1$), this yields that $a = b$. Hence, $\mu > a = b = t$ (since $t = b$).

Hence, there exists a letter $z \in \mathbb{A}$ such that $z > t$ (namely, $z = \mu$). Thus, (11.959) is proven in Case 1.

Let us now consider Case 2. In this case, we have $\ell(t) > 1$. Let $g$ be the last letter of the word $t$ (this is well-defined since $t$ is nonempty). Then, the one-letter word $g$ is a suffix of the word $t$, and therefore there exists a word $t' \in \mathbb{A}^*$ such that $t = t'g$. Consider this $t'$. Since $g$ is a one-letter words, we have $\ell(g) = 1$.

Thus, $\ell(t') = \ell(t') + \ell(g) = \ell(t') + 1$, so that $\ell(t') = \ell(t) - 1 > 1 - 1 = 0$. Hence, the word $t'$ is nonempty. Thus, $g$ is a proper suffix of $t$ (since $t = t'g$). Also, $g$ is nonempty (since $\ell(g) = 1$). But recall that the word $t$ is Lyndon. By the definition of a Lyndon word, this yields that every nonempty proper suffix $v$ of $t$ satisfies $v > t$. Applying this to $v = g$, we obtain $g > t$. Thus, there exists a letter $z \in \mathbb{A}$ such that $z > t$ (namely, $z = g$). Thus, (11.959) is proven in Case 2.

We have thus proven (11.959) in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that (11.959) always holds.

Now, (11.959) shows that there exists a letter $z \in \mathbb{A}$ such that $z > t$. Consider this letter $z$. We have $z > t$, thus $t < z$, hence $t \leq z$. The one-letter word $z$ is Lyndon (since every one-letter word is Lyndon). Thus, $t, t, \ldots, t, z$ are $\ell + 2$ Lyndon words (since both $t$ and $z$ are Lyndon) satisfying $t \leq t \leq \cdots \leq t \leq z$.

(since $t \leq z$) and $t < z$. Hence, Exercise 6.23 (applied to $n = \ell + 2$ and $(w_1, w_2, \ldots, w_n) = (t, t, \ldots, t, z)\)\) yields that $t^{\ell+1}z$ is a Lyndon word. In other words, $t^{\ell+1}z$ is a Lyndon word (since $t^{\ell+1}z = t^{\ell+1}$).

But $p$ is a prefix of $t$. In other words, there exists a $q \in \mathbb{A}^*$ such that $t = pq$. Consider this $q$. We then have $t^{\ell+1}z = t^{\ell+1}t^{\ell}z = t^{\ell+1}p quz = wqz = w(\ell+1)z$. Hence, $w$ is a prefix of the word $t^{\ell+1}z$. Thus, $w$ is a prefix of a Lyndon word in $\mathbb{A}^*$ (because $t^{\ell+1}z$ is a Lyndon word in $\mathbb{A}^*$). In other words, Assertion $F'$ holds. This proves the implication $F'' \implies F'$. This solves Exercise 6.32(f) in the partial-order case.

**Remark:** Of course, for a letter $\mu \in \mathbb{A}$, if we have $(\mu > a$ for every letter $a$ of $w$), then we also have $(\mu > a$ for some letter $a$ of $w$) (since $w$ is nonempty and thus has at least one letter). Hence, Exercise 6.32(e) is a particular case of Exercise 6.32(f).

Altogether, we have now solved all parts of Exercise 6.32 in the partial-order case. Exercise 6.33(g) is thus solved.

[Remark: The validity of some parts of Exercise 6.32 in the partial-order case can also be deduced from their validity in the total-order case using Proposition 11.31. For example, the equivalence $B' \iff C' \iff E'$ can be treated this way. We shall not give any details of this alternative approach, however.]
11.15. **Solution to Exercise 6.34.** **Solution to Exercise 6.34.** In the following, whenever $G$ is a group and $P$ is a (left) $G$-set, we denote by $G_u$ the stabilizer of $u$ in $G$, that is, the subgroup $\{g \in G \mid gu = u\}$ of $G$.

Let $C^{(n)}$ denote the subgroup $(c^n)$ of the infinite cyclic group $C$. Then, $C/C^{(n)}$ is a cyclic group with $n$ elements. Hence, $|C/C^{(n)}| = n$.

Also, recall that $c$ acts on $\mathbb{Z}^n$ by cyclically rotating $n$-tuples one step to the left $(c \cdot (a_1, a_2, \ldots, a_n) = (a_2, a_3, \ldots, a_n, a_1))$. Thus, $c^n$ acts trivially on $\mathbb{Z}^n$ (since cyclically rotating an $n$-tuple $n$ steps to the left does nothing). Therefore, the whole subgroup $C^{(n)}$ acts trivially on $\mathbb{Z}^n$ (since $C^{(n)} = \langle c^n \rangle$).

The word “divisor” shall mean “positive divisor” throughout this solution.

(a) Let $N$ be any $n$-necklace. Then, $N$ is an orbit of the $C$-action, thus a nonempty set. Fix an element $w$ of $N$ (such a $w$ exists since $N$ is nonempty).

Recall that the subgroup $C^{(n)}$ acts trivially on $\mathbb{Z}^n$. In particular, $C^{(n)}$ stabilizes $w$. Hence, $C^{(n)} \subseteq C_w$. Thus, $[C : C^{(n)}] = [C : C_w] \cdot [C_w : C^{(n)}]$. As a consequence, $[C : C^{(n)}] = [C : C^{(n)}] = |C/C^{(n)}| = n$, so that $|C : C_w| < \infty$.

But $N$ is an orbit of the $C$-action containing $w$. Hence, $N$ is the $C$-orbit of $w$. Thus, by the orbit-stabilizer theorem, we have $|N| = |C : C_w|$. This yields $|N| = |C : C_w| < \infty$, so that $N$ is a finite set. Also, $|N| = |C : C_w| / n$. This solves Exercise 6.34(a).

(b) Let $w = (w_1, w_2, \ldots, w_n) \in \mathbb{Z}^n$ be an $n$-tuple. We need to prove the equivalence between the following two assertions:

- **Assertion A:** The $n$-necklace $[w]$ is aperiodic.
- **Assertion B:** Every $k \in \{1, 2, \ldots, n-1\}$ satisfies $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$.

We will achieve this by proving the implications $\mathcal{A} \implies \mathcal{B}$ and $\mathcal{B} \implies \mathcal{A}$. Let us first do some preparatory work.

Let $N$ denote the $n$-necklace $[w]$. Thus, $w \in N$ and therefore (as we have shown in the solution of Exercise 6.34(a)) we have $C^{(n)} \subseteq C_w$ and $|N| = |C : C_w|$.

We are now ready to prove implications $\mathcal{A} \implies \mathcal{B}$ and $\mathcal{B} \implies \mathcal{A}$.

**Proof of the implication $\mathcal{A} \implies \mathcal{B}$:** Assume that Assertion A holds. In other words, the $n$-necklace $[w]$ is aperiodic.

In other words, $N$ is aperiodic (since $N = [w]$). In other words, the period of $N$ is $n$ (by the definition of “aperiodic”). In other words, $|N|$ is $n$ (since the period of $N$ is defined as $|N|$). In other words, $|N| = n$. Thus, $n = |N| = |C : C_w|$. But

$$n = [C : C^{(n)}] = [C : C_w] \cdot [C_w : C^{(n)}] \quad (\text{since } C^{(n)} \subseteq C_w)$$

$$= n \cdot [C_w : C^{(n)}].$$

Solving this for $[C_w : C^{(n)}]$, we obtain $[C_w : C^{(n)}] = 1$, so that $C^{(n)} = C_w$.

Let now $k \in \{1, 2, \ldots, n-1\}$. We are going to prove that $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$. Indeed, assume the contrary. Thus, $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) = w$.

Recall that $c$ acts on $\mathbb{Z}^n$ by cyclically rotating $n$-tuples one step to the left. Hence, $c^k$ acts on $\mathbb{Z}^n$ by cyclically rotating $n$-tuples $k$ steps to the left. Hence,

$$c^k w = (w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) = w.$$ 

Thus, $c^k \in C_w = C^{(n)} = \langle c^n \rangle$. But if an integer $a \in \mathbb{Z}$ satisfies $c^a \in \langle c^n \rangle$, then we must have $n \mid a$ (this follows from the structure of the infinite cyclic group $C$). Applied to $a = k$, this yields $n \mid k$. But this is absurd, since $k \in \{1, 2, \ldots, n-1\}$. This contradiction proves that our assumption was wrong. Hence, we have proven $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$.

Now, let us forget that we fixed $k$. We thus have shown that every $k \in \{1, 2, \ldots, n-1\}$ satisfies $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$. In other words, Assertion B holds. This proves the implication $\mathcal{A} \implies \mathcal{B}$.

**Proof of the implication $\mathcal{B} \implies \mathcal{A}$:** Assume that Assertion B holds. In other words, every $k \in \{1, 2, \ldots, n-1\}$ satisfies $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$.

We will prove that $C^{(n)} = C_w$. Indeed, assume the contrary. Thus, $C^{(n)} \neq C_w$. Since $C^{(n)} \subseteq C_w$, this yields that $C^{(n)}$ is a proper subset of $C_w$. Hence, there exists some $d \in C_w$ such that $d \notin C^{(n)}$. Consider
this $d$. Write $d$ in the form $d = c^h$ for some $h \in \mathbb{Z}$ (this is possible since $c$ generates $C$). Let $k$ denote the remainder of $h$ modulo $n$. Then, $h - k$ is divisible by $n$, and thus $c^{h-k} \in \langle c^n \rangle = C(n) \subset C_w$, so that $c^{h-k}w = w$. But also, $c^hw = w$ (since $c^h = d \in C_w$). Hence, $w = c^h w = c^k c^{h-k}w = c^k w$.

But $k \in \{0, 1, \ldots, n-1\}$ (since $k$ is a remainder modulo $n$). We have $c^{h-k} \neq c^h$ (since $c^{h-k} \in C(n)$ whereas $c^h \notin C(n)$). Hence, $h - k \neq h$, so that $k \neq 0$. Combined with $k \in \{0, 1, \ldots, n-1\}$, this yields $k \in \{0, 1, \ldots, n-1\} \setminus \{0\} = \{1, 2, \ldots, n-1\}$. Thus, Assertion $B$ yields that $(w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w$.

Recall that $c$ acts on $A^n$ by cyclically rotating $n$-tuples one step to the left. Hence, $c^k$ acts on $A^n$ by cyclically rotating $n$-tuples $k$ steps to the left. Hence,

$$c^k w = (w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) \neq w = c^k w.$$  

This is absurd. This contradiction shows that our assumption was wrong. Hence, we have shown that $\langle c^n \rangle = C_w$. Now, $|N| = \left[ C : \frac{C_w}{\langle c^n \rangle} \right] = \left[ C : C(n) \right] = n$. In other words, the period of $N$ is $n$ (since the period of $N$ is defined as $|N|$). In other words, $N$ is aperiodic (by the definition of “aperiodic”). In other words, $[w]$ is aperiodic (since $N = [w]$). In other words, Assertion $A$ holds. This proves the implication $B \implies A$.

Now we have proven both implications $A \implies B$ and $B \implies A$. Hence, the Assertions $A$ and $B$ are equivalent. Exercise 6.34(b) is solved.

Before we come to the solution of Exercise 6.34(c), let us state a simple lemma:

**Lemma 11.36.** Let $n$ be a positive integer. Assume that the set $A$ is totally ordered. Let $w \in A^n$ be a Lyndon word. Let $k \in \{1, 2, \ldots, n-1\}$. Then, $c^k w > w$ in the lexicographic order.

**Proof of Lemma 11.36.** We have $w = (w_1, w_2, \ldots, w_n)$ (since $w \in A^n$). Let $u = (w_1, w_2, \ldots, w_k)$ and $v = (w_{k+1}, w_{k+2}, \ldots, w_n)$. These words $u$ and $v$ are well-defined and nonempty (since $k \in \{1, 2, \ldots, n-1\}$) and satisfy

$$uv = (w_1, w_2, \ldots, w_k) (w_{k+1}, w_{k+2}, \ldots, w_n) = (w_1, w_2, \ldots, w_k, w_{k+1}, w_{k+2}, \ldots, w_n) = (w_1, w_2, \ldots, w_n) = w.$$  

But the $n$-tuple $c^k w$ is obtained from $w$ by $k$-fold cyclic rotation to the left (since $c$ acts on $A^n$ by cyclically rotating $n$-tuples one step to the left). In other words,

$$c^k w = (w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k) = (w_{k+1}, w_{k+2}, \ldots, w_n) (w_1, w_2, \ldots, w_k) = uv > wv \quad \text{(by Proposition 6.14(c))}$$  

$$= w.$$  

This proves Lemma 11.36. $\square$

(c) Let $N$ be any aperiodic $n$-necklace. We need to show that $N$ contains exactly one Lyndon word. Since the necklace $N$ is aperiodic, we know that the period of $N$ is $n$. In other words, $|N| = n$.

Let $w$ be the lexicographically smallest word contained in $N$. The word $w$ has $n$ letters (since $w \in N \subset A^n$), and thus is nonempty. We will prove that the word $w$ is Lyndon.

Clearly, $N$ is a $C$-orbit (since $N$ is an $n$-necklace), and thus $N$ is the orbit of the word $w$ (since $w \in N$). In other words, $N = Cw$.

We can see (as in the solution of Exercise 6.34(a)) that $|N| = [C : C_w]$. Now, $n = [C : C(n)] = \left[ \frac{C : C(n)}{C : C_w} \right] \left[ \frac{C_w : C(n)}{C_w : C_w} \right] = n \cdot \left[ C_w : C(n) \right]$. Thus, $[C_w : C(n)] = 1$, so that $C_w = C(n)$.

Indeed, let $u$ and $v$ be two nonempty words satisfying $w = uv$. We will prove that $vuv > wv$.

Assume the contrary. Thus, $vuv \leq wv$.

We have $w = (w_1, w_2, \ldots, w_n)$ (since $w \in N \subset A^n$). Thus, there exists some $k \in \{0, 1, \ldots, n\}$ such that $u = (w_1, w_2, \ldots, w_k)$ and $v = (w_{k+1}, w_{k+2}, \ldots, w_n)$ (since $w = uv$). Consider this $k$. We have $0 < k < n$ (since $u$ and $v$ are nonempty). Since $v = (w_{k+1}, w_{k+2}, \ldots, w_n)$ and $u = (w_1, w_2, \ldots, w_k)$, we have $vuv = (w_{k+1}, w_{k+2}, \ldots, w_n, w_1, w_2, \ldots, w_k)$. In other words, the $n$-tuple $vuv$ is obtained from $w$ by $k$-fold cyclic rotation to the left. In yet other words, $vuv = c^k w$ (since $c$ acts on $A^n$ by cyclically rotating $n$-tuples one step to the left). However, $vuv > wv$ (by Lemma 11.36), so that $uv > wv$, which contradicts our assumption. Therefore, $vuv > wv$, as desired.
step to the left). Hence, $vu = \ell C^k w \in Cw = N$. Thus, $vu \geq w$ (since $w$ is the lexicographically smallest word contained in $N$). Combined with $vu \leq uw = w$, this yields $vu = w$. Hence, $\ell C^k w = vu = w$, so that $\ell C^k$ stabilizes $w$. In other words, $\ell C^k \subseteq C^\alpha(w)$. But this is impossible, since $0 < k < n$ (and since $C^\alpha(n)$ is the subgroup $(C^n)$ of $C$). This contradiction proves that our assumption was wrong. Thus, we have proven that $vu > w$.

Let us now forget that we fixed $u$ and $v$. We thus have proven that any nonempty words $u$ and $v$ satisfying $w = uw$ satisfy $vu > w$. In other words, the word $w$ satisfies Assertion $\mathcal{D}$ of Theorem 6.20. Consequently, the word $w$ satisfies Assertion $\mathcal{A}$ of Theorem 6.20 as well (since Theorem 6.20 yields that these two assertions are equivalent); in other words, $w$ is Lyndon. The orbit $N$ thus contains at least one Lyndon word (namely, $w$).

We shall next prove that $w$ is the only Lyndon word in $N$. Indeed, let $p$ be any Lyndon word in $N$ distinct from $w$. We will derive a contradiction.

It is easy to see that there exists a $k \in \{1, 2, \ldots, n - 1\}$ satisfying $p = \ell C^k w$. Consider this $k$. Then, $w = \ell C^{n-k} p$. Notice that $n-k \in \{1, 2, \ldots, n - 1\}$ (since $k \in \{1, 2, \ldots, n - 1\}$) and $p \in N \subset \mathfrak{A}^n$. Hence, Lemma 11.36 (applied to $p$ and $n-k$ instead of $w$ and $k$) yields that $\ell C^{n-k} p > p$. Hence, $p < \ell C^{n-k} p = w$. Since $p \in N$, this shows that there exists an element of $N$ which is lexicographically smaller than $w$ (namely, $p$). This contradicts the fact that $w$ is the lexicographically smallest word contained in $N$.

Now, let us forget that we fixed $w$. We thus have obtained a contradiction for every Lyndon word $p$ in $N$ distinct from $w$. Thus, there exists no Lyndon word $p$ in $N$ distinct from $w$. Hence, $w$ is the only Lyndon word in $N$. Since we already know that $w$ is a Lyndon word in $N$, this yields that $N$ contains exactly one Lyndon word (namely, $w$). This solves Exercise 6.34(c).

(d) Let $N$ be an $n$-necklace which is not aperiodic. We shall prove that $N$ contains no Lyndon word.

Indeed, assume the contrary. Then, $N$ contains a Lyndon word. Let $w$ be this Lyndon word.

But the $n$-necklace $N$ is not aperiodic. Thus, $|N| \neq n$ (since $N$ is aperiodic if and only if $|N| = n$). We can see (as in the solution of Exercise 6.34(a)) that $|N| = |C : Cw|$. Thus, $|C : Cw| = |N| \neq n = |C : C^\alpha(n)|$, so that $Cw \neq C^\alpha(n)$. But $C^\alpha(n) \subset Cw$ (this can be proven just as in the solution of Exercise 6.34(a)). Hence, $C^\alpha(n)$ is a proper subset of $Cw$. Hence, there exists some $e \in Cw$ such that $e \notin C^\alpha(n)$. Consider this $e$.

We have $ew = w$ (since $e \in Cw$). It is now easy to see that there exists a $k \in \{1, 2, \ldots, n - 1\}$ such that $\ell C^k w = w$. Consider this $k$. Lemma 11.36 yields $\ell C^k w > w$ (since $w \in N \subset \mathfrak{A}^n$), which contradicts $\ell C^k w > w$. This contradiction proves that our assumption was false. Thus, $N$ contains no Lyndon word. This solves Exercise 6.34(d).

---

**Proof.** We have $p \in N = Cw$. Hence, there exists some $e \in C$ satisfying $p = ew$. Consider this $e$.

We know that $C^\alpha(n)$ acts trivially on $\mathfrak{A}^n$, and thus stabilizes $w$. Hence, $C^\alpha(n)w = \{w\}$.

By the well-known properties of the infinite cyclic group, we know that the elements $e^{0}, e, e^{1}, \ldots, e^{n-1}$ of $C$ form a system of unique representatives for the cosets of $C^\alpha(n)$ in $C$. In other words, every $d \in C$ belongs to the coset $e^\alpha C^\alpha(n)$ for some $\ell \in \{0, 1, \ldots, n - 1\}$. Applied to $d = e$, this yields that $e$ belongs to the coset $e^\alpha C^\alpha(n)$ for some $\ell \in \{0, 1, \ldots, n - 1\}$. Consider this $\ell$. We have $e \in e^\alpha C^\alpha(n)$, thus $p = e^{\ell} = \ell C^{n-k} p$.

If $\ell = 0$, then $p = \ell C^{n-k} w = 1 Cw = w$, which contradicts the fact that $p$ is distinct from $w$. Hence, we cannot have $\ell = 0$.

**Proof.** We know that $e^{n}$ acts trivially on $\mathfrak{A}^n$. Thus, $\ell C^\alpha w = w$. Hence, $w = e^{\ell} = \ell C^{n-k} p = \ell C^{n-k} p$, qed.

**Proof.** We know that $C^\alpha(n)$ acts trivially on $\mathfrak{A}^n$, and thus stabilizes $w$. Hence, $C^\alpha(n)w = \{w\}$.

By the well-known properties of the infinite cyclic group, we know that the elements $e^{0}, e, e^{1}, \ldots, e^{n-1}$ of $C$ form a system of unique representatives for the cosets of $C^\alpha(n)$ in $C$. In other words, every $d \in C$ belongs to the coset $e^\alpha C^\alpha(n)$ for some $\ell \in \{0, 1, \ldots, n - 1\}$. Applied to $d = e$, this yields that $e$ belongs to the coset $e^\alpha C^\alpha(n)$ for some $\ell \in \{0, 1, \ldots, n - 1\}$. Consider this $\ell$. We have $e \in e^\alpha C^\alpha(n)$, thus $w = e^{\ell} = \ell C^{n-k} p = \ell C^{n-k} p$, so that $w = \ell C^{n-k} p$. Thus, $\ell C^\alpha w = w$.

If $\ell = 0$, then $e \in \ell C^\alpha(n) = 0 C^\alpha(n) = C^\alpha(n)$, which contradicts $e \notin C^\alpha(n)$. Hence, we cannot have $\ell = 0$. Thus, $\ell \in \{0, 1, \ldots, n - 1\} \setminus \{0\} = \{1, 2, \ldots, n - 1\}$. Hence, there exists a $k \in \{1, 2, \ldots, n - 1\}$ satisfying $\ell C^k w = w$ (namely, $k = \ell$), qed.
(c) If \( w \) is a Lyndon word of length \( n \), then \([w]\) is an aperiodic \( n \)-necklace\(^{751}\). Hence, the map

\[
\text{the set of all Lyndon words of length } n \to \text{(the set of all aperiodic } n\text{-necklaces)},
\]

\[
w \mapsto [w]
\]
is well-defined. Denote this map by \( \Phi \). This map \( \Phi \) is injective\(^{752}\) and surjective\(^{753}\). Hence, \( \Phi \) is bijective.

In other words, \( \Phi \) is a bijection between the set of all Lyndon words of length \( n \) and the set of all aperiodic \( n \)-necklaces. Thus, the aperiodic \( n \)-necklaces are in bijection with Lyndon words of length \( n \). This solves Exercise 6.34(e).

Before we start solving Exercise 6.34(f), we state a lemma about words:

**Lemma 11.37.** Let \( N \) be a positive integer, and let \( w \in \mathfrak{A}^N \). Let \( p \) be a positive divisor of \( N \). Assume that \( c^pw = w \). Then, there exists a word \( q \in \mathfrak{A}^p \) such that \( w = q^{N/p} \).

**Proof of Lemma 11.37.** (The following proof is overkill, but it is the simplest proof to formalize.)

Notice that \( N/p \) is a positive integer (since \( p \) is a positive divisor of \( N \)), so that \( N/p - 1 \in \mathbb{N} \).

We have \( w = (w_1, w_2, \ldots, w_N) \) (since \( w \in \mathfrak{A}^N \)). Let \( u = (w_1, w_2, \ldots, w_p) \) and \( v = (w_{p+1}, w_{p+2}, \ldots, w_N) \). (These \( u \) and \( v \) are well-defined since \( p \in \{0,1,\ldots,N\} \).) Then,

\[
\begin{align*}
\Phi (u) &= (w_1, w_2, \ldots, w_p, w_{p+1}, w_{p+2}, \ldots, w_N) \\
\Phi (v) &= (w_1, w_2, \ldots, w_p, w_{p+1}, w_{p+2}, \ldots, w_N) \\
\Phi (w) &= (w_1, w_2, \ldots, w_N)
\end{align*}
\]

On the other hand, \( c \) acts on \( \mathfrak{A}^N \) by cyclically rotating \( N \)-tuples one step to the left. Hence, \( c^p w \) is the result of cyclically rotating the \( N \)-tuple \( w \) to the left \( p \) times. In other words,

\[
c^p w = (w_{p+1}, w_{p+2}, \ldots, w_N, w_1, w_2, \ldots, w_p) = (w_{p+1}, w_{p+2}, \ldots, w_N) (w_1, w_2, \ldots, w_p) = vu = \Phi (w).
\]

Compared with \( c^p w = w = uv \), this yields \( uv = vu \). Hence, Proposition 6.4 yields that there exist a \( t \in \mathfrak{A}^* \) and two nonnegative integers \( n \) and \( m \) such that \( u = t^n \) and \( v = t^m \). Consider this \( t \) and these \( n \) and \( m \).

We have \( u = (w_1, w_2, \ldots, w_p) \), so that \( u \in \mathfrak{A}^p \) and thus \( \ell (u) = p \). Hence, \( p = \ell \left( \frac{u}{t^n} \right) = \ell (t^n) = n \ell (t) \), so that \( n \ell (t) = p \). On the other hand, \( v = (w_{p+1}, w_{p+2}, \ldots, w_N) \), so that \( \ell (v) = N - p \) and thus \( N - p = \ell \left( \frac{v}{t^m} \right) = \ell (t^m) = m \ell (t) \) and thus \( m \ell (t) = N - p \). Now, \( n \ell (t) = p \neq 0 \). Hence, \( n \neq 0 \) and \( \ell (t) \neq 0 \). Now,

---

\(^{751}\) Proof. Let \( w \) be a Lyndon word of length \( n \). Then, \([w]\) is an aperiodic \( n \)-necklace. If \([w]\) was not aperiodic, then the necklace \([w]\) would contain no Lyndon word (by Exercise 6.34(d), applied to \([w]\)) which would contradict the fact that this necklace \([w]\) contains the Lyndon word \( w \). Hence, \([w]\) must be aperiodic, qed.

\(^{752}\) Proof. Let \( u \) and \( u' \) be two Lyndon words of length \( n \) such that \( \Phi (w) = \Phi (w') \). We shall prove that \( w = w' \).

We have \( \Phi (w) = [w] \) (by the definition of \( \Phi \)) and \( \Phi (w') = [w'] \) (similarly). Thus, \([w'] = \Phi (w') = \Phi (w) = [w] \).

The word \( w \) is contained in \([w]\) (obviously), and the word \( w' \) is contained in \([w'] = [w] \). Thus, both \( w \) and \( w' \) are contained in \([w]\).

We know that \([w]\) is an aperiodic \( n \)-necklace. Hence, \([w]\) contains exactly one Lyndon word (by Exercise 6.34(c)). Hence, any two Lyndon words contained in \([w]\) must be identical. Applying this to the two Lyndon words \( w \) and \( w' \) both contained in \([w]\), we obtain that \( w \) and \( w' \) are identical, i.e., we have \( w = w' \).

Let us now forget that we fixed \( w \) and \( w' \). We have thus proven that any two Lyndon words \( w \) and \( w' \) of length \( n \) satisfying \( \Phi (w) = \Phi (w') \) must satisfy \( w = w' \). In other words, the map \( \Phi \) is injective, qed.

\(^{753}\) Proof. Let \( N \) be an aperiodic \( n \)-necklace. Hence, \( N \) contains exactly one Lyndon word (by Exercise 6.34(c)). Let \( w \) be this Lyndon word. The definition of \( \Phi \) yields \( \Phi (w) = [w] = N \) (since \( N \) is the aperiodic \( n \)-necklace containing \( w \)). Thus, \( N = \Phi (w) \in \text{Im } \Phi \).

Now, let us forget that we fixed \( N \). We have thus shown that \( N \in \text{Im } \Phi \) for every aperiodic \( n \)-necklace \( N \). In other words, the map \( \Phi \) is surjective, qed.
let $$\frac{m}{n} = \frac{m \ell(t)}{n \ell(t)} = \frac{N - p}{p} = \frac{N/p - 1}{1},$$

so that $$m = n \cdot (N/p - 1).$$ Hence, $$t^m = t^{n \cdot (N/p - 1)} = \left(\frac{t^n}{u}\right)^{N/p - 1} = u^{N/p - 1}.$$ Now, $$w = u \underbrace{w}_{= t^m = u^{N/p - 1}} = uu^{N/p - 1} = u^{N/p}.$$ Hence, there exists a word $$q \in \mathbb{A}^p$$ such that $$w = q^{N/p}$$ (namely, $$q = u$$). Lemma 11.37 is proven.

We also recall a lemma about the functions $$\mu$$ and $$\phi:

**Lemma 11.38.** Every positive integer $$n$$ satisfies

\[(11.960) \sum_{d \mid n} \mu(d) = \delta_{n,1};\]

\[(11.961) \sum_{d \mid n} \mu(d) \frac{n}{d} = \phi(n).\]

**Proof of Lemma 11.38.** Both equalities (11.960) and (11.961) have been proven in the solution of Exercise 2.82. (Indeed, (11.960) is (11.337), and (11.961) is (11.339).) $\square$

(f) Exercise 6.34(a) yields that for any $$n$$-necklace $$N$$, we have $$|N| \mid n$$. Hence, for any $$n$$-necklace $$N$$, the cardinality $$|N|$$ is a divisor of $$n$$ (since $$|N| \in \mathbb{N}$$). Now, recall that the $$n$$-necklaces are the orbits of the $$C$$-action on $$\mathbb{A}^n$$, and therefore form a set partition of the set $$\mathbb{A}^n$$. Hence,

$$|\mathbb{A}^n| = \sum_{N \text{ is an } n\text{-necklace}} |N| = \sum_{d \mid n} \sum_{N \text{ is an } n\text{-necklace; } |N|=d} d$$

(because for any $$n$$-necklace $$N$$, the cardinality $$|N|$$ is a divisor of $$n$$). Hence,

$$|\mathbb{A}| = |\mathbb{A}^n| = \sum_{d \mid n} \sum_{N \text{ is an } n\text{-necklace; } |N|=d} d = \sum_{d \mid n} d \left| \{N \text{ is an } n\text{-necklace } \mid |N|=d \} \right|$$

Now, for every positive integer $$e$$, let $$Aper(e)$$ denote the set of all aperiodic $$e$$-necklaces. We shall now prove that

$$|\{N \text{ is an } n\text{-necklace } \mid |N|=n/d \}| = |Aper(n/d)|$$

for every divisor $$d$$ of $$n$$.

**Proof of (11.963):** We first notice that every positive integer $$m$$, every letter $$u$$ and every word $$v \in \mathbb{A}^{m-1}$$ satisfy

$$e(uv) = vu$$

(where we identify the letter $$u$$ with the one-letter word $$(u)$$).

**Proof of (11.964):** Let $$m$$ be a positive integer. Let $$u$$ be a letter. Let $$v \in \mathbb{A}^{m-1}$$ be a word. Then, $$v = (v_1, v_2, \ldots, v_{m-1})$$, so that

$$\underbrace{uv}_{= (u)} = (v_1, v_2, \ldots, v_{m-1}) = (u, v_1, v_2, \ldots, v_{m-1}).$$

But recall that $$c$$ acts on $$\mathbb{A}^n$$ by cyclically rotating $$n$$-tuples one step to the left. Thus,

$$c(u, v_1, v_2, \ldots, v_{m-1}) = (v_1, v_2, \ldots, v_{m-1}, u) = \underbrace{(v_1, v_2, \ldots, v_{m-1})}_{= v} (u) = vu.$$ 

Hence,

$$c(uv) = c(u, v_1, v_2, \ldots, v_{m-1}) = vu.$$
Let \( d \) be a divisor of \( n \). Then, \( n/d \) is a positive integer. We can thus define a map
\[
\Delta : \mathfrak{A}^{n/d} \to \mathfrak{A}^n,
\]
\[
w \mapsto w^d
\]
(where, as we recall, \( w^d \) means the \( d \)-fold concatenation \( w \cdot w \cdots w \) \( d \) times of \( w \) with itself). It is easy to see that this map \( \Delta \) is \( C \)-equivariant (meaning that \( \Delta(gw) = g \cdot \Delta(w) \) for every \( g \in C \) and \( w \in \mathfrak{A}^{n/d} \)) \(^{755}\) and injective\(^{756}\).

Since the map \( \Delta \) is \( C \)-equivariant, it gives rise to a map
\[
\overline{\Delta} : \left( \text{the set of all } C\text{-orbits on } \mathfrak{A}^{n/d} \right) \to \left( \text{the set of all } C\text{-orbits on } \mathfrak{A}^n \right),
\]
\[N \mapsto \Delta(N).
\]
Consider this map \( \overline{\Delta} \). The map \( \overline{\Delta} \) is injective\(^{757}\). Furthermore, \( \overline{\Delta} \) is a map from the set of all \( C \)-orbits on \( \mathfrak{A}^{n/d} \) to the set of all \( C \)-orbits on \( \mathfrak{A}^n \). In other words, \( \overline{\Delta} \) is a map from the set of all \( (n/d) \)-necklaces to the set of all \( n \)-necklaces (since the \( C \)-orbits on \( \mathfrak{A}^{n/d} \) are the \( (n/d) \)-necklaces, and the \( C \)-orbits on \( \mathfrak{A}^n \) are the \( n \)-necklaces). It is now easy to see that
\[
(11.965) \quad \overline{\Delta}(\text{Aper}(n/d)) \subset \{ N \text{ is an } n\text{-necklace } | |N| = n/d \}.
\]
and thus (11.964) is proven.

\(^{755}\)Proof. We need to show that \( \Delta \) is \( C \)-equivariant. Clearly, it is enough to prove that \( \Delta(cw) = c \cdot \Delta(w) \) for every \( w \in \mathfrak{A}^{n/d} \) (since \( c \) generates the group \( C \)). So let us prove this.

Let \( w \in \mathfrak{A}^{n/d} \). Thus, \( w = (w_1, w_2, \ldots, w_{n/d}) \). Let \( \overline{w} \) denote the word \( (w_2, w_3, \ldots, w_{n/d}) \). Then, (identifying the letter \( w_1 \) with the one-letter word \( (w_1) \)) we have
\[
\overline{w}_1 = (w_1, \overline{w}) = (w_1, (w_2, w_3, \ldots, w_{n/d})).
\]
Thus, \( w = w_1 \overline{w} \), so that
\[
cw = c(w_1 \overline{w}) = \overline{w}w_1 \quad \text{(by } (11.964)\text{), applied to } m = n/d, u = w_1 \text{ and } v = \overline{w}.\]

Also, taking both sides of the equality \( w = w_1 \overline{w} \) to the \( d \)-th power, we obtain
\[
\begin{align*}
w^d &= (w_1 \overline{w})^d \cdot \overline{w}w_1 \\
&= (w_1 \overline{w})^d \cdot \overline{w}w_1 \cdot \overline{w}(w_1 \overline{w}) \cdot \overline{w}(w_1 \overline{w}) \cdot \overline{w}(w_1 \overline{w}) = \overline{w}(w_1 \overline{w})^d = \overline{w}w_1 \cdot \overline{w}(w_1 \overline{w})^d \cdot \overline{w}w_1 \cdot \overline{w}(w_1 \overline{w})^d \cdot \overline{w}w_1.
\end{align*}
\]

The definition of \( \Delta(w) \) yields \( \Delta(w) = w^d = w_1 (\overline{w}w_1)^{d-1} \overline{w} \). Applying \( c \) to both sides of this identity, we obtain
\[
c \cdot \Delta(w) = c \cdot (w_1 (\overline{w}w_1)^{d-1} \overline{w}) = (\overline{w}w_1)^{d-1} w_1 \overline{w} \quad \text{(by } (11.964)\text{), applied to } m = n, u = w_1 \text{ and } v = (\overline{w}w_1)^{d-1} \overline{w} \)
\]
\[
= (\overline{w}w_1)^{d-1} (\overline{w}w_1) = (\overline{w}w_1)^{d-1} \overline{w} = (cw)^d.
\]

Compared with \( \Delta(cw) = (cw)^d \) (by the definition of \( \Delta \)), this yields \( \Delta(cw) = c \cdot \Delta(w) \). We thus have shown that \( \Delta \) is \( C \)-equivariant, qed.

\(^{756}\)Proof. We need to prove that \( \Delta \) is injective. In other words, we need to prove that every \( w \in \mathfrak{A}^{n/d} \) can be reconstructed from \( \Delta(w) \). But this is easy: Since \( \Delta(w) = w^d \) (by the definition of \( \Delta \)), the word \( w \) can be obtained from \( \Delta(w) \) by taking the first \( n/d \) letters of \( \Delta(w) \). Hence, \( w \) can be reconstructed from \( \Delta(w) \), qed.

\(^{757}\)Proof. Let \( P \) and \( Q \) be two \( C \)-orbits on \( \mathfrak{A}^{n/d} \) such that \( \overline{\Delta}(P) = \overline{\Delta}(Q) \). We want to show that \( P = Q \).

By the definition of \( \overline{\Delta} \), we have \( \overline{\Delta}(P) = \Delta(P) \) and \( \overline{\Delta}(Q) = \Delta(Q) \). Thus, \( \Delta(P) = \overline{\Delta}(P) = \overline{\Delta}(Q) = \Delta(Q) \).

The orbit \( P \) is nonempty, and thus contains an element. Let \( p \) be such an element.

We have \( \begin{pmatrix} p \\ \in P \end{pmatrix} \in \Delta(P) = \Delta(Q) \). Hence, there exists some \( q \in Q \) such that \( \Delta(p) = \Delta(q) \). Consider this \( q \). We have
\[
\Delta(p) = \Delta(q), \text{ and thus } p = q \text{ (since } \Delta \text{ is injective). The element } p \text{ belongs to } P \cap Q \text{ (since } p \in P \text{ and } p = q \in Q \text{), and thus the two orbits } P \text{ and } Q \text{ have an element in common (namely, } p \text{). But any two orbits which have an element in common must be identical. Thus, } P \text{ and } Q \text{ are identical, i.e., we have } P = Q.\]

Let us forget that we fixed \( P \) and \( Q \). We thus have shown that any two \( C \)-orbits \( P \) and \( Q \) on \( \mathfrak{A}^{n/d} \) which satisfy \( \overline{\Delta}(P) = \overline{\Delta}(Q) \) must satisfy \( P = Q \). In other words, the map \( \overline{\Delta} \) is injective.
But we also have

(11.966) \{ N \text{ is an } n\text{-necklace} \mid |N| = n/d \} \subset \overline{\Delta} \left( \text{Aper} \left( n/d \right) \right).

Proof of (11.966): Let \( P \in \{ N \text{ is an } n\text{-necklace} \mid |N| = n/d \} \). Thus, \( P \) is an \( n\text{-necklace} \) such that \(|P| = n/d\). We shall now prove that \( P \in \overline{\Delta} \left( \text{Aper} \left( n/d \right) \right) \).

The set \( P \) is an \( n\text{-necklace} \), thus a \( C\)-orbit on \( \mathfrak{A}^n \), thus nonempty. Pick some \( w \in P \) (this clearly exists since \( P \) is nonempty). Then \( P \) is the \( C\)-orbit of \( w \). Hence, by the orbit-stabilizer theorem, we have \(|P| = [C : C_w]\). Hence, \([C : C_w] = |P| = n/d\). Hence, \( C_w \) is a subgroup of \( C \) having index \( n/d \). Since the only subgroup of \( C \) having index \( n/d \) is \( \langle c^{n/d} \rangle \), this yields that \( C_w = \langle c^{n/d} \rangle \). Hence, \( C_w = \langle c^{n/d} \rangle \), so that \( c^{n/d} \in \langle c^{n/d} \rangle = C_w \), and thus \( c^{n/d}w = w \). Now, Lemma 11.37 (applied to \( N = n \) and \( p = n/d \)) yields that there exists a word \( q \in \mathfrak{A}^{n/d} \) such that \( w = q^{n/(n/d)} \). Consider this \( q \). We have \( w = q^{n/(n/d)} = q^d \), whereas the definition of \( \Delta \) yields \( \Delta(q) = q^d \). Thus, \( w = q^d = \Delta \left( \frac{q}{\in [q]} \right) \in \Delta ([q]) \). The definition of \( \overline{\Delta} \) yields \( \overline{\Delta} ([q]) = \Delta ([q]) \). Hence, \( w \in \Delta ([q]) = \overline{\Delta} ([q]) \). Thus, \( \overline{\Delta} ([q]) \) is the \( C\)-orbit of \( w \) (since \( \overline{\Delta} ([q]) \) is a \( C\)-orbit on \( \mathfrak{A}^n \) (since \( \overline{\Delta} ([q]) \) is an \( n\text{-necklace} \)). In other words, \( \overline{\Delta} ([q]) \) is the \( C\)-orbit of \( w \). Hence, \( \overline{\Delta} ([q]) = P \) (since \( P \) is the \( C\)-orbit of \( w \)).

But \( P = \overline{\Delta} ([q]) = \Delta ([q]) \) and thus \(|P| = |\Delta ([q])| = |[q]| \) (since \( \Delta \) is injective), so that \(|[q]| = |P| = n/d\).

By the definition of the period of an \((n/d)\)-necklace, we see that the period of the \((n/d)\)-necklace \([q]\) is \(|[q]| = n/d\). In other words, the \((n/d)\)-necklace \([q]\) is aperiodic (by the definition of “aperiodic”). In other words, \( [q] \in \text{Aper} \left( n/d \right) \) (since \( \text{Aper} \left( n/d \right) \) is the set of all aperiodic \((n/d)\)-necklaces).

Let us now forget that we fixed \( P \). We thus have proven that every \( P \in \{ N \text{ is an } n\text{-necklace} \mid |N| = n/d \} \) satisfies \( \overline{\Delta} \left( \text{Aper} \left( n/d \right) \right) \). In other words, \( \{ N \text{ is an } n\text{-necklace} \mid |N| = n/d \} \subset \overline{\Delta} \left( \text{Aper} \left( n/d \right) \right) \). This proves (11.966).

Combining (11.965) with (11.966), we obtain

\[ \overline{\Delta} \left( \text{Aper} \left( n/d \right) \right) = \{ N \text{ is an } n\text{-necklace} \mid |N| = n/d \}. \]

Hence,

\[ |\overline{\Delta} \left( \text{Aper} \left( n/d \right) \right)| = |\{ N \text{ is an } n\text{-necklace} \mid |N| = n/d \}|,
\]

so that

\[ |\{ N \text{ is an } n\text{-necklace} \mid |N| = n/d \}| = |\overline{\Delta} \left( \text{Aper} \left( n/d \right) \right)| = |\text{Aper} \left( n/d \right)|
\]

(since \( \overline{\Delta} \) is injective). This proves (11.963).

---

758 Proof. Let \( M \in \text{Aper} \left( n/d \right) \). Then, \( M \) is an aperiodic \((n/d)\)-necklace (since \( \text{Aper} \left( n/d \right) \) is the set of all aperiodic \((n/d)\)-necklaces). In other words, \( M \) is an \((n/d)\)-necklace whose period is \( n/d \).

The period of \( M \) is defined to be \(|M| \). Thus, \(|M| = n/d \) (since the period of \( M \) is \( n/d \)). In other words, \(|M| = n/d \). But the definition of \( \overline{\Delta} \) yields \( \overline{\Delta} (M) = \Delta (M) \), thus

\[ |\overline{\Delta} (M)| = |\Delta (M)| = |M| \] (since \( \Delta \) is injective)

\[ = n/d. \]

Hence, \( \overline{\Delta} (M) \in \{ N \text{ is an } n\text{-necklace} \mid |N| = n/d \} \) (since \( \overline{\Delta} (M) \) is an \( n\text{-necklace} \) (because \( \overline{\Delta} \) is a map from the set of all \((n/d)\)-necklaces to the set of all \((n/d)\)-necklaces)).

Now, let us forget that we fixed \( M \). We thus have proven that \( \overline{\Delta} (M) \in \{ N \text{ is an } n\text{-necklace} \mid |N| = n/d \} \) for every \( M \in \text{Aper} \left( n/d \right) \). In other words, \( \overline{\Delta} \left( \text{Aper} \left( n/d \right) \right) \subset \{ N \text{ is an } n\text{-necklace} \mid |N| = n/d \}. \) This proves (11.965).

759 This is a particular case of the following elementary fact: If \( e \) is a positive integer, then the only subgroup of \( C \) having index \( e \) is \( \langle c^e \rangle \). (This is because \( C \cong \mathbb{Z}^+ \), and because the only subgroup of \( \mathbb{Z}^+ \) having index \( e \) is \( e\mathbb{Z}^+ \).)
Now, every divisor $d$ of $n$ satisfies

\[ \begin{align*}
\left\{ N \text{ is an } n\text{-necklace} \mid |N| = \frac{d}{n/(n/d)} \right\} & = \{N \text{ is an } n\text{-necklace} \mid |N| = n/(n/d)\} \\
\quad & = \text{Aper} \left( \frac{n/(n/d)}{=d} \right) \quad \text{(by (11.963), applied to } n/d \text{ instead of } d) \\
\quad & = |\text{Aper} (d)|.
\end{align*} \]

(11.967)

Now, (11.962) becomes

\[ |A|_n = \sum_{d|n} d \left\{ N \text{ is an } n\text{-necklace} \mid |N| = d \right\} = \sum_{d|n} d |\text{Aper} (d)| = \sum_{e|n} e |\text{Aper} (e)| \]

(here, we renamed the summation index $d$ as $e$). Now,

\[ \sum_{d|n} \mu (d) |A|^{n/d} = \sum_{d|n} \mu (d) \sum_{e|n/d} e |\text{Aper} (e)| = \sum_{e|n} \sum_{d|n/e} \mu (d) e |\text{Aper} (e)| \]

\[ = \sum_{e|n} \sum_{d|n/e} \mu (d) e |\text{Aper} (e)| = \sum_{e|n} \sum_{d|n/e} \mu (d) e |\text{Aper} (e)| \quad \text{(by (11.968), applied to } n/d \text{ instead of } n) \]

\[ = \sum_{e|n} \delta_{n/(e,1)} e |\text{Aper} (e)| = \sum_{e|n} \delta_{n,e} e |\text{Aper} (e)| = n |\text{Aper} (n)|. \]

Solving this for $|\text{Aper} (n)|$, we obtain

\[ |\text{Aper} (n)| = \frac{1}{n} \sum_{d|n} \mu (d) |A|^{n/d}. \]

(11.969)

Now, let us recall that $\text{Aper} \{n\}$ is the set of all aperiodic $n$-necklaces. Thus,

\[ |\text{Aper} (n)| = \text{(the number of all aperiodic } n\text{-necklaces).} \]

Hence,

\[ \text{(the number of all aperiodic } n\text{-necklaces)} = |\text{Aper} (n)| = \frac{1}{n} \sum_{d|n} \mu (d) |A|^{n/d}. \]

This solves Exercise 6.34(f).

(g) We shall use the notations we introduced in the solution of Exercise 6.34(f).
We have seen (in the solution of Exercise 6.34(f)) that for any \( n \)-necklace \( N \), the cardinality \(|N|\) is a divisor of \( n \). Hence,

\[
\text{(the number of all } n \text{-necklaces)} = \sum_{d|n} \text{(the number of all } n \text{-necklaces } N \text{ such that } |N| = d)
\]

\[
= |\{ \text{N is an } n \text{-necklace } | |N| = d \}| = |\text{Aper}(d)|
\]

(by (11.967))

\[
= \sum_{d|n} |\text{Aper}(d)| = \sum_{e|n} |\text{Aper}(e)|
\]

(here, we renamed the summation index \( d \) as \( e \))

\[
\text{(11.970)}
\]

From (11.961), we have

\[
(11.971) \quad \phi(n) = \sum_{d|n} \mu(d) \frac{n}{d}.
\]

Now,

\[
\frac{1}{n} \sum_{d|n} \phi(d) |\mathcal{A}|^{n/d} = \frac{1}{n} \sum_{d|n} \phi(n/d) |\mathcal{A}|^{d}
\]

(here, we have substituted \( n/d \) for \( d \) in the sum)

\[
= \frac{1}{n} \sum_{e|n} \phi(n/e) |\mathcal{A}|^{e}
\]

(here, we have renamed the summation index \( d \) as \( e \))

\[
(11.971)
\]

\[
\sum_{e|n} \sum_{d|n/e} \frac{1}{de} \mu(d) |\mathcal{A}|^{e}
\]

(here, we have substituted \( de \) for \( e \) in the inner sum)

\[
= \sum_{e|n} \sum_{d|n/e} \frac{1}{de} \mu(d) |\mathcal{A}|^{e}.
\]

Compared with (11.970), this yields

\[
\text{(the number of all } n \text{-necklaces)} = \frac{1}{n} \sum_{d|n} \phi(d) |\mathcal{A}|^{n/d}.
\]

This solves Exercise 6.34(g).

\( h \) Alternative solution of Exercise 6.29: Let \( n \) be a positive integer. Exercise 6.34(e) yields that the aperiodic \( n \)-necklaces are in bijection with Lyndon words of length \( n \). Hence,

\[
\text{(the number of all aperiodic } n \text{-necklaces)} = \text{(the number of all Lyndon words of length } n),
\]
so that

\[
\text{(the number of all Lyndon words of length } n) = \frac{1}{n} \sum_{d|n} \mu(d) |A|^{n/d}
\]

(by Exercise 6.34(f)). Thus,

\[
\text{(the number of all Lyndon words of length } n) = \frac{1}{n} \sum_{d|n} \mu(d) \phi(d) q^{n/d}.
\]

This solves Exercise 6.29 again. Thus, Exercise 6.34(h) is solved.

(i) \textit{First solution of Exercise 6.34(i):} Let } q \in \mathbb{Z}. \text{ We need to show that } n \mid \sum_{d|n} \mu(d) q^{n/d} \text{ and } n \mid \sum_{d|n} \phi(d) q^{n/d}.

Let } r \text{ be the remainder of } q \text{ modulo } n. \text{ Then, } r \in \{0, 1, \ldots, n-1\} \subset \mathbb{N}. \text{ Fix a finite set } \mathfrak{A} \text{ containing } r \text{ elements. (Such an } \mathfrak{A} \text{ exists, since } r \in \mathbb{N}.) \text{ Exercise 6.34(f) (applied to } r \text{ instead of } q) \text{ shows that the number of all aperiodic } n\text{-necklaces is } \frac{1}{n} \sum_{d|n} \mu(d) |\mathfrak{A}|^{n/d}. \text{ Hence, } \frac{1}{n} \sum_{d|n} \mu(d) |\mathfrak{A}|^{n/d} \text{ is an integer (since the number of all aperiodic } n\text{-necklaces is an integer). In other words, } n \mid \sum_{d|n} \mu(d) |\mathfrak{A}|^{n/d}, \text{ so that } \sum_{d|n} \mu(d) |\mathfrak{A}|^{n/d} \equiv 0 \text{ mod } n. \text{ But } r \text{ is the remainder of } q \text{ modulo } n. \text{ Thus, } n \equiv q \text{ mod } n, \text{ so that } n \equiv r = |\mathfrak{A}| \text{ mod } n (\text{since the set } \mathfrak{A} \text{ has } r \text{ elements). Now, } \sum_{d|n} \mu(d) \phi(d) q^{n/d} \equiv \sum_{d|n} \mu(d) |\mathfrak{A}|^{n/d} \equiv 0 \text{ mod } n, \text{ so that } n \mid \sum_{d|n} \mu(d) \phi(d) q^{n/d}.

Also, Exercise 6.34(g) (applied to } r \text{ instead of } q) \text{ shows that the number of all } n\text{-necklaces is } \frac{1}{n} \sum_{d|n} \phi(d) |\mathfrak{A}|^{n/d}. \text{ Hence, } \frac{1}{n} \sum_{d|n} \phi(d) |\mathfrak{A}|^{n/d} \text{ is an integer (since the number of all } n\text{-necklaces is an integer). In other words, } n \mid \sum_{d|n} \phi(d) |\mathfrak{A}|^{n/d}, \text{ so that } \sum_{d|n} \phi(d) |\mathfrak{A}|^{n/d} \equiv 0 \text{ mod } n. \text{ Now, } \sum_{d|n} \phi(d) \phi(d) q^{n/d} \equiv \sum_{d|n} \phi(d) |\mathfrak{A}|^{n/d} \equiv 0 \text{ mod } n, \text{ so that } n \mid \sum_{d|n} \phi(d) q^{n/d}. \text{ The solution of Exercise 6.34(i) is thus complete.}

\textit{Second solution of Exercise 6.34(i):} \text{ Forget that we fixed } n. \text{ Let } q \in \mathbb{Z}. \text{ Let } A \text{ denote the ring } \mathbb{Z}. \text{ For every } n \in \{1, 2, 3, \ldots\}, \text{ let } \varphi_n \text{ denote the identity endomorphism id of } A. \text{ Exercise 2.84 yields (among other things) that the seven equivalent assertions } C, D, E, F, G, H \text{ and } J \text{ of Exercise 2.82 are satisfied for the family } (b_n)_{n \geq 1} = (q^n)_{n \geq 1}. \text{ In particular, Assertion F of Exercise 2.82 is satisfied for the family } (b_n)_{n \geq 1} = (q^n)_{n \geq 1}. \text{ In other words, every positive integer } n \text{ satisfies}

\begin{equation}
\sum_{d|n} \mu(d) \varphi_d \left( q^{n/d} \right) \in n\mathbb{Z}.
\end{equation}

Also, Assertion G of Exercise 2.82 is satisfied for the family } (b_n)_{n \geq 1} = (q^n)_{n \geq 1} \text{ (since the seven equivalent assertions } C, D, E, F, G, H \text{ and } J \text{ of Exercise 2.82 are satisfied for the family } (b_n)_{n \geq 1} = (q^n)_{n \geq 1}. \text{ In other words, every positive integer } n \text{ satisfies}

\begin{equation}
\sum_{d|n} \phi(d) \varphi_d \left( q^{n/d} \right) \in n\mathbb{Z}.
\end{equation}

Now, fix a positive integer } n. \text{ We can rewrite (11.972) as } n \mid \sum_{d|n} \mu(d) \varphi_d (b_n/d). \text{ Since } \sum_{d|n} \mu(d) \varphi_d \left( q^{n/d} \right) = \sum_{d|n} \mu(d) q^{n/d}, \text{ this simplifies to } n \mid \sum_{d|n} \mu(d) q^{n/d}. \text{ Also, we can rewrite (11.973) as } n \mid \sum_{d|n} \phi(d) \varphi_d (b_n/d). \text{ Since } \sum_{d|n} \phi(d) \varphi_d \left( q^{n/d} \right) = \sum_{d|n} \phi(d) q^{n/d}, \text{ this simplifies to } n \mid \sum_{d|n} \phi(d) q^{n/d}. \text{ Exercise 6.34(i) is thus solved again.
11.116. Solution to Exercise 6.35. Solution to Exercise 6.35. There exists a \( u \in \mathcal{A}^* \) such that \( w = uv \) (since \( v \) is a suffix of \( w \)). Consider this \( u \).

The word \( v \) is Lyndon and thus nonempty. Hence, \( \ell(v) > 0 \).

The exercise asks us to prove the logical equivalence

\[
(11.974) \quad (t \text{ is the longest Lyndon suffix of } wt) \iff (\text{we do not have } v < t).
\]

We shall prove the \( \implies \) and \( \iff \) parts of this equivalence separately:

\( \implies \): Assume that \( t \) is the longest Lyndon suffix of \( wt \). We need to prove that we do not have \( v < t \).

Assume the contrary. Thus, \( v < t \). Now, both \( v \) and \( t \) are Lyndon words. Hence, Proposition 6.16(a) (applied to \( v \) and \( t \) instead of \( u \) and \( v \)) yields that the word \( vt \) is Lyndon. But \( vt \) is a suffix of \( wt \) (since \( v \) is a suffix of \( w \)). Hence, \( vt \) is a Lyndon suffix of \( wt \). This Lyndon suffix is clearly longer than \( t \) (since \( \ell(vt) = \ell(v) + \ell(t) > \ell(t) \)), which flies in the face of the fact that \( t \) is the longest Lyndon suffix of \( wt \). This contradiction shows that our assumption was wrong. Hence, we have proven that we do not have \( v < t \).

This proves the \( \implies \) part of the equivalence (11.974).

\( \iff \): Assume that we do not have \( v < t \). We must show that \( t \) is the longest Lyndon suffix of \( wt \).

Assume the contrary. Then, \( t \) is not the longest Lyndon suffix of \( wt \). Since \( t \) is a Lyndon suffix of \( wt \), this means that there exists a Lyndon suffix \( r \) of \( wt \) which satisfies \( \ell(r) > \ell(t) \). Let \( q \) be the shortest suffix such that \( \ell(q) > \ell(t) \). Thus, \( q \) is a Lyndon suffix of \( wt \) and satisfies \( \ell(q) > \ell(t) \). Moreover, \( q \) is the shortest such Lyndon suffix. Hence, if \( r \) is a Lyndon suffix of \( wt \) which satisfies \( \ell(r) > \ell(t) \), then

\[
\ell(r) \geq \ell(q).
\]

Both \( q \) and \( t \) are suffixes of \( wt \), and the suffix \( q \) begins earlier (since \( \ell(q) > \ell(t) \)). Thus, there exists a nonempty suffix \( g \) of \( w \) such that \( q = gt \). Consider this \( g \).

The word \( t \) is nonempty (since it is Lyndon) and a suffix of \( q \) (since \( q = gt \)). Thus, Corollary 6.15 (applied to \( q \) and \( t \) instead of \( w \) and \( v \)) yields \( t \geq q \) (since \( q \) is Lyndon).

The word \( t \) is Lyndon (by assumption) and a proper suffix of \( q \) (since \( q = gt \) and since \( g \) is nonempty). In other words, \( t \) is a Lyndon proper suffix of \( q \). Moreover, the word \( t \) is the longest Lyndon proper suffix of \( q \) (since \( t \) is nonempty). Thus, \( q \) is a Lyndon word of length \( > 1 \). Hence, Exercise 6.31(a) (applied to \( q \) and \( g \) and \( t \) instead of \( w \), \( u \) and \( v \)) shows that the words \( g \) and \( t \) are Lyndon. In particular, \( g \) is a Lyndon suffix of \( w \) (since \( g \) is Lyndon and a suffix of \( w \)). Since \( v \) is the longest Lyndon suffix of \( w \), this shows that \( g \) is at most as long as \( v \). In other words, \( \ell(g) \leq \ell(v) \).

Since \( t \) is nonempty, we have \( g < gt \).

Both \( g \) and \( v \) are suffixes of \( w \), and the suffix \( g \) begins no earlier than \( v \) (since \( \ell(g) \leq \ell(v) \)). Therefore, \( g \) is a suffix of \( v \). Hence, Corollary 6.15 (applied to \( v \) and \( g \) instead of \( w \) and \( v \)) yields \( g \geq v \). Hence, \( v \leq g < gt = q \leq t \) (since \( t \geq g \)). This contradicts the fact that we do not have \( v < t \). Thus, we have obtained a contradiction. Our assumption was therefore wrong, and we have shown that \( t \) is the longest Lyndon suffix of \( wt \). This proves the \( \iff \) part of the equivalence (11.974).

We have now proven both parts of the equivalence (11.974), and thus solved Exercise 6.35.

[Remark: Exercise 6.35 still holds in the partial-order setting. In fact, the solution we have given above still applies in this setting.]
11.117. **Solution to Exercise 6.36.** Solution to Exercise 6.36. Let $a$ be the first letter of the word $w$. (This is well-defined, since $w$ has length $> 1 > 0$.) We consider $a$ as a one-letter word; thus, $\ell(a) = 1$. Clearly, $a$ is a prefix of $w$ (since $a$ is the first letter of $w$). Hence, there exists a word $w'$ such that $w = aw'$. Consider this $w'$. The word $w'$ is nonempty (since $w$ has length $> 1$).

Now, if $h$ is any word, then

\begin{equation}
(11.976) \quad \text{the suffixes of } h \text{ are precisely the proper suffixes of } ah
\end{equation}

the suffixes of $h$ are precisely the proper suffixes of $ah$ (since $a$ is a single letter). Applying this to $h = w'$, we see that the suffixes of $w'$ are precisely the proper suffixes of $aw'$. In other words, the suffixes of $w'$ are precisely the proper suffixes of $w$ (since $aw' = w$). Thus, $v$ is the longest Lyndon suffix of $w'$ (since $v$ is the longest Lyndon proper suffix of $w'$).

On the other hand, applying (11.976) to $h = w't$, we see that the suffixes of $w't$ are precisely the proper suffixes of $aw't$. In other words, the suffixes of $w't$ are precisely the proper suffixes of $wt$ (since $aw't = w$).

But Exercise 6.35 (applied to $w'$ instead of $w$) shows that

\begin{equation}
(t \text{ is the longest Lyndon suffix of } w't \text{ if and only if we do not have } v < t).
\end{equation}

Since the suffixes of $w't$ are precisely the proper suffixes of $wt$, this result rewrites a follows:

\begin{equation}
(t \text{ is the longest Lyndon proper suffix of } wt \text{ if and only if we do not have } v < t).
\end{equation}

This solves Exercise 6.36.

[Remark: Exercise 6.36 still holds in the partial-order settingootnote{See Exercise 6.33 for an explanation of what the partial-order setting is.}. In fact, the solution we have given above still applies in this setting.]

11.118. **Solution to Exercise 6.39.** Solution to Exercise 6.39. Recall the definition of stf $w$. It says that stf $w = (u, v)$, where $u$ and $v$ are defined as follows:

- The word $v$ is defined as the longest proper suffix of $w$ such that $v$ is Lyndon.
- The word $u$ is defined as the nonempty word such that $w = uv$.

Consider these $u$ and $v$. Thus, $(u, v) = \text{stf } w = (g, h)$. In other words, $u = g$ and $v = h$.

We know that $v$ is the longest proper suffix of $w$ such that $v$ is Lyndon. In other words, $v$ is the longest Lyndon proper suffix of $w$. In other words, $h$ is the longest Lyndon proper suffix of $w$ (since $v = h$). This solves Exercise 6.39(a).

We have $w = \underbrace{u}_{=g} \underbrace{v}_{=h} = gh$. This solves Exercise 6.39(b).

From Exercise 6.31(b), we conclude that $u < w < v$. Since $u = g$, $w = gh$ and $v = h$, this rewrites as $g < gh < h$. This solves Exercise 6.39(c).

From Exercise 6.31(a), we conclude that the words $u$ and $v$ are Lyndon. Thus, the word $w$ is Lyndon. In other words, the word $g$ is Lyndon (since $u = g$). This solves Exercise 6.39(d).

The word $v$ is the longest proper suffix of $w$ such that $v$ is Lyndon. In particular, the word $v$ is Lyndon. In other words, $v \in \mathcal{L}$.

We have $g \in \mathcal{L}$ (since the word $g$ is Lyndon) and $h = v \in \mathcal{L}$. Also, the word $g$ is nonempty (since $g$ is Lyndon), and thus we have $\ell(g) \geq 1$. Furthermore, the word $h$ is Lyndon (since $h \in \mathcal{L}$) and thus nonempty.

Hence, $\ell(h) \geq 1$. Now, $\ell\left(\underbrace{w}_{=gh}\right) = \ell(gh) = \ell(g) + \ell(h) \geq \ell(g) + 1 > \ell(g)$, so that $\ell(g) < \ell(w)$. Also,

\begin{equation}
\ell\left(\underbrace{w}_{=gh}\right) = \ell(gh) = \ell(g) + \ell(h) \geq 1 + \ell(h) > \ell(h), \text{ so that } \ell(g) < \ell(w) \text{. Thus, Exercise 6.39(e) is solved.}
\end{equation}

(f) Exercise 6.36 (applied to $v = h$) shows that $t$ is the longest Lyndon proper suffix of $wt$ if and only if we do not have $h < t$ (since $h$ is the longest Lyndon proper suffix of $w$). This solves Exercise 6.39(f).
11.19. Solution to Exercise 6.40. Solution to Exercise 6.40. We define a binary relation \( \sim \) on the set \( \mathbb{A}^* \) as follows: If \( w \) and \( w' \) are two words, then we write \( w \sim w' \) if and only if \( w' \) is a permutation of the word \( w \) (that is, if and only if there exists a permutation \( \sigma \in S_k \) satisfying \( w' = (w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(k)}) \)), where \( k = \ell(w) \). We notice the following properties of the relation \( \sim \):

- The relation \( \sim \) is an equivalence relation; in other words, it is reflexive, symmetric and transitive.
- If \( w \) and \( w' \) are two words satisfying \( w \sim w' \), then \( \ell(w) = \ell(w') \).
- If \( u, v, u', v' \) are four words satisfying \( u \sim u' \) and \( v \sim v' \), then \( uv \sim u'v' \). We shall refer to this fact as the \textit{monoidality} of the relation \( \sim \).
- If \( u \) and \( v \) are any two words, then \( uv \sim vu \).

We also recall a fundamental property of Lie algebras (one of the forms of the Jacobi identity):

- Every three elements \( x, y \) and \( z \) of a Lie algebra \( \mathfrak{t} \) satisfy

\[
[[x, y], z] = [[x, z], y] - [[y, z], x].
\]

We can now finally come to the solution of Exercise 6.40.

For every \( h \in \mathbb{A}^* \) and \( s \in \mathbb{A}^* \), we define a subset \( \mathcal{L}_{h,s} \) of \( \mathcal{L} \) by

\[
\mathcal{L}_{h,s} = \{ w \in \mathcal{L} \mid w \sim h \text{ and } w \leq s \}.
\]

For every \( h \in \mathbb{A}^* \) and \( s \in \mathbb{A}^* \), we define a \( \mathbb{A} \)-module \( B_{h,s} \) of \( B \) by

\[
B_{h,s} = \sum_{w \in \mathcal{L}_{h,s}} kw_w.
\]

(In other words, for every \( h \in \mathbb{A}^* \) and \( s \in \mathbb{A}^* \), we define \( B_{h,s} \) as the \( \mathbb{A} \)-linear span of the elements \( kw_w \) with \( w \in \mathcal{L}_{h,s} \)). If \( h, s, g \) and \( t \) are four words satisfying \( h \sim g \) and \( s < t \), then

\[
B_{h,s} \subseteq B_{g,t}
\]

(a) We claim that

\[
[b_p, b_q] \in B_{pq,q} \quad \text{for every } (p, q) \in \mathcal{L} \times \mathcal{L} \text{ satisfying } p < q.
\]

Proof of (11.979): We can WLOG assume that the alphabet \( \mathbb{A} \) is finite. Assume this. We shall prove (11.979) by strong induction over \( \ell(pq) \):

\[\text{Proof of (11.978): Let } h, s \text{ and } t \text{ be four words satisfying } h \sim g \text{ and } s < t.\]

Let \( v \in \mathcal{L}_{h,s} \). Thus, \( v \in \mathcal{L}_{h,s} = \{ w \in \mathcal{L} \mid w \sim h \text{ and } w \leq s \} \) (by the definition of \( \mathcal{L}_{h,s} \)). In other words, \( v \) is an element of \( \mathcal{L} \) and satisfies \( v \sim h \) and \( v \leq s \). From \( v \sim h \) and \( h \sim g \), we obtain \( v \sim g \) (since the relation \( \sim \) is transitive). Thus, \( v \) is an element of \( \mathcal{L} \) and satisfies \( v \sim g \) and \( v < t \). In other words, \( v \in \{ w \in \mathcal{L} \mid v \sim g \text{ and } w \leq t \} = \mathcal{L}_{g,t} \) (since \( \mathcal{L}_{g,t} \) is defined to be \( \{ w \in \mathcal{L} \mid v \sim g \text{ and } (v, t) = (w, u) \} \)).

Now, let us forget that we fixed \( v \). We thus have proven that every \( v \in \mathcal{L}_{h,s} \) satisfies \( v \in \mathcal{L}_{g,t} \). In other words, \( \mathcal{L}_{h,s} \subseteq \mathcal{L}_{g,t} \).

Now, the definition of \( B_{h,s} \) yields \( B_{h,s} = \sum_{w \in \mathcal{L}_{h,s}} kw_w \subseteq \sum_{w \in \mathcal{L}_{g,t}} kw_w \) (since \( \mathcal{L}_{h,s} \subseteq \mathcal{L}_{g,t} \)). Since \( B_{g,t} = \sum_{w \in \mathcal{L}_{g,t}} kw_w \) (by the definition of \( B_{g,t} \)), this rewrites as \( B_{h,s} \subseteq B_{g,t} \). This proves (11.978).

\[\text{Proof of (11.979): Let } h, s \text{ and } t \text{ be four words satisfying } h \sim g \text{ and } s < t.\]

We shall prove (11.979) by strong induction over \( \ell(pq) \):

But we have \( \mathcal{L}' = \mathcal{L} \cap \mathbb{B}^* \subseteq \mathcal{L} \), thus \( \mathcal{L}_{pq',q} \subseteq \mathcal{L}_{pq,q} \) and therefore \( B_{pq',q} \subseteq B_{pq,q} \) (actually, a little thought shows that \( B_{pq',q} = B_{pq,q} \)), so that we have \( b_p, b_q \in B_{pq,q} \subseteq B_{pq,q} \), and thus (11.979) is proven.
**Induction step:** Let $N$ be a nonnegative integer. Assume that (11.979) holds in the case when $\ell(pq) < N$. We need to prove that (11.979) holds in the case when $\ell(pq) = N$.

We have assumed that (11.979) holds in the case when $\ell(pq) < N$. In other words, we have

(11.980) $[b_p, b_q] \in B_{pq,q}$ for every $(p, q) \in \mathcal{L} \times \mathcal{L}$ satisfying $p < q$ and $\ell(pq) < N$.

We now must prove that (11.979) holds in the case when $\ell(pq) = N$. In other words, we must prove that

(11.981) $[b_p, b_q] \in B_{pq,q}$ for every $(p, q) \in \mathcal{L} \times \mathcal{L}$ satisfying $p < q$ and $\ell(pq) = N$.

Let $\mathfrak{S}$ be the set $\bigcup_{i=0}^{N} \mathfrak{A}^i$. This set $\mathfrak{S}$ is finite (since the set $\mathfrak{A}$ is finite) and totally ordered (by the lexicographic order). Thus, for every $w \in \mathfrak{S}$, we can define a nonnegative integer $\rho(w)$ by

\[ \rho(w) = |\{g \in \mathfrak{S} \mid g < w\}|. \]

In other words, for every $w \in \mathfrak{S}$, we define $\rho(w)$ to be the number of all $g \in \mathfrak{S}$ which are smaller than $w$. It is clear that if $w$ and $w'$ are two elements of $\mathfrak{S}$ satisfying $w < w'$, then

(11.982) $\rho(w) < \rho(w')$.

For every $(p, q) \in \mathcal{L} \times \mathcal{L}$ satisfying $p < q$ and $\ell(pq) = N$, we have $q \in \mathfrak{S}$, and thus $\rho(q)$ is well-defined. We are thus going to prove (11.981) by strong induction over $\rho(q)$.

**Induction step:** Let $K$ be a nonnegative integer. Assume that (11.981) holds in the case when $\rho(q) < K$. We need to prove that (11.981) holds in the case when $\rho(q) = K$.

We have assumed that (11.981) holds in the case when $\rho(q) < K$. In other words, we have

(11.983) $[b_p, b_q] \in B_{pq,q}$ for every $(p, q) \in \mathcal{L} \times \mathcal{L}$ satisfying $p < q$ and $\ell(pq) = N$ and $\rho(q) < K$.

Now, let $(p, q) \in \mathcal{L} \times \mathcal{L}$ be such that $p < q$ and $\ell(pq) = N$ and $\rho(q) = K$. We are going to prove that $[b_p, b_q] \in B_{pq,q}$.

We have $(p, q) \in \mathcal{L} \times \mathcal{L}$. In other words, the words $p$ and $q$ are Lyndon. Proposition 6.16(a) (applied to $u = p$ and $v = q$) thus shows that the word $pq$ is Lyndon. In other words, $pq \in \mathcal{L}$.

Furthermore, Proposition 6.16(b) (applied to $u = p$ and $v = q$) shows that $pq < q$.

The definition of $\mathcal{E}_{pq,q}$ yields $\mathcal{E}_{pq,q} = \{w \in \mathcal{L} \mid w \sim pq \text{ and } w < q\}$. The definition of $B_{pq,q}$ yields $B_{pq,q} = \sum_{w \in \mathcal{E}_{pq,q}} k_b w$.

The words $p$ and $q$ are nonempty (since they are Lyndon), and thus have length $\geq 1$ each. Hence, the word $pq$ has length $\geq 1 + 1 > 1$. Hence, $pq$ is a Lyndon word of length $> 1$ (since $pq$ is Lyndon and since $pq$ has length $\ell(pq) > 1$). Therefore, $\text{stf}(pq)$ is well-defined. If $\text{stf}(pq) = (p, q)$, then it is easy to see that $[b_p, b_q] \in B_{pq,q}$.

Hence, for the rest of the proof of $[b_p, b_q] \in B_{pq,q}$, we WLOG assume that we don’t have $\text{stf}(pq) = (p, q)$. Thus, $q$ is not the longest Lyndon proper suffix of $pq$. In other words, the statement that $q$ is the longest Lyndon proper suffix of $pq$ is false.

---

Proof: Let $w$ and $w'$ be two elements of $\mathfrak{S}$ satisfying $w < w'$. Then, $\{g \in \mathfrak{S} \mid g < w\}$ is a proper subset of $\{g \in \mathfrak{S} \mid g < w'\}$ (proper because $w$ belongs to the latter set but not to the former set). Hence, $|\{g \in \mathfrak{S} \mid g < w\}| < |\{g \in \mathfrak{S} \mid g < w'\}|$. Since $\rho(w) = |\{g \in \mathfrak{S} \mid g < w\}|$ and $\rho(w') = |\{g \in \mathfrak{S} \mid g < w'\}|$ (similarly), this rewrites as $\rho(w) < \rho(w')$. qed.

Proof: Let $(p, q) \in \mathcal{L} \times \mathcal{L}$ be such that $p < q$ and $\ell(pq) = N$. Then, $N = \ell(pq) = \ell(p) + \ell(q) \geq \ell(q)$, so that $\ell(q) \leq N$.

and thus $q \in \bigcup_{i=0}^{N} \mathfrak{A}^i = \mathfrak{S}$, qed.

Proof of course will be an induction within our current induction step, so the reader should try not to confuse the two inductions going on.

Proof: Assume that $\text{stf}(pq) = (p, q)$. Thus, (6.2) (applied to $w = pq$, $u = p$ and $v = q$) shows that $b_{pq} = [b_p, b_q]$. But $pq$ is an element of $\mathcal{L}$ and satisfies $pq \sim pq$ (since the relation $\sim$ is reflexive) and $pq < q$. In other words, $pq \in \{w \in \mathcal{L} \mid w \sim pq \text{ and } w < q\} = \mathcal{E}_{pq,q}$. Thus, $k_b pq \subseteq \sum_{w \in \mathcal{E}_{pq,q}} k_b w$.

Now, from $b_{pq} = [b_p, b_q]$, we obtain

$[b_p, b_q] = b_{pq} \in k_b pq \subseteq \sum_{w \in \mathcal{E}_{pq,q}} k_b w = B_{pq,q}$

(since $B_{pq,q} = \sum_{w \in \mathcal{E}_{pq,q}} k_b w$), qed.

Proof: Assume the contrary. Thus, $q$ is the longest Lyndon proper suffix of $pq$.

Let $w = pq$. Recall that $q$ is the longest Lyndon proper suffix of $pq$. In other words, $q$ is the longest Lyndon proper suffix of $w$ (since $w = pq$).
We have $\ell(p) > 1$. Thus, $p$ is a Lyndon word of length $> 1$ (since $p$ is Lyndon and since $p$ has length $\ell(p) > 1$). Therefore, $\text{stf} p$ is well-defined. Set $(u, v) = \text{stf} p$. Then, Exercise 6.39(a) (applied to $w = p$, $g = u$ and $h = v$) says that $v$ is the longest Lyndon proper suffix of $p$. In particular, $v$ is a Lyndon proper suffix of $p$. Moreover, Exercise 6.39(b) (applied to $w = p$, $g = u$ and $h = v$) says that we have $p = uv$. Also, Exercise 6.39(c) (applied to $w = p$, $g = u$ and $h = v$) says that we have $u < uv < v$. Finally, Exercise 6.39(d) (applied to $w = p$, $g = u$ and $h = v$) says that the word $u$ is Lyndon. In other words, $u \in \Sigma$. The words $u$ and $v$ are nonempty (since they are Lyndon).

Exercise 6.36 (applied to $w = p$ and $t = q$) now shows that $q$ is the longest Lyndon proper suffix of $pq$ if and only if we do not have $v < q$. Thus, the statement that we do not have $v < q$ is false (because the statement that $q$ is the longest Lyndon proper suffix of $pq$ is false). Thus, we have $v < q$. Thus, $u < v < q$.

Recall that $(u, v) = \text{stf} p$. Thus, (6.2) (applied to $w = p$) shows that $b_p = [b_u, b_v]$. Thus,

\begin{align*}
(11.984) & \quad \begin{bmatrix} b_p, b_q \end{bmatrix} = \begin{bmatrix} [b_u, b_v], b_q \end{bmatrix} = \begin{bmatrix} [b_u, b_q], b_v \end{bmatrix} - \begin{bmatrix} [b_v, b_q], b_u \end{bmatrix} \\
\text{(by (11.977), applied to } & \quad \ell = g, \ x = b_u, \ y = b_v \text{ and } z = b_q) \end{align*}

Now,

(11.985) $[b_u, b_q] \in B_{u q, q}$

772 and

(11.986) $[b_v, b_q] \in B_{v q, q}$

773. Hence, (11.984) becomes

\begin{align*}
(11.987) & \quad [b_p, b_q] = \begin{bmatrix} [b_u, b_q], b_v \end{bmatrix} - \begin{bmatrix} [b_v, b_q], b_u \end{bmatrix} \in [B_{u q, q}, b_v] - [B_{v q, q}, b_u].
\end{align*}

Let $(g, h) = \text{stf} (pq)$. Then, $(g, h) = \text{stf} \left( \begin{array}{c} pq \\ \ell = u \end{array} \right) = \text{stf} w$. Thus, $h$ is the longest Lyndon proper suffix of $w$ (by Exercise 6.39(a)). Comparing this with the fact that $q$ is the longest Lyndon proper suffix of $w$, we obtain that $h = q$.

But Exercise 6.39(b) shows that $w = g \cdot h = gq$. Thus, $gq = w = pq$. Cancelling $q$ from this equality, we obtain $g = p$. Now,

\begin{align*}
(11.988) & \quad \ell(u) > \ell(pq). \text{ But } \ell(pq) = N, \text{ so that } N = \ell(pq) = \ell(p) + \ell(q) > \ell(u) + \ell(q) = \ell(uq). \text{ Thus, } \ell(uq) < N. \\
\text{(since } v \text{ is nonempty)}
\end{align*}

Therefore, (11.980) (applied to $(u, q)$ instead of $(p, q)$) yields $[b_u, b_q] \in B_{u q, q}$. This proves (11.985).

773 Proof of (11.986): We have $(v, q) \in \Sigma \times \Sigma$ (since $v \in \Sigma$ and $q \in \Sigma$) and $u < q$. Also, $\ell \left( \begin{array}{c} p \\ \ell = u \end{array} \right) = \ell(uv) = \ell(u) + \ell(v) > \ell(u).$ \text{(since } v \text{ is nonempty)}

Since $v$ is nonempty, therefore, (11.980) (applied to $(u, q)$ instead of $(p, q)$) yields $[b_u, b_q] \in B_{u q, q}$. This proves (11.986).
On the other hand, if $g$ and $h$ are two Lyndon words satisfying $g < q$, $h < q$ and $gh \sim p$, then

\[(11.988)\]

\[
\text{(every } r \in \mathcal{L}_{gq,q} \text{ satisfies } [b_r, b_h] \in B_{pq,q})
\]

and therefore

\[(11.989)\]

\[
[B_{gq,q}, b_h] \subset B_{pq,q}
\]

\[774\text{Proof of (11.988): Let } g \text{ and } h \text{ be two Lyndon words satisfying } g < q, h < q \text{ and } gh \sim p. \text{ Let } r \in \mathcal{L}_{gq,q}. \text{ We must prove that } [b_r, b_h] \in B_{pq,q}. \]

The words $g$ and $h$ are Lyndon. In other words, $g \in \mathcal{L}$ and $h \in \mathcal{L}$.

We have $gh \sim p$ and thus $\ell(gh) = \ell(p)$. Hence, $\ell(p) = \ell(g) + \ell(h)$.

We have $gh \sim p$ and $g \sim q$ (since the relation $\sim$ is reflexive). Hence, $ghq \sim pq$ (by the monoidality of the relation $\sim$).

We have $N = \ell(pq) = \ell(p) + \ell(q) \geq \ell(q)$, whence $\ell(q) \leq N$ and thus $q \in \bigcup_{i=0}^{N} \mathfrak{A}^i = \emptyset$.

We have $r \in \mathcal{L}_{gq,q} = \{w \in \mathcal{L} \mid w \sim gq \text{ and } w \not< q\}$ (by the definition of $\mathcal{L}_{gq,q}$). In other words, $r$ is an element of $\mathcal{L}$ and satisfies $r \sim gq$ and $r < q$. From $r \sim gq$, we obtain $\ell(r) = \ell(gq) = \ell(g) + \ell(q)$.

If $r = h$, then $b_r \sim b_h$, so $[b_r, b_h] = 0 \in B_{pq,q}$ (since $B_{pq,q}$ is a $\mathbf{k}$-module). Hence, for the rest of the proof of $[b_r, b_h] \in B_{pq,q}$, we WLOG assume that we don’t have $r = h$.

Thus, we have $r \neq h$. Hence, we have either $r < h$ or $r > h$ (since the lexicographic order on $\mathfrak{A}^*$ is a total order). In other words, we are in one of the following two cases:

- **Case 1**: We have $r < h$.
- **Case 2**: We have $r > h$.

Let us first consider Case 1. In this case, we have $r < h$. We have $\ell(rh) = \ell(r) + \ell(h) = \ell(g) + \ell(h) + \ell(q)$. Comparing this with $\ell(pq) = \ell(p) + \ell(q) = \ell(g) + \ell(q) + \ell(h)$, we obtain $\ell(rh) = \ell(pq) = N$.

We have $r \sim gq$ and $h \sim h$ (since the relation $\sim$ is reflexive). Thus, $rh \sim ghq$ (by the monoidality of the relation $\sim$). But $ghq \sim ghq$ (by the monoidality of the relation $\sim$ again, since $g \sim g$ and $gh \sim hq$). From $rh \sim ghq$ and $ghq \sim ghq$, we obtain $rh \sim ghq$ (since the relation $\sim$ is transitive).

On the other hand, $N = \ell(rh) = \ell(r) + \ell(h) \geq \ell(h)$, whence $\ell(h) \leq N$ and thus $h \in \bigcup_{i=0}^{N} \mathfrak{A}^i = \emptyset$. Since $h < q$, we have $\rho(h) < \rho(q)$ (by (11.982), applied to $h$ and $q$ instead of $w$ and $w'$), so that $\rho(h) < \rho(q) = K$.

Now, we have $(r, h) \in \mathfrak{L} \times \mathfrak{L}$ (since $r \in \mathcal{L}$ and $h \in \mathcal{L}$) and $r < h$ and $\ell(rh) = N$ and $\rho(h) < K$. Therefore, (11.983) (applied to $r$ and $h$ instead of $p$ and $q$) shows that $[b_r, b_h] \in B_{hr,h} \subset B_{pq,q}$. Thus, $[b_r, b_h] \in B_{pq,q}$ is proven in Case 1.

Let us now consider Case 2. In this case, we have $h < r$. We have $\ell(hr) = \ell(h) + \ell(r) = \ell(h) + \ell(g) + \ell(q) = \ell(g) + \ell(h) + \ell(q)$. Comparing this with $\ell(pq) = \ell(p) + \ell(q) = \ell(g) + \ell(q) + \ell(h)$, we obtain $\ell(hr) = \ell(pq) = N$.

We have $h \sim h$ (since the relation $\sim$ is reflexive) and $r \sim gq$. Thus, $hr \sim hqg$ (by the monoidality of the relation $\sim$). But $hqg \sim hqg$ (by the monoidality of the relation $\sim$ again, since $hqg \sim ghq$ and $q \sim q$). From $hr \sim hqg$ and $hqg \sim hqg$, we obtain $hr \sim hqg$ (since the relation $\sim$ is transitive).

On the other hand, $N = \ell(hr) = \ell(h) + \ell(r) \geq \ell(r)$, whence $\ell(r) \leq N$ and thus $r \in \bigcup_{i=0}^{N} \mathfrak{A}^i = \emptyset$. Since $r < q$, we have $\rho(r) < \rho(q)$ (by (11.982), applied to $r$ and $q$ instead of $w$ and $w'$), so that $\rho(r) < \rho(q) = K$.

Now, we have $(h, r) \in \mathfrak{L} \times \mathfrak{L}$ (since $h \in \mathcal{L}$ and $r \in \mathcal{L}$) and $h < r$ and $\ell(hr) = N$ and $\rho(r) < K$. Therefore, (11.983) (applied to $h$ and $r$ instead of $p$ and $q$) shows that $[b_h, b_r] \in B_{hr,r} \subset B_{pq,q}$. Thus, $[b_r, b_h] = -[b_h, b_r] \in -B_{pq,q} \subset B_{pq,q}$ (since $B_{pq,q}$ is a $\mathbf{k}$-module). Hence, $[b_r, b_h] \in B_{pq,q}$ is proven in Case 2.

Thus, $[b_r, b_h] \in B_{pq,q}$ is proven in each of the two Cases 1 and 2. This completes the proof of (11.988).

\[775\text{Proof of (11.989): Let } g \text{ and } h \text{ be two Lyndon words satisfying } g < q, h < q \text{ and } gh \sim p. \text{ The definition of } B_{pq,q} \text{ shows that } B_{pq,q} = \sum_{w \in \mathcal{L}_{pq,q}} \mathbf{k} b_w. \text{ In other words, } B_{pq,q} \text{ is the } \mathbf{k}\text{-linear span of the elements } b_w \text{ with } w \in \mathcal{L}_{pq,q}. \text{ Hence, in order to prove the relation (11.988), it suffices to show that } [b_w, b_h] \in B_{pq,q} \text{ for every } w \in \mathcal{L}_{pq,q}. \text{ But this follows from (11.988) (applied to } r = w). \text{ This proves (11.989).}\]
We have \( p \sim p \) (since the relation \( \sim \) is reflexive). In other words, \( uv \sim p \) (since \( p = uv \)). Also, \( vu \sim uv \).

Since \( p = uv \), this rewrites as \( vu \sim p \).

Now, the words \( u \) and \( v \) are Lyndon and satisfy \( u < q \), \( v < q \) and \( uv \sim p \). Thus, we can apply (11.989) to \( q = u \) and \( h = v \). As a result, we obtain \([B_{uq,q}, b_u] \subset B_{pq,q}\).

Furthermore, the words \( v \) and \( u \) are Lyndon and satisfy \( v < q \), \( u < q \) and \( vu \sim p \). Thus, we can apply (11.989) to \( g = v \) and \( h = u \). As a result, we obtain \([B_{vq,q}, b_v] \subset B_{pq,q}\).

Now, (11.987) becomes
\[
[b_p, b_q] \in [B_{uq,q}, b_u] - [B_{vq,v}, b_v] \subset B_{pq,q} - B_{pq,q} \subset B_{pq,q}
\]
(since \( B_{pq,q} \) is a \( k \)-module).

Let us now forget that we fixed \((p, q)\). We thus have proven that
\[
[b_p, b_q] \in B_{pq,q} \quad \text{for every } (p, q) \in L \times L \text{ satisfying } p < q \text{ and } \ell(pq) = N \text{ and } \rho(q) = K.
\]

In other words, (11.981) holds in the case when \( \rho(q) = K \). This completes the induction step (in the induction proof of (11.981)). The induction proof of (11.981) is thus finished.

We thus have proven (11.981). In other words, (11.979) holds in the case when \( \ell(pq) = N \). This completes the induction step (in the induction proof of (11.979)). The induction proof of (11.979) is thus complete.

Now, we recall that \( B \) is the \( k \)-submodule of \( g \) spanned by the family \((b_w)_{w \in \mathfrak{L}}\). In other words,
\[
\text{(11.990)} \quad B = \sum_{w \in \mathfrak{L}} kb_w.
\]

Thus,
\[
\text{(11.991)} \quad B_{h,s} \subset B \quad \text{for every } h \in \mathfrak{A}^* \text{ and } s \in \mathfrak{A}^*
\]

\[776\text{Proof of (11.991): Let } h \in \mathfrak{A}^* \text{ and } s \in \mathfrak{A}^*.
\]

Clearly, \( \mathfrak{L}_{h,s} \subset \mathfrak{L} \). Thus, \( \sum_{w \in \mathfrak{L}_{h,s}} kb_w \subset \sum_{w \in \mathfrak{L}} kb_w \). Since \( B_{h,s} = \sum_{w \in \mathfrak{L}_{h,s}} kb_w \) and \( B = \sum_{w \in \mathfrak{L}} kb_w \) (by (11.990)), this rewrites as \( B_{h,s} \subset B \). This proves (11.991).

\[776\text{Proof of (11.992): Let } h \in \mathfrak{A}^* \text{ and } s \in \mathfrak{A}^*.
\]

Clearly, \( \mathfrak{L}_{h,s} \subset \mathfrak{L} \). Thus, \( \sum_{w \in \mathfrak{L}_{h,s}} kb_w \subset \sum_{w \in \mathfrak{L}} kb_w \). Since \( B_{h,s} = \sum_{w \in \mathfrak{L}_{h,s}} kb_w \) and \( B = \sum_{w \in \mathfrak{L}} kb_w \) (by (11.990)), this rewrites as \( B_{h,s} \subset B \). This proves (11.991).
Now, (11.990) yields $B = \sum_{w \in \mathcal{L}} k_b w = \sum_{p \in \mathcal{L}} k_b p$ (here, we renamed the summation index $w$ as $p$) and $B = \sum_{w \in \mathcal{L}} k_b w = \sum_{q \in \mathcal{L}} k_b q$ (here, we renamed the summation index $w$ as $q$). Thus,

$$
\begin{bmatrix}
\sum_{p \in \mathcal{L}} k_b p \\
\sum_{q \in \mathcal{L}} k_b q
\end{bmatrix}
= \sum_{p \in \mathcal{L}} k_b p = \sum_{q \in \mathcal{L}} k_b q
$$

(since the Lie bracket on $\mathfrak{g}$ is $k$-bilinear)

$$
\subset \sum_{p \in \mathcal{L}} k \in B \subset B
$$

(b) We shall first prove that

\begin{equation}
(11.993) \quad f ([b_p, b_q]) = [f (b_p), f (b_q)] \quad \text{for every } (p, q) \in \mathcal{L} \times \mathcal{L} \text{ satisfying } p < q.
\end{equation}

The proof of (11.993) is very similar to our above proof of (11.979); it proceeds using the same kind of double induction, with almost the same computations. Here are the details of this proof:

**Proof of (11.993):** We can WLOG assume that the alphabet $\mathcal{A}$ is finite. Assume this.

We shall prove (11.993) by strong induction over $\ell (pq)$:

**Induction step:** Let $N$ be a nonnegative integer. Assume that (11.993) holds in the case when $\ell (pq) < N$.

We need to prove that (11.993) holds in the case when $\ell (pq) = N$.

We have assumed that (11.993) holds in the case when $\ell (pq) < N$. In other words, we have

\begin{equation}
(11.994) \quad f ([b_p, b_q]) = [f (b_p), f (b_q)] \quad \text{for every } (p, q) \in \mathcal{L} \times \mathcal{L} \text{ satisfying } p < q \text{ and } \ell (pq) < N.
\end{equation}

We now must prove that (11.993) holds in the case when $\ell (pq) = N$. In other words, we must prove that

\begin{equation}
(11.995) \quad f ([b_p, b_q]) = [f (b_p), f (b_q)] \quad \text{for every } (p, q) \in \mathcal{L} \times \mathcal{L} \text{ satisfying } p < q \text{ and } \ell (pq) = N.
\end{equation}

We define a set $\mathfrak{G}$ in the same fashion as in our proof of (11.979). Likewise, we define a nonnegative integer $\rho (w)$ for every $w \in \mathfrak{G}$ in the same way as we did in our proof of (11.979).

For every $(p, q) \in \mathcal{L} \times \mathcal{L}$ satisfying $p < q$ and $\ell (pq) = N$, we have $q \in \mathfrak{G}$, and thus $\rho (q)$ is well-defined. We are thus going to prove (11.995) by strong induction over $\rho (q)$.

---

**Proof of (11.992):** Let $p \in \mathcal{L}$ and $q \in \mathcal{L}$. We need to prove that $[b_p, b_q] \in B$. If $p = q$, then this is clear (because if $p = q$, then $[b_p, b_q] = 0 \in B$ (since $B$ is a $k$-module)). Thus, for the rest of this proof, we WLOG assume that we don’t have $p = q$.

We have $(p, q) \in \mathcal{L} \times \mathcal{L}$ (since $p \in \mathcal{L}$ and $q \in \mathcal{L}$) and $(q, p) \in \mathcal{L} \times \mathcal{L}$ (since $q \in \mathcal{L}$ and $p \in \mathcal{L}$). We have $p \neq q$ (since we don’t have $p = q$). Thus, we have either $p < q$ or $p > q$ (since the lexicographic order on $\mathcal{N}^*$ is a total order). In other words, we are in one of the following two Cases:

**Case 1:** We have $p < q$.

**Case 2:** We have $p > q$.

Let us first consider Case 1. In this case, we have $p < q$. Thus, from (11.979), we obtain $[b_p, b_q] \in B_{pq}, q \subset B$ (by (11.991) (applied to $h = pq$ and $s = q$)). Thus, $[b_p, b_q] \in B$ is proven in Case 1.

Let us now consider Case 2. In this case, we have $p > q$. In other words, $q < p$. Thus, (11.979) (applied to $(q, p)$ instead of $(p, q)$) shows that $[b_q, b_p] \in B_{pq}, p \subset B$ (by (11.991) (applied to $h = qp$ and $s = p$)). Now, $[b_p, b_q] = -[b_q, b_p] \in -B \subset B$ (since $B$ is a $k$-module). Thus, $[b_p, b_q] \in B$ is proven in Case 2.

Now, $[b_p, b_q] \in B$ is proven in each of the two Cases 1 and 2. Thus, (11.992) is proven.
Therefore, \((11.994)\) applied to \((p, q) \in \mathcal{L} \times \mathcal{L}\) for every \((p, q) \in \mathcal{L} \times \mathcal{L}\) satisfying \(p < q\) and \(\ell(pq) = N\) and \(\rho(q) < K\).

We need to prove that \((11.995)\) holds in the case when \(\rho(q) < K\). In other words, we have \((11.996)\)
\[
f ([b_p, b_q]) = [f (b_p), f (b_q)]
\]
for every \((p, q) \in \mathcal{L} \times \mathcal{L}\) satisfying \(p < q\) and \(\ell(pq) = N\) and \(\rho(q) < K\).

Now, let \((p, q) \in \mathcal{L} \times \mathcal{L}\) be such that \(p < q\) and \(\ell(pq) = N\) and \(\rho(q) = K\). We are going to prove that \(f ([b_p, b_q]) = [f (b_p), f (b_q)]\).

As in our proof of \((11.981)\), we can prove the following facts:

- The words \(p\) and \(q\) are Lyndon.
- The word \(pq\) is Lyndon. In other words, \(pq \in \mathcal{L}\).
- We have \(pq < q\).
- The word \(pq\) is a Lyndon word of length \(> 1\). Therefore, \(\text{stf}(pq)\) is well-defined.

If \(\text{stf}(pq) = (p, q)\), then it is easy to see that \(f ([b_p, b_q]) = [f (b_p), f (b_q)]\). Hence, for the rest of the proof of \((11.997)\) \((11.998)\) and \((11.100)\), we WLOG assume that we don’t have \(\text{stf}(pq) = (p, q)\). As in our proof of \((11.981)\), we can see that \(p\) is a Lyndon word of length \(> 1\). Therefore, \(\text{stf} p\) is well-defined. Set \((u, v) = \text{stf} p\).

As in our proof of \((11.981)\), we can see the following facts:

- The word \(v\) is a Lyndon proper suffix of \(p\).
- We have \(p = uv\).
- We have \(u < uv < v\).
- The word \(u\) is Lyndon. In other words, \(u \in \mathcal{L}\).
- The words \(u\) and \(v\) are nonempty.
- We have \(u < v < q\).
- We have \(b_p = [b_u, b_v]\).
- The equality \((11.984)\) holds.

On the other hand, \((6.3)\) (applied to \(w = p\)) shows that \(f ([b_u, b_v]) = [f (b_u), f (b_v)]\) (since \(p\) is a Lyndon word of length \(> 1\), and since \((u, v) = \text{stf} p\)). Now, applying the map \(f\) to both sides of the equality \(b_p = [b_u, b_v]\), we obtain \((11.997)\)
\[
f (b_p) = f ([b_u, b_v]) = [f (b_u), f (b_v)].
\]

Now, \((11.998)\)
\[
f ([b_u, b_q]) = [f (b_u), f (b_q)]
\]
and \((11.999)\)
\[
f ([b_v, b_q]) = [f (b_v), f (b_q)]
\]

On the other hand, if \(g\) and \(h\) are two Lyndon words satisfying \(g < q\), \(h < q\) and \(gh \sim p\), then \((11.100)\)
\[
(\text{every } r \in \mathcal{L}_{gq,q} \text{ satisfies } f ([b_r, b_h]) = [f (b_r), f (b_h)])
\]

---

**Proof.** Assume that \(\text{stf}(pq) = (p, q)\). Thus, \((p, q) = \text{stf}(pq)\). Hence, \((6.3)\) (applied to \(w = pq\), \(u = p\) and \(v = q\)) shows that \(f ([b_p, b_q]) = [f (b_p), f (b_q)]\), qed.

**Proof of \((11.998)\):** We have \((u, q) \in \mathcal{L} \times \mathcal{L}\) (since \(u \in \mathcal{L}\) and \(q \in \mathcal{L}\)) and \(u < q\). Also, \(\ell\left(\frac{p}{uv}\right) = \ell(uv) = \ell(u) + \ell(v) > \ell(u)\). But \(\ell(pq) = N\), so that \(N = \ell(pq) = \ell(p) + \ell(q) > \ell(u) + \ell(q) = \ell(uq)\). Thus, \(\ell(uq) < N\).

*Proof of \((11.999)\):* We have \((v, q) \in \mathcal{L} \times \mathcal{L}\) (since \(v \in \mathcal{L}\) and \(q \in \mathcal{L}\)) and \(v < q\). Also, \(\ell\left(\frac{p}{uv}\right) = \ell(uv) = \ell(u) + \ell(v) > \ell(u)\). But \(\ell(pq) = N\), so that \(N = \ell(pq) = \ell(p) + \ell(q) > \ell(v) + \ell(q) = \ell(vq)\). Thus, \(\ell(vq) < N\).

*Proof of \((11.100)\):* We have \((u, q) \in \mathcal{L} \times \mathcal{L}\) (since \(u \in \mathcal{L}\) and \(q \in \mathcal{L}\)) and \(u < q\). Also, \(\ell\left(\frac{p}{uv}\right) = \ell(uv) = \ell(u) + \ell(v) > \ell(u)\). But \(\ell(pq) = N\), so that \(N = \ell(pq) = \ell(p) + \ell(q) > \ell(v) + \ell(q) = \ell(vq)\). Thus, \(\ell(vq) < N\).
and therefore

\[(11.1001) \quad (\text{every } x \in B_{q,q} \text{ satisfies } f([x,b_h]) = [f(x), f(b_h)])\]

Recall that (11.979) holds. (This was proven in our solution to Exercise 6.40(a).) Now, \((u, q) \in \mathcal{L} \times \mathcal{L}\) (since \(u \in \mathcal{L}\) and \(q \in \mathcal{L}\) and \(u < q\). Thus, (11.979) (applied to \((u, q)\) instead of \((p, q)\)) shows that \([b_u, b_q] \in B_{u,q,q}\). Also, \((v, q) \in \mathcal{L} \times \mathcal{L}\) (since \(v \in \mathcal{L}\) and \(q \in \mathcal{L}\) and \(v < q\). Thus, (11.979) (applied to \((v, q)\) instead of \((p, q)\)) shows that \([b_v, b_q] \in B_{v,q,q}\).

As in our proof of (11.981), we can show that \(uv \sim p\) and \(vu \sim p\).

Now, the words \(u, v\) are Lyndon and satisfy \(u < q, v < q, uv \sim p\) and \([b_u, b_q] \in B_{u,q,q}\). Thus, we can apply (11.1001) to \(g = u, h = v\) and \(x = [b_u, b_q]\). As a result, we obtain

\[(11.1002) \quad f([b_u, b_q], b_v) = \left[ f([b_u, b_q]), f(b_v) \right] = [[f(b_u), f(b_q)], f(b_v)].\]

Furthermore, the words \(v\) and \(u\) are Lyndon and satisfy \(v < q, u < q, vu \sim p\) and \([b_v, b_q] \in B_{v,q,q}\). Thus, we can apply (11.1001) to \(g = v, h = u\) and \(x = [b_v, b_q]\). As a result, we obtain

\[(11.1003) \quad f([b_v, b_q], b_u) = \left[ f([b_v, b_q]), f(b_u) \right] = [[f(b_v), f(b_q)], f(b_u)].\]

Now, applying the map \(f\) to both sides of the equality (11.984), we obtain

\[f([b_p, b_q]) = f([b_u, b_q], b_v) - [b_v, b_q], b_u)\]

\[= \left[ f([b_u, b_q]), f(b_v) \right] - \left[ f([b_v, b_q]), b_u)\right] \quad \text{(since the map } f \text{ is } k\text{-linear)}\]

\[= [[f(b_u), f(b_q)], f(b_v)] - [[f(b_v), f(b_q)], f(b_u)]\]

\[= [[f(b_u), f(b_q)], f(b_v)] - [f(b_v), f(b_q)], f(b_u)].\]

Comparing this with

\[f([b_p, b_q]) = [f(b_p), f(b_q)], \quad \quad \text{(by (11.997))}\]

\[= [f(b_u), f(b_v)], f(b_v)] - [f(b_v), f(b_q)], f(b_u)].\]

we obtain \(f([b_p, b_q]) = [f(b_p), f(b_q)]\).

Let us now forget that we fixed \((p, q)\). We thus have proven that

\[f([b_p, b_q]) = [f(b_p), f(b_q)] \quad \text{for every } (p, q) \in \mathcal{L} \times \mathcal{L} \text{ satisfying } p < q \text{ and } \ell(pq) = N \text{ and } |\mathfrak{P}(pq)| = K.\]

In other words, (11.995) holds in the case when \(\rho(q) = K\). This completes the induction step (in the induction proof of (11.995)). The induction proof of (11.995) is thus finished.
We thus have proven (11.995). In other words, (11.993) holds in the case when \( \ell(\ell pq) = N \). This completes the induction step (in the induction proof of (11.993)). The induction proof of (11.993) is thus complete.

Using (11.993), we can easily see the following fact: For any \( p \in \mathcal{L} \) and \( q \in \mathcal{L} \), we have
\[
(11.1004) \quad f([b_p, b_q]) = [f(b_p), f(b_q)]
\]

Now, we recall that \( B \) is the \( k \)-submodule of \( \mathfrak{g} \) spanned by the family \((b_w)_{w \in \mathcal{L}}\). Thus, the family \((b_w)_{w \in \mathcal{L}}\) spans the \( k \)-module \( B \).

Now, we have
\[
(11.1005) \quad f([x, y]) = [f(x), f(y)] \quad \text{for every } x \in B \text{ and } y \in B
\]

In other words, the map \( f : B \rightarrow \mathfrak{h} \) is a Lie algebra homomorphism. This solves Exercise 6.40(b).

### 11.120. Solution to Exercise 6.41

**Solution to Exercise 6.41.** The definition of \( \mathfrak{g} \) shows that \( \mathfrak{g} = \mathfrak{g}_1 + \mathfrak{g}_2 + \mathfrak{g}_3 + \cdots = \sum_{i \geq 1} \mathfrak{g}_i \). Thus, for every positive integer \( k \), we have
\[
(11.1006) \quad \mathfrak{g}_k \subset \mathfrak{g}
\]

Notice that \( \mathfrak{g}_1 = V \), so that
\[
(11.1007) \quad V = \mathfrak{g}_1 \subset \mathfrak{g} \quad \text{(by (11.1006), applied to } k = 1).}
\]

We also recall a fundamental property of Lie algebras (one of the forms of the Jacobi identity):

- Every three elements \( x, y \) and \( z \) of a Lie algebra \( \mathfrak{t} \) satisfy
\[
(11.1008) \quad [[x, y], z] = [[x, z], y] - [[y, z], x].
\]

(a) For every two positive integers \( i \) and \( j \), we have
\[
(11.1009) \quad [\mathfrak{g}_i, \mathfrak{g}_j] \subset \mathfrak{g}_{i+j}.
\]

**Proof of (11.1009):** We shall prove (11.1009) by induction over \( i \):

\[\text{Proof of (11.1004): Let } p \in \mathcal{L} \text{ and } q \in \mathcal{L}. \text{ We need to prove that } f([b_p, b_q]) = [f(b_p), f(b_q)]. \text{ If } p = q, \text{ then this is clear (because if } p = q, \text{ then } f\left(\begin{bmatrix} b_p & b_q \\ \ell(pq) = 0 & \ell(pq) \neq 0 \end{bmatrix}\right) = f\left(\begin{bmatrix} b_q \\ \ell(pq) = 0 \end{bmatrix}\right) = f(0) = 0 \text{ and } f\left(\begin{bmatrix} b_p \\ \ell(pq) = 0 \end{bmatrix}\right) = f(b_q). \text{ Thus, for the rest of this proof, we WLOG assume that } p \neq q. \text{ We have } (p, q) \in \mathcal{L} \times \mathcal{L} \text{ (since } p \in \mathcal{L} \text{ and } q \in \mathcal{L}) \text{ and } (q, p) \in \mathcal{L} \times \mathcal{L} \text{ (since } q \in \mathcal{L} \text{ and } p \in \mathcal{L}). \text{ We have } p \neq q \text{ (since we don't have } p = q). \text{ Thus, we have either } p < q \text{ or } p > q \text{ (since the lexicographic order on } \mathcal{L}^* \text{ is a total order). In other words, we are in one of the following two Cases:}
\]

- **Case 1:** We have \( p < q. \)
- **Case 2:** We have \( p > q. \)

Let us first consider Case 1. In this case, we have \( p < q. \) Thus, from (11.993), we obtain \( f([b_p, b_q]) = [f(b_p), f(b_q)]. \) Thus, \( f([b_p, b_q]) = f(b_p), f(b_q) \) is proven in Case 1.

Let us now consider Case 2. In this case, we have \( p > q. \) In other words, \( q < p. \) Thus, (11.993) (applied to \( (q, p) \) instead of \( (p, q) \)) shows that \( f([b_q, b_p]) = [f(b_q), f(b_p)]. \) Now,
\[
\quad f\left(\begin{bmatrix} b_p & b_q \\ \ell(pq) = 0 & \ell(pq) \neq 0 \end{bmatrix}\right) = f(-[b_q, b_p]) = -f\left(\begin{bmatrix} b_q & b_p \\ \ell(pq) = 0 & \ell(pq) \neq 0 \end{bmatrix}\right) = -[f(b_q), f(b_p)] = [f(b_p), f(b_q)].
\]

Thus, \( f([b_p, b_q]) = [f(b_p), f(b_q)] \) is proven in Case 2.

Now, \( f([b_p, b_q]) = [f(b_p), f(b_q)] \) is proven in each of the two Cases 1 and 2. This completes the proof of (11.1004).

\[\text{Proof of (11.1005): Let } x \in B \text{ and } y \in B. \text{ We must prove the equality } f([x, y]) = [f(x), f(y)]. \text{ Both sides of this equality are } k\text{-linear in each of } x \text{ and } y. \text{ Hence, we can WLOG assume that both } x \text{ and } y \text{ belong to the family } (b_w)_{w \in \mathcal{L}} \text{ (since the family } (b_w)_{w \in \mathcal{L}} \text{ spans the } k\text{-module } B). \text{ In other words, we can WLOG assume that there exist } p \in \mathcal{L} \text{ and } q \in \mathcal{L} \text{ satisfying } x = b_p \text{ and } y = b_q. \text{ Assume this, and consider these } p \text{ and } q. \]

From (11.1004), we obtain \( f([b_p, b_q]) = [f(b_p), f(b_q)]. \) This rewrites as \( f([x, y]) = [f(x), f(y)] \) (since \( x = b_p \) and \( y = b_q \)). This proves (11.1005).
**Induction base:** For every positive integer \( j \), we have \([g_1, g_j] \subset g_{1+j} \). In other words, (11.1009) holds for \( i = 1 \). This completes the induction base.

**Induction step:** Let \( I \) be a positive integer. Assume that (11.1009) holds for \( i = I \). We must prove that (11.1009) holds for \( i = I + 1 \).

Now, we must prove the relation (11.1010)

\[
[g_I, g_j] \subset g_{I+j}.
\]

Now, let \( j \) be a positive integer. Thus, the recursive definition of \( g_{j+1} \) yields \( g_{j+1} = \left[ V, g_{(j+1)-1} \right]_{=g_{j+1}} \).

Also, the recursive definition of \( g_{I+j+1} \) yields \( g_{I+j+1} = \left[ V, g_{(I+j+1)-1} \right]_{=g_{I+j+1}} \).

Now, let \( z \in g_j \). Then, \( z \in g_j \subset g \) (by (11.1006), applied to \( k = j \)).

Now, let \( p \in g_{I+1} \). We are going to prove that \([p, z] \in g_{I+j+1} \).

Recall that \([V, g_I] \) is the \( k \)-linear span of all elements of the form \([x, y]\) with \((x, y) \in V \times g_I \) (indeed, this is how \([V, g_I] \) is defined). Thus, \( p \) is a \( k \)-linear combination of elements of the form \([x, y]\) with \((x, y) \in V \times g_I \) (since \( p \in g_{I+1} = [V, g_I] \)).

Now, we must prove the relation \([p, z] \in g_{I+j+1} \). But this relation is \( k \)-linear in \( p \). Thus, we WLOG assume that \( p \) is an element of the form \([x, y]\) for with \((x, y) \in V \times g_I \) (since \( p \) is a \( k \)-linear combination of elements of the form \([x, y]\) with \((x, y) \in V \times g_I \)). In other words, there exists an \((x, y) \in V \times g_I \) such that \( p = [x, y] \). Consider this \((x, y) \).

We have \((x, y) \in V \times g_I \). In other words, \( x \in V \) and \( y \in g_I \). Thus, \( y \in g_I \subset g \) (by (11.1006), applied to \( k = I \)). Also, \( x \in V \subset g \) (by (11.1007)). Now,

\[
\begin{align*}
\left[ \begin{array}{c} p \\ =_{[x, y]} \\
\end{array} \right] & = \left[ [x, y], z \right] = \left[ [x, z], y \right] - \left[ [y, z], x \right] \\
& = -[y, [x, z]] - ([x, y], z] = \left[ \begin{array}{ccc} y & z & x \\ x & y & \in V \\
\in g_I \\
\in g_I \\
\end{array} \right] - \left[ \begin{array}{ccc} y & z & x \\ x & y & \in V \\
\in g_I \\
\in g_I \\
\end{array} \right] \\
\end{align*}
\]

\[
\begin{align*}
& \in \left[ V, g_I, g_J \right]_{g_{I+j+1}} - \left[ V, g_J \right]_{g_{j+1}} \\
& \subset \left[ V, g_{I+j+1} \right]_{g_{I+j+1}} - \left[ V, g_{I+j+1} \right]_{g_{j+1}} \\
& = g_{I+1+j} \\
\end{align*}
\]

Now, let us forget that we fixed \( z \) and \( p \). We thus have proven that \([p, z] \in g_{I+1+j} \) for any \( p \in g_{I+1} \) and \( z \in g_j \). Thus, \([g_{I+1}, g_j] \subset g_{I+1+j} \) (since \( g_{I+1+j} \) is a \( k \)-module).

---

788 **Proof.** Let \( j \) be a positive integer. Thus, the recursive definition of \( g_{1+j} \) yields \( g_{1+j} = \left[ V, g_{(1+j)-1} \right]_{=g_{1+j}} \). In other words, (11.1009) holds for \( i = 1 \). This completes the induction base.

Now, we must prove the relation \([V, g_j] = g_{1+j} \). In other words, \([V, g_j] = g_{1+j} \subset g_{1+j}, \text{ qed.} \)
Let us now forget that we fixed $j$. We thus have proven that, for every positive integer $j$, we have $[g_{i+1}, g_j] \subset g_{i+j}$. In other words, (11.1009) holds for $i = I + 1$. This completes the induction step. The induction proof of (11.1009) is thus complete.

Now, for every two positive integers $i$ and $j$, we have
\[(11.1011)\] holds for $g_{i+j}$. (by (11.1009))

But $g = \sum_{i \geq 1} g_i = \sum_{j \geq 1} g_j$ (here, we have renamed the summation index $i$ as $j$). Thus,
\[\sum_{i \geq 1} \sum_{j \geq 1} [g_i, g_j] \subset g \quad \text{(since the Lie bracket on $T(V)$ is $k$-bilinear)}\]
\[\subset \sum_{i \geq 1} \sum_{j \geq 1} g \subset g \quad \text{(since $g$ is a $k$-module)}.}\]

Thus, $g$ is a Lie subalgebra of $T(V)$. This solves Exercise 6.41(a).

(b) Let $\mathfrak{k}$ be any Lie subalgebra of $T(V)$ satisfying $V \subset \mathfrak{k}$. We must prove that $g \subset \mathfrak{k}$.

We claim that
\[(11.1012)\] holds for every positive integer $k$.

Proof of (11.1012): We shall prove (11.1012) by induction over $k$:

Induction base: We have $g_1 = V \subset \mathfrak{k}$. In other words, (11.1012) holds for $i = 1$. This completes the induction base.

Induction step: Let $K$ be a positive integer. Assume that (11.1012) holds for $k = K$. We must prove that (11.1012) holds for $k = K + 1$.

We have assumed that (11.1012) holds for $k = K$. In other words, we have $g_K \subset \mathfrak{k}$.

But $\mathfrak{k}$ is a Lie subalgebra of $T(V)$. Hence, $[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}$.

Now, the recursive definition of $g_{K+1}$ yields $g_{K+1} = \left[ V \sum_{\mathfrak{k} \subset \mathfrak{k}} g_{(K+1)-1} \right] \subset [\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}$. In other words,
\[(11.1012)\] holds for $k = K + 1$. This completes the induction step. The induction proof of (11.1012) is thus complete.

Now,
\[g = \sum_{i \geq 1} \sum_{\mathfrak{k} \subset \mathfrak{k}} g_i \subset \sum_{i \geq 1} \mathfrak{k} \subset \mathfrak{k}\]
\[\text{in the $k$-algebra $T(V)$} \quad ^{789}\].

\[789\text{Proof of (11.1013): Let } u \in A^* \text{ and } v \in A^*. \text{ We have } uv = \left( \left( (uv)_1, (uv)_2, \ldots, (uv)_{(uv)} \right) \right) \text{ and thus}\]
\[\left( (uv)_1, (uv)_2, \ldots, (uv)_{(uv)} \right) = \sum_{u_1, u_2, \ldots, u_{(uv)}} \left( u_1, u_2, \ldots, u_{(uv)} \right) \]
\[= \left( u_1, u_2, \ldots, u_{(uv)} \right) \left( v_1, v_2, \ldots, v_{(uv)} \right) = \left( u_1, u_2, \ldots, u_{(uv)}, v_1, v_2, \ldots, v_{(uv)} \right) \]
\[\text{and therefore}\]
\[x_{(uv)_1} \otimes x_{(uv)_2} \otimes \cdots \otimes x_{(uv)_{(uv)}} = x_{u_1} \otimes x_{u_2} \otimes \cdots \otimes x_{u_{(uv)}} \otimes x_{v_1} \otimes x_{v_2} \otimes \cdots \otimes x_{v_{(uv)}}.\]

The definition of $x_{uv}$ now yields
\[(11.1014)\]
\[x_{uv} = x_{(uv)_1} \otimes x_{(uv)_2} \otimes \cdots \otimes x_{(uv)_{(uv)}} = x_{u_1} \otimes x_{u_2} \otimes \cdots \otimes x_{u_{(uv)}} \otimes x_{v_1} \otimes x_{v_2} \otimes \cdots \otimes x_{v_{(uv)}}.\]
(c) We shall solve Exercise 6.41(c) by strong induction on $\ell(w)$:

**Induction step:** Let $N \in \mathbb{N}$. Assume that Exercise 6.41(c) holds under the condition that $\ell(w) < N$. We need to show that Exercise 6.41(c) also holds under the condition that $\ell(w) = N$.

We have assumed that Exercise 6.41(c) holds under the condition that $\ell(w) < N$. In other words, we have

$$b_w \in x_w + \sum_{v \in \mathcal{A}(w), v > w} kx_v \quad \text{for every } w \in \mathcal{L} \text{ satisfying } \ell(w) < N. \quad (11.1015)$$

For every $w \in \mathcal{L}$ satisfying $\ell(w) < N$, we have

$$b_w \in x_w + \sum_{v \in \mathcal{A}(w), v > w} kx_v$$

$$= x_w + \sum_{p \in \mathcal{A}(w), p > w} kx_p \quad \text{(here, we renamed the summation index } v \text{ as } p) \quad (11.1016)$$

$$= x_w + \sum_{q \in \mathcal{A}(w), q > w} kx_q \quad \text{(here, we renamed the summation index } p \text{ as } q). \quad (11.1017)$$

From this, it is easy to conclude the following: For every $w \in \mathcal{L}$ satisfying $\ell(w) < N$, we have

$$b_w \in \sum_{q \in \mathcal{A}(w), q \geq w} kx_q. \quad (11.1018)$$

Now, let $w \in \mathcal{L}$ be such that $\ell(w) = N$. We shall prove that $b_w \in x_w + \sum_{r \in \mathcal{A}(w), r > w} kx_r$.

If $\ell(w) \leq 1$, then $b_w \in x_w + \sum_{r \in \mathcal{A}(w), r > w} kx_r$ holds\(^{791}\). Hence, for the rest of this proof, we can WLOG assume that we don’t have $\ell(w) \leq 1$. Assume this.

But the definition of $x_u$ yields $x_u = x_{u_1} \otimes x_{u_2} \otimes \cdots \otimes x_{u_{\ell(u)}}$, and the definition of $x_v$ yields $x_v = x_{v_1} \otimes x_{v_2} \otimes \cdots \otimes x_{v_{\ell(v)}}$.

Thus,

$$x_u = x_{u_1} \otimes x_{u_2} \otimes \cdots \otimes x_{u_{\ell(u)}} = x_{v_1} \otimes x_{v_2} \otimes \cdots \otimes x_{v_{\ell(v)}}$$

$$= \left(x_{u_1} \otimes x_{u_2} \otimes \cdots \otimes x_{u_{\ell(u)}}\right) \left(x_{v_1} \otimes x_{v_2} \otimes \cdots \otimes x_{v_{\ell(v)}}\right)$$

$$= x_{u_1} \otimes x_{u_2} \otimes \cdots \otimes x_{u_{\ell(u)}} \otimes x_{v_1} \otimes x_{v_2} \otimes \cdots \otimes x_{v_{\ell(v)}}$$

$$= x_{uw} \quad \text{(by } (11.1014)) \text{.}$$

This proves (11.1013).

**Proof of (11.1018):** Let $w \in \mathcal{L}$ be such that $\ell(w) < N$.

Every $q \in \mathcal{A}(w)$ satisfying $q > w$ must also satisfy $q \geq w$. Thus, $\sum_{q \in \mathcal{A}(w), q \geq w} kx_q \subseteq \sum_{q \in \mathcal{A}(w), q > w} kx_q$.

On the other hand, $w$ is an element of $\mathcal{A}(w)$ (since $\ell(w) = \ell(w)$) and satisfies $w \geq w$. In other words, $w$ is a $q \in \mathcal{A}(w)$ satisfying $q \geq w$. Thus, $kx_w$ is an addend of the sum $\sum_{q \in \mathcal{A}(w), q \geq w} kx_q$. Hence, $kx_w \subseteq \sum_{q \in \mathcal{A}(w), q \geq w} kx_q$.

Now, $x_w \in kx_w \subseteq \sum_{q \in \mathcal{A}(w), q \geq w} kx_q$. But (11.1017) becomes

$$b_w \in \sum_{q \in \mathcal{A}(w), q \geq w} kx_q \subseteq \sum_{q \in \mathcal{A}(w), q \geq w} kx_q$$

(since $\sum_{q \in \mathcal{A}(w), kx_q \text{ is a } k\text{-module}$). This proves (11.1018).

\(^{791}\) **Proof.** Assume that $\ell(w) \leq 1$. The word $w$ is Lyndon (since $w \in \mathcal{L}$) and thus nonempty. Hence, $\ell(w) \geq 1$. Combined with $\ell(w) \leq 1$, this yields $\ell(w) = 1$. In other words, the word $w$ consists of a single letter. In other words, $w = (a)$ for some $a \in \mathcal{A}$. Consider this $a$. The definition of $b_w$ then yields $b_w = x_a$ (since $\ell(w) = 1$ and $w = (a)$).
The word \( w \) is Lyndon (since \( w \in \mathcal{L} \)) and satisfies \( \ell (w) > 1 \) (since we don’t have \( \ell (w) \leq 1 \)). Thus, \( w \) is a Lyndon word of length \( > 1 \). Therefore, \( \text{stf} \ w \) is well-defined. Let \((u,v) = \text{stf} \ w\). The recursive definition of \( b_w \) yields \( b_w = [b_u,b_v] \) (since \( \ell (w) > 1 \) and \((u,v) = \text{stf} \ w\)).

From Exercise 6.39(a) (applied to \((g,h) = (u,v)\)), we see that \( v \) is the longest Lyndon proper suffix of \( w \). In particular, \( v \) is a Lyndon proper suffix of \( \ell (w) \). Thus, \( \text{stf} \ v = v \). Hence, \( \ell (v) \geq 1 \). Also, Exercise 6.39(d) (applied to \((g,h) = (u,v)\)) shows that the word \( u \) is Lyndon. From Exercise 6.39(c) (applied to \((g,h) = (u,v)\)), we obtain \( u < uv < v \). Finally, Exercise 6.39(b) (applied to \((g,h) = (u,v)\)) shows that \( w = uv \). Exercise 6.39(c) (applied to \((g,h) = (u,v)\)) shows that \( w \in \mathcal{L}, u \in \mathcal{L}, u \subseteq \mathcal{L} \) and \( \ell (u) < \ell (w) \) and \( \ell (v) < \ell (w) \).

The word \( u \) is Lyndon and thus nonempty. Hence, \( \ell (u) \geq 1 \). Also, the word \( v \) is Lyndon and thus nonempty. Thus, \( \ell (v) \geq 1 \).

Now, \( \ell (u) < \ell (w) = N \). Thus, (11.1016) (applied to \( u \)) instead of \( w \) shows that

\[
(11.1019) \quad b_u \in x_u + \sum_{p \in \mathfrak{A}^{\ell (u)} ; \, p > u} kx_p.
\]

Also, (11.1018) (applied to \( u \)) instead of \( w \) shows that

\[
(11.1020) \quad b_u \in \sum_{q \in \mathfrak{A}^{\ell (u)} ; \, q \geq u} kx_q = \sum_{p \in \mathfrak{A}^{\ell (u)} ; \, p \geq u} kx_p
\]

(here, we renamed the summation index \( q \) as \( p \)).

Also, \( \ell (v) < \ell (w) = N \). Thus, (11.1017) (applied to \( v \)) instead of \( w \) shows that

\[
(11.1021) \quad b_v \in x_v + \sum_{q \in \mathfrak{A}^{\ell (v)} ; \, q > u} kx_q.
\]

Also, (11.1018) (applied to \( v \)) instead of \( w \) shows that

\[
(11.1022) \quad b_v \in \sum_{q \in \mathfrak{A}^{\ell (v)} ; \, q \geq u} kx_q.
\]

Now, let \( G = \sum_{r \in \mathfrak{A}^{\ell (w)} , \, r > w} kx_r \). Thus, \( G = \sum_{p \in \mathfrak{A}^{\ell (u)} ; \, p \geq u} kx_p \) is a \( k \)-submodule of \( T(V) \).

If \( p \in \mathfrak{A}^{\ell (u)} \) and \( q \in \mathfrak{A}^{\ell (v)} \) are such that \( q \geq v \), then

\[
(11.1023) \quad x_q x_p \in G.
\]

Hence,

\[
(11.1024) \quad b_v b_u \in G.
\]

On the other hand, the definition of \( x_w \) yields \( x_w = x_a \) (since \( w = (a) \)). Compared with \( b_w = x_a \), this yields

\[
b_w = x_w = x_u + \sum_{r \in \mathfrak{A}^{\ell (w)} ; \, r > w} kx_r
\]

(since \( \sum_{r \in \mathfrak{A}^{\ell (w)} ; \, r > w} kx_r \) is a \( k \)-module)

\[
\ell (qp) = \ell (w) \text{ in other words, } qp \in \mathfrak{A}^{\ell (w)}.
\]

But the word \( w \) is Lyndon and satisfies \( w = uv \). Thus, Proposition 6.14(a) shows that \( v \geq w \) (since \( v \) is nonempty). Since \( \ell (v) \neq \ell (w) \) (because \( \ell (v) < \ell (w) \)), we have \( v \neq w \). Combined with \( v \geq w \), this yields \( v > w \). Now, \( qp \geq q \geq v > w \).

So we know that \( qp \in \mathfrak{A}^{\ell (w)} \) and \( qp > w \). In other words, \( qp \) is an \( r \in \mathfrak{A}^{\ell (w)} \) satisfying \( r > w \). Thus, \( kx_{qp} \) is an addend of the sum \( \sum_{r \in \mathfrak{A}^{\ell (w)} , \, r > w} kx_r \). Therefore, \( kx_{qp} + \sum_{r \in \mathfrak{A}^{\ell (w)} , \, r > w} kx_r = G \) (since \( G = \sum_{r \in \mathfrak{A}^{\ell (w)} , \, r > w} kx_r \)).

But (11.1013) (applied to \( q \) and \( p \) instead of \( u \) and \( v \)) shows that \( x_q x_p = x_{qp} \in kx_{qp} \subset G \). This proves (11.1023).
Furthermore, if \( p \in \mathfrak{A}^{(u)} \) and \( q \in \mathfrak{A}^{(v)} \) are such that \( p > u \), then
\[
(11.1025) \quad x_p x_q \in G
\]

Hence,
\[
(11.1026) \quad (b_u - x_u) b_v \in G
\]

Furthermore, if \( q \in \mathfrak{A}^{(v)} \) is such that \( q > v \), then
\[
(11.1027) \quad x_u x_q \in G
\]

\textbf{Proof of (11.1024):} We have
\[
\sum_{q \in \mathfrak{A}^{(v)}, \ell \geq v} b_u \sum_{p \in \mathfrak{A}^{(u)}, \ell \geq u} k_{x_q} + b_u \sum_{p \in \mathfrak{A}^{(u)}, \ell \geq u} k_{x_p} \subseteq \sum_{q \in \mathfrak{A}^{(v)}, p \in \mathfrak{A}^{(u)}, q \geq v} k_{x_q x_p} \subseteq \sum_{q \in \mathfrak{A}^{(v)}, p \in \mathfrak{A}^{(u)}, q \geq v} k_{G \subseteq G}
\]
(since \( G \) is a \( k \)-module). This proves (11.1024).

\textbf{Proof of (11.1025):} Let \( p \in \mathfrak{A}^{(u)} \) and \( q \in \mathfrak{A}^{(v)} \) be such that \( p > u \).

We have \( \ell(pq) = \ell(p) + \ell(q) \) (since \( p \in \mathfrak{A}^{(u)} \), so that \( \ell(u) = \ell(p) \)).

Assume (for the sake of contradiction) that the word \( u \) is a prefix of \( p \). Since the word \( u \) has the same length as \( p \) (since \( \ell(u) = \ell(p) \)), this shows that \( u = p \). Thus, \( u > p \), which is absurd. This contradiction shows that our assumption (that the word \( u \) is a prefix of \( p \)) was false. In other words, the word \( u \) is not a prefix of \( p \).

We have \( u < p \) (since \( p > u \)), thus \( u \leq p \). Thus, Proposition 6.2(d) (applied to \( u, v, p \) and \( \emptyset \) instead of \( a, b, c \) and \( d \)) shows that either we have \( u \leq p \emptyset \) or the word \( u \) is a prefix of \( p \). Since the word \( u \) is not a prefix of \( p \), we thus conclude that \( u \leq p \emptyset \).

Hence, \( w = w \leq p \emptyset = p \).

Now, \( q \in \mathfrak{A}^{(v)} \), so that \( \ell(q) \geq \ell(v) \geq 1 \). The word \( q \) is thus nonempty. Hence, \( p < pq \). Thus, \( w \leq p < pq \), so that \( pq > w \).

So we know that \( pq \in \mathfrak{A}^{(w)} \) and \( pq > w \). In other words, \( pq \) is an \( r \in \mathfrak{A}^{(w)} \) satisfying \( r > w \). Thus, \( k_{x_{pq}} \) is an addend of the sum \( \sum_{r > w} k_{x_r} \). Therefore, \( k_{x_{pq}} \subseteq \sum_{r > w} k_{x_r} = G \) (since \( G = \sum_{r > w} k_{x_r} \)).

But (11.1013) (applied to \( p \) and \( q \) instead of \( u \) and \( v \)) shows that \( x_p x_q = x_{pq} \in k_{x_{pq}} \in G \). This proves (11.1025).

\textbf{Proof of (11.1026):} Subtracting \( x_u \) from both sides of the relation (11.1019), we obtain
\[
(b_u - x_u) \subseteq \sum_{p \in \mathfrak{A}^{(u)}, p > u} k_{x_p}
\]

Now,
\[
\left( \sum_{p \in \mathfrak{A}^{(u)}, p > u} k_{x_p} \right) \left( \sum_{q \in \mathfrak{A}^{(v)}, q > v} k_{x_q} \right) \subseteq \sum_{p \in \mathfrak{A}^{(u)}, q \in \mathfrak{A}^{(v)}, q > v} k_{x_p x_q} \subseteq \sum_{p \in \mathfrak{A}^{(u)}, q \in \mathfrak{A}^{(v)}, q > v} k_{G \subseteq G}
\]
(since \( G \) is a \( k \)-module). This proves (11.1026).
In other words, we have proven that Exercise 6.41(c) holds under the condition that \( \ell \) (11.1028)

(hence, \( x_u (b_v - x_v) \in G \))

Now,

\[
b_w = [b_u, b_v] = b_u b_v - b_v b_u \quad \text{(by the definition of the Lie bracket on } T(V) )
\]

\[
= \left( \sum_{u \in G} b_u \right) b_v + b_u \left( \sum_{u \in G} b_v - x_v^- x_v \right) - b_v b_u
\]

(by (11.1026)) \( \text{ (by (11.1028))} \) \( \text{ (by (11.1024))} \)

(by straightforward computation)

\[
\in G + G + x_u x_v - G = \sum_{u \in G} x_u x_v + \sum_{u \in G} x_u x_v - \sum_{u \in G} x_u x_v
\]

(11.1013) \( \text{ (since } G \text{ is a } k\text{-module}) \)

\[= x_w + \sum_{r \in A(w)} k x_r. \]

Thus, \( b_w \in x_w + \sum_{r \in A(w)} k x_r \) is proven.

Let us now forget that we defined \((u, v)\). We have

\[
b_w \in x_w + \sum_{r \in A(w)}, k x_r = x_w + \sum_{v \in A(w)} k x_v
\]

(here, we have renamed the summation index \( r \) as \( v \)).

Now, let us forget that we fixed \( w \). We thus have proven that

\[
b_w \in x_w + \sum_{v \in A(w)}, k x_v \quad \text{for every } v \in A \text{ satisfying } \ell(w) = N.
\]

In other words, we have proven that Exercise 6.41(c) holds under the condition that \( \ell(w) = N \). Thus, our induction is complete, and Exercise 6.41(c) is solved.

(d) We know that \( g \) is a Lie subalgebra of \( T(V) \) (by Exercise 6.41(a)). Thus, \([g, g] \subset g\).

---

Proof of (11.1027): Let \( q \in A(w) \) be such that \( q > v \).

We have \( \ell(uq) = \ell(u) + \ell(q) = \ell(u) + \ell(v) \).

Compared with \( \ell(w) = \ell(u) + \ell(v) \), this yields \( \ell(uq) = \ell(w) \).

In other words, \( uq \in A(w) \).

We have \( q > v \), thus \( v < q \), and \( u \leq q \). Hence, Proposition 6.2(b) (applied to \( u \), \( v \) and \( q \) instead of \( a \), \( c \), and \( d \)) shows that \( uv \leq uq \).

If we had \( uw = uq \), then we would have \( v = q \) (since we could cancel \( u \) from the equality \( uw = uq \)), which would contradict \( v < q \).

Thus, we cannot have \( uv = uq \). Hence, we have \( uv \neq uq \).

In other words, \( uq \neq uq \).

So we know that \( uq \in A(w) \) and \( uq \neq w \). In other words, \( uq \) is an \( r \in A(w) \) satisfying \( r > w \). Thus, \( k x_{uq} \) is an addend of the sum \( \sum_{r \in A(w)} k x_r \).

Therefore, \( k x_{uq} \subset \sum_{r \in A(w)} k x_r \).

But (11.1013) (applied to \( u \) and \( q \) instead of \( u \) and \( v \)) shows that \( x_u x_q = x_{uq} \in k x_{uq} \subset G \). This proves (11.1027).

---

Proof of (11.1028): Subtracting \( x_v \) from both sides of the relation (11.1021), we obtain

\[
b_w - x_v \in \sum_{q \in A(v)} k x_q
\]

Now,

\[
x_u \left( b_v - x_v \right) = x_u \left( \sum_{q \in A(v)} k x_q \right) \subset \sum_{q \in A(v)} k x_q
\]

(by (11.1027)) \( \subset \sum_{q \in A(v)} k G \subset G \) \( \text{ (since } G \text{ is a } k\text{-module}) \).

This proves (11.1028).
We first notice that
\begin{equation}
(11.1029) \quad b_w \in \mathfrak{g} \quad \text{for every } w \in \mathfrak{L}.
\end{equation}

**Proof of (11.1029):** We shall prove (11.1029) by strong induction on \(\ell(w)\):

**Induction step:** Let \(N \in \mathbb{N}\). Assume that (11.1029) holds under the condition that \(\ell(w) < N\). We need to show that (11.1029) also holds under the condition that \(\ell(w) = N\).

We have assumed that (11.1029) holds under the condition that \(\ell(w) < N\). In other words, we have
\begin{equation}
(11.1030) \quad b_w \in \mathfrak{g} \quad \text{for every } w \in \mathfrak{L} \text{ satisfying } \ell(w) < N.
\end{equation}

Now, let \(w \in \mathfrak{L}\) be such that \(\ell(w) = N\). We shall prove that \(b_w \in \mathfrak{g}\).

If \(\ell(w) \leq 1\), then \(b_w \in \mathfrak{g}\) holds\(^{798}\). Hence, for the rest of this proof, we can WLOG assume that we don’t have \(\ell(w) \leq 1\). Assume this.

The word \(w\) is Lyndon (since \(w \in \mathfrak{L}\)) and satisfies \(\ell(w) > 1\) (since we don’t have \(\ell(w) \leq 1\)). Thus, \(w\) is a Lyndon word of length \(> 1\). Therefore, \(\text{stf}\, w\) is well-defined. Let \((u, v) = \text{stf}\, w\). The recursive definition of \(b_w\) yields \(b_w = [b_u, b_v]\) (since \(\ell(w) > 1\) and \((u, v) = \text{stf}\, w\)).

Exercise 6.39(e) (applied to \((g, h) = (u, v)\)) shows that \(u \in \mathfrak{L}, v \in \mathfrak{L}, \ell(u) < \ell(w)\) and \(\ell(v) < \ell(w)\).

Now, \(\ell(u) < \ell(w) = N\). Thus, (11.1030) (applied to \(u\) instead of \(w\)) shows that \(b_u \in \mathfrak{g}\).

Also, \(\ell(v) < \ell(w) = N\). Thus, (11.1030) (applied to \(v\) instead of \(w\)) shows that \(b_v \in \mathfrak{g}\).

Now, \(b_w = \left[ b_u, b_v \right] \in [\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{g}\).

Let us now forget that we fixed \(w\). We thus have shown that \(b_w \in \mathfrak{g}\) for every \(w \in \mathfrak{L}\) satisfying \(\ell(w) = N\). In other words, (11.1029) holds under the condition that \(\ell(w) = N\). This completes the induction step. The induction proof of (11.1029) is thus complete.

Now, (11.1029) shows that \(\{b_w\}_{w \in \mathfrak{L}}\) is a family of elements of \(\mathfrak{g}\). Using Exercise 6.40(a), it is easy to conclude that this family \(\{b_w\}_{w \in \mathfrak{L}}\) spans the \(\mathbf{k}\)-module \(\mathfrak{g}\)\(^{799}\).

We shall now show that the family \(\{b_w\}_{w \in \mathfrak{L}}\) is \(\mathbf{k}\)-linearly independent. In order to do so, we will prove a more general result:

**Proposition 11.39.** Let \(\mathfrak{g}\) be a \(\mathbf{k}\)-module, and let \(W\) be a totally ordered set. Let \(\{p_w\}_{w \in \mathfrak{L}}\) be a \(\mathbf{k}\)-linearly independent family of elements of \(\mathfrak{G}\). Let \(L\) be a subset of \(W\). Let \(\{s_w\}_{w \in L}\) be a family of elements of \(\mathfrak{G}\).

---

\(^{798}\)Proof: Assume that \(\ell(w) \leq 1\). The word \(w\) is Lyndon (since \(w \in \mathfrak{L}\)) and thus nonempty. Hence, \(\ell(w) \geq 1\). Combined with \(\ell(w) \leq 1\), this yields \(\ell(w) = 1\). In other words, the word \(w\) consists of a single letter. In other words, \(w = (a)\) for some \(a \in \mathfrak{A}\). Consider this \(a\). The definition of \(b_w\) then yields \(b_w = x_a\) (since \(\ell(w) = 1\) and \(w = (a)\)).

Now, \(b_w = x_a \in V \subset \mathfrak{g}\). qed.

\(^{799}\)Proof: Let \(B\) be the \(\mathbf{k}\)-submodule of \(\mathfrak{g}\) spanned by the family \(\{b_w\}_{w \in \mathfrak{L}}\) (this is well-defined since \(\{b_w\}_{w \in \mathfrak{L}}\) is a family of elements of \(\mathfrak{g}\)). Then, clearly,
\begin{equation}
(11.1031) \quad b_w \in B \quad \text{for every } w \in \mathfrak{L}.
\end{equation}

Whenever \(w\) is a Lyndon word of length \(\geq 1\), we have
\[b_w = [b_u, b_v], \quad \text{where } (u, v) = \text{stf}\, w\]
(due to the recursive definition of \(b_w\)). Thus, Exercise 6.40(a) shows that \(B\) is a Lie subalgebra of \(\mathfrak{g}\). Hence, \(B\) is a Lie subalgebra of \(T(V)\).

Recall that \((x_a)_{a \in \mathfrak{A}}\) is a basis of the \(\mathbf{k}\)-module \(V\). Hence,
\begin{equation}
(11.1032) \quad \text{(the } \mathbf{k}\text{-linear span of the family } (x_a)_{a \in \mathfrak{A}}\text{)} = V.
\end{equation}

Now, let \(a \in \mathfrak{A}\). We shall show that \(x_a \in B\).

Indeed, let \(w\) be the one-letter word \((a)\). Then, \(w\) is Lyndon (since \(w\) is a one-letter word) and has length 1. Thus, \(b_w = x_a\) (by the definition of \(b_w\), since \(w = (a)\)). Hence, \(x_a = b_w \in B\) (by (11.1031)).

Now, let us forget that we fixed \(a\). We thus have shown that \(x_a \in B\) for every \(a \in \mathfrak{A}\). Since \(B\) is a \(\mathbf{k}\)-module, this entails that \((\text{the } \mathbf{k}\text{-linear span of the family } (x_a)_{a \in \mathfrak{A}}\) \subset B\). Because of (11.1032), this rewrites as \(V \subset B\).

Now, we know that \(B\) is a Lie subalgebra of \(T(V)\) satisfying \(V \subset B\). Thus, \(\mathfrak{g} \subset B\) (by Exercise 6.41(b), applied to \(t = B\)). Combined with \(B \subset \mathfrak{g}\) (since \(B\) is a \(\mathbf{k}\)-submodule of \(\mathfrak{g}\)), this yields \(\mathfrak{g} = B\). But the family \(\{b_w\}_{w \in \mathfrak{L}}\) spans the \(\mathbf{k}\)-module \(B\) (since \(B\) is the \(\mathbf{k}\)-submodule of \(\mathfrak{g}\) spanned by the family \(\{b_w\}_{w \in \mathfrak{L}}\)). Since \(\mathfrak{g} = B\), this rewrites as follows: The family \(\{b_w\}_{w \in \mathfrak{L}}\) spans the \(\mathbf{k}\)-module \(\mathfrak{g}\). Qed.
Assume that every \( w \in L \) satisfies
\[
\tag{11.1033}
 s_w \in p_w + \sum_{v \in W; \ v \succ w} k p_v.
\]

(Here, the "\( \succ \)" sign under the sum refers to the total order on \( W \).) Then, the family \( (s_w)_{w \in L} \) is \( k \)-linearly independent.

Proposition 11.39 is actually a standard criterion for linear independence. It is often summarized by the motto "vectors with distinct leading coordinates are linearly independent". For the sake of completeness, we shall nevertheless prove it. First, let us state an obvious lemma:

**Lemma 11.40.** Let \( \mathfrak{F} \) be a \( k \)-module, and let \( W \) be a set. Let \( (p_w)_{w \in W} \) be a \( k \)-linearly independent family of elements of \( \mathfrak{F} \). Let \( x \in W \) and \( \lambda \in k \) be such that
\[
\tag{11.1034}
 0 \in \lambda p_x + \sum_{v \in W \setminus \{x\}} k p_v.
\]
Then, \( \lambda = 0 \).

**Proof of Lemma 11.40.** The equality (11.1034) shows that we can write 0 as a \( k \)-linear combination of the elements \( p_v \) with \( v \in W \) in such a way that the coefficient in front of \( p_x \) is \( \lambda \). But since \( (p_w)_{w \in W} \) is \( k \)-linearly independent, the only such linear combination which gives 0 is the trivial one (i.e., the one where all the coefficients are 0). Hence, all coefficients in our linear combination are 0. In particular, the coefficient in front of \( p_x \) is 0. Since this coefficient is \( \lambda \), this means that \( \lambda = 0 \). This proves Lemma 11.40.

**Proof of Proposition 11.39.** Let us first prove the following fact: If \( S \) is a finite subset of \( L \), then
\[
\tag{11.1035}
 \text{the family } (s_w)_{w \in S} \text{ is } k \text{-linearly independent.}
\]

**Proof of (11.1035):** We shall prove (11.1035) by induction over \( |S| \):

1. **Induction base:** If \( S \) is a finite subset of \( L \) satisfying \( |S| = 0 \), then the family \( (s_w)_{w \in S} \) is \( k \)-linearly independent (because \( |S| = 0 \) entails \( S = \emptyset \), and thus the family \( (s_w)_{w \in S} \) is empty). In other words, (11.1035) holds in the case when \( |S| = 0 \). This completes the induction base.
2. **Induction step:** Let \( K \) be a positive integer. Assume that (11.1035) is proven in the case when \( |S| = K - 1 \). We must prove that (11.1035) holds in the case when \( |S| = K \).

We have assumed that (11.1035) is proven in the case when \( |S| = K - 1 \). In other words, if \( S \) is a finite subset of \( L \) satisfying \( |S| = K - 1 \), then
\[
\tag{11.1036}
 \text{the family } (s_w)_{w \in S} \text{ is } k \text{-linearly independent.}
\]

Now, let \( S \) be a finite subset of \( L \) satisfying \( |S| = K \). Let \( (\lambda_w)_{w \in S} \in k^S \) be a family of elements of \( k \) such that \( \sum_{w \in S} \lambda_w s_w = 0 \). We shall prove that \( (\lambda_w)_{w \in S} = (0)_{w \in S} \).

We have \( |S| = K > 0 \). Thus, the set \( S \) is nonempty. Also, \( S \subset L \subset W \), and thus the set \( S \) is totally ordered (since it is a subset of the totally ordered set \( W \)). Therefore, this set \( S \) has a smallest element (since every nonempty finite totally ordered set has a smallest element). Let \( x \) be this smallest element. Thus,
\[
\tag{11.1037}
 x \leq w \quad \text{for every } w \in S
\]
(by the definition of the smallest element). Also, \( x \in S \) (since \( x \) is the smallest element of \( S \)), and thus \( |S \setminus \{x\}| = |S| - 1 = K - 1 \). Hence, we can apply (11.1036) to \( S \setminus \{x\} \) instead of \( S \). As a result, we see that the family \( (s_w)_{w \in S \setminus \{x\}} \) is \( k \)-linearly independent. In other words, if \( (\mu_w)_{w \in S \setminus \{x\}} \in k^{S \setminus \{x\}} \) is a family of elements of \( k \) such that \( \sum_{w \in S \setminus \{x\}} \mu_w s_w = 0 \), then
\[
\tag{11.1038}
 (\mu_w)_{w \in S \setminus \{x\}} = (0)_{w \in S \setminus \{x\}}.
\]

On the other hand, every \( w \in S \setminus \{x\} \) satisfies
\[
\tag{11.1039}
 s_w \in \sum_{v \in W \setminus \{x\}} k p_v
\]
Also,\footnote{Proof of (11.1039): Let \(w \in S \setminus \{x\}\). Thus, \(w \in S\) and \(w \neq x\). Let \(v \in \{q \in W \mid q > w\}\). We shall show that \(v \in W \setminus \{x\}\). Indeed, we have \(v \in \{q \in W \mid q > w\}\). In other words, \(v\) is an element of \(W\) and satisfies \(v > w\). But \(w \in S \setminus \{x\} \subset S\), so that (11.1037) shows that \(x \leq w\). Hence, \(w \geq x\), so that \(v > w \geq x\). Thus, \(v \neq x\). Combining \(v \in W\) with \(v \neq x\), we obtain \(v \in W \setminus \{x\}\). Let us now forget that we fixed \(v\). We thus have proven that every \(v \in \{q \in W \mid q > w\}\) satisfies \(v \in W \setminus \{x\}\). In other words, \(\{q \in W \mid q > w\} \subset W \setminus \{x\}\). Therefore, \(\sum_{v \in \{q \in W \mid q > w\}} k_{pv} \subset \sum_{v \in W \setminus \{x\}} k_{pv}\). Also, combining \(w \in W\) with \(w \neq x\), we obtain \(w \in W \setminus \{x\}\). Hence, \(k_{pw}\) is an addend of the sum \(\sum_{v \in W \setminus \{x\}} k_{pv}\). Thus, \(k_{pw} \subset \sum_{v \in W \setminus \{x\}} k_{pv}\). Now, \(p_w \in k_{pw} \subset \sum_{v \in W \setminus \{x\}} k_{pv}\). Now, \(w \in S \subset L \subset W\). Hence, (11.1033) yields
\[
 s_w \in \sum_{\{q \in W \mid q > w\}} k_{pv} \subset \sum_{v \in W \setminus \{x\}} k_{pv} \subset \sum_{v \in W \setminus \{x\}} k_{pv}
\]
(since \(\sum_{v \in W \setminus \{x\}} k_{pv}\) is a \(k\)-module). This proves (11.1039).\footnote{Proof of (11.1040): Let \(v \in \{q \in W \mid q > x\}\). We shall show that \(v \in W \setminus \{x\}\). Indeed, we have \(v \in \{q \in W \mid q > x\}\). In other words, \(v\) is an element of \(W\) and satisfies \(v > x\). Hence, \(v \neq x\) (since \(v > x\)). Combining \(v \in W\) with \(v \neq x\), we obtain \(v \in W \setminus \{x\}\). Let us now forget that we fixed \(v\). We thus have proven that every \(v \in \{q \in W \mid q > x\}\) satisfies \(v \in W \setminus \{x\}\). In other words, \(\{q \in W \mid q > x\} \subset W \setminus \{x\}\). Therefore, \(\sum_{v \in \{q \in W \mid q > x\}} k_{pv} \subset \sum_{v \in W \setminus \{x\}} k_{pv}\). Now, \(x \in S \subset L \subset W\). Hence, (11.1033) (applied to \(w = x\)) yields
\[
 s_x \in p_x + \sum_{v \in W \setminus \{x\}} k_{pv}
\]
Subtracting \(p_x\) from both sides of this relation, we obtain
\[
 s_x - p_x \in \sum_{v \in W \setminus \{x\}} k_{pv} \subset \sum_{v \in W \setminus \{x\}} k_{pv}
\]
This proves (11.1040).}
Lemma 11.40 (applied to $\lambda = \lambda_x$) thus shows that $\lambda_x = 0$. Thus,

$$0 = \lambda_x s_x + \sum_{w \in S; \ w \neq x} \lambda_w s_w = 0 s_x + \sum_{w \in S \setminus \{x\}} \lambda_w s_w = \sum_{w \in S \setminus \{x\}} \lambda_w s_w.$$ 

Thus, the family $(\lambda_w)_{w \in S \setminus \{x\}} \in k^S \setminus \{x\}$ satisfies $\sum_{w \in S \setminus \{x\}} \lambda_w s_w = 0$. Hence, $(\lambda_w)_{w \in S \setminus \{x\}} = (0)_{w \in S \setminus \{x\}}$ (according to (11.1038), applied to $(\mu_w)_{w \in S \setminus \{x\}} = (\lambda_w)_{w \in S \setminus \{x\}}$). In other words,

$(11.1041) \quad \lambda_w = 0 \quad \text{for every } w \in S \setminus \{x\}.$

Combining this with the fact that $\lambda_x = 0$, we conclude that $\lambda_w = 0$ for every $w \in S$. In other words, $(\lambda_w)_{w \in S} = (0)_{w \in S}$.

Now, let us forget that we fixed $(\lambda_w)_{w \in S}$. We thus have shown that if $(\lambda_w)_{w \in S} \in k^S$ is a family of elements of $k$ such that $\sum_{w \in S} \lambda_w s_w = 0$, then $(\lambda_w)_{w \in S} = (0)_{w \in S}$. In other words, the family $(s_w)_{w \in S}$ is $k$-linearly independent.

Let us now forget that we fixed $S$. We thus have proven that if $S$ is a finite subset of $L$ satisfying $|S| = K$, then the family $(s_w)_{w \in S}$ is $k$-linearly independent. In other words, (11.1035) holds in the case when $|S| = K$.

This completes the induction step. Thus, we have proven (11.1035) by induction.

Now, a family $f$ of vectors in a $k$-module is $k$-linearly independent if every finite subfamily of $f$ is $k$-linearly independent (because every linear dependence relation for $f$ has only finitely many nonzero coefficients, and can thus be recast as a linear dependence relation for some finite subfamily of $f$). Now, (11.1035) shows that every finite subfamily of the family $(s_w)_{w \in L}$ is $k$-linearly independent; therefore, the previous sentence shows that the family $(s_w)_{w \in L}$ is $k$-linearly independent. This proves Proposition 11.39. \qed

Let us now return to the solution of Exercise 6.41(d). The set $A^*$ is totally ordered (by the lexicographic order), and the set $L$ is a subset of $A^*$. The family $(x_w)_{w \in A^*}$ is a basis of the $k$-module $T(V)$, and thus is a $k$-linearly independent family of elements of this $k$-module. The family $(b_w)_{w \in L}$ is a family of elements of $T(V)$. Every $w \in L$ satisfies

$$b_w \in x_w + \sum_{v \in A^* \setminus \{v \in A^*: v > w\}} k x_v \quad \text{(by Exercise 6.41(c))}$$

$$\subset x_w + \sum_{v \in A^*: v > w} k x_v \quad \text{(since $A^* \setminus \{v \in A^*: v > w\}$)}$$

Hence, Proposition 11.39 (applied to $A = T(V)$, $W = A^*$, $(p_w)_{w \in A^*} = (x_w)_{w \in A^*}$, $L = L$ and $(s_w)_{w \in L} = (b_w)_{w \in L}$) shows that the family $(b_w)_{w \in L}$ is $k$-linearly independent. Combining this with the fact that this family $(b_w)_{w \in L}$ spans the $k$-module $g$, we therefore conclude that the family $(b_w)_{w \in L}$ is a basis of the $k$-module $g$. This solves Exercise 6.41(d).

(c) Let us first show a simple lemma:

**Lemma 11.41.** Let $a$ and $b$ be two $k$-Lie algebras. Let $p : a \to b$ and $q : a \to b$ be two Lie algebra homomorphisms. Then, $Ker (p - q)$ is a Lie subalgebra of $a$.

**Proof of Lemma 11.41.** Let $x \in Ker (p - q)$ and $y \in Ker (p - q)$. Then, $(p-q)(x) = 0$ (since $x \in Ker (p-q)$), so that $0 = (p-q)(x) = p(x) - q(x)$, and thus $p(x) = q(x)$. The same argument (applied to $y$ instead of $x$) shows that $p(y) = q(y)$.

But $p$ is a Lie algebra homomorphism, and thus satisfies $p([x,y]) = \begin{bmatrix} p(x), p(y) \end{bmatrix}_{[q(x), q(y)]} = q([x,y])$ (since $q$ is a Lie algebra homomorphism). Now, $(p-q)([x,y]) = p([x,y]) - q([x,y]) = q([x,y]) - q([x,y]) = 0$. In other words, $[x,y] \in Ker (p-q)$.
Let us now forget that we fixed $x$ and $y$. We thus have proven that $[x, y] \in \text{Ker} (p - q)$ whenever $x \in \text{Ker} (p - q)$ and $y \in \text{Ker} (p - q)$. In other words, the set $\text{Ker} (p - q)$ is closed under the Lie bracket. Also, clearly, $p - q$ is a $\mathbb{k}$-linear map (since both maps $p$ and $q$ are $\mathbb{k}$-linear), and therefore its kernel $\text{Ker} (p - q)$ is a $\mathbb{k}$-submodule of $\mathfrak{a}$. Thus, $\text{Ker} (p - q)$ is a $\mathbb{k}$-submodule of $\mathfrak{a}$ which is closed under the Lie bracket. In other words, $\text{Ker} (p - q)$ is a Lie subalgebra of $\mathfrak{a}$. This proves Lemma 11.41. □

Now, it is easy to see that if $\Xi_1$ and $\Xi_2$ are two Lie algebra homomorphisms $\Xi : \mathfrak{g} \to \mathfrak{h}$ such that every $a \in \mathfrak{A}$ satisfies $\Xi (x_a) = \xi (a)$, then $\Xi_1 = \Xi_2$. In other words, there exists at most one Lie algebra homomorphism $\Xi : \mathfrak{g} \to \mathfrak{h}$ such that every $a \in \mathfrak{A}$ satisfies $\Xi (x_a) = \xi (a)$. We shall now show that there actually exists such a homomorphism.

For every $w \in \mathfrak{L}$, we define an element $z_w$ of $\mathfrak{h}$ as follows: We define $z_w$ by recursion on the length of $w$.

If the length of $w$ is 1, then we have $w = (a)$ for some letter $a \in \mathfrak{A}$, and we set $z_w = \xi (a)$ for this letter $a$. If the length of $w$ is $> 1$, then we set $z_w = [z_{w_1}, z_{w_2}]$, where $(u, v) = \text{stf} w$.

Thus, we have defined a $z_w \in \mathfrak{h}$ for every $w \in \mathfrak{L}$. Now, Exercise 6.41(d) shows that the family $(b_w)_{w \in \mathfrak{L}}$ is a basis of the $\mathbb{k}$-module $\mathfrak{g}$. Thus, we can define a $\mathbb{k}$-module homomorphism $f : \mathfrak{g} \to \mathfrak{h}$ by requiring that

$$(f(b_w)) = z_w \quad \text{for every } w \in \mathfrak{L}.$$

Consider this $f$. Whenever $w$ is a Lyndon word of length $> 1$, we have

$$(11.1043) \quad (f([b_u, b_v])) = (f(b_u), f(b_v)), \quad \text{where } (u, v) = \text{stf} w.$$

---

\textbf{Proof}. Let $\Xi_1$ and $\Xi_2$ be two Lie algebra homomorphisms $\Xi : \mathfrak{g} \to \mathfrak{h}$ such that every $a \in \mathfrak{A}$ satisfies $\Xi (x_a) = \xi (a)$. We must prove that $\Xi_1 = \Xi_2$.

Both $\Xi_1$ and $\Xi_2$ are $\mathbb{k}$-linear maps, and therefore $\Xi_1 - \Xi_2$ is a $\mathbb{k}$-linear map as well. Hence, its kernel $\text{Ker} (\Xi_1 - \Xi_2)$ is a $\mathbb{k}$-module of $\mathfrak{g}$.

We know that $\Xi_1$ is a Lie algebra homomorphisms $\Xi : \mathfrak{g} \to \mathfrak{h}$ such that every $a \in \mathfrak{A}$ satisfies $\Xi (x_a) = \xi (a)$. Thus, every $a \in \mathfrak{A}$ satisfies $\Xi (x_a) = \xi (a)$. Similarly, every $a \in \mathfrak{A}$ satisfies $\Xi (x_a) = \xi (a)$. Hence, every $a \in \mathfrak{A}$ satisfies $\Xi (x_a) = \xi (a)$. Thus, the set $\text{Ker} (\Xi_1 - \Xi_2)$ contains $x_a$ for every $a \in \mathfrak{A}$. Since $\text{Ker} (\Xi_1 - \Xi_2)$ is a $\mathbb{k}$-submodule of $\mathfrak{g}$, this shows that

$$\text{Ker} (\Xi_1 - \Xi_2) \supset (\text{the } \mathbb{k}\text{-linear span of the family } (x_a)_{a \in \mathfrak{A}}).$$

Since (the $\mathbb{k}$-linear span of the family $(x_a)_{a \in \mathfrak{A}}$) = $V$ (because the family $(x_a)_{a \in \mathfrak{A}}$ is a basis of the $\mathbb{k}$-module $V$), this rewrites as

$\text{Ker} (\Xi_1 - \Xi_2) \supset V$. In other words, $V \subset \text{Ker} (\Xi_1 - \Xi_2)$.

From Lemma 11.41 (applied to $\mathfrak{a} = \mathfrak{g}$, $\mathfrak{b} = \mathfrak{h}$, $p = \Xi_1$ and $q = \Xi_2$), we see that $\text{Ker} (\Xi_1 - \Xi_2)$ is a Lie subalgebra of $\mathfrak{g}$. Since $\mathfrak{g}$ (in turn) is a Lie subalgebra of $T (V)$, this shows that $\text{Ker} (\Xi_1 - \Xi_2)$ is a Lie subalgebra of $T (V)$.

Thus, we know that $\text{Ker} (\Xi_1 - \Xi_2)$ is a Lie subalgebra of $T (V)$ satisfying $V \subset \text{Ker} (\Xi_1 - \Xi_2)$. Exercise 6.41(b) thus shows that $\mathfrak{g} \subset \text{Ker} (\Xi_1 - \Xi_2)$. Hence, $\Xi_1 - \Xi_2 = 0$, so that $\Xi_1 = \Xi_2$. Qed.

The length of any $w \in \mathfrak{L}$ must be at least 1. (Indeed, if $w \in \mathfrak{L}$, then the word $w$ is Lyndon and thus nonempty, and hence its length must be at least 1.)

This is well-defined, because $z_u$ and $z_v$ have already been defined. □

Proof. Let $(u, v) = \text{stf} w$. Then, Exercise 6.39(e) (applied to $(\mathfrak{g}, \mathfrak{h}) = (u, v)$) shows that $u \in \mathfrak{L}$, $v \in \mathfrak{L}$, $\ell (u) < \ell (w)$ and $\ell (v) < \ell (w)$. Recall that we are defining $z_w$ by recursion on the length of $w$. Hence, $z_{u^p}$ is already defined for every $p \in \mathfrak{L}$ satisfying $\ell (p) < \ell (w)$. Applying this to $p = u$, we see that $z_u$ is already defined (since $u \in \mathfrak{L}$ and $\ell (u) < \ell (w)$). The same argument (but applied to $v$ instead of $u$) shows that $z_v$ is already defined. Hence, $z_u$ and $z_v$ have already been defined. Thus, $z_w$ is well-defined by $z_w = [z_u, z_v]$, qed.]
But recall that the family \((b_w)_{w \in \mathcal{L}}\) is a basis of the \(k\)-module \(\mathfrak{g}\). Thus, this family spans the \(k\)-module \(\mathfrak{g}\). In other words, \(\mathfrak{g}\) is the \(k\)-submodule of \(\mathfrak{g}\) spanned by the family \((b_w)_{w \in \mathcal{L}}\). Thus, Exercise 6.40(b) (applied to \(B = \mathfrak{g}\)) shows that \(f\) is a Lie algebra homomorphism.

We notice that

\[
every \ a \in \mathfrak{A} \ satisfies \ f(x_a) = \xi(a)
\]

Thus, \(f\) is a Lie algebra homomorphism \(\mathfrak{g} \to \mathfrak{h}\) having the property that every \(a \in \mathfrak{A}\) satisfies \(f(x_a) = \xi(a)\). Therefore, there exists at least one Lie algebra homomorphism \(\Xi : \mathfrak{g} \to \mathfrak{h}\) such that every \(a \in \mathfrak{A}\) satisfies \(\Xi(x_a) = \xi(a)\) (namely, \(\Xi = f\)). Combining this with the fact that there exists at most one Lie algebra homomorphism \(\Xi : \mathfrak{g} \to \mathfrak{h}\) such that every \(a \in \mathfrak{A}\) satisfies \(\Xi(x_a) = \xi(a)\). This solves Exercise 6.41(e).

### 11.121. Solution to Exercise 6.49. Solution to Exercise 6.49.

**Proof of Remark 6.48.** (a) Let \(I\) and \(J\) be two nonempty intervals of \(\mathbb{Z}\) satisfying \(I < J\). Then,

\[
(11.1044) \quad \text{every } i \in I \text{ and } j \in J \text{ satisfy } i < j
\]

(by the definition of \(I < J\), since \(I < J\)).

Let \(p \in I \cap J\). Then, \(p \in I \cap J \subseteq I\) and \(p \in I \cap J \subseteq J\). Hence, \(p < p\) (by (11.1044)), which is absurd.

Now, let us forget that we fixed \(p\). Thus, we have obtained a contradiction for every \(p \in I \cap J\). Hence, there exists no \(p \in I \cap J\). In other words, \(I \cap J = \emptyset\), so that the sets \(I\) and \(J\) are disjoint. This proves Remark 6.48(a).

(b) Let \(I\) and \(J\) be two disjoint nonempty intervals of \(\mathbb{Z}\). We need to prove that \(I < J\) or \(J < I\).

Let \(i_0\) be the smallest element of \(I\) (this exists, since \(I\) is nonempty), and let \(j_0\) be the smallest element of \(J\) (this exists, since \(J\) is nonempty). We WLOG assume that \(i_0 \leq j_0\) (since otherwise, we can simply interchange \(I\) with \(J\)).

The intervals \(I\) and \(J\) are disjoint, and thus \(I \cap J = \emptyset\).

Now, let \(i \in I\) and \(j \in J\) be arbitrary. Assume (for the sake of contradiction) that \(i \geq j\). Notice that \(j \geq j_0\) (since \(j\) is an element of \(J\), whereas \(j_0\) is the smallest element of \(J\)), so that \(i \geq j \geq j_0\), so that \(j_0 \leq i\).

Write the interval \(I\) in the form \([p : q]^+\) for some \(p \in \mathbb{Z}\) and \(q \in \mathbb{Z}\). Since \(i \in I = [p : q]^+ = \{p + 1, p + 2, \ldots, q\}\), we have \(p < i \leq q\). Since \(i_0 \in I = [p : q]^+ = \{p + 1, p + 2, \ldots, q\}\), we have \(p < i_0 \leq q\).

Now, \(p < i_0 \leq j_0\) and \(j_0 \leq i \leq q\). Hence, \(p < j_0 \leq q\) and thus \(j_0 \in \{p + 1, p + 2, \ldots, q\} = [p : q]^+ = I\). Combined with \(j_0 \in J\) (since \(j_0\) is the smallest element of \(J\)), this yields \(j_0 \in I \cap J = \emptyset\), which is absurd. Hence, our assumption (that \(i \geq j\)) was wrong. We thus have \(i < j\).

**Proof of (11.1043):** Let \(w\) be a Lyndon word of length \(> 1\). Let \((u, v) = \text{stf} \ w\). Thus, \((u, v) = \text{stf} \ w \in \mathcal{L} \times \mathcal{L}\).

The recursive definition of \(z_w = [z_u, z_v]\) (since \(w\) is a Lyndon word of length \(> 1\), and since \((u, v) = \text{stf} \ w\)). The recursive definition of \(b_w\) shows that \(b_w = [b_u, b_v]\) (for the same reasons). Thus, \([b_u, b_v] = b_w\), so that

\[
f\left(\begin{array}{c}
[b_u, b_v] \\
\mathcal{L}
\end{array}\right) = f(b_w) = z_w \quad \text{(by (11.1042))}
\]

Comparing this with

\[
\begin{pmatrix}
\begin{bmatrix} f(b_u) \end{bmatrix} \\
\mathcal{L}
\end{pmatrix} = \begin{bmatrix} z_u, z_v \end{bmatrix},
\begin{pmatrix} f(b_v) \end{bmatrix} = \begin{bmatrix} z_u, z_v \end{bmatrix},
\end{pmatrix}
\]

we obtain \(f([b_u, b_v]) = [f(b_u), f(b_v)]\). This proves (11.1043).

**Proof.** Let \(a \in \mathfrak{A}\). Let \(w\) be the one-letter word \((a)\). Then, \(w\) is a Lyndon word (since \(w\) is a one-letter word) and has length \(1\). Thus, the definition of \(z_w\) shows that \(z_w = \xi(a)\) (since \(w = (a)\)).

But the definition of \(b_w\) yields \(b_w = x_a\) (since \(\ell(w) = 1\) and \(w = (a)\)). Hence, \(f\left(\begin{bmatrix} b_w \end{bmatrix} = f(x_a)\right)\), so that \(f(x_a) = f(b_w) = z_w\) (by (11.1042)) and thus \(f(x_a) = z_w = \xi(a)\), qed.
Now, forget that we have fixed $i$ and $j$. We thus have shown that every $i \in I$ and $j \in J$ satisfy $i < j$. In other words, $I < J$ (by the definition of $I < J$). Hence, $I < J$ or $J < I$. This proves Remark 6.48(b).

(c) For any $\ell$-tuple $(I_1, I_2, \ldots, I_\ell)$ of nonempty intervals of $\mathbb{Z}$, we can state the following three properties (which might and might not be satisfied):

- **Property C1:** The intervals $I_1, I_2, \ldots, I_\ell$ form a set partition of the set $[0 : n]^+$, where $n = |\alpha|$.
- **Property C2:** We have $I_1 < I_2 < \cdots < I_\ell$.
- **Property C3:** We have $|I_i| = \alpha_i$ for every $i \in \{1, 2, \ldots, \ell\}$. We have to prove that the interval system intsys $\alpha$ is the unique $\ell$-tuple $(I_1, I_2, \ldots, I_\ell)$ of nonempty intervals of $\mathbb{Z}$ satisfying these three properties C1, C2 and C3.

First, it is easy to check that the interval system intsys $\alpha$ is an $\ell$-tuple $(I_1, I_2, \ldots, I_\ell)$ of nonempty intervals of $\mathbb{Z}$ satisfying these three properties C1, C2 and C3. It remains to prove that it is the only such $\ell$-tuple.

So, let us fix any $\ell$-tuple $(I_1, I_2, \ldots, I_\ell)$ of nonempty intervals of $\mathbb{Z}$ satisfying the three properties C1, C2 and C3. We need to prove that this $\ell$-tuple $(I_1, I_2, \ldots, I_\ell)$ is intsys $\alpha$. In order to prove this, it is enough to show that

\[
(11.1046) \quad I_i = \left[ \sum_{k=1}^{i-1} \alpha_k : \sum_{k=1}^i \alpha_k \right]^+ \quad \text{for every } i \in \{1, 2, \ldots, \ell\}
\]

(according to the definition of intsys $\alpha$). So it remains to prove (11.1046).

Set $n = |\alpha|$. The intervals $I_1, I_2, \ldots, I_\ell$ form a set partition of the set $[0 : n]^+$ (according to Property C1). Thus, the intervals $I_1, I_2, \ldots, I_\ell$ are pairwise disjoint, and satisfy $I_1 \cup I_2 \cup \cdots \cup I_\ell = [0 : n]^+$.}

807 **Proof.** Let $(I_1, I_2, \ldots, I_\ell)$ denote the interval system intsys $\alpha$. We need to prove that the Properties C1, C2 and C3 are satisfied.

Let us first recall that

\[
(11.1045) \quad I_i = \left[ \sum_{k=1}^{i-1} \alpha_k : \sum_{k=1}^i \alpha_k \right]^+ \quad \text{for every } i \in \{1, 2, \ldots, \ell\}
\]

(by the definition of intsys $\alpha$).

**Proof that Property C1 is satisfied:** Let $n = |\alpha|$. Then, $\sum_{i=1}^\ell \alpha_i = |\alpha| = n$. Clearly, $I_i \in [0 : n]^+$ for every $i \in \{1, 2, \ldots, \ell\}$. We have

\[
0 = \sum_{k=1}^{\ell} \alpha_k \leq \sum_{k=1}^i \alpha_k \leq \cdots \leq \sum_{k=1}^{\ell} \alpha_k = n
\]

(since $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ are positive integers). Hence, for every $x \in [0 : n]^+$, there exists precisely one $i \in \{1, 2, \ldots, \ell\}$ satisfying $\sum_{k=1}^{i-1} \alpha_k < x \leq \sum_{k=1}^i \alpha_k$. In other words, for every $x \in [0 : n]^+$, there exists precisely one $i \in \{1, 2, \ldots, \ell\}$ satisfying $x \in \left[ \sum_{k=1}^{i-1} \alpha_k : \sum_{k=1}^i \alpha_k \right]^+$ (since $x \in \left[ \sum_{k=1}^{i-1} \alpha_k : \sum_{k=1}^i \alpha_k \right]^+$ is equivalent to $\sum_{k=1}^{i-1} \alpha_k < x \leq \sum_{k=1}^i \alpha_k$). In other words, for every $i \in \{1, 2, \ldots, \ell\}$ satisfying $x \in I_i$ (since $I_i = \left[ \sum_{k=1}^{i-1} \alpha_k : \sum_{k=1}^i \alpha_k \right]^+$). In other words, the subsets $I_1, I_2, \ldots, I_\ell$ of $[0 : n]^+$ are pairwise disjoint, and their union is $[0 : n]^+$. In other words, the subsets $I_1, I_2, \ldots, I_\ell$ of $[0 : n]^+$ form a set partition of the set $[0 : n]^+$. This proves that Property C1 is satisfied.

**Proof that Property C2 is satisfied:** Let $u \in \{1, 2, \ldots, \ell - 1\}$. Let $i \in I_u$ and $j \in I_{u+1}$. Since $i \in I_u = \left[ \sum_{k=1}^{u-1} \alpha_k : \sum_{k=1}^u \alpha_k \right]^+$ (by (11.1045), applied to $u$ instead of $i$), we have $\sum_{k=1}^{u-1} \alpha_k < i \leq \sum_{k=1}^u \alpha_k$. Since $j \in I_{u+1} = \left[ \sum_{k=1}^{u+1} \alpha_k : \sum_{k=1}^{u+1} \alpha_k \right]^+$ (by (11.1045), applied to $u + 1$ instead of $i$), we have $\sum_{k=1}^{u+1} \alpha_k < j \leq \sum_{k=1}^{u+1} \alpha_k$. Now, $i \leq \sum_{k=1}^{u+1} \alpha_k = \sum_{k=1}^{u+1} \alpha_k < j$. Now, let us forget that we fixed $i$ and $j$. We thus have shown that every $i \in I_u$ and $j \in I_{u+1}$ satisfy $i < j$. In other words, $I_u < I_{u+1}$ (by the definition of $I_u < I_{u+1}$).

Now, let us forget that we fixed $u$. We thus have shown that $I_u < I_{u+1}$ for every $u \in \{1, 2, \ldots, \ell - 1\}$. In other words, $I_1 < I_2 < \cdots < I_\ell$. This proves that Property C2 is satisfied.

**Proof that Property C3 is satisfied:** Let $i \in \{1, 2, \ldots, \ell\}$. We have $\sum_{k=1}^i \alpha_k = \sum_{k=1}^{i-1} \alpha_k + \alpha_i > \sum_{k=1}^{i-1} \alpha_k$. Now, (11.1045) yields

\[
|I_i| = \left[ \sum_{k=1}^{i-1} \alpha_k : \sum_{k=1}^i \alpha_k \right]^+ = \sum_{k=1}^{i-1} \alpha_k - \sum_{k=1}^{i-1} \alpha_k = \alpha_i
\]

This proves that Property C3 is satisfied.

We thus have shown that the Properties C1, C2 and C3 are satisfied, qed.
Now, let us show that

(11.1047) \[ I_1 \cup I_2 \cup \cdots \cup I_u = \left[ 0 : \sum_{k=1}^{u} \alpha_k \right]^+ \quad \text{for every } u \in \{0, 1, \ldots, \ell \}. \]

**Proof of (11.1047):** We will prove (11.1047) by induction over \( u \).

**Induction base:** For \( u = 0 \), we have \( I_1 \cup I_2 \cup \cdots \cup I_u = I_1 \cup I_2 \cup \cdots \cup I_0 = (\text{empty union}) = \emptyset \) and

\[
[0 : \sum_{k=1}^{u} \alpha_k]^+ = \left[ 0 : \sum_{k=1}^{0} \alpha_k \right]^+ = [0 : 0]^+ = \emptyset. \]

Hence, for \( u = 0 \), both sides of the equality (11.1047) are \( \emptyset \). Thus, for \( u = 0 \), the equality (11.1047) holds. This completes the induction base.

**Induction step:** Let \( U \in \{1, 2, \ldots, \ell \} \). Assume that (11.1047) holds for \( u = U - 1 \). We now need to prove that (11.1047) holds for \( u = U \) as well.

We have assumed that (11.1047) holds for \( u = U - 1 \). In other words,

(11.1048) \[ I_1 \cup I_2 \cup \cdots \cup I_{U-1} = \left[ 0 : \sum_{k=1}^{U-1} \alpha_k \right]^+. \]

Let \( \xi = \sum_{k=1}^{U-1} \alpha_k \). Then, \( \xi \geq 0 \), so that \( \xi + 1 \geq 1 \). Then,

\[
n = |\alpha| = \sum_{k=1}^{\ell} \alpha_k = \sum_{k=1}^{U-1} \alpha_k + \sum_{k=U}^{\ell} \alpha_k > \sum_{k=1}^{U-1} \alpha_k = \xi. \]

Hence, \( n \geq \xi + 1 \) (since \( \xi \) and \( n \) are integers), so that \( 1 \leq \xi + 1 \leq n \). In other words, \( \xi + 1 \in \{1, 2, \ldots, n\} = [0 : n]^+ = I_1 \cup I_2 \cup \cdots \cup I_\ell \) (since \( I_1 \cup I_2 \cup \cdots \cup I_\ell = [0 : n]^+ \)). In other words, there exists some \( v \in \{1, 2, \ldots, \ell \} \) such that \( \xi + 1 \in I_v \). Consider this \( v \). If we had \( v < U \), then we would have

\[
\xi + 1 \in I_v \subset I_1 \cup I_2 \cup \cdots \cup I_{U-1} = \left[ 0 : \sum_{k=1}^{U-1} \alpha_k \right]^+, \]

which is absurd. Hence, we cannot have \( v < U \). Thus, we have \( v \geq U \).

Recall that

(11.1049) \[ \left[ 0 : \sum_{k=1}^{U-1} \alpha_k \right]^+ = \left[ 0 : \sum_{k=1}^{U-1} \alpha_k \right]^+ = I_1 \cup I_2 \cup \cdots \cup I_{U-1} \]

(by (11.1048)). Using this, it is easy to see that

(11.1050) \[ \text{every } p \in I_U \text{ satisfies } p \geq \xi + 1. \]

Proof of (11.1050): Let \( p \in I_U \). Assume (for the sake of contradiction) that \( p \leq \xi \). But \( p \in I_U \subset I_1 \cup I_2 \cup \cdots \cup I_\ell = [0 : n]^+ = \{1, 2, \ldots, n\} \), so that \( 0 < p \leq n \). Since \( 0 < p \leq \xi \), we have \( p \in \{1, 2, \ldots, \xi\} = [0 : \xi]^+ = I_1 \cup I_2 \cup \cdots \cup I_{U-1} \), which shows that there exists some \( r \in \{1, 2, \ldots, U - 1\} \) such that \( p \in I_r \). Consider this \( r \). Since \( r \neq U \) (because \( r \in \{1, 2, \ldots, U - 1\} \)), the intervals \( I_r \) and \( I_U \) are pairwise disjoint (since the intervals \( I_1, I_2, \ldots, I_\ell \) are pairwise disjoint). In other words, \( I_r \cap I_U = \emptyset \). But combining \( p \in I_r \) with \( p \in I_U \), we obtain \( p \in I_r \cap I_U = \emptyset \), which is absurd. This contradiction proves that our assumption (that \( p \leq \xi \)) was wrong. Hence, we have \( p > \xi \). Thus, \( p \geq \xi + 1 \) (since \( p \) and \( \xi \) are integers). This proves (11.1050).
We now assume (for the sake of contradiction) that \( v \neq U \). Then, \( v > U \) (since \( v \geq U \) and \( v \neq U \)), whence \( I_U < I_v \) (since Property C2 yields \( I_1 < I_2 < \cdots < I_n \)).

The interval \( I_U \) is nonempty, and thus there exists some \( p \in I_U \). Consider such a \( p \). Recall that \( I_U < I_v \). Thus, every \( i \in I_U \) and \( j \in I_v \) satisfy \( i < j \) (by the definition of \( I_U < I_v \)). Applying this to \( i = p \) and \( j = \xi + 1 \), we obtain \( p < \xi + 1 \). But (11.1050) yields \( p \geq \xi + 1 \), which contradicts \( p < \xi + 1 \). This contradiction proves that our assumption (that \( v \neq U \)) was wrong. Hence, we have \( v = U \).

Now, \( \xi + 1 \in I_v = I_U \) (since \( v = U \)). Hence, \( \xi + 1 \) is an element of \( I_U \). Due to (11.1050), this element \( \xi + 1 \) is the smallest element of \( I_U \).

But since property C3 is satisfied, we have \( |I_U| = \alpha_U \) (by Property C3, applied to \( i = U \)). So we know that the interval \( I_U \) has length \( \alpha_U \) (since \( |I_U| = \alpha_U \)) and smallest element \( \xi + 1 \). Therefore, \( I_U = [\xi : \xi + \alpha_U]^+ \) (since the only interval having length \( \alpha_U \) and smallest element \( \xi + 1 \) is the interval \([\xi : \xi + \alpha_U]^+ \)). Now,

\[
I_1 \cup I_2 \cup \cdots \cup I_U = (I_1 \cup I_2 \cup \cdots \cup I_{U-1}) \cup I_U = [0 : \xi]^+ = [0 : \xi + \alpha_U]^+ = [0 : \xi + \alpha_U]^+ (\text{by (11.1049)})
\]

(since \( 0 \leq \xi \leq \xi + \alpha_U \) (since \( \xi \geq 0 \) and \( \alpha_U \geq 0 \)).

Let \( \alpha_U = \sum_{k=1}^{U-1} \alpha_k > 0 \). Then, \( \alpha_U = \sum_{k=1}^{U-1} \alpha_k + \alpha_U = \sum_{k=1}^{U} \alpha_k \).

This rewrites as \( I_1 \cup I_2 \cup \cdots \cup I_U = [0 : \sum_{k=1}^{i} \alpha_k]^+ \). In other words, (11.1047) holds for \( u = U \). This completes the induction step. The induction proof of (11.1047) is thus complete.

Proof of (11.1046): Let \( i \in \{1, 2, \ldots , \ell \} \). Notice that \( \sum_{k=1}^{i} \alpha_k \geq 0 \) (since \( \alpha_k > 0 \) for every \( k \)) and \( \sum_{k=1}^{i} \alpha_k = \sum_{k=1}^{i-1} \alpha_k + \alpha_U > \sum_{k=1}^{i-1} \alpha_k \). Hence, \( 0 \leq \sum_{k=1}^{i-1} \alpha_k \leq \sum_{k=1}^{i} \alpha_k \).

The intervals \( I_1, I_2, \ldots , I_{\ell} \) are pairwise disjoint. Hence, \( I_i \) is disjoint from \( I_1 \cup I_2 \cup \cdots \cup I_{i-1} \). Thus,

\[
I_i = ((I_1 \cup I_2 \cup \cdots \cup I_{i-1}) \cup I_i) \backslash (I_1 \cup I_2 \cup \cdots \cup I_{i-1}) = I_1 \cup I_2 \cup \cdots \cup I_i = (I_1 \cup I_2 \cup \cdots \cup I_{i-1}) \cup I_i = [0 : \sum_{k=1}^{i-1} \alpha_k]^+ \] (by (11.1047), applied to \( u = i-1 \))

(since \( 0 \leq \sum_{k=1}^{i-1} \alpha_k \leq \sum_{k=1}^{i} \alpha_k \)). This proves (11.1046). As explained above, this completes our proof of Remark 6.48(c). \( \square \)


Proof of Lemma 6.50. We have \( \sigma \in \mathbb{R}_{n,m} \), so that \( \sigma^{-1}(1) < \sigma^{-1}(2) < \cdots < \sigma^{-1}(n) \) and \( \sigma^{-1}(n+1) < \sigma^{-1}(n+2) < \cdots < \sigma^{-1}(n+m) \). In other words, the restriction of the map \( \sigma^{-1} \) to the interval \([0 : n]^+ \) is strictly increasing, and so is the restriction of the map \( \sigma^{-1} \) to the interval \([n : n+m]^+ \).

(a) Let \( I \) be an interval of \( \mathbb{Z} \) such that \( I \subset [0 : n+m]^+ \). We will only show that \( \sigma(I) \cap [0 : n]^+ \) is an interval; the proof that \( \sigma(I) \cap [n : n+m]^+ \) is an interval is completely analogous.

It is known that

\[
0 \leq \sum_{k=1}^{i-1} \alpha_k \leq \sum_{k=1}^{i} \alpha_k \] (if \( \mathcal{R} \) is a finite subset of \( \mathbb{Z} \) such that every \( \alpha \in \mathcal{R} \), \( \gamma \in \mathcal{R} \) and \( \beta \in \mathbb{Z} \) satisfying \( \alpha < \beta < \gamma \) satisfy \( \beta \in \mathcal{R} \), then \( \mathcal{R} \) is an interval of \( \mathbb{Z} \)).
In other words, the sets $P_\sigma$ and $Z_\gamma$ form $I_\sigma$ and $\sigma I$ (indeed, it is clear that $\sigma I \cap [0 : n]^+ \subset [0 : n]^+$ and similarly $\gamma \in [0 : n]^+$. Combined with $\alpha < \beta < \gamma$, these yield $\beta \in [0 : n]^+$). But recall that the restriction of the map $\sigma^{-1}$ to the interval $[0 : n]^+$ is strictly increasing. Hence, from $\alpha < \beta < \gamma$, we obtain $\sigma^{-1}(1) < \sigma^{-1}(\beta) < \sigma^{-1}(\gamma)$ (since $\alpha$ and $\beta$ belong to $[0 : n]^+$). In other words, the integer $\sigma^{-1}(\beta)$ lies strictly between the integers $\sigma^{-1}(1)$ and $\sigma^{-1}(\gamma)$. Since $\sigma^{-1}(1) \in I$ (because $\alpha \in \sigma I \cap [0 : n]^+ \subset \sigma I$) and $\sigma^{-1}(\gamma) \in I$ (for similar reasons), this entails $\sigma^{-1}(\beta) \in I$ (because $I$ is an interval, and thus any integer lying between two elements of $I$ must also belong to $I$). Hence, $\beta \in \sigma I$. Combined with $\beta \in [0 : n]^+$, this yields $\beta \in \sigma I \cap [0 : n]^+ = \mathfrak{R}$.

Now, forget that we fixed $\alpha$, $\gamma$ and $\beta$. We thus have shown that every $\alpha \in \mathfrak{R}$, $\gamma \in \mathfrak{R}$ and $\beta \in \mathbb{Z}$ satisfying $\alpha < \beta < \gamma$ satisfy $\beta \in \mathfrak{R}$. Thus, (11.1051) shows that $\mathfrak{R}$ is an interval of $Z$. In other words, $\sigma I \cap [0 : n]^+$ is an interval of $Z$ (since $\mathfrak{R} = \sigma I \cap [0 : n]^+$). This completes the proof of Lemma 6.50(a).

(b) The intervals $K$ and $L$ both are subsets of $[0 : n]^+$. Therefore, from $K \subset L$, we obtain $\sigma^{-1}(K) \subset \sigma^{-1}(L)$ (because the restriction of the map $\sigma^{-1}$ to the interval $[0 : n]^+$ is strictly increasing).

Set $K = \sigma^{-1}(K)$ and $L = \sigma^{-1}(L)$. Then, $K$ is a nonempty interval, and thus can be written in the form $K = [x_K : y_K]^+$ for two elements $x_K$ and $y_K$ of $[0 : n + m]^+$ satisfying $x_K < y_K$. Consider these $x_K$ and $y_K$. Also, $L$ is a nonempty interval, and thus can be written in the form $L = [x_L : y_L]^+$ for two elements $x_L$ and $y_L$ of $[0 : n + m]^+$ satisfying $x_L < y_L$. Consider these $x_L$ and $y_L$. Since $[x_K : y_K]^+ = K = \sigma^{-1}(K) \subset \sigma^{-1}(L) = L = [x_L : y_L]^+$, we have $y_K \leq x_L$.

Notice that $\sigma(K) = K$ (since $\sigma^{-1}(K)$) and $\sigma(L) = L$ (since $L = \sigma^{-1}(L)$).

If we had $y_K = x_L$, then

$$\sigma^{-1}(K) \cup \sigma^{-1}(L) = [x_K : x_L]^+ \cup [x_L : y_L]^+ = [x_K : y_L]^+$$

would be an interval, which would contradict the assumption that $\sigma^{-1}(K) \cup \sigma^{-1}(L)$ is not an interval. Hence, we cannot have $y_K = x_L$. Thus, $y_K \neq x_L$. Therefore, $y_K < x_L$ (since $y_K \leq x_L$). Hence, we can define a nonempty interval $P \subset [0 : n + m]^+$ by $P = [y_K : x_L]^+$. Consider this $P$. It satisfies $|P| \neq 0$ (since it is nonempty) and $K < P < L$ (since $K = [x_K : y_K]^+$, $P = [y_K : x_L]^+$ and $L = [x_L : y_L]^+$), so that the sets $K$, $P$ and $L$ are pairwise disjoint. As a consequence, the sets $\sigma(K)$, $\sigma(P)$ and $\sigma(L)$ are pairwise disjoint. In other words, the sets $K$, $\sigma(P)$ and $L$ are pairwise disjoint (since $\sigma(K) = K$ and $\sigma(L) = L$).

It is easy to see that $\sigma(P) \cap [n : n + m]^+ \neq 0$. Thus, $\sigma(P) \cap [n : n + m]^+ = \sigma(P)$. But Lemma 6.50(a) (applied to $P$ instead of $I$) shows that $\sigma(P) \cap [0 : n]^+$ and $\sigma(P) \cap [n : n + m]^+$ are intervals. In particular, $\sigma(P) \cap [n : n + m]^+$ is an interval. In other words, $\sigma(P)$ is an interval (since $\sigma(P) \cap [n : n + m]^+ = \sigma(P)$).

---

809 Because $K = \sigma^{-1}(K)$ and because we know that $\sigma^{-1}(K)$ is an interval and $K$ is nonempty

810 Because $\mathfrak{L} = \sigma^{-1}(L)$ and because we know that $\sigma^{-1}(L)$ is an interval and $L$ is nonempty

811 Proof. Assume the contrary. Then, there exists some $q \in \sigma(P)$ such that $q \notin [n : n + m]^+$. Consider this $q$. Since $q \notin [n : n + m]^+$, we must have $q \in [0 : n]^+$. Notice that $q \in \sigma(P)$, so that $\sigma^{-1}(q) \notin \mathfrak{P}$.

The elements $\min K$ and $\min L$ of $K$ and $L$ are well-defined, since $K$ and $L$ are nonempty.

Now, $\max K \leq q - 1$. (To prove this, assume the contrary. Thus, $\max K > q - 1$, so that $\max K \geq q$. But $\max K \in K \supseteq [0 : n]^+$ and $q \in [0 : n]^+$. Therefore, $\sigma^{-1}(\max K) \geq \sigma^{-1}(q)$ (because $\max K \geq q$, and since the restriction of the map $\sigma^{-1}$ to the interval $[0 : n]^+$ is strictly increasing). But since $K \subset P$, we have $\sigma^{-1}(\max K) < \sigma^{-1}(q)$ (since $\sigma^{-1}(\max K) \in K$ and $\sigma^{-1}(q) \in P$), which contradicts $\sigma^{-1}(\max K) \geq \sigma^{-1}(q)$. This contradiction shows that our assumption was wrong, and we have shown that $\max K \leq q - 1$.

Furthermore, $q \leq \min L - 1$. (To prove this, assume the contrary. Thus, $q > \min L - 1$. Hence, $q \geq \min L$. We have $\min L \in L \subset [0 : n]^+$ and $q \in [0 : n]^+$. Therefore, $\sigma^{-1}(q) \geq \sigma^{-1}(\min L)$ (because $q \geq \min L$, and since the restriction of the map $\sigma^{-1}$ to the interval $[0 : n]^+$ is strictly increasing). But since $P \subset L$, we have $\sigma^{-1}(q) < \sigma^{-1}(\min L)$ (since $\sigma^{-1}(q) \in P$ and $\sigma^{-1}(\min L) \in \sigma^{-1}(L) = \mathfrak{L}$), which contradicts $\sigma^{-1}(q) \geq \sigma^{-1}(\min L)$. This contradiction shows that our assumption was wrong, and we have shown that $q \leq \min L - 1$.)
So we know that $\sigma(P) \subset [n : n + m]^+$ is a nonempty interval\footnote{nonempty because $P$ is nonempty}, and $\sigma^{-1}(\sigma(P))$ is also an interval (since $\sigma^{-1}(\sigma(P)) = P$). Moreover, we have $\sigma^{-1}(K) < \sigma^{-1}(\sigma(P))$ (since $\sigma^{-1}(K) = K < P = \sigma^{-1}(\sigma(P))$). Also, $\sigma^{-1}(K) \cup \sigma^{-1}(\sigma(P)) = [x_K : y_K]^+ \cup [y_K : x_L]^+ = [x_K : x_L]^+$ is an interval. Furthermore, $\sigma^{-1}(\sigma(P)) < \sigma^{-1}(L)$ (since $\sigma^{-1}(\sigma(P)) = P < L = \sigma^{-1}(L)$), and the set $\sigma^{-1}(\sigma(P)) \cup \sigma^{-1}(L) = [y_K : x_L]^+ \cup [x_L : y_K]^+$ is an interval.

Altogether, we now know that $\sigma(P) \subset [n : n + m]^+$ is a nonempty interval such that $\sigma^{-1}(\sigma(P))$, $\sigma^{-1}(K) \cup \sigma^{-1}(\sigma(P))$ and $\sigma^{-1}(\sigma(P)) \cup \sigma^{-1}(L)$ are intervals and such that $\sigma^{-1}(K) < \sigma^{-1}(\sigma(P)) < \sigma^{-1}(L)$. Thus, there exists a nonempty interval $P \subset [n : n + m]^+$ such that $\sigma^{-1}(P)$, $\sigma^{-1}(K) \cup \sigma^{-1}(P)$ and $\sigma^{-1}(P) \cup \sigma^{-1}(L)$ are intervals and such that $\sigma^{-1}(K) < \sigma^{-1}(P) < \sigma^{-1}(L)$ (namely, $P = \sigma(P)$). This proves Lemma 6.50(b).

(c) The proof of Lemma 6.50(c) is analogous to the proof of Lemma 6.50(b).
But let \( w = (w_1, w_2, \ldots, w_{n+m}) \) denote the word \( uv \). The definition of \( u \sqcup v \) then yields
\[
\begin{align*}
\sigma(11.1053) & \quad \sigma \in [0 : n ] \quad a < b \\
\sigma(11.1054) & \quad \tau \quad \text{(where we consider the permutation} \\
& \quad \text{and these intervals satisfy} \\
& \quad \text{This word} \\
& \quad \text{by writing it in one-line notation).} \\
& \quad \text{Now, define a word} \ \\
\sigma(11.1054) & \quad \tau \quad \text{by}
\end{align*}
\]
for every word \( p \in \mathfrak{B}^{n+m} \)

\[
\begin{align*}
p & = p [A] \cdot p [\sigma^{-1} (I)] \cdot p [\sigma^{-1} (J)] \cdot p [Z] \\
\end{align*}
\]

This word \( \tau \) is obtained from the word \( \sigma \) by switching the two factors \( \sigma [\sigma^{-1} (I)] \) and \( \sigma [\sigma^{-1} (J)] \) (this is clear by comparing (11.1054) with (11.1053)), and thus is a permutation written in one-line notation (because \( \sigma \) is a permutation). In other words, \( \tau \in \mathfrak{S}_{n+m} \).
From (11.1054), we have

$$\tau = \sigma \left[ \frac{A}{[0:a]^{+}} \right] \cdot \sigma \left[ \frac{\sigma^{-1}(J)}{[a:d]^{+}} \right] \cdot \sigma \left[ \frac{\sigma^{-1}(I)}{[a:b]^{+}} \right] \cdot \sigma \left[ \frac{Z}{[d:n+m]^{+}} \right]$$

$$= \sigma \left[ \frac{0:a}{[0:a]^{+}} \right] \cdot \sigma \left[ \frac{b:d}{[a:d]^{+}} \right] \cdot \sigma \left[ \frac{a:b}{[a:b]^{+}} \right] \cdot \sigma \left[ \frac{d:n+m}{[d:n+m]^{+}} \right]$$

$$= (\sigma(1), \sigma(2), \ldots, \sigma(a)) \cdot (\sigma(b+1), \sigma(b+2), \ldots, \sigma(d)) \cdot \sigma(a+1, \sigma(a+2), \ldots, \sigma(b)) \cdot (\sigma(d+1), \sigma(d+2), \ldots, \sigma(n+m))$$

$$= (\sigma(1), \sigma(2), \ldots, \sigma(a), \sigma(b+1), \sigma(b+2), \ldots, \sigma(d), \sigma(a+1), \sigma(a+2), \ldots, \sigma(b), \sigma(d+1), \sigma(d+2), \ldots, \sigma(n+m)).$$

Thus,

$$\tau = (\tau(1), \tau(2), \ldots, \tau(a)) \quad (11.1055)$$

$$\tau = (\tau(a+1), \tau(a+2), \ldots, \tau(a+d-b)) \quad (11.1056)$$

$$\tau = (\tau(a+d-b+1), \tau(a+d-b+2), \ldots, \tau(d)) \quad (11.1057)$$

$$\tau = (\tau(d+1), \tau(d+2), \ldots, \tau(n+m)) \quad (11.1058)$$

From this, it is easy to see that $\tau^{-1}(J) = [a:a+d-b]^{1}$ and $\tau^{-1}(I) = [a+d-b:d]^{1}$ in particular, $\tau^{-1}(J)$ and $\tau^{-1}(I)$ are intervals. Now, obviously, the intervals $[0:a]^{+}$, $[a:a+d-b]^{+}$, $[a+d-b:d]^{+}$ and $[d:n+m]^{+}$ are disjoint intervals having union $[0:n+m]^{+}$ and satisfying $[0:a]^{+} < [a:a+d-b]^{+} < [a+d-b:d]^{+} < [d:n+m]^{+}$. Since $A = [0:a]^{+}$, $\tau^{-1}(J) = [a:a+d-b]^{+}$, $\tau^{-1}(I) = [a+d-b:d]^{+}$ and $Z = [d:n+m]^{+}$, this rewrites as follows: The intervals $A$, $\tau^{-1}(J)$, $\tau^{-1}(I)$ and $Z$ are disjoint intervals having union $[0:n+m]^{+}$ and satisfying $A < \tau^{-1}(J) < \tau^{-1}(I) < Z$. Therefore,

$$p = p[A] \cdot p[\tau^{-1}(J)] \cdot p[\tau^{-1}(I)] \cdot p[Z] \quad \text{for every word } p \in \mathfrak{A}^{n+m}$$

for every alphabet $\mathfrak{A}$.

Next, we claim that $\tau$ belongs to $S_{h_{n,m}}$. In fact, let $i$ and $j$ be two elements of $\{1, 2, \ldots, n\}$ such that $i < j$. Our next goal is to prove that $\tau^{-1}(i) < \tau^{-1}(j)$.

Indeed, assume the contrary. Thus, $\tau^{-1}(i) \geq \tau^{-1}(j)$, so that $\tau^{-1}(i) > \tau^{-1}(j)$ (since $\tau$ is a permutation). In other words, the letter $i$ lies further right than the letter $j$ in the word $\tau$. But we also have $\sigma^{-1}(i) < \sigma^{-1}(j)$ (since $\sigma^{-1}(1) < \sigma^{-1}(2) < \cdots < \sigma^{-1}(n)$ and $i < j$ and since $i, j \in \{1, 2, \ldots, n\}$), which means that the letter $i$ lies further left than the letter $j$ in the word $\sigma$.

So the order in which the letters $i$ and $j$ appear in the word $\tau$ is different from that in $\sigma$. Since the word $\tau$ is obtained from the word $\sigma$ by switching the two adjacent factors $\sigma[\sigma^{-1}(J)]$ and $\sigma[\sigma^{-1}(I)]$, this is only possible if one of the letters $i$ and $j$ is contained in one of these two factors and the other is contained in the other factor. Hence, this is what must be happening. In particular, one of the letters $i$ and $j$ must be contained in the factor $\sigma[\sigma^{-1}(J)]$. Since all letters of $\sigma[\sigma^{-1}(J)]$ belong to the set $[n:n+m]^{+}$ (since the letters of $\sigma[\sigma^{-1}(J)]$ are precisely the elements of $J$, but we have $J \subseteq [n:n+m]^{+}$), this yields that

$^{813}$In fact, this follows from

$$\tau \left( \frac{[a:a+d-b]^{+}}{[a+1,a+2,\ldots,a+d-b]} \right) = \tau \left( \{a+1,a+2,\ldots,a+d-b\} \right) = \{\tau(a+1), \tau(a+2), \ldots, \tau(a+d-b)\}$$

$$= \{\sigma(b+1), \sigma(b+2), \ldots, \sigma(d)\} \quad \text{(by (11.1056))}$$

$$= \sigma \left( \frac{b+1, b+2, \ldots, d}{=[b:d]^{+} \circ \sigma^{-1}(J)} \right) = \sigma \left( \sigma^{-1}(J) \right) = J.$$

$^{814}$The proof of this is similar to that of $\tau^{-1}(J) = [a:a+d-b]^{+}$, but we have to use (11.1057) this time.
one of the letters $i$ and $j$ must belong to the set $[n : n + m]^+$. But this is impossible, since the letters $i$ and $j$ belong to $\{1, 2, \ldots, n\}$ and thus neither of them can be an element of $[n : n + m]^+ (because [n : n + m]^+$ is disjoint with $\{1, 2, \ldots, n\})$. This contradiction shows that our assumption was wrong, and so we do have $\tau^{-1}(i) < \tau^{-1}(j)$.

Now, forget that we fixed $i$ and $j$. We have shown that any two elements $i$ and $j$ of $\{1, 2, \ldots, n\}$ such that $i < j$ must satisfy $\tau^{-1}(i) < \tau^{-1}(j)$. Thus, $\tau^{-1}(1) < \tau^{-1}(2) \cdots < \tau^{-1}(n)$. Similarly, $\tau^{-1}(n + 1) < \tau^{-1}(n + 2) \cdots < \tau^{-1}(n + m)$. These two chains of inequalities, together, yield that $\tau \in \text{Sh}_{n,m}$. Hence, $u \sqcup v$ is a well-defined element of the multiset $u \sqcup v$. Since $u \sqcup v$ is the lexicographically highest element of this multiset, this shows that

(11.1060) \[ u \sqcup v \geq u \sqcup v. \]

Let $a$ denote the word $\left( \left( u \sqcup v \right) \right) [A]$, and let $\zeta$ denote the word $\left( \left( u \sqcup v \right) \right) [Z]$.

We can now apply (11.1052) to $p = u \sqcup \sigma$ and $B = A$. As a result, we obtain

(11.1061) \[ u \sqcup \sigma = \left( \left( u \sqcup v \right) \right) [A] \cdot \left( \left( u \sqcup v \right) \right) \left[ \sigma^{-1} (I) \right] \cdot \left( \left( u \sqcup v \right) \right) \left[ \sigma^{-1} (J) \right] \cdot \left( \left( u \sqcup v \right) \right) [Z]. \]

Let now $(w_1, w_2, \ldots, w_{n+m})$ denote the word $uv$. The definition of $u \sqcup \sigma$ then yields

\[ u \sqcup \sigma = (w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}). \]

Now, it is easy to see that

\[ \left( \left( u \sqcup \sigma \right) \right) [A] = a \]

and

\[ \left( \left( u \sqcup \sigma \right) \right) [Z] = \zeta. \]

Thus, (11.1061) becomes

\[ u \sqcup \sigma = \left( \left( u \sqcup v \right) \right) [A] \cdot \left( \left( u \sqcup v \right) \right) \left[ \sigma^{-1} (I) \right] \cdot \left( \left( u \sqcup v \right) \right) \left[ \sigma^{-1} (J) \right] \cdot \left( \left( u \sqcup v \right) \right) [Z] \]

\[ \quad = a \cdot (uv) [I] \cdot (uv) [J] \cdot \zeta. \]

Proof. We have

\[ \left( \left( u \sqcup \sigma \right) \right) = \left( \left( u \sqcup v \right) \right) \left[ \sigma^{-1} (I) \right] \cdot \left( \left( u \sqcup v \right) \right) \left[ \sigma^{-1} (J) \right] \cdot \left( \left( u \sqcup v \right) \right) [Z] \]

The same argument, applied to $\tau$ instead of $\sigma$, shows that $\left( \left( u \sqcup \tau \right) \right) [A] = (w_{\tau(1)}, w_{\tau(2)}, \ldots, w_{\tau(a)}).$ But (11.1055) yields $(w_{\tau(1)}, w_{\tau(2)}, \ldots, w_{\tau(a)}) = (w_{\tau(1)}, w_{\tau(2)}, \ldots, w_{\tau(a)}).$

QL. We have

\[ \left( \left( u \sqcup \tau \right) \right) = \left( \left( u \sqcup v \right) \right) \left[ \sigma^{-1} (I) \right] \cdot \left( \left( u \sqcup v \right) \right) \left[ \sigma^{-1} (J) \right] \cdot \left( \left( u \sqcup v \right) \right) [Z] \]

The same argument, applied to $\tau$ instead of $\sigma$, shows that

\[ \left( \left( u \sqcup \tau \right) \right) [Z] = (w_{\tau(d+1)}, w_{\tau(d+2)}, \ldots, w_{\tau(n+m)}). \]

But (11.1058) yields $(w_{\tau(d+1)}, w_{\tau(d+2)}, \ldots, w_{\tau(n+m)}) = (w_{\tau(d+1)}, w_{\tau(d+2)}, \ldots, w_{\tau(n+m)}).$

\[ \left( \left( u \sqcup \tau \right) \right) [Z] = (w_{\tau(d+1)}, w_{\tau(d+2)}, \ldots, w_{\tau(n+m)}). \]

QED.
Thus,
\[
\mathbf{a} \cdot (uv)[I] \cdot (uv)[J] \cdot \mathbf{3} = u \uplus v \geq u \uplus v \quad \text{(by } \ref{eq:11.1060})
\]
\[
= (u \uplus v)[A] \cdot (u \uplus v)[\tau^{-1}(J)] \cdot (u \uplus v)[\tau^{-1}(J)] \cdot (u \uplus v)[Z] = a \quad \text{(by } \ref{eq:6.4'), applied to } \tau \text{ and } J \text{ instead of } \sigma \text{ and } I)
\]
\[
(\text{by } \ref{eq:11.1059}, \text{ applied to } p = u \uplus v \text{ and } \mathcal{B} = \mathcal{A})
\]
\[
= \mathbf{a} \cdot (uv)[I] \cdot (uv)[J] \cdot \mathbf{3}.
\]
In other words, \( \mathbf{a} \cdot (uv)[I] \cdot (uv)[J] \cdot \mathbf{3} \leq \mathbf{a} \cdot (uv)[I] \cdot (uv)[J] \cdot \mathbf{3} \). Hence, Proposition 6.2(c) (applied to \( \mathbf{a}, (uv)[J], (uv)[I], \mathbf{3} \) and \( (uv)[I], (uv)[J], \mathbf{3} \) instead of \( \alpha, c \) and \( \mathcal{D} \) yields \( (uv)[I] \cdot (uv)[J] \cdot \mathbf{3} \leq (uv)[I] \cdot (uv)[J] \cdot \mathbf{3} \).

Thus, Proposition 6.2(d) (applied to \( (uv)[I], (uv)[J], \mathbf{3} \), \( (uv)[I], (uv)[J], \mathbf{3} \) and \( (uv)[I], (uv)[J], \mathbf{3} \) instead of \( \alpha, c \) and \( \mathcal{D} \) yields \( (uv)[J] \cdot (uv)[I] \leq (uv)[I] \cdot (uv)[J] \) (since \( \ell((uv)[J], (uv)[I]) = \ell((uv)[J]) + \ell((uv)[I]) = \ell((uv)[I]) + \ell((uv)[J]) \)). In other words, \( (uv)[I] \cdot (uv)[J] \geq (uv)[I] \cdot (uv)[J] \). This proves Lemma 6.52(b).

(c) The proof of Lemma 6.52(c) is analogous to that of Lemma 6.52(b).

\[\square\]


Proof of Proposition 6.56. (a) Let \( j \in \{1, 2, \ldots, \ell\} \). We have \( \sigma = \text{iper}(\alpha, \tau) = I_{(1)} \cdot I_{(2)} \cdot \cdots \cdot I_{(\ell)} \) (in one-line notation). Therefore, the word \( I_{(j)} \) appears as a factor in this word \( \sigma \), starting at position \( \sum_{k=1}^{j-1} \alpha_{(k)} + 1 \) and ending at position \( \sum_{k=1}^{j} \alpha_{(k)} \). In other words, the letters of the word \( I_{(j)} \) are the letters \( \sigma_j \) of \( \sigma \) for \( j \in \left[ \sum_{k=1}^{j-1} \alpha_{(k)} : \sum_{k=1}^{j} \alpha_{(k)} \right]^+ \). Since the letters of the word \( I_{(j)} \) are precisely the elements of \( I_{(j)} \), this rewrites as follows: The elements of \( I_{(j)} \) are the letters \( \sigma_j \) of \( \sigma \) for \( j \in \left[ \sum_{k=1}^{j-1} \alpha_{(k)} : \sum_{k=1}^{j} \alpha_{(k)} \right]^+ \). In other words,
\[
I_{(j)} = \left\{ \sigma_j \mid j \in \left[ \sum_{k=1}^{j-1} \alpha_{(k)} : \sum_{k=1}^{j} \alpha_{(k)} \right]^+ \right\} = \sigma \left( \left[ \sum_{k=1}^{j-1} \alpha_{(k)} : \sum_{k=1}^{j} \alpha_{(k)} \right]^+ \right).
\]
Hence, \( \sigma^{-1}(I_{(j)}) = \left[ \sum_{k=1}^{j-1} \alpha_{(k)} : \sum_{k=1}^{j} \alpha_{(k)} \right]^+ \), so that Proposition 6.56(a) is proven.

(b) Let \( j \in \{1, 2, \ldots, \ell\} \). We need to show that the restriction of the map \( \sigma^{-1} \) to the interval \( I_{(j)} \) is increasing. In other words, we need to prove that the elements of \( I_{(j)} \) occur in increasing order in the word \( \sigma \). But this is clear, because these elements all occur in the factor \( I_{(j)} \) of the word \( \sigma \), and this factor has them in increasing order (by its definition). This proves Proposition 6.56(b).

(c) Let \( i \in \{1, 2, \ldots, \ell\} \). Then, \( \sigma^{-1}(I_i) = \sigma^{-1}(I_{(\tau^{-1}(i))}) = \left[ \sum_{k=1}^{\tau^{-1}(i)-1} \alpha_{(k)} : \sum_{k=1}^{\tau^{-1}(i)} \alpha_{(k)} \right]^+ \) (according to Proposition 6.56(a), applied to \( j = \tau^{-1}(i) \)). Hence, \( \sigma^{-1}(I_i) \) is an interval. Furthermore, the restriction of the map \( \sigma^{-1} \) to the interval \( I_i = I_{(\tau^{-1}(i))} \) is increasing (according to Proposition 6.56(b), applied to \( j = \tau^{-1}(i) \)).

Now, forget that we fixed \( i \). We thus have proven that every \( i \in \{1, 2, \ldots, \ell\} \) has the two properties that:
- the set \( \sigma^{-1}(I_i) \) is an interval;
- the restriction of the map \( \sigma^{-1} \) to the interval \( I_i \) is increasing.

In other words, the permutation \( \sigma \) is \( \alpha \)-clumping. Since \( \sigma = \text{iper}(\alpha, \tau) \), this shows that \( \text{iper}(\alpha, \tau) \) is \( \alpha \)-clumping. Proposition 6.56(c) is proven.
(d) Let \( i \in \{1, 2, \ldots, \ell - 1\} \). Then,

\[
\sigma^{-1}(I_{\tau(i)}) = \left[ \sum_{k=1}^{i-1} \alpha_{\tau(k)} : \sum_{k=1}^{i} \alpha_{\tau(k)} \right]^+
\]

(by Proposition 6.56(a), applied to \( j = i \), so that \( \sigma^{-1}(I_{\tau(i)}) \) is an interval. Thus, \( \sigma^{-1}(I_{\tau(i)}) \) is a nonempty interval (nonempty because \( I_{\tau(i)} \) is nonempty). Similarly, \( \sigma^{-1}(I_{\tau(i+1)}) \) is a nonempty interval.

Also, Proposition 6.56(a) (applied to \( j = i + 1 \)) yields

\[
\sigma^{-1}(I_{\tau(i+1)}) = \left[ \sum_{k=1}^{i} \alpha_{\tau(k)} : \sum_{k=1}^{i+1} \alpha_{\tau(k)} \right]^+.
\]

Now,

\[
\sigma^{-1}(I_{\tau(i)}) = \left[ \sum_{k=1}^{i-1} \alpha_{\tau(k)} : \sum_{k=1}^{i} \alpha_{\tau(k)} \right]^+ < \left[ \sum_{k=1}^{i} \alpha_{\tau(k)} : \sum_{k=1}^{i+1} \alpha_{\tau(k)} \right]^+ = \sigma^{-1}(I_{\tau(i+1)}).
\]

Also,

\[
\sigma^{-1}(I_{\tau(i)}) \cup \sigma^{-1}(I_{\tau(i+1)}) = \left[ \sum_{k=1}^{i-1} \alpha_{\tau(k)} : \sum_{k=1}^{i} \alpha_{\tau(k)} \right]^+ \cup \left[ \sum_{k=1}^{i} \alpha_{\tau(k)} : \sum_{k=1}^{i+1} \alpha_{\tau(k)} \right]^+ = \left[ \sum_{k=1}^{i-1} \alpha_{\tau(k)} : \sum_{k=1}^{i+1} \alpha_{\tau(k)} \right]^+,
\]

which is obviously an interval. Proposition 6.56(d) is proven. \( \square \)

### 11.125. Solution to Exercise 6.59

**Solution to Exercise 6.59.**

**Proof of Proposition 6.58.** We first recall a known fact from elementary combinatorics. For any \( N \in \mathbb{N} \) and any \( \varphi \in \mathfrak{S}_N \), define the *inversion set* \( \text{Inv} \varphi \) of \( \varphi \) to be the set \( \\{(i, j) \in \{1, 2, \ldots, N\}^2 \mid i < j; \ \varphi(i) > \varphi(j)\} \).

(This notation is not completely standard.) Then, a well-known fact states that if \( N \in \mathbb{N} \), and if \( \varphi \) and \( \psi \) are two elements of \( \mathfrak{S}_N \) satisfying \( \text{Inv} \varphi = \text{Inv} \psi \), then

\[
\varphi = \psi.
\]

(In other words, a permutation in \( \mathfrak{S}_N \) is uniquely determined by its inversion set.)

(a) The interval system corresponding to \( \alpha \) is an \( \ell \)-tuple of intervals (since \( \ell(\alpha) = \ell \)) denoted by \( (I_1, I_2, \ldots, I_\ell) \). Then, the intervals \( I_1, I_2, \ldots, I_\ell \) form a set partition of \( [0 : n]^+ \) (according to Remark 6.48(c)) and are nonempty (also according to Remark 6.48(c)).

We define a map \( \text{iper}_\alpha : \{\omega \in \mathfrak{S}_N \mid \omega \text{ is } \alpha\text{-clumping}\} \to \mathfrak{S}_\ell \) as follows: Let \( \omega \) be an \( \alpha \)-clumping element of \( \mathfrak{S}_N \). Then, every \( i \in \{1, 2, \ldots, \ell\} \) has the property that \( \omega^{-1}(I_i) \) is an interval (since \( \omega \) is an \( \alpha \)-clumping). These intervals \( \omega^{-1}(I_1), \omega^{-1}(I_2), \ldots, \omega^{-1}(I_\ell) \) form a set partition of \( [0 : n]^+ \) (since the intervals \( I_1, I_2, \ldots, I_\ell \) form a set partition of \( [0 : n]^+ \)), and thus are pairwise disjoint (and nonempty).\(^{817}\) Hence, these intervals form a totally ordered set with respect to the relation \( < \) (by Remark 6.48(b)). Thus, there exists a unique permutation \( \tau \in \mathfrak{S}_\ell \) such that \( \omega^{-1}(I_{\tau(1)}) < \omega^{-1}(I_{\tau(2)}) < \cdots < \omega^{-1}(I_{\tau(\ell)}) \). We define \( \text{iper}_\alpha(\omega) \) to be this permutation \( \tau \).

\(^{817}\)Their nonemptiness follows from the fact that the intervals \( I_1, I_2, \ldots, I_\ell \) are nonempty.
Thus, we have defined a map $iper'_{\alpha} : \{\omega \in \mathfrak{S}_n \mid \omega \text{ is } \alpha\text{-clumping} \} \to \mathfrak{S}_\ell$. It is easy to see that $iper_{\alpha} \circ iper'_{\alpha} = \text{id}^{818}$ and that $iper'_{\alpha} \circ iper_{\alpha} = \text{id}^{819}$. Hence, the maps $iper_{\alpha}$ and $iper'_{\alpha}$ are mutually inverse, and thus the map $iper_{\alpha}$ is bijective. This completes the proof of Proposition 6.58(a). 

---

818 Proof. Let $\omega \in \mathfrak{S}_n$ be $\alpha$-clumping. We are going to prove that $(iper_{\alpha} \circ iper'_{\alpha})(\omega) = \text{id}(\omega)$.

Indeed, let us recall how $iper'_{\alpha}(\omega)$ was defined: Every $i \in \{1, 2, \ldots, \ell\}$ has the property that $\omega^{-1}(I_i)$ is an interval. These intervals $\omega^{-1}(I_1), \omega^{-1}(I_2), \ldots, \omega^{-1}(I_{\ell})$ form a totally ordered set with respect to the relation $<$. Then, $iper'_{\alpha}(\omega)$ is defined as the unique permutation $\tau \in \mathfrak{S}_\ell$ such that $\omega^{-1}(I_1) < \omega^{-1}(I_2) < \cdots < \omega^{-1}(I_{\ell})$. Thus, we have $iper'_{\alpha}(\omega) \in \mathfrak{S}_\ell$ and

$$\omega^{-1}(I_{iper'_{\alpha}(\omega)}(1)) < \omega^{-1}(I_{iper'_{\alpha}(\omega)}(2)) < \cdots < \omega^{-1}(I_{iper'_{\alpha}(\omega)}(\ell)) \quad (11.1065)$$

Denote the permutation $iper'_{\alpha}(\omega) \in \mathfrak{S}_\ell$ by $\tau$. Then, $(11.1065)$ rewrites as

$$\omega^{-1}(I_{\tau(1)}) < \omega^{-1}(I_{\tau(2)}) < \cdots < \omega^{-1}(I_{\tau(\ell)}) \quad (11.1066)$$

(since $\tau = iper'_{\alpha}(\omega)$). The definition of $iper(\alpha, \tau)$ shows that $iper(\alpha, \tau)$ is the permutation in $\mathfrak{S}_n$ which (in one-line notation) is the word $I_{\tau(1)}I_{\tau(2)}I_{\tau(3)} \cdots I_{\tau(\ell)}$ (a concatenation of $\ell$ words). In other words, $iper(\alpha, \tau) = I_{\tau(1)}I_{\tau(2)}I_{\tau(3)} \cdots I_{\tau(\ell)}$. Denote this permutation $iper(\alpha, \tau)$ by $\eta$. Then, $\eta = iper(\alpha, \tau) = I_{\tau(1)}I_{\tau(2)}I_{\tau(3)} \cdots I_{\tau(\ell)}$ and

$$(iper_{\alpha} \circ iper'_{\alpha})(\omega) = iper_{\alpha}\left(iper'_{\alpha}(\omega)\right) = iper_{\alpha}\tau = iper_{\alpha}(\alpha, \tau) = iper_{\alpha}\tau = iper(\alpha, \tau) = \eta. \quad (11.1067)$$

Our next goal is to prove that $\eta = \omega$.

We know that $\omega$ is $\alpha$-clumping. In other words, every $i \in \{1, 2, \ldots, \ell\}$ has the two properties that:

$$\omega^{-1}(I_i) \text{ is an interval, and}$$

$$\text{the set } \omega^{-1}(I_i) \text{ is increasing (by the definition of “} \alpha\text{-clumping”).}$$

We shall now prove that $\text{Inv}(\eta^{-1}) \subseteq \text{Inv}(\omega^{-1})$. Indeed, let $(i, j)$ be some element of $\text{Inv}(\eta^{-1})$. We want to prove that $(i, j) \in \text{Inv}(\omega^{-1})$.

We have $(i, j) \in \text{Inv}(\eta^{-1})$. By the definition of $\text{Inv}(\eta^{-1})$, this yields that $(i, j)$ is an element of $\{1, 2, \ldots, n\}^2$ satisfying $i < j$ and $\eta^{-1}(i) > \eta^{-1}(j)$. In particular, $\eta^{-1}(i) > \eta^{-1}(j)$. In other words, the letter $i$ must appear after the letter $j$ in the word $\eta$ (where our use of the word “after” does not imply “immediately after”). Since $\eta = I_{\tau(1)}I_{\tau(2)}I_{\tau(3)} \cdots I_{\tau(\ell)}$, this rewrites as follows: The letter $i$ must appear after the letter $j$ in the word $I_{\tau(1)}I_{\tau(2)}I_{\tau(3)} \cdots I_{\tau(\ell)}$. Thus, we must be in one of the following two cases:

Case 1: There exist some elements $i$ and $j$ of $\{1, 2, \ldots, \ell\}$ such that $i > j$ and such that the letter $i$ appears in the word $I_{\tau(i)}$, whereas the letter $j$ appears in the word $I_{\tau(j)}$.

Case 2: There exist some element $i$ of $\{1, 2, \ldots, \ell\}$ such that both letters $i$ and $j$ appear in the word $I_{\tau(i)}$, and the letter $i$ appears after the letter $j$ in this word.

Let us first consider Case 1. In this case, there exist some elements $i$ and $j$ of $\{1, 2, \ldots, \ell\}$ such that $i > j$ and such that the letter $i$ appears in the word $I_{\tau(i)}$, whereas the letter $j$ appears in the word $I_{\tau(j)}$. Consider these $i$ and $j$. The letter $i$ appears in the word $I_{\tau(i)}$, and thus is an element of $I_{\tau(i)}$ (since the letters of the word $I_{\tau(i)}$ are precisely the elements of $I_{\tau(i)}$). In other words, $i \in I_{\tau(i)}$. Hence, $\omega^{-1}(i) \in \omega^{-1}(I_{\tau(i)})$. Similarly, $\omega^{-1}(j) \in \omega^{-1}(I_{\tau(j)})$. Now, $i > j$, so that $j < i$ and thus $\omega^{-1}(I_{\tau(j)}) < \omega^{-1}(I_{\tau(i)})$ (due to $(11.1066)$). In other words, every $j' \in \omega^{-1}(I_{\tau(j)})$ and $i' \in \omega^{-1}(I_{\tau(i)})$ satisfy $j' < i'$ (by the definition of $\omega^{-1}(I_{\tau(j)}) < \omega^{-1}(I_{\tau(i)})$). Applying this to $j' = \omega^{-1}(j)$ and $i' = \omega^{-1}(i)$, we obtain $\omega^{-1}(j) < \omega^{-1}(i)$, so that $\omega^{-1}(i) > \omega^{-1}(j)$.

Now, we know that $(i, j)$ is an element of $\{1, 2, \ldots, n\}^2$ satisfying $i < j$ and $\omega^{-1}(i) > \omega^{-1}(j)$. In other words, $(i, j) \in \text{Inv}(\omega^{-1})$ (by the definition of $\text{Inv}(\omega^{-1})$). Hence, $(i, j) \in \text{Inv}(\omega^{-1})$ is proven in Case 1.

Let us now consider Case 2. In this case, there exists some element $i$ of $\{1, 2, \ldots, \ell\}$ such that both letters $i$ and $j$ appear in the word $I_{\tau(i)}$, and the letter $i$ appears after the letter $j$ in this word. Consider this $i$. The letters of the word $I_{\tau(i)}$ are in increasing order (since $I_{\tau(i)}$ is defined as the list of all elements of $I_{\tau(i)}$ in increasing order). Since the letter $i$ appears after the letter $j$ in this word, we must therefore have $i > j$. But this contradicts $i < j$. This contradiction shows that Case 2 cannot occur. Hence, the only possible case is Case 1. Since we have proven $(i, j) \in \text{Inv}(\omega^{-1})$ in this case, we therefore conclude that $(i, j) \in \text{Inv}(\omega^{-1})$ always holds.

Now, let us forget that we fixed $(i, j)$. We thus have proven that $(i, j) \in \text{Inv}(\omega^{-1})$ for every $(i, j) \in \text{Inv}(\eta^{-1})$. In other words,

$$\text{Inv}(\eta^{-1}) \subseteq \text{Inv}(\omega^{-1}).$$
(b) Consider the map $\text{iper}_\alpha$ defined in Proposition 6.58(a). Since $\sigma$ is $\alpha$-clumping, we have $\sigma \in \{ \omega \in \mathcal{S}_n | \omega$ is $\alpha$-clumping$\}$. In other words, $\sigma$ belongs to the target of the map $\text{iper}_\alpha$. Thus, $\sigma$ has a unique preimage under the map $\text{iper}_\alpha$ (since the map $\text{iper}_\alpha$ is bijective (by Proposition 6.58(a))). In other words, there exists a unique $\tau \in \mathcal{S}_\ell$ satisfying $\sigma = \text{iper}_\alpha \tau$. Since $\text{iper}_\alpha \tau = \text{iper}(\alpha, \tau)$ for every $\tau \in \mathcal{S}_{p+q}$ (by the definition of $\text{iper}_\alpha$), this rewrites as follows: There exists a unique $\tau \in \mathcal{S}_\ell$ satisfying $\sigma = \text{iper}(\alpha, \tau)$. This proves Proposition 6.58(b).

Next, let us prove that $\text{Inv}(\omega^{-1}) \subset \text{Inv}(\eta^{-1})$.

Indeed, let $(i, j)$ be some element of $\text{Inv}(\omega^{-1})$. We want to prove that $(i, j) \in \text{Inv}(\eta^{-1})$.

We have $(i, j) \in \text{Inv}(\omega^{-1})$. By the definition of $\text{Inv}(\omega^{-1})$, this yields that $(i, j)$ is an element of $\{1, 2, \ldots, n\}^2$ satisfying $i < j$ and $\omega^{-1}(i) > \omega^{-1}(j)$.

The intervals $I_1, I_2, \ldots, I_\ell$ form a set partition of $[0 : n]^+$ (according to Remark 6.48(c)). Hence, there exists some $u \in \{1, 2, \ldots, \ell\}$ such that $i \in I_u$ (because $i \in \{1, 2, \ldots, n\} = [0 : n]^+$). Similarly, there exists some $v \in \{1, 2, \ldots, \ell\}$ such that $j \in I_v$. Consider these $u$ and $v$.

First, let us assume (for the sake of contradiction) that $u = v$. Then, $i \in I_u = I_v$ (since $u = v$) and $j \in I_u$. But (11.1069) (applied to $v$ instead of $i$) yields that the restriction of the map $\omega^{-1}$ to the interval $I_v$ is increasing. Hence, $\omega^{-1}(i) \leq \omega^{-1}(j)$ (since the elements $i$ and $j$ both lie in the interval $I_v$ and satisfy $i < j$), which contradicts $\omega^{-1}(i) > \omega^{-1}(j)$. This contradiction shows that our assumption (that $u = v$) was wrong. Hence, we have $u \neq v$. Thus, $(i, j) \notin \text{Inv}(\omega^{-1})$ (since $i \neq j$ and $\omega^{-1}$ is a permutation).

Define two elements $i$ and $j$ of $\{1, 2, \ldots, \ell\}$ by $i = \tau^{-1}(u)$ and $j = \tau^{-1}(v)$. Then, $\tau(i) = u$ (since $i = \tau^{-1}(u)$), so that $I_{\tau(i)} = I_u$ and thus $i \in I_u = I_{\tau(i)}$ (since $I_{\tau(i)} = I_u$). Similarly, $j \in I_{\tau(j)}$. We have $\tau^{-1}(u) \neq \tau^{-1}(v)$.

Now, let us assume (for the sake of contradiction) that $i \leq j$. Combined with $i \neq j$, this yields $i < j$. Hence, $\omega^{-1}(I_{\tau(i)}) < \omega^{-1}(I_{\tau(j)})$ (due to (11.1066)). Thus, every $i' \in \omega^{-1}(I_{\tau(i)})$ and $j' \in \omega^{-1}(I_{\tau(j)})$ satisfy $i' < j'$ (by the definition of $\omega^{-1}(I_{\tau(i)}) < \omega^{-1}(I_{\tau(j)})$). Applying this to $i' = \omega^{-1}(i)$ and $j' = \omega^{-1}(j)$, we obtain $\omega^{-1}(i) < \omega^{-1}(j)$ (since $\omega^{-1}(\tau(i)) < \omega^{-1}(\tau(j))$ and $\omega^{-1}(\tau(i)) \in \omega^{-1}(I_{\tau(i)})$, which contradicts $\omega^{-1}(i) > \omega^{-1}(j)$. This contradiction proves that our assumption (that $i \leq j$) was wrong. Hence, we have $i > j$.

Now, recall that the word $I_{\tau(i)}$ is defined as the list of all elements of $I_{\tau(i)}$ in increasing order. Hence, $i$ is a letter of the word $I_{\tau(i)}$ (since $i$ is an element of $I_{\tau(i)}$). Similarly, $j$ is a letter of the word $I_{\tau(j)}$. Both words $I_{\tau(i)}$ and $I_{\tau(j)}$ are factors of the concatenation $I_{\tau(1)}I_{\tau(2)}\cdots I_{\tau(\ell)}$, with the factor $I_{\tau(i)}$ appearing after the factor $I_{\tau(j)}$ (since $i > j$). Thus, if $i_0$ is any letter of the word $I_{\tau(i)}$, and if $j_0$ is any letter of the word $I_{\tau(j)}$, then the letter $i_0$ appears after the letter $j_0$ in the concatenation $I_{\tau(i)}I_{\tau(2)}\cdots I_{\tau(\ell)}$. Applying this to $i_0 = i$ and $j_0 = j$, we conclude that the letter $i$ appears after the letter $j$ in the concatenation $I_{\tau(1)}I_{\tau(2)}\cdots I_{\tau(\ell)}$ (since $i$ is a letter of the word $I_{\tau(i)}$, and since $j$ is a letter of the word $I_{\tau(j)}$). In other words, the letter $i$ appears after the letter $j$ in the word $\eta$ (since $\eta = I_{\tau(1)}I_{\tau(2)}\cdots I_{\tau(\ell)}$). In other words, $\eta^{-1}(i) > \eta^{-1}(j)$.

So $(i, j)$ is an element of $\{1, 2, \ldots, n\}^2$ satisfying $i < j$ and $\eta^{-1}(i) > \eta^{-1}(j)$. In other words, $(i, j) \in \text{Inv}(\eta^{-1})$ (by the definition of $\text{Inv}(\eta^{-1})$).

Now, let us forget that we fixed $(i, j)$, we thus prove that $(i, j) \in \text{Inv}(\eta^{-1})$ for every $(i, j) \in \text{Inv}(\omega^{-1})$. In other words,$$
\text{Inv}(\omega^{-1}) \subset \text{Inv}(\eta^{-1}).
$$
Combined with (11.1070), this yields $\text{Inv}(\omega^{-1}) = \text{Inv}(\eta^{-1})$. Thus, (11.1064) (applied to $n = n$, $\varphi = \omega^{-1}$ and $\psi = \eta^{-1}$) yields $\omega^{-1} = \eta^{-1}$, whence $\omega = \eta = (\text{iper}_\alpha \circ \text{iper}_\alpha')(\omega)$ (by (11.1067)), so that $(\text{iper}_\alpha \circ \text{iper}_\alpha')(\omega) = \omega = \text{id}(\omega)$.

Now, let us forget that we fixed $\omega$. We thus have proven that $(\text{iper}_\alpha \circ \text{iper}_\alpha')(\omega) = \text{id}(\omega)$ for every $\alpha$-clumping permutation $\omega \in \mathcal{S}_n$. In other words, $\text{iper}_\alpha \circ \text{iper}_\alpha' = \text{id}$. Q.E.D.

819 Proof. Let $\pi \in \mathcal{S}_\ell$. We shall show that that $(\text{iper}_\alpha' \circ \text{iper}_\alpha')(\pi) = \text{id}(\pi)$.

Indeed, let $\omega = \text{iper}_\alpha \pi$. Then, $\omega = \text{iper}_\alpha \pi = \text{iper}(\alpha, \pi)$ (by the definition of $\text{iper}_\alpha \pi$). The permutation $\omega = \text{iper}(\alpha, \pi)$ is $\alpha$-clumping (according to Proposition 6.56(c), applied to $\pi = \tau$).

Let us now recall how $\text{iper}_\alpha'(\omega)$ was defined: Every $i \in \{1, 2, \ldots, \ell\}$ has the property that $\omega^{-1}(I_i)$ is an interval. These intervals $\omega^{-1}(I_1), \omega^{-1}(I_2), \ldots, \omega^{-1}(I_\ell)$ form a totally ordered set with respect to the relation $<$. Then, $\text{iper}_\alpha'(\omega)$ is defined as the unique permutation $\tau \in \mathcal{S}_\ell$ such that $\omega^{-1}(I_{\tau(1)}) < \omega^{-1}(I_{\tau(2)}) < \cdots < \omega^{-1}(I_{\tau(\ell)})$. Hence, if $\tau \in \mathcal{S}_\ell$ is a permutation satisfying $\omega^{-1}(I_{\tau(1)}) < \omega^{-1}(I_{\tau(2)}) < \cdots < \omega^{-1}(I_{\tau(\ell)})$, then

(11.1071)
$$
\tau = \text{iper}_\alpha'(\omega).
$$

We shall now use this to prove that $\pi = \text{iper}_\alpha'(\omega)$.

We have $\omega = \text{iper}(\alpha, \pi) = I_{\tau(1)}I_{\tau(2)}\cdots I_{\tau(\ell)}$ (in one-line notation), according to the definition of $\text{iper}(\alpha, \pi)$.

The intervals $\omega^{-1}(I_1), \omega^{-1}(I_2), \ldots, \omega^{-1}(I_\ell)$ form a set partition of $[0 : n]^+$ (since the intervals $I_1, I_2, \ldots, I_\ell$ form a set partition of $[0 : n]^+$), and thus are pairwise disjoint. These intervals $\omega^{-1}(I_1), \omega^{-1}(I_2), \ldots, \omega^{-1}(I_\ell)$ are also nonempty (since the intervals $I_1, I_2, \ldots, I_\ell$ are nonempty).

Proof of Proposition 6.60. We know that $\alpha \beta$ is a composition of $n + m$ having length $\ell(\alpha \beta) = \ell(\alpha) + \ell(\beta) = p + q$. Hence, the interval system corresponding to $\alpha \beta$ is a $(p + q)$-tuple of intervals which covers $[0 : n + m]^+$. Denote this $(p + q)$-tuple by $(I_1, I_2, \ldots, I_{p+q})$. It is clear that $I_1 \cup I_2 \cup \cdots \cup I_p = [0 : n]^+$ and $I_{p+1} \cup I_{p+2} \cup \cdots \cup I_{p+q} = [n : n + m]^+$ (since the first $p$ parts of the composition $\alpha \beta$ form the composition $\alpha$ of $n$). Moreover, $I_1 < I_2 < \cdots < I_{p+q}$ (since $(I_1, I_2, \ldots, I_{p+q})$ is the interval system corresponding to $\alpha \beta$).

The definition of $\text{iper}(\alpha \beta, \tau)$ yields that $\text{iper}(\alpha \beta, \tau) = \frac{I_{\tau(1)}}{I_{\tau(2)}} \cdots \frac{I_{\tau(p+q)}}{I_{\tau(p+q)}}$ (in one-line notation). Denote the permutation $\text{iper}(\alpha \beta, \tau)$ by $\omega$; then, this becomes $\omega = \frac{I_{\tau(1)}}{I_{\tau(2)}} \cdots \frac{I_{\tau(p+q)}}{I_{\tau(p+q)}}$. For every $j \in \{1, 2, \ldots, p + q\}$, the restriction of the map $\omega^{-1}$ to the interval $I_{\tau(j)}$ is increasing (by Proposition 6.56(b), applied to $n + m$, $\alpha \beta$, $p + q$ and $\omega$ instead of $n$, $\alpha$, $\ell$ and $\sigma$). Substituting $k$ for $\tau(j)$ here, we obtain: For every $k \in \{1, 2, \ldots, p + q\}$, the restriction of the map $\omega^{-1}$ to the interval $I_k$ is strictly increasing (because $\omega^{-1}$ is injective).

Now, we need to prove that $\tau \in \text{Sh}_{p,q}$ if and only if $\text{iper}(\alpha \beta, \tau) \in \text{Sh}_{n,m}$. In other words, we need to prove that $\tau \in \text{Sh}_{p,q}$ if and only if $\omega \in \text{Sh}_{n,m}$ (since $\omega = \text{iper}(\alpha \beta, \tau)$). We shall prove the $\Rightarrow$ and $\Leftrightarrow$ directions of this statement separately:

$\Rightarrow$: Assume that $\tau \in \text{Sh}_{p,q}$. We need to show that $\omega \in \text{Sh}_{n,m}$.

We have $\tau \in \text{Sh}_{p,q}$. Thus, $\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p)$ and $\tau^{-1}(p + 1) < \tau^{-1}(p + 2) < \cdots < \tau^{-1}(p + q)$.

Now, let $i$ and $j$ be two elements of $\{1, 2, \ldots, n\}$ such that $i < j$. We are going to prove that $\omega^{-1}(i) < \omega^{-1}(j)$.

Indeed, we have $i \in \{1, 2, \ldots, n\} = [0 : n]^+ = I_1 \cup I_2 \cup \cdots \cup I_p$, so that $i \in I_{\tau(i)}$ for some $i' \in \{1, 2, \ldots, p\}$. Similarly, $j \in I_{\tau(j)}$ for some $j' \in \{1, 2, \ldots, p\}$. Consider these $i'$ and $j'$. We cannot have $j' < i'$ (because this would entail $I_{\tau(j')} < I_{\tau(i')}$, which would lead to $j < i$, since $\omega^{-1}(i) < j < j'$). Hence, we must have either $j' = j'$ or $j' > i'$. If $j' = i'$, then $i$ and $j$ lie in one and the same $I_k$ (namely, the one with $k = j' = i'$), and so $\omega^{-1}(i) < \omega^{-1}(j)$ follows from the fact that the map $\omega^{-1}$ restricted to $I_k$ is strictly increasing (and from the inequality $i < j$). Hence, it only remains to consider the case when $j' > i'$. Assume WLOG that we are in this case. Then, $i' < j'$, and so $\tau^{-1}(i') < \tau^{-1}(j')$ (since $\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p)$ and $i'$ and $j'$ belong to $\{1, 2, \ldots, p\}$, and thus the word $I_{\tau(j')}$ appears before the word $I_{\tau(j)}$ in the concatenation $I_{\tau(1)}I_{\tau(2)} \cdots I_{\tau(p+q)}$, hence, the letter $i$ appears before the letter $j$.

Now, let $u \in \{1, 2, \ldots, \ell - 1\}$. The intervals $\omega^{-1}(I_{\tau(u)})$ and $\omega^{-1}(I_{\tau(u+1)})$ are nonempty (since the intervals $\omega^{-1}(I_1)$, $\omega^{-1}(I_2)$, \ldots, $\omega^{-1}(I_t)$ are nonempty). We will show that $\omega^{-1}(I_{\tau(u)}) < \omega^{-1}(I_{\tau(u+1)})$.

Indeed, let $i \in \omega^{-1}(I_{\tau(u)})$ and $j \in \omega^{-1}(I_{\tau(u+1)})$ be arbitrary. We shall prove that $i < j$.

We have $i \in \omega^{-1}(I_{\tau(u)})$, so that $\omega(i) \in I_{\tau(u)}$. Hence, $\omega(i)$ is a letter of the word $I_{\tau(u)}$ (since the word $I_{\tau(u)}$ is defined as the list of all elements of $\omega(I_{\tau(u)})$ in increasing order). Similarly, $\omega(j)$ is a letter of the word $I_{\tau(u+1)}$ (since $j \in \omega^{-1}(I_{\tau(u+1)})$).

Both words $I_{\tau(u)}$ and $I_{\tau(u+1)}$ are factors of the concatenation $I_{\tau(1)}I_{\tau(2)} \cdots I_{\tau(\ell)}$, with the factor $I_{\tau(u)}$ appearing after the factor $I_{\tau(u+1)}$. Thus, if $i_0$ is any letter of the word $I_{\tau(u)}$, and if $j_0$ is any letter of the word $I_{\tau(u+1)}$, then the letter $j_0$ appears after the letter $i_0$ in the concatenation $I_{\tau(u)}I_{\tau(u+1)} \cdots I_{\tau(\ell)}$. Applying this to $i_0 = \omega(i)$ and $j_0 = \omega(j)$, we conclude that the letter $\omega(j)$ appears after the letter $\omega(i)$ in the concatenation $I_{\tau(u)}I_{\tau(u+1)} \cdots I_{\tau(\ell)}$ (since $\omega(i)$ is a letter of the word $I_{\tau(u)}$, and since $\omega(j)$ is a letter of the word $I_{\tau(u+1)}$). In other words, the letter $\omega(j)$ appears after the letter $\omega(i)$ in the word $\omega$ (since $\omega = I_{\tau(1)}I_{\tau(2)} \cdots I_{\tau(\ell)}$). In other words, $\omega^{-1}(\omega(j)) > \omega^{-1}(\omega(i))$. In other words, $j > i$, so that $i < j$.

Now, let us forget that we fixed $i$ and $j$. We thus have proven that every $i \in \omega^{-1}(I_{\tau(u)})$ and $j \in \omega^{-1}(I_{\tau(u+1)})$ satisfy $i < j$. In other words, $\omega^{-1}(I_{\tau(u)}) < \omega^{-1}(I_{\tau(u+1)})$ (by the definition of $\omega^{-1}(I_{\tau(u)}) < \omega^{-1}(I_{\tau(u+1)})$).

Now, let us forget that we fixed $u$. We thus have proven that $\omega^{-1}(I_{\tau(u)}) < \omega^{-1}(I_{\tau(u+1)})$ for every $u \in \{1, 2, \ldots, \ell - 1\}$. In other words, $\omega^{-1}(I_{\tau(1)}) < \omega^{-1}(I_{\tau(2)}) < \cdots < \omega^{-1}(I_{\tau(\ell)})$. Therefore, (11.1071) (applied to $\tau = \pi$) yields $\pi = \text{iper}^u_{\omega}(\text{iper}_u \pi) = (\text{iper}_u \circ \text{iper}_u) \pi$. Hence, $(\text{iper}_u \circ \text{iper}_u) \pi = \pi = \text{id}(\pi)$.

Now, let us forget $u$ and $\pi$. We thus have proven that $(\text{iper}_u \circ \text{iper}_u) \pi = \text{id}(\pi)$ for every $\pi \in \varpi$. In other words, $\text{iper}_u \circ \text{iper}_u = \text{id}$, qed.
appears in the word \( \omega \) appears in \( \mathcal{I}_\tau \). Since \( I_{\tau(1)} \rightarrow I_{\tau(2)} \rightarrow I_{\tau(p+q)} = \omega \), this shows that the letter \( i \) appears before the letter \( j \) in the word \( \omega \). In other words, \( \omega^{-1}(i) < \omega^{-1}(j) \).

So we have shown that if two elements \( i \) and \( j \) of \( \{1, 2, \ldots, n\} \) satisfy \( i < j \), then \( \omega^{-1}(i) < \omega^{-1}(j) \). In other words, \( \omega^{-1}(1) < \omega^{-1}(2) < \cdots < \omega^{-1}(n) \). Similarly, \( \omega^{-1}(n+1) < \omega^{-1}(n+2) < \cdots < \omega^{-1}(n+m) \).

Combining these two chains of inequalities, we conclude that \( \omega \in \mathcal{S}_{n,m} \). This proves the \( \implies \) direction.

\[ \iff: \] Assume that \( \omega \in \mathcal{S}_{n,m} \). We need to prove that \( \tau \in \mathcal{S}_{p,q} \).

We have \( \omega \in \mathcal{S}_{n,m} \). Thus, \( \omega^{-1}(1) < \omega^{-1}(2) < \cdots < \omega^{-1}(n) \) and \( \omega^{-1}(n+1) < \omega^{-1}(n+2) < \cdots < \omega^{-1}(n+m) \).

Let \( i' \) and \( j' \) be two elements of \( \{1, 2, \ldots, p\} \) such that \( i' < j' \). We are going to prove that \( \tau^{-1}(i') < \tau^{-1}(j') \).

Indeed, assume the contrary. Then, \( \tau^{-1}(i') \geq \tau^{-1}(j') \), thus \( \tau^{-1}(i') > \tau^{-1}(j') \) (since \( \tau \) is a permutation).

Hence, the word \( \mathcal{I}_\tau \) appears after\(^{20} \) the word \( \mathcal{I}_{i'} \) in the concatenation \( I_{\tau(1)} \rightarrow I_{\tau(2)} \rightarrow \cdots \rightarrow I_{\tau(p+q)} \) (because the letter \( i \) appears in the word \( \mathcal{I}_\tau \), whereas the letter \( j \) appears in the word \( \mathcal{I}_{i'} \), and we know that the word \( \mathcal{I}_{i'} \) appears after the word \( \mathcal{I}_\tau \) in the concatenation \( I_{\tau(1)} \rightarrow I_{\tau(2)} \rightarrow \cdots \rightarrow I_{\tau(p+q)} \)). In other words, the letter \( i \) appears after the letter \( j \) in the word \( \omega \) (since \( I_{\tau(1)} \rightarrow I_{\tau(2)} \rightarrow \cdots \rightarrow I_{\tau(p+q)} = \omega \)). In other words, \( \omega^{-1}(i) > \omega^{-1}(j) \).

But \( i, j \in \{1, 2, \ldots, p\} \), so that \( I_{\tau(1)} \supset I_{\tau(2)} \supset \cdots \supset I_{\tau(p+q)} = [0 : n]^+ \) and thus \( i \in I_{\tau(1)} \cap [0 : n]^+ \). Similarly, \( j \in [0 : n]^+ \). Since the map \( \omega^{-1} \) restricted to \( [0 : n]^+ \) is strictly increasing (since \( \omega^{-1}(1) < \omega^{-1}(2) < \cdots < \omega^{-1}(n) \)), we thus have \( \omega^{-1}(i) < \omega^{-1}(j) \) (since \( i < j \)), contradicting \( \omega^{-1}(i) > \omega^{-1}(j) \). This contradiction shows that our assumption was wrong, and so we have shown that \( \tau^{-1}(i') < \tau^{-1}(j') \).

Thus, we have proven that if two elements \( i' \) and \( j' \) of \( \{1, 2, \ldots, p\} \) satisfy \( i' < j' \), then \( \tau^{-1}(i') < \tau^{-1}(j') \). In other words, \( \tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p) \).

The combination of these two chains of inequalities shows that \( \tau \in \mathcal{S}_{p,q} \). Thus, the \( \iff \) direction is proven. The proof of Proposition 6.60 is thus complete. \( \square \)

---


Proof of Lemma 6.62. We have \( \tau \in \mathcal{S}_{p,q} \) if and only if \( \text{iper}(\alpha \beta, \tau) \in \mathcal{S}_{n,m} \) (by Proposition 6.60). Since we know that \( \tau \in \mathcal{S}_{p,q} \), we thus conclude that \( \text{iper}(\alpha \beta, \tau) \in \mathcal{S}_{n,m} \). In other words, \( \sigma \in \mathcal{S}_{n,m} \) (since \( \sigma = \text{iper}(\alpha \beta, \tau) \)).

Write the composition \( \alpha \beta \) in the form \( (\gamma_1, \gamma_2, \ldots, \gamma_{p+q}) \). For every \( j \in \{0, 1, \ldots, p+q\} \), let \( s_j \) denote the integer \( \sum_{k=1}^j \gamma_{\tau(k)} \). Then, \( 0 = s_0 < s_1 < s_2 < \cdots < s_{p+q} = n + m \).

Let \( w \) denote the word \( u \cup v \).

The first \( p \) parts of the composition \( \alpha \beta \) form the composition \( \alpha \) of \( n \). Hence, the interval system \( (I_1, I_2, \ldots, I_{p+q}) \) corresponding to \( \alpha \beta \) satisfies \( I_1 \cup I_2 \cup \cdots \cup I_p = [0 : n]^+ \) and \( I_{p+1} \cup I_{p+2} \cup \cdots \cup I_{p+q} = [n : n+m]^+ \).

Let \( i \in \{1, 2, \ldots, p+q\} \). We know that \( I_{\tau(i)} \) is an interval of \( \mathbb{Z} \) satisfying either \( I_{\tau(i)} \subset [0 : n]^+ \) or \( I_{\tau(i)} \subset [n : n+m]^+ \) (in fact, if \( \tau(i) \leq p \), then \( I_{\tau(i)} \subset I_1 \cup I_2 \cup \cdots \cup I_p = [0 : n]^+ \), whereas otherwise, \( I_{\tau(i)} \subset I_{p+1} \cup I_{p+2} \cup \cdots \cup I_{p+q} = [n : n+m]^+ \)). Also, Proposition 6.56(a) (applied to \( n+m \), \( \alpha \beta \), \( (\gamma_1, \gamma_2, \ldots, \gamma_{p+q}) \)),

\(^{20}\)"after" does not imply "immediately after".
We have

\begin{equation}
\sigma^{-1}(I_{\tau(i)}) = \left[ \sum_{k=1}^{i-1} \gamma_{\tau(k)} : \sum_{k=1}^{i} \gamma_{\tau(k)} \right] = [s_{i-1} : s_i]^+.
\end{equation}

Consequently, \( \sigma^{-1}(I_{\tau(i)}) \) is an interval. Therefore, we can apply Lemma 6.52(a) to \( I_{\tau(i)} \) instead of \( I \). As a result, we obtain

\[ (u \shuffle_\sigma v) \left[ \sigma^{-1}(I_{\tau(i)}) \right] = (uv) \left[ I_{\tau(i)} \right]. \]

Hence,

\begin{equation}
(uv) \left[ I_{\tau(i)} \right] = \left( u \shuffle_\sigma v \right) \left[ \sigma^{-1}(I_{\tau(i)}) \right] = w \left[ [s_{i-1} : s_i]^+ \right].
\end{equation}

Now, forget that we fixed \( i \). We thus have proven (11.1073) for every \( i \in \{1, 2, \ldots, p+q\} \). Now,

\[ (uv) \left[ I_{\tau(1)} \right] \cdot (uv) \left[ I_{\tau(2)} \right] \cdots (uv) \left[ I_{\tau(p+q)} \right] = w[s_0 : s_1]^+ \cdot w[s_1 : s_2]^+ \cdots w[s_{p+q-1} : s_{p+q}]^+ \]

(by (11.1073))

\[ = w \left[ [s_0 : s_1]^+ \right] \cdots w \left[ [s_0 : s_{p+q}]^+ \right] = w \left[ [0 : n+m]^+ \right] (\text{since } s_0 = 0 \text{ and } s_{p+q} = n+m) \]

\[ = u \shuffle_\sigma v. \]

This proves Lemma 6.62. \( \square \)


**Proof of Proposition 6.65.** We have \( h(1) \geq h(2) \geq \cdots \geq h(p) \). Hence, for every \( w \in \mathfrak{W} \), the set \( \{ i \in \{1, 2, \ldots, p\} \mid h(i) = w \} \) is a (possibly empty) interval of \( \{1, 2, \ldots, p\} \), and will be denoted by \( P_w \).

Similarly, for every \( w \in \mathfrak{W} \), the set \( \{ i \in \{p+1, p+2, \ldots, p+q\} \mid h(i) = w \} \) is a (possibly empty) interval of \( \{p+1, p+2, \ldots, p+q\} \), and will be denoted by \( Q_w \). Every \( w \in \mathfrak{W} \) satisfies \( h^{-1}(w) = P_w \sqcup Q_w \). Notice that \( |P_w| = a(w) \) and \( |Q_w| = b(w) \) for all \( w \in \mathfrak{W} \).

Let \( (g_1, g_2, \ldots, g_{p+q}) \) be the result of sorting the list \( (h(1), h(2), \ldots, h(p+q)) \) in decreasing order. Then, \( g_1 \geq g_2 \geq \cdots \geq g_{p+q} \). Hence, for every \( w \in \mathfrak{W} \), the set \( \{ j \in \{1, 2, \ldots, p+q\} \mid g_j = w \} \) is a (possibly empty) interval of \( \{1, 2, \ldots, p+q\} \). Denote this interval by \( I_w \). The size of this interval is

\[ |I_w| = |\{ j \in \{1, 2, \ldots, p+q\} \mid g_j = w \}| (\text{since } I_w = \{ j \in \{1, 2, \ldots, p+q\} \mid g_j = w \}) \]

\[ = | \{ j \in \{1, 2, \ldots, p+q\} \mid h(j) = w \}| (\text{since } (g_1, g_2, \ldots, g_{p+q}) \text{ is the result of sorting the list } (h(1), h(2), \ldots, h(p+q))) \]

\[ = |P_w \sqcup Q_w| = |P_w| + |Q_w|. \]

Notice that \( \{1, 2, \ldots, p+q\} = \bigsqcup_{w \in \mathfrak{W}} I_w \) (by the definition of the \( I_w \)).

Fix \( w \in \mathfrak{W} \). Let us define an \((I_w, P_w, Q_w)\)-shuffle to mean a bijection \( \kappa : I_w \to P_w \sqcup Q_w \) having the property that the maps \( \kappa^{-1}|_{P_w} : P_w \to I_w \) and \( \kappa^{-1}|_{Q_w} : Q_w \to I_w \) are strictly increasing. It is easy to see...
that such a \((I_w, P_w, Q_w)\)-shuffle \(\kappa\) is uniquely determined by the subset \(\kappa^{-1}(P_w)\) of \(I_w\), and that for a given subset \(U\) of \(I_w\), such a \((I_w, P_w, Q_w)\)-shuffle \(\kappa\) satisfying \(\kappa^{-1}(P_w) = U\) exists if and only if \(|U| = |P_w|\). Hence, there are as many \((I_w, P_w, Q_w)\)-shuffles as there are subsets \(U\) of \(I_w\) satisfying \(|U| = |P_w|\). In other words,

\[
\text{(the number of } (I_w, P_w, Q_w)\text{-shuffles)} = \left(\frac{|I_w|}{|P_w|}\right) = \binom{a(w) + b(w)}{a(w)}
\]

(since \(|I_w| = |P_w| + |Q_w| = a(w) + b(w)\) and \(|P_w| = a(w))\).

Now, forget that we fixed \(w\).

Consider any \(\tau \in \text{Sh}_{p,q}\) satisfying \(h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))\). Then, \(h(\tau(1)), h(\tau(2)), \ldots, h(\tau(p+q))\) is the result of sorting the list \((g_{1}, g_{2}, \ldots, g_{p+q})\) \(821\). In other words, \(h(\tau(j)) = g_{j}\) for every \(j \in \{1, 2, \ldots, p+q\}\). Thus,

\[
\tau(I_w) = P_w \sqcup Q_w \quad \text{for every } w \in \mathfrak{W}
\]

\(822\). As a consequence, for every \(w \in \mathfrak{W}\), the bijection \(\tau : \{1, 2, \ldots, p+q\} \to \{1, 2, \ldots, p+q\}\) restricts to a bijection \(\tau_w : I_w \to P_w \sqcup Q_w\). This bijection \(\tau_w\) is an \((I_w, P_w, Q_w)\)-shuffle \(823\). Thus, we have obtained a family \((\tau_w)_{w \in \mathfrak{W}}\) of \((I_w, P_w, Q_w)\)-shuffles parametrized over all \(w \in \mathfrak{W}\).

Now, let us forget that we fixed \(w\). We thus have constructed, for every \(\tau \in \text{Sh}_{p,q}\) satisfying \(h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))\), a family \((\tau_w)_{w \in \mathfrak{W}}\) of \((I_w, P_w, Q_w)\)-shuffles parametrized over all \(w \in \mathfrak{W}\). We thus obtain a map

\[
\{\tau \in \text{Sh}_{p,q} \mid h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))\} \to \prod_{w \in \mathfrak{W}} \left(\text{the set of all } (I_w, P_w, Q_w)\text{-shuffles}, \quad \tau \mapsto (\tau_w)_{w \in \mathfrak{W}}\right).
\]

This map is injective \(824\) and surjective \(825\). Hence, it is bijective, and this yields that the sets

\[
\{\tau \in \text{Sh}_{p,q} \mid h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))\}
\]

\(821\)Proof. The list \((h(\tau(1)), h(\tau(2)), \ldots, h(\tau(p+q)))\) is a permutation of the list \((h(1), h(2), \ldots, h(p+q))\) (since \(\tau \in \text{Sh}_{p,q}\), but is weakly decreasing (since \(h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))\)). Hence, the list \((h(\tau(1)), h(\tau(2)), \ldots, h(\tau(p+q)))\) is the result of sorting the list \((h(1), h(2), \ldots, h(p+q))\) in decreasing order. But this yields that \((h(\tau(1)), h(\tau(2)), \ldots, h(\tau(p+q))) = (g_{1}, g_{2}, \ldots, g_{p+q})\) (since \((g_{1}, g_{2}, \ldots, g_{p+q})\), too, is the result of sorting the list \((h(1), h(2), \ldots, h(p+q))\) in decreasing order), qed.

\(822\)Proof of (11.1075): Let \(w \in \mathfrak{W}\). Then,

\[
I_w = \left\{ j \in \{1, 2, \ldots, p+q\} \mid g_{j} = w \right\} = \left\{ j \in \{1, 2, \ldots, p+q\} \mid (h \circ \tau)(j) = w \right\}
\]

\(823\)Proof. We have \(\tau \in \text{Sh}_{p,q}\), so that \(\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p)\) and \(\tau^{-1}(p+1) < \tau^{-1}(p+2) < \cdots < \tau^{-1}(p+q)\). In other words, the restriction of \(\tau^{-1}\) to \([1, 2, \ldots, p]\) is strictly increasing, and the restriction of \(\tau^{-1}\) to \([p+1, p+2, \ldots, p+q]\) is strictly increasing. Since the restriction of \(\tau^{-1}\) to \([1, 2, \ldots, p]\) is strictly increasing, the restriction of \(\tau^{-1}\) to \(P_w\) must also be strictly increasing (because \(P_w \subset [1, 2, \ldots, p]\)). But this restriction is \(\tau_{w}^{-1}|_{P_w}: P_w \to I_w\). Hence, \(\tau_{w}^{-1}|_{P_w}: P_w \to I_w\) is strictly increasing. Similarly, the same can be said about \(\tau_{w}^{-1}|_{Q_w}: Q_w \to I_w\). Since the maps \(\tau_{w}^{-1}|_{P_w}: P_w \to I_w\) and \(\tau_{w}^{-1}|_{Q_w}: Q_w \to I_w\) are strictly increasing, we conclude that \(\tau_{w}\) is an \((I_w, P_w, Q_w)\)-shuffle (by the definition of a \((I_w, P_w, Q_w)\)-shuffle), qed.

\(824\)In fact, any \(\tau \in \text{Sh}_{p,q}\) is uniquely determined by \((\tau_{w})_{w \in \mathfrak{W}}\) (because \(\tau_{w}\) is the restriction of \(\tau\) to \(I_w\) (with a restricted codomain, but this doesn’t matter right now), and so knowing \((\tau_{w})_{w \in \mathfrak{W}}\) means knowing the values \(\tau\) on each of the intervals \(I_w\); but this means knowing all values of \(\tau\), because \([1, 2, \ldots, p+q] = \bigcup_{w \in \mathfrak{W}} I_w\).

\(825\)Proof. We need to show that for every

\[
(\sigma_{w})_{w \in \mathfrak{W}} \in \prod_{w \in \mathfrak{W}} \left(\text{the set of all } (I_w, P_w, Q_w)\text{-shuffles}\right),
\]

there exists a \(\tau \in \text{Sh}_{p,q}\) satisfying \(h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))\) and \((\tau_{w})_{w \in \mathfrak{W}} = (\sigma_{w})_{w \in \mathfrak{W}}\).
and
\[\prod_{w \in \mathfrak{M}} \text{(the set of all } (I_w, P_w, Q_w)\text{-shuffles)}\]
are in bijection. Thus,
\[\begin{align*}
|\{\tau \in \mathbb{S}_{p,q} \mid h(\tau(1)) &\geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))\}| \\
= &\prod_{w \in \mathfrak{M}} \text{(the set of all } (I_w, P_w, Q_w)\text{-shuffles)} \\
= &\prod_{w \in \mathfrak{M}} \text{(the set of all } (I_w, P_w, Q_w)\text{-shuffles)} \\
= &\text{(the number of } (I_w, P_w, Q_w)\text{-shuffles)} \left(\frac{a(w) + b(w)}{a(w)}\right) \\
= &\prod_{w \in \mathfrak{M}} \left(\frac{a(w) + b(w)}{a(w)}\right).
\end{align*}\]
(by (11.1074))

This is precisely the statement of Proposition 6.65. \(\square\)

So fix some \((\sigma_w)_{w \in \mathfrak{M}} \in \prod_{w \in \mathfrak{M}}\) (the set of all \((I_w, P_w, Q_w)\)-shuffles). For every \(w \in \mathfrak{M}\), the map \(\sigma_w\) is an \((I_w, P_w, Q_w)\)-shuffle, hence a bijection from \(I_w\) to \(P_w \cup Q_w\). Since \(\prod_{w \in \mathfrak{M}} I_w = \{1, 2, \ldots, p+q\}\) and \(\prod_{w \in \mathfrak{M}} (P_w \cup Q_w) = \bigcup_{w \in \mathfrak{M}} h^{-1}(w) = \{1, 2, \ldots, p+q\}\), we can piece these bijections \(\sigma_w\) together to a bijection
\[\prod_{w \in \mathfrak{M}} \sigma_w : \{1, 2, \ldots, p+q\} \to \{1, 2, \ldots, p+q\},\]
whose restriction to each interval \(I_w\) coincides with the respective \(\sigma_w\) (except that the codomains of the maps are different). Let \(\tau\) be this bijection \(\prod_{w \in \mathfrak{M}} \sigma_w\). Clearly, \(\tau \in \mathfrak{S}_{p,q}\). We will now show that \(\tau \in \mathbb{S}_{p,q}\) (because \(h(1) \geq h(2) \geq \cdots \geq h(p+q)\)) and \((\tau_w)_w \in \mathfrak{M} = \{\sigma_w\}_{w \in \mathfrak{M}}\). Once this is done, the required surjectivity will clearly follow.

Since \(\tau = \prod_{w \in \mathfrak{M}} \sigma_w\), we have \(\tau|_{I_w} = \sigma_w\) for every \(w \in \mathfrak{M}\) (up to the fact that the maps \(\tau|_{I_w}\) and \(\sigma_w\) have different codomains). More precisely, \(\tau_w = \sigma_w\) for every \(w \in \mathfrak{M}\). Hence, for every \(w \in \mathfrak{M}\), the map \(\sigma_w\) is a restriction of \(\tau\) (with appropriately restricted codomain).

The map \(\tau = \prod_{w \in \mathfrak{M}} \sigma_w\) is pieced together from bijections \(\sigma_w : I_w \to P_w \cup Q_w\). Thus, \(\tau(I_w) = P_w \cup Q_w\) for every \(w \in \mathfrak{M}\).

In other words, \(\tau(I_w) = h^{-1}(w)\) for every \(w \in \mathfrak{M}\) (since every \(w \in \mathfrak{M}\) satisfies \(h^{-1}(w) = P_w \cup Q_w\)). Hence,
\[(11.1076) \quad \text{every } i \in \{1, 2, \ldots, p+q\} \text{ satisfies } h(\tau(i)) = g(i).\]

Proof of (11.1076): Let \(i \in \{1, 2, \ldots, p+q\}\). Set \(w = h(\tau(i))\). Then, \(\tau(i) \in h^{-1}(w) = \tau(I_w)\) (since \(\tau(I_w) = h^{-1}(w)\)) and thus \(i \in I_w\) (since \(\tau\) is a bijection), so that \(i \in I_w = \{j \in \{1, 2, \ldots, p+q\} \mid g_j = w\}\) and thus \(g_i = w = h(\tau(i))\). This proves (11.1076).

We have \(g_1 \geq g_2 \geq \cdots \geq g_{p+q}\). Due to (11.1076), this rewrites as \(h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))\).

We now are going to prove that \(\tau \in \mathbb{S}_{p,q}\). In order to prove this, it is enough to show that \(\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p+q)\) and \(\tau^{-1}(p+1) < \tau^{-1}(p+2) < \cdots < \tau^{-1}(p+q)\). We shall only verify \(\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p)\) (the proof of \(\tau^{-1}(p+1) < \tau^{-1}(p+2) < \cdots < \tau^{-1}(p+q)\) being analogous).

So let \(i\) and \(j\) be elements of \(\{1, 2, \ldots, p\}\) such that \(i < j\). We will prove that \(\tau^{-1}(i) < \tau^{-1}(j)\). Indeed, assume the contrary. Then, \(\tau^{-1}(i) \geq \tau^{-1}(j)\). In other words, \(\tau^{-1}(j) \leq \tau^{-1}(i)\). Thus, \(h(\tau^{-1}(j)) \geq h(\tau^{-1}(i))\) (since \(h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))\)). In other words, \(h(j) \geq h(i)\). Combined with \(h(i) = h(j)\) (since \(h(1) \geq h(2) \geq \cdots \geq h(p+q)\)) and \(i < j\), this yields \(h(i) = h(j)\). So we can define \(a \in \mathfrak{M}\) by \(w = h(i) = h(j)\). Consider this \(w\). We have \(h(i) = h(j)\), so that \(i \in h^{-1}(w) = P_w \cup Q_w\). But we cannot have \(i \in Q_w\) (because \(i\) lies in the set \(\{1, 2, \ldots, p\}\), which is disjoint to \(Q_w\) (since \(Q_w \subset \{p+1, p+2, \ldots, p+q\}\)). Since we have \(i \in P_w \cup Q_w\) but not \(i \in Q_w\), we must have \(i \in P_w\). Similarly, \(j \in P_w\). But recall that \(\sigma_w\) is a \((I_w, P_w, Q_w)\)-shuffle. In other words, \(\sigma_w : I_w \to P_w \cup Q_w\) is a bijection having the property that the maps \(\sigma^{-1}_w|_{P_w} : P_w \to I_w\) and \(\sigma^{-1}_w|_{Q_w} : Q_w \to I_w\) are strictly increasing. Since \(\sigma^{-1}_w|_{P_w} : \sigma^{-1}_w|_{P_w} < \sigma^{-1}_w|_{Q_w} : \sigma^{-1}_w|_{Q_w}\) (since \(i < j\) and \(i \in P_w\) and \(j \in P_w\)). Since \(\sigma^{-1}_w|_{P_w} : \sigma^{-1}_w|_{P_w} = \sigma^{-1}_w|_{Q_w} : \sigma^{-1}_w|_{Q_w}\) (because \(\sigma_w\) is a restriction of \(\tau\) and \(\sigma^{-1}_w|_{P_w} : \sigma^{-1}_w|_{P_w} = \sigma^{-1}_w|_{Q_w} : \sigma^{-1}_w|_{Q_w}\) (similarly), this rewrites as \(\tau^{-1}(i) < \tau^{-1}(j)\), which contradicts \(\tau^{-1}(i) \geq \tau^{-1}(j)\). This contradiction proves our assumption wrong, and so we have \(\tau^{-1}(i) < \tau^{-1}(j)\).

Let us forget that we fixed \(i\) and \(j\). We thus have seen that \(\tau^{-1}(i) < \tau^{-1}(j)\) for any elements \(i\) and \(j\) of \(\{1, 2, \ldots, p\}\) such that \(i < j\). In other words, \(\tau^{-1}(1) < \tau^{-1}(2) < \cdots < \tau^{-1}(p)\). Similarly, \(\tau^{-1}(p+1) < \tau^{-1}(p+2) < \cdots < \tau^{-1}(p+q)\). Thus, \(\tau \in \mathbb{S}_{p,q}\).

We now know that \(\tau \in \mathbb{S}_{p,q}\) and \(h(\tau(1)) \geq h(\tau(2)) \geq \cdots \geq h(\tau(p+q))\). Finally, \((\tau_w)_{w \in \mathfrak{M}} = (\sigma_w)_{w \in \mathfrak{M}}\) follows from the very definition of \(\tau_w\) (since \(\tau_w = \sigma_w\) for every \(w \in \mathfrak{M}\)). This completes the proof.
11.129. Solution to Exercise 6.67. Solution to Exercise 6.67. \( \implies \): Assume that \( w \) is Lyndon. We need to prove that for any two nonempty words \( u \in \mathfrak{A}^* \) and \( v \in \mathfrak{A}^* \) satisfying \( w = uv \), there exists at least one \( s \in u \sqcup v \) satisfying \( s > w \).

Let \( u \in \mathfrak{A}^* \) and \( v \in \mathfrak{A}^* \) be two nonempty words satisfying \( w = uv \). Then, \( uv \) is an element of \( u \sqcup v \) and satisfies \( vu \geq w \) (by Proposition 6.14(c)). Thus, \( vu > uv = w \). Hence, there exist at least one \( s \in u \sqcup v \) satisfying \( s > w \) (namely, \( s = vu \)). Thus, the \( \implies \) direction of Exercise 6.67 is solved.

\( \iff \): Assume that for any two nonempty words \( u \in \mathfrak{A}^* \) and \( v \in \mathfrak{A}^* \) satisfying \( w = uv \),

\[
\tag{11.1077}
\text{there exists at least one } s \in u \sqcup v \text{ satisfying } s > w.
\]

We need to prove that \( w \) is Lyndon.

In fact, assume the contrary. Thus, \( w \) is not Lyndon. Let \( (a_1, a_2, \ldots, a_p) \) be the CFL factorization of \( w \); then, \( a_1 \geq a_2 \geq \cdots \geq a_p \) and \( a_1 a_2 \cdots a_p = w \). Also, \( p \neq 0 \) (since \( a_1 a_2 \cdots a_p = w \) is nonempty). Since \( a_1 \) is Lyndon but \( w \) is not, we have \( a_1 \neq w \). Thus, \( p \neq 1 \) (because otherwise, we would have \( a_1 = a_1 a_2 \cdots a_p = w \), contradicting \( a_1 \neq w \)). Combined with \( p \neq 0 \), this yields \( p \geq 2 \).

Let \( u = a_1 \) and \( v = a_2 a_3 \cdots a_p \). Then, \( u \) is Lyndon (since \( u = a_1 \)), thus nonempty. Also, \( v \) is a nonempty product of Lyndon words (nonempty because \( p \geq 2 \)), and hence nonempty itself (since Lyndon words are nonempty). Clearly, \( \overbrace{u = a_1 a_2 a_3 \cdots a_p}^{a_1 = a_2 a_3 \cdots a_p} \implies (a_1 a_2 a_3 \cdots a_p) = a_1 a_2 \cdots a_p = w \). Thus, \( (11.1077) \) yields that there exists at least one \( s \in u \sqcup v \) satisfying \( s > w \). Thus, \( w \) is not the lexicographically highest element of the multiset \( u \sqcup v \).

Now, notice that \( a_2, a_3, \ldots, a_p \) are Lyndon words satisfying \( a_2 \geq a_3 \geq \cdots \geq a_p \) (since \( a_1 \geq a_2 \geq \cdots \geq a_p \)) and \( a_2 a_3 \cdots a_p = v \). Hence, \( (a_2, a_3, \ldots, a_p) \) is the CFL factorization of \( v \). We have \( u = a_1 \geq a_{j+1} \) for every \( j \in \{1, 2, \ldots, p-1\} \) (since \( a_1 \geq a_2 \geq \cdots \geq a_p \)). Thus, Theorem 6.44(e) (applied to \( p-1 \) and \( (a_2, a_3, \ldots, a_p) \) instead of \( q \) and \( (b_1, b_2, \ldots, b_m) \)) yields that the lexicographically highest element of the multiset \( u \sqcup v \) is \( w \), and the multiplicity with which this word \( wv \) appears in the multiset \( u \sqcup v \) is \( \max_n v + 1 \).

Now, we know that the lexicographically highest element of the multiset \( u \sqcup v \) is \( w = wv \). This contradicts the fact that \( w \) is not the lexicographically highest element of the multiset \( u \sqcup v \). This contradiction shows that our assumption was wrong. Thus, \( w \) is Lyndon. This completes the solution of the \( \iff \) direction of Exercise 6.67.

Thus, both the \( \implies \) and \( \iff \) directions of Exercise 6.67 are proven. Exercise 6.67 is thus solved.

[Remark: Exercise 6.67 still holds in the partial-order setting\(^{827}\). To prove this, we can use Proposition 11.31 along with the fact that Exercise 6.67 holds in the total-order setting. Here are the details:

\textbf{Solution to Exercise 6.67 in the partial-order setting.} \( \implies \): In the partial-order setting, the \( \implies \) direction of Exercise 6.67 can be proved in the same way as it was proven in the total-order setting.

\( \iff \): Assume that for any two nonempty words \( u \in \mathfrak{A}^* \) and \( v \in \mathfrak{A}^* \) satisfying \( w = uv \),

\[
\tag{11.1078}
\text{there exists at least one } s \in u \sqcup v \text{ satisfying } s > w.
\]

We need to prove that \( w \) is Lyndon.

Let \( \mathfrak{B} \) be any linear extension of the alphabet \( \mathfrak{A} \). Then, \( \mathfrak{B}^* \) is an extension of \( \mathfrak{A}^* \) (according to Proposition 11.31(a)). Thus, \( \mathfrak{B}^* = \mathfrak{A}^* \) as sets. It is easy to see that for any two nonempty words \( u \in \mathfrak{B}^* \) and \( v \in \mathfrak{B}^* \) satisfying \( w = uv \),

\[
\tag{11.1079}
\text{there exists at least one } s \in u \sqcup v \text{ satisfying } s > w \text{ in } \mathfrak{B}^*.
\]

\(^{826}\)\textit{Proof}: Let \( n = \ell(u) \) and \( m = \ell(v) \). Then, the permutation in \( S_{n+m} \) which is written as \((n+1, n+2, \ldots, n+m, 1, 2, \ldots, n)\) in one-line notation belongs to \( S_{n+m} \). If we denote this permutation by \( \sigma \), then \( uv = u \sqcup v \in u \sqcup v \), qed.

\(^{827}\)\textit{Proof of (11.1079)}: Let \( u \in \mathfrak{B}^* \) and \( v \in \mathfrak{B}^* \). Then, \( u \in \mathfrak{B}^* = \mathfrak{A}^* \) and \( v \in \mathfrak{B}^* = \mathfrak{A}^* \). Hence, \( (11.1078) \) shows that there exists at least one \( s \in u \sqcup v \) satisfying \( s > w \) in \( \mathfrak{A}^* \). Let \( t \) be such an \( s \). Then, \( t \) is an element of \( u \sqcup v \) satisfying \( t > w \) in \( \mathfrak{A}^* \). Hence, \( w < t \) in \( \mathfrak{A}^* \) (since \( t > w \) in \( \mathfrak{A}^* \)), so that \( w \leq t \) in \( \mathfrak{A}^* \).
Now, we can apply Exercise 6.67 to \( \mathcal{B} \) instead of \( \mathcal{A} \) (since \( \mathcal{B} \) is totally ordered). As a consequence, we see that \( w \) is Lyndon as a word in \( \mathcal{B}^* \) if and only if for any two nonempty words \( u \in \mathcal{B}^* \) and \( v \in \mathcal{B}^* \) satisfying \( w = uv \), there exists at least one \( s \in u \sqcup v \) satisfying \( s > w \) in \( \mathcal{B}^* \). Thus, \( w \) is Lyndon as a word in \( \mathcal{B}^* \) (because we know that for any two nonempty words \( u \in \mathcal{B}^* \) and \( v \in \mathcal{B}^* \) satisfying \( w = uv \), there exists at least one \( s \in u \sqcup v \) satisfying \( s > w \) in \( \mathcal{B}^* \)).

Now, the definition of a Lyndon word shows the following: The word \( w \) (as a word in \( \mathcal{B}^* \)) is Lyndon if and only if it is nonempty and satisfies the following property:

\[
\text{(11.1080)} \quad \text{Every nonempty proper suffix } v \text{ of } w \text{ satisfies } v > w \text{ in } \mathcal{B}^*.
\]

Since the word \( w \) is Lyndon, this shows that the word \( w \) is nonempty and satisfies the property (11.1080).

Let now \( v \) be a nonempty proper suffix of \( w \) (as a word in \( \mathcal{A}^* \)). Then, \( v \) is a nonempty proper suffix of \( w \) (as a word in \( \mathcal{B}^* \)). Thus, (11.1080) shows that \( v \) satisfies \( v > w \) in \( \mathcal{B}^* \).

Now, let us forget that we fixed \( \mathcal{B} \) and \( v \). We thus have shown that if \( v \) is any nonempty proper suffix of \( w \) and if \( \mathcal{B} \) is any linear extension of the alphabet \( \mathcal{A} \), then

\[
\text{(11.1081)} \quad v > w \text{ in } \mathcal{B}^*.
\]

But the definition of a Lyndon word shows the following: The word \( w \) (as a word in \( \mathcal{A}^* \)) is Lyndon if and only if it is nonempty and satisfies the following property:

\[
\text{(11.1082)} \quad \text{Every nonempty proper suffix } v \text{ of } w \text{ satisfies } v > w \text{ (in } \mathcal{A}^*\text{)}.
\]

We already know that \( w \) is nonempty. We are now going to prove (11.1082):

Let \( v \) be a nonempty proper suffix of \( w \). We have \( v > w \) in \( \mathcal{B}^* \) for every linear extension \( \mathcal{B} \) of \( \mathcal{A} \) (according to (11.1081)). In other words, \( w < v \) in \( \mathcal{B}^* \) for every linear extension \( \mathcal{B} \) of \( \mathcal{A} \).

But Proposition 11.31(b) (applied to \( w \) instead of \( u \)) yields that \( w < v \) holds in \( \mathcal{A}^* \) if and only if we have \( w < v \) in \( \mathcal{B}^* \) for every linear extension \( \mathcal{B} \) of \( \mathcal{A} \).

We thus conclude that \( w < v \) holds in \( \mathcal{A}^* \) (since we know that \( w < v \) in \( \mathcal{B}^* \) for every linear extension \( \mathcal{B} \) of \( \mathcal{A} \)). In other words, \( v > w \) in \( \mathcal{A}^* \). This proves (11.1082).

So we know that the word \( w \) is nonempty and satisfies the property (11.1082). Thus, the word \( w \) (as a word in \( \mathcal{A}^* \)) is Lyndon (since the word \( w \) (as a word in \( \mathcal{A}^* \)) is Lyndon and only if it is nonempty and satisfies the property (11.1082)). Thus, the \( \iff \) direction of Exercise 6.67 is proven in the partial-order setting.

Thus, both the \( \implies \) and \( \iff \) directions of Exercise 6.67 are proven in the partial-order setting. Hence, Exercise 6.67 is solved in the partial-order setting. \( \square \)

11.130. **Solution to Exercise 6.70.** Solution to Exercise 6.70.

**Proof of Remark 6.69.** Let \((w_1, w_2, \ldots, w_{n+m})\) denote the concatenation \(u \cdot v = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)\). Then, \((u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m) = (w_1, w_2, \ldots, w_{n+m})\), so that \((u_1, u_2, \ldots, u_n) = (w_1, w_2, \ldots, w_n)\) and \((v_1, v_2, \ldots, v_m) = (w_{n+1}, w_{n+2}, \ldots, w_{n+m})\).

By the definition of \( b_u \), we have \( b_u = b_{u_1} b_{u_2} \cdots b_{u_n} = b_{w_1} b_{w_2} \cdots b_{w_n} \) (since \((u_1, u_2, \ldots, u_n) = (w_1, w_2, \ldots, w_n)\)).

By the definition of \( b_v \), we have \( b_v = b_{v_1} b_{v_2} \cdots b_{v_m} = b_{w_{n+1}} b_{w_{n+2}} \cdots b_{w_{n+m}} \) (since \((v_1, v_2, \ldots, v_m) = (w_{n+1}, w_{n+2}, \ldots, w_{n+m})\)). Now,

\[
\begin{align*}
(b_u & \downarrow b_v) \downarrow (b_{w_1} b_{w_2} \cdots b_{w_n}) \downarrow (b_{w_{n+1}} b_{w_{n+2}} \cdots b_{w_{n+m}}) \\
& = \sum_{\sigma \in \mathcal{S}_{n,m}} b_{w_{\sigma(1)}} b_{w_{\sigma(2)}} \cdots b_{w_{\sigma(n+m)}}
\end{align*}
\]

We know that \( \mathcal{B}^* \) is an extension of \( \mathcal{A}^* \). Thus, any two elements \( a \) and \( b \) of \( \mathcal{A}^* \) satisfying \( a \leq b \) in \( \mathcal{A}^* \) satisfy \( a \leq b \) in \( \mathcal{B}^* \) (according to the definition of an “extension”). Applying this to \( a = w \) and \( b = t \), we obtain \( w \leq t \) in \( \mathcal{B}^* \). Combined with \( w \neq t \) (since \( w < t \) in \( \mathcal{A}^* \)), this yields \( w < t \) in \( \mathcal{B}^* \). In other words, \( t > w \) in \( \mathcal{B}^* \).

Hence, there exists at least one \( s \in u \sqcup v \) satisfying \( s > w \) in \( \mathcal{B}^* \) (namely, \( s = t \)). This proves (11.1079).
(by the definition of $\mu$).

But for every $\sigma \in Sh_{n,m}$, we have $u_{\mu} v = (w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)})$ (by the definition of $u_{\mu} v$), and therefore $b_{u_{\mu} v} = b_{w_{\sigma(1)}} b_{w_{\sigma(2)}} \cdots b_{w_{\sigma(n+m)}}$ (by the definition of $b_{u_{\mu} v}$). Hence, $\sum_{\sigma \in Sh_{n,m}} b_{u_{\mu} v} = \sum_{\sigma \in Sh_{n,m}} b_{w_{\sigma(1)}} b_{w_{\sigma(2)}} \cdots b_{w_{\sigma(n+m)}}$. Compared with (11.1083), this yields $b_{u_{\mu} v} = \sum_{\sigma \in Sh_{n,m}} b_{u_{\mu} v}$. This proves Remark 6.69.


Proof of Lemma 6.73. (a) Let $\mathfrak{M}$ denote the set of all finite multisets of Lyndon words over $\mathfrak{A}$. In other words, $\mathfrak{M}$ is the set of all finite multisets of elements of $\mathfrak{L}$. We can then define a map $\mathbf{m} : \mathfrak{M} \rightarrow \mathfrak{A}^*$ as follows: Given an $M \in \mathfrak{M}$, we set $\mathbf{m}(M) = a_1 a_2 \cdots a_k$, where $a_1, a_2, \ldots, a_k$ denote the elements of $M$ listed in decreasing order. Then, $\mathbf{m}(M)$ is an algebraically independent generating set of the $k$-algebra $A$ if and only if $(\mathbf{b}_M)_{M \in \mathfrak{M}}$ is a basis of the $k$-module $A$.

829 When we say “the elements of $M$ listed in decreasing order”, we mean a weakly decreasing listing of all elements of $M$ which has the property that every element of $M$ appears in the list exactly as often as it appears in $M$.

830 Proof. Let $\mathfrak{M}$ be the set of all families $(k_w)_{w \in \mathfrak{L}} \in \mathfrak{N}^2$ of nonnegative integers (indexed by the Lyndon words) such that all but finitely many $w \in \mathfrak{L}$ satisfy $k_w = 0$. We can now define a bijection $\text{mult} : \mathfrak{M} \rightarrow \mathfrak{M}$ by sending every multiset $M \in \mathfrak{M}$ to the family $(\{w \in \mathfrak{L} : w \in M\})_{w \in \mathfrak{L}} \in \mathfrak{M}$.

For every family $f = (k_w)_{w \in \mathfrak{L}} \in \mathfrak{M}$, we can define an element $b_f$ of $A$ by $b_f = \prod_{w \in \mathfrak{L}} (b_w)^{k_w}$. This is well-defined, since $A$ is commutative and since all but finitely many $w \in \mathfrak{L}$ satisfy $k_w = 0$. Then, $\{b_f\}_{f \in \mathfrak{M}} \in \mathfrak{M}$ is the family of all possible monomials in the “variables” $b_w$ (that is, of all possible finite products of elements of the family $(b_w)_{w \in \mathfrak{L}}$, with multiplicities allowed). Hence,

\[
(11.1084) \quad \{b_w\}_{w \in \mathfrak{L}} \text{ is an algebraically independent generating set of the } \mathfrak{k}\text{-algebra } A \text{ if and only if } \{b_f\}_{f \in \mathfrak{M}} \text{ is a basis of the } \mathfrak{k}\text{-module } A
\]

Now, we claim that the family $\{b_f\}_{f \in \mathfrak{M}}$ is a reindexing of the family $(\mathbf{b}_M)_{M \in \mathfrak{M}}$. Indeed, since mult is a bijection, it is clear that the family $\{b_f\}_{f \in \mathfrak{M}}$ is a reindexing of the family $(\mathbf{b}_{\text{mult} M})_{M \in \mathfrak{M}}$. We now will prove that every $M \in \mathfrak{M}$ satisfies $\mathbf{b}_{\text{mult} M} = \mathbf{b}_M$.

Indeed, let $M \in \mathfrak{M}$. Let $a_1, a_2, \ldots, a_k$ denote the elements of $M$ listed in decreasing order. Then, every $w \in \mathfrak{L}$ satisfies

\[
(11.1085) \quad \text{(multiplicity of } w \text{ in the multiset } M) = \text{(the number of } i \in \{1, 2, \ldots, k\} \text{ satisfying } a_i = w).
\]

But by the definition of mult $M$, we have

\[
\text{mult } M = \left( \begin{array}{c}
(\text{multiplicity of } w \text{ in the multiset } M) \\
= \text{(the number of } i \in \{1, 2, \ldots, k\} \text{ satisfying } a_i = w) \\
\text{by (11.1085)} \\
\end{array} \right)_{w \in \mathfrak{L}}
\]

so that the definition of $\mathbf{b}_{\text{mult} M}$ becomes

\[
\mathbf{b}_{\text{mult} M} = \prod_{w \in \mathfrak{L}} (b_w)^{\text{(the number of } i \in \{1, 2, \ldots, k\} \text{ satisfying } a_i = w)} = \prod_{w \in \mathfrak{L}} \prod_{i = 1}^k b_{a_i} = \prod_{w \in \mathfrak{L}} b_{a_i} \quad \text{(since } w = a_i) = \prod_{i \in \{1, 2, \ldots, k\}} b_{a_i} = \mathbf{b}_M.
\]

(since $\mathbf{b}_M$ was defined as $b_{a_1} b_{a_2} \cdots b_{a_k}$).

Forget now that we fixed $M$. We thus have shown that every $M \in \mathfrak{M}$ satisfies $\mathbf{b}_{\text{mult} M} = \mathbf{b}_M$. Hence, $(\mathbf{b}_{\text{mult} M})_{M \in \mathfrak{M}} = (\mathbf{b}_M)_{M \in \mathfrak{M}}$. But we know that the family $\{b_f\}_{f \in \mathfrak{M}}$ is a reindexing of the family $(\mathbf{b}_{\text{mult} M})_{M \in \mathfrak{M}}$. Since $(\mathbf{b}_{\text{mult} M})_{M \in \mathfrak{M}} = (\mathbf{b}_M)_{M \in \mathfrak{M}}$, this rewrites as follows: The family $\{b_f\}_{f \in \mathfrak{M}}$ is a reindexing of the family $(\mathbf{b}_M)_{M \in \mathfrak{M}}$. Hence, $\{b_f\}_{f \in \mathfrak{M}}$ is a basis
Hence, in order to prove Lemma 6.73(a), it remains to prove that \((b_M)_{M \in \mathfrak{M}}\) is a basis of the \(k\)-module \(A\) if and only if \((b_u)_{u \in \mathfrak{A}^*}\) is a basis of the \(k\)-module \(A\).

Let us show that

\[
(b_{m(M)}) = b_M \quad \text{for every } M \in \mathfrak{M}.
\]

Proof of (11.1086): Let \(M \in \mathfrak{M}\). Let \(a_1, a_2, \ldots, a_k\) denote the elements of \(M\) listed in decreasing order. Then, \(m(M) = a_1 a_2 \cdots a_k\) (by the definition of \(m(M)\)). Combining this with the fact that \(a_1, a_2, \ldots, a_k\) are Lyndon words (since they are elements of \(M\), which is a multiset of Lyndon words) and satisfy \(a_1 \geq a_2 \geq \cdots \geq a_k\) (since \(a_1, a_2, \ldots, a_k\) are the elements of \(M\) listed in decreasing order), this yields that \((a_1, a_2, \ldots, a_k)\) is the CFL factorization of \(m(M)\). Hence, the definition of \(b_{m(M)}\) says that \(b_{m(M)} = b_{a_1} b_{a_2} \cdots b_{a_k}\). Compared with \(b_M = b_{a_1} b_{a_2} \cdots b_{a_k}\) (which is just the definition of \(b_M\)), this yields \(b_{m(M)} = b_M\). This proves (11.1086).

Now, the family \((b_{m(M)})_{M \in \mathfrak{M}}\) is a reindexing of the family \((b_u)_{u \in \mathfrak{A}^*}\) (since \(m\) is a bijection). Since \((b_{m(M)})_{M \in \mathfrak{M}} = (b_M)_{M \in \mathfrak{M}}\) (by (11.1086)), this rewrites as follows: The family \((b_M)_{M \in \mathfrak{M}}\) is a reindexing of the family \((b_u)_{u \in \mathfrak{A}^*}\). Hence, \((b_M)_{M \in \mathfrak{M}}\) is a basis of the \(k\)-module \(A\) if and only if \((b_u)_{u \in \mathfrak{A}^*}\) is a basis of the \(k\)-module \(A\). As we said above, this completes our proof of Lemma 6.73(a).

(b) The proof of Lemma 6.73(b) is analogous to the above proof of Lemma 6.73(a).

(c) We assumed that the family \((b_u)_{u \in \mathfrak{A}^*}\) generates the \(k\)-algebra \(A\). By Lemma 6.73(b), this yields that the family \((b_u)_{u \in \mathfrak{A}^*}\) spans the \(k\)-module \(A\). Recall also that the family \((g_u)_{u \in \mathfrak{A}^*}\) is a basis of the \(k\)-module \(A\), and thus spans this \(k\)-module.

We need to prove that the family \((b_u)_{u \in \mathfrak{A}^*}\) is an algebraically independent generating set of the \(k\)-algebra \(A\). According to Lemma 6.73(a), this is equivalent to proving that the family \((b_u)_{u \in \mathfrak{A}^*}\) is a basis of the \(k\)-module \(A\). We are going to prove the latter statement.

Let us first notice that \(b_u\) is a homogeneous element of \(A\) of degree \(Wt(u)\) for every \(u \in \mathfrak{A}^*\).

Now, let \(n \in \mathbb{N}\). It is easy to see that the family \((b_u)_{u \in Wt^{-1}(n)}\) spans the \(k\)-module \(A_n\) (that is, the \(n\)-th homogeneous component of the \(k\)-module \(A\))

The definition of \(Wt\) easily yields that

\[
Wt(s_1 s_2 \cdots s_k) = Wt(s_1) + Wt(s_2) + \cdots + Wt(s_k)
\]

for any \(k \in \mathbb{N}\) and any \(k\) words \(s_1, s_2, \ldots, s_k\) in \(\mathfrak{A}^*\).

Applying this to \(k = p\) and \(s_i = a_i\), we obtain \(Wt(a_1 a_2 \cdots a_p) = Wt(a_1) + Wt(a_2) + \cdots + Wt(a_p)\), so that

\[
Wt(a_1) + Wt(a_2) + \cdots + Wt(a_p) = Wt(u).
\]

But the definition of \(b_u\) yields \(b_u = b_{a_1} b_{a_2} \cdots b_{a_p}\). Thus, the element \(b_u\) is homogeneous of degree \(Wt(a_1) + Wt(a_2) + \cdots + Wt(a_p)\) (since for every \(u \in \mathfrak{A}^\ast\), the element \(b_u\) of \(A\) is homogeneous of degree \(Wt(u)\)). Since \(Wt(a_1) + Wt(a_2) + \cdots + Wt(a_p) = Wt(u)\), this rewrites as follows: The element \(b_u\) is homogeneous of degree \(Wt(u)\), qed.

Proof. For every \(u \in \mathfrak{A}^*\), the element \(b_u\) is a homogeneous element of \(A\) of degree \(Wt(u) = n\) (since \(u \in Wt^{-1}(n)\)).

In other words, for every \(u \in Wt^{-1}(n)\), we have \(b_u \in A_n\).

Let \(\xi \in A_n\). Then, \(\xi \in A_n \subset A\), so that \(\xi\) is a \(k\)-linear combination of the elements \(b_u\) for \(u \in \mathfrak{A}^*\) (since the family \((b_u)_{u \in \mathfrak{A}^*}\) spans the \(k\)-module \(A\)). In other words, \(\xi = \sum_{u \in \mathfrak{A}^*} \lambda_u b_u\) for some family \((\lambda_u)_{u \in \mathfrak{A}^*} \in k^{\mathfrak{A}^*}\) of elements of \(k\) such that all but finitely many \(u \in \mathfrak{A}^*\) satisfy \(\lambda_u = 0\). Consider this \((\lambda_u)_{u \in \mathfrak{A}^*}\).

What happens if we apply the canonical projection \(A \rightarrow A_n\) (which projects \(A\) onto its \(n\)-th homogeneous component \(A_n\), annihilating all other components) to both sides of the equality \(\xi = \sum_{u \in \mathfrak{A}^*} \lambda_u b_u\)? The left hand side remains \(\xi\) (since \(\xi \in A_n\)).

On the right hand side, all addends of the sum in which the \(u\) satisfies \(Wt(u) = n\) stay fixed (because we know that for each such \(u\), the element \(b_u\) is a homogeneous element of \(A\) of degree \(Wt(u) = n\)), whereas all other addends become 0 (for a similar reason); therefore, the sum on the right hand side becomes \(\sum_{u \in Wt^{-1}(n)} \lambda_u b_u\). Thus, the equality becomes \(\xi = \sum_{Wt(u) = n} \lambda_u b_u\).

Hence, \(\xi\) is a \(k\)-linear combination of the elements \(b_u\) with \(u \in Wt^{-1}(n)\).
independent (because it is a subfamily of the basis \( (g_u)_{u \in A} \) of \( A \)), this shows that the family \( (g_u)_{u \in W^{-1}(n)} \) is a basis of the \( k \)-module \( A_n \).

But the set \( W^{-1}(n) \) is finite\(^ {833} \). Hence, \( A_n \) is a finite free \( k \)-module (since \( (g_u)_{u \in W^{-1}(n)} \) is a basis of the \( k \)-module \( A_n \)), and Exercise 2.46(b) (applied to \( A_n, W^{-1}(n), u, g_u \) and \( b_u \) instead of \( A, I, i, \gamma_i \) and \( \beta_i \)) yields that \( (b_u)_{u \in W^{-1}(n)} \) is a \( k \)-basis of \( A_n \).

Now, let us forget that we fixed \( n \). We thus have shown that for every \( n \in \mathbb{N} \), the family \( (b_u)_{u \in W^{-1}(n)} \) is a \( k \)-basis of \( A_n \). Hence, the family \( (b_u)_{u \in \mathbb{N}} \) (being the disjoint union of the families \( (b_u)_{u \in W^{-1}(n)} \) over all \( n \in \mathbb{N} \)) is a \( k \)-basis of \( \bigoplus_{n \in \mathbb{N}} A_n = A \). In other words, the family \( (b_u)_{u \in \mathbb{N}} \) is a basis of the \( k \)-module \( A \).

As we know, this completes the proof of Lemma 6.73(c). \( \square \)


Proof of Lemma 6.76. (a) Let \( u, v \) and \( v' \) be three words satisfying \( \ell(u) = n, \ell(v) = m, \ell(v') = m \) and \( v' < v \). We must show that \( u \sqcup v' < u \sqcup v \).

We have \( v' < v \), thus \( v' \leq v \). By the definition of the relation \( \leq \), this means that

- either there exists an \( i \in \{1, 2, \ldots, \min\{\ell(v'), \ell(v)\}\} \) such that \( (v')_i < v_i \), and every \( j \in \{1, 2, \ldots, i - 1\} \) satisfies \( (v')_j = v_j \),
- or the word \( v' \) is a prefix of \( v \).

Since the word \( v' \) cannot be a prefix of \( v \) (because this would entail that \( v' = v \) (because \( \ell(v') = m = \ell(v) \), so that \( v' \) has the same length as \( v \)), which would contradict \( v' < v \), this shows that there exists an \( i \in \{1, 2, \ldots, \min\{\ell(v'), \ell(v)\}\} \) such that \( (v')_i < v_i \), and every \( j \in \{1, 2, \ldots, i - 1\} \) satisfies \( (v')_j = v_j \).

Denote this \( i \) by \( k \). Thus, \( k \in \{1, 2, \ldots, \min\{\ell(v'), \ell(v)\}\} \) has the property that \( (v')_k < v_k \), and

\[
\text{every } j \in \{1, 2, \ldots, k - 1\} \text{ satisfies } (v')_j = v_j.
\]

We have \( k \in \{1, 2, \ldots, \min\{\ell(v'), \ell(v)\}\} = \{1, 2, \ldots, m\} \).

---

\(^{833}\) Proof. We have

\[
W^{-1}(n) = \{ w \in \mathbb{N}^* \mid Wt(w) = n \}
\]

\[
= \bigcup_{k \in \mathbb{N}} \left\{ (w_1, w_2, \ldots, w_k) \in \mathbb{N}^* \mid \frac{Wt\left(\left(\begin{array}{c} w_1, w_2, \ldots, w_k \end{array}\right)\right)}{\text{wt}(w_1) + \text{wt}(w_2) + \cdots + \text{wt}(w_k)} = n \right\}
\]

(since every word \( w \in \mathbb{N}^* \) has the form \( (w_1, w_2, \ldots, w_k) \) for some \( k \in \mathbb{N} \))

\[
= \bigcup_{k \in \mathbb{N}} \left\{ (w_1, w_2, \ldots, w_k) \in \mathbb{N}^* \mid wt(w_1) + \text{wt}(w_2) + \cdots + \text{wt}(w_k) = n \right\}
\]

\[
= \bigcup_{\{1, 2, \ldots, k\} \subseteq \{2, 3, \ldots, k\}} \left\{ (w_1, w_2, \ldots, w_k) \in \mathbb{N}^* \mid wt(w_1) = 1, \text{wt}(w_2) = 2, \ldots, \text{wt}(w_k) = k \right\}
\]

\[
= \bigcup_{(i_1, i_2, \ldots, i_k) \in \mathbb{N}^*} \left\{ (w_1, w_2, \ldots, w_k) \in \mathbb{N}^* \mid wt(w_1) = i_1, \text{wt}(w_2) = i_2, \ldots, \text{wt}(w_k) = i_k \right\}
\]

\[
= \bigcup_{(i_1, i_2, \ldots, i_k) \in \mathbb{N}^*} \left\{ (w_1, w_2, \ldots, w_k) \in \mathbb{N}^* \mid wt^{-1}(i_1) \times wt^{-1}(i_2) \times \cdots \times wt^{-1}(i_k) \right\}
\]

(a finite set

(since for every \( N \in \{2, 3, \ldots\} \), the set \( wt^{-1}(N) \) is finite)

Hence, \( W^{-1}(n) \) is a union of finitely many finite sets, and thus itself finite, qed.

---
Now, let \((w_1, w_2, \ldots, w_{n+m})\) denote the concatenation \(u \cdot v = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)\), and let \((w'_1, w'_2, \ldots, w'_{n+m})\) denote the concatenation \(u \cdot v' = (u_1, u_2, \ldots, u_n, (v')_1, (v')_2, \ldots, (v')_m)\). We have \(u \upsilon v = (w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)})\) (by the definition of \(u \upsilon v\)) and \(u \upsilon v' = (w'_{\sigma(1)}, w'_{\sigma(2)}, \ldots, w'_{\sigma(n+m)})\) (by the definition of \(u \upsilon v'\)). Hence, we have \(u \upsilon v' \neq u \upsilon v\). But our goal is to prove that \(u \upsilon v' < u \upsilon v\). Hence, it remains to prove that \(u \upsilon v' \leq u \upsilon v\) (since we have shown that \(u \upsilon v' \neq u \upsilon v\)). Due to the definition of the relation \(\leq\), this amounts to proving that

\[
\begin{align*}
\text{either} & \quad \text{there exists an } i \in \{1, 2, \ldots, \min \{\ell(u \upsilon v'), \ell(u \upsilon v)\}\} \text{ such that } \\
& \quad \left( (u \upsilon v')_i < (u \upsilon v)_i \right), \text{ and every } j \in \{1, 2, \ldots, i-1\} \text{ satisfies } (u \upsilon v')_j = (u \upsilon v)_j, \\
& \text{or the word } u \upsilon v' \text{ is a prefix of } u \upsilon v.
\end{align*}
\]

We shall prove that the first of these two alternatives holds, i.e., that there exists an \(i \in \{1, 2, \ldots, \min \{\ell(u \upsilon v'), \ell(u \upsilon v)\}\}\) such that

\[
\left( (u \upsilon v')_i < (u \upsilon v)_i \right), \text{ and every } j \in \{1, 2, \ldots, i-1\} \text{ satisfies } (u \upsilon v')_j = (u \upsilon v)_j.
\]

Indeed, we claim that \(\sigma^{-1}(n+k)\) is such an \(i\). In order to conclude the proof, we then need to show that

\[
\sigma^{-1}(n+k) \in \{1, 2, \ldots, \min \{\ell(u \upsilon v'), \ell(u \upsilon v)\}\},
\]

and that

\[
every \ j \in \{1, 2, \ldots, \sigma^{-1}(n+k) - 1\} \text{ satisfies } (u \upsilon v')_j = (u \upsilon v)_j.
\]

**Proof of (11.1088):** We have \(\ell(u \upsilon v) = \ell(u) + \ell(v) = n + m\) and similarly \(\ell(u \upsilon v') = n + m\). Thus,

\[
\min \left\{ \ell(u \upsilon v'), \ell(u \upsilon v) \right\} = \min(n + m, n + m) = n + m.
\]

But since \(\sigma \in S_{n+m} \subset G_{n+m}\), we have \(\sigma^{-1}(n+k) \in \{1, 2, \ldots, n+m\}\). In other words, \(\sigma^{-1}(n+k) \in \{1, 2, \ldots, \min \{\ell(u \upsilon v'), \ell(u \upsilon v)\}\}\) (since \(n + m + m = n + m\)). This proves (11.1088).

**Proof of (11.1089):** Since \(u \upsilon v = (w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)})\), we have \((u \upsilon v)_j = w_{\sigma(j)}\) for every \(j \in \{1, 2, \ldots, n+m\}\). Applying this to \(j = \sigma^{-1}(n+k)\), we obtain \((u \upsilon v)_{\sigma^{-1}(n+k)} = w_{\sigma(\sigma^{-1}(n+k))} = w_{n+k} = v_k\) (since \(k \in \{1, 2, \ldots, m\}\) and \((w_1, w_2, \ldots, w_{n+m}) = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)\)). Similarly, \((u \upsilon v')_{\sigma^{-1}(n+k)} = (v')_k\). Recalling that \((v')_k < v_k\), we now see that \((u \upsilon v')_{\sigma^{-1}(n+k)} = (v')_k < v_k\). This proves (11.1089).

**Proof of (11.1090):** Let \(j \in \{1, 2, \ldots, \sigma^{-1}(n+k) - 1\}\). We need to prove that \((u \upsilon v')_j = (u \upsilon v)_j\).

Assume the contrary. Then, \((u \upsilon v')_j \neq (u \upsilon v)_j\).

\[\text{$834$ Proof. Assume the contrary. Then, } u \upsilon v' = u \upsilon v, \text{ thus } (w'_{\sigma(1)}, w'_{\sigma(2)}, \ldots, w'_{\sigma(n+m)}) = u \upsilon v' = u \upsilon v = (w_{\sigma(1)}, w_{\sigma(2)}, \ldots, w_{\sigma(n+m)}). \text{ Hence, } (w'_1, w'_2, \ldots, w'_{n+m}) = (w_1, w_2, \ldots, w_{n+m}) \text{ (since } \sigma \text{ is a permutation), and thus } u \cdot v' = (w'_1, w'_2, \ldots, w'_{n+m}) = (w_1, w_2, \ldots, w_{n+m}) = u \cdot v \text{, so that } v' = v. \text{ This contradicts } v' < v. \text{ This contradiction shows that our assumption was wrong, qed.}\]
We have $j \in \{1, 2, \ldots, n + m\}$ and $j < \sigma^{-1}(n + k)$ (since $j \in \{1, 2, \ldots, \sigma^{-1}(n + k) - 1\}$).

Since $u \uplus v = (w_\sigma(1), w_\sigma(2), \ldots, w_\sigma(n + m))$, we have $(u \uplus v)_j = w_\sigma(j)$. Similarly, $(u \uplus v')_j = w'_\sigma(j)$.

Hence, $w'_\sigma(j) = (u \uplus v')_j \neq (u \uplus v)_j = w_\sigma(j)$.

If we had $\sigma(j) \leq n$, then we would have $w'_\sigma(j) = w_\sigma(j)$ (since $(w_1, w_2, \ldots, w_{n+m}) = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)$) and $w'_\sigma(j) = u_\sigma(j)$ (for similar reasons), which would yield that $w'_\sigma(j) = w_\sigma(j)$, which would contradict $w'_\sigma(j) \neq w_\sigma(j)$. Hence, we cannot have $\sigma(j) \leq n$. We thus must have $\sigma(j) > n$, whence $\sigma(j) \in \{n + 1, n + 2, \ldots, n + m\}$. Consequently, $w'_\sigma(j) = v_\sigma(j) - n$ (because $(w_1, w_2, \ldots, w_{n+m}) = (u_1, u_2, \ldots, u_n, v_1, v_2, \ldots, v_m)$ and $(w'_\sigma)_{\sigma(j)} = (v_\sigma)_{\sigma(j)} - n$ (similarly).

If we had $\sigma(j) - n \in \{1, 2, \ldots, k - 1\}$, then we would have $(v'_\sigma(j) - n = v_\sigma(j) - n$ (by (11.1087), applied to $\sigma(j) - n$ instead of $j$). Hence, if we had $\sigma(j) - n \in \{1, 2, \ldots, k - 1\}$, then we would have $(w'_\sigma)_{\sigma(j)} = (v'_\sigma)_{\sigma(j)} - n = v_\sigma(j) - n = w_\sigma(j)$, which would contradict $w'_\sigma(j) \neq w_\sigma(j)$. Hence, we cannot have $\sigma(j) - n \in \{1, 2, \ldots, k - 1\}$.

In other words, we have $\sigma(j) - n \notin \{1, 2, \ldots, k - 1\}$, so that $\sigma(j) - n \geq k$. Hence, $\sigma(j) \geq n + k$.

We have $\sigma \in Sh_{n,m}$, so that $\sigma^{-1}(1) < \sigma^{-1}(2) < \cdots < \sigma^{-1}(n)$ and $\sigma^{-1}(n + 1) < \sigma^{-1}(n + 2) < \cdots < \sigma^{-1}(n + m)$. In particular, the restriction of the map $\sigma^{-1}$ to the set $\{n + 1, n + 2, \ldots, n + m\}$ is strictly increasing (since $\sigma^{-1}(n + 1) < \sigma^{-1}(n + 2) < \cdots < \sigma^{-1}(n + m)$). Since $\sigma(j)$ and $n + k$ both lie in this set $\{n + 1, n + 2, \ldots, n + m\}$, we thus have $\sigma^{-1}(\sigma(j)) \geq \sigma^{-1}(n + k)$ (because $\sigma(j) \geq n + k$). Hence, $j = \sigma^{-1}(\sigma(j)) \geq \sigma^{-1}(n + k)$, which contradicts $j < \sigma^{-1}(n + k)$. This contradiction shows that our assumption was wrong. Hence, (11.1090) is proven.

Now that (11.1088), (11.1089) and (11.1090) are all proven, we conclude that there exists an $i \in \{1, 2, \ldots, \min\{\ell(u \uplus v'), \ell(u \uplus v)\}\}$ such that

$$(u \uplus v')_i < (u \uplus v)_i \quad \text{and every } j \in \{1, 2, \ldots, i - 1\} \quad \text{satisfies} \quad (u \uplus v')_j = (u \uplus v)_j \quad \text{(namely, } i = \sigma^{-1}(n + k)).$$

This concludes the proof of Lemma 6.76(a).

(b) The proof of Lemma 6.76(b) is similar to the proof of Lemma 6.76(a) above. (One of the changes necessary is to replace $\sigma^{-1}(n + k)$ by $\sigma^{-1}(k)$.) (Alternatively, it is not hard to derive Lemma 6.76(b) from Lemma 6.76(a), because if $\tau \in S_{n+m}$ denotes the permutation which is written $(m+1, m+2, \ldots, n+m, 1, 2, \ldots, m)$ in one-line notation, then the permutation $\tau \circ \sigma$ belongs to $Sh_{m,n}$ and satisfies $u \uplus v = v \uplus u$ and $u \uplus v' = v' \uplus u$, and therefore we can obtain Lemma 6.76(b) by applying Lemma 6.76(a) to $m, n, v, u, u'$ and $\tau \circ \sigma$ instead of $n, m, u, v, v'$ and $\sigma$.)

(c) Let $u, v$ and $v'$ be three words satisfying $\ell(u) = n, \ell(v) = m, \ell(v') = m$ and $v' \leq v$. We must show that $u \uplus v' \leq u \uplus v$. This is obvious if $v' = v$ (in fact, if $v' = v$, then $u \uplus v' = u \uplus v \leq u \uplus v$). Hence, we can WLOG assume that $v' \neq v$. Assuming this, we immediately obtain $v' < v$ (since $v' \neq v$ and $v' \leq v$), so that $u \uplus v' < u \uplus v$ (by Lemma 6.76(a)). Thus, of course, $u \uplus v' \leq u \uplus v$, and Lemma 6.76(c) is proven.

11.133. Solution to Exercise 6.78. Solution to Exercise 6.78.

Proof of Proposition 6.75. We shall prove Proposition 6.75 by strong induction over $\ell$. So we fix some $L \in \mathbb{N}$, and we assume that Proposition 6.75 holds whenever $\ell < L$. We now need to prove that Proposition 6.75 holds for $\ell = L$. In other words, we need to prove that for every word $x \in \mathcal{A}^L$, there is a family $(\eta_x,y)_{y \in \mathcal{A}^L} \in \mathbb{N}^{\mathcal{A}^L}$ of elements of $\mathbb{N}$ satisfying

$$(11.1091) \quad b_x = \sum_{y \in \mathcal{A}^L; y \leq x} \eta_x,y b_y$$

and $\eta_x,x \neq 0$ (in $\mathbb{N}$).

Let $x \in \mathcal{A}^L$ be a word. We need to prove that there is a family $(\eta_x,y)_{y \in \mathcal{A}^L} \in \mathbb{N}^{\mathcal{A}^L}$ of elements of $\mathbb{N}$ satisfying (11.1091) and $\eta_x,x \neq 0$. 

Let \((a_1, a_2, \ldots, a_p)\) be the CFL factorization of \(x\). Then, \(a_1, a_2, \ldots, a_p\) are Lyndon words satisfying \(x = a_1 a_2 \cdots a_p\) and \(a_1 \geq a_2 \geq \cdots \geq a_p\).

If \(x\) is the empty word, then our claim is trivial (in fact, we can just set \(\eta_{x,x} = 1\) in this case, and (11.1091) holds obviously). Hence, for the rest of this proof, we WLOG assume that \(x\) is not the empty word. Thus, \(p \neq 0\) (because otherwise, \(x = a_1 a_2 \cdots a_p\) would be an empty product and thus the empty word, contradicting the assumption that \(x\) is not the empty word). Hence, we can define two words \(u \in \mathfrak{A}^*\) and \(v \in \mathfrak{A}^*\) by \(u = a_1\) and \(v = a_2 a_3 \cdots a_p\). The word \(u\) is Lyndon (since \(u = a_1\) and since \(a_1\) is Lyndon), and thus has \((u)\) as its CFL factorization. The CFL factorization of the word \(v\) is \((a_2, a_3, \ldots, a_p)\) (since the words \(a_2, a_3, \ldots, a_p\) are Lyndon and satisfy \(v = a_2 a_3 \cdots a_p\) and \(a_2 \geq a_3 \geq \cdots \geq a_p\) (because \(a_1 \geq a_2 \geq \cdots \geq a_p\)).

Also, \(u = a_1 \geq a_{j+1}\) for every \(i \in \{1, 2, \ldots, 1\}\) and \(j \in \{1, 2, \ldots, p - 1\}\) (because \(a_1 \geq a_2 \geq \cdots \geq a_p\)). Hence, we can apply Theorem 6.44(c) to \((u)\) and \((a_2, a_3, \ldots, a_p)\) instead of \((p, (a_1, a_2, \ldots, a_p), q\) and \((b_1, b_2, \ldots, b_q)\). As a result, we conclude that the lexicographically highest element of the multiset \(u \cup v\) is

\[
\begin{align*}
\text{but } b_u &= b_u \text{ (by the definition of } b_u, \text{ since } u \text{ has CFL factorization } (u), \text{ and } b_x &= b_u \sqcup b_v. \\
\end{align*}
\]

The word \(u\) is Lyndon and thus nonempty, so that \(\ell(u) > 0\). Now, \(\ell(\frac{x}{\Rightarrow u}) = \ell(uv) = \ell(u) + \ell(v) > \ell(v)\), so that \(\ell(v) < \ell(x) = L\) (since \(x \in \mathfrak{A}^L\)). Hence, the induction hypothesis tells us that we can apply Proposition 6.75 to \(\ell(v)\) and \(v\) instead of \(\ell\) and \(x\) (since \(v \in \mathfrak{A}^{\ell(v)}\)). As a result, we see that there is a family \((\eta_{v,v})_{v \in \mathfrak{A}^{\ell(v)}} \in \mathbb{N}^{\mathfrak{A}^{\ell(v)}}\) of elements of \(\mathbb{N}\) satisfying

\[
\eta_{v,v} = \sum_{y \in \mathfrak{A}^{\ell(v)}; y \leq v} \eta_{v,y} b_y
\]
and \(\eta_{v,v} \neq 0\) (in \(\mathbb{N}\)). Consider this family \((\eta_{v,y})_{y \in \mathfrak{A}^{\ell(v)}}\).

Now,

\[
\begin{align*}
b_x &= b_u \sqcup b_v = b_u \sqcup \left( \sum_{y \in \mathfrak{A}^{\ell(v)}; y \leq v} \eta_{v,y} b_y \right) = \sum_{y \in \mathfrak{A}^{\ell(v)}; y \leq v} \eta_{v,y} b_y \sqcup b_v \\
&= \sum_{y \in \mathfrak{A}^{\ell(v)}; y \leq v} \eta_{v,z} b_y \sqcup b_z \quad \text{(by Remark 6.69, applied to } y, \ell(u) \text{ and } \ell(v) \text{ instead of } v, n \text{ and } m) \\
\end{align*}
\]

(11.1092)

(here, we renamed the summation index \(y\) as \(z\)).

Now, let \(z \in \mathfrak{A}^{\ell(v)}\) and \(\sigma \in \mathfrak{S}_{\ell(u), \ell(v)}\) be such that \(z \leq v\). We will prove that \(u \sqcup z \in \mathfrak{A}^L\) and \(u \sqcup z \leq x\).

First of all, it is clear that \(\ell(\frac{u \sqcup z}{\sigma}) = \ell(u) + \ell(z) = \ell(u) + \ell(v) = \ell(\frac{uv}{\sigma}) = \ell(x) = L\), so that \(\sigma u \sqcup z \in \mathfrak{A}^L\). Applying Lemma 6.76(c) to \(\ell(u), \ell(v)\) and \(z\) instead of \(n, m\) and \(v'\), we obtain \(u \sqcup z \leq u \sqcup v\).

\[\text{Proof.}\] Since the CFL factorization of \(v\) is \((a_2, a_3, \ldots, a_p)\), we have \(b_v = b_{a_2} \sqcup b_{a_3} \sqcup \cdots \sqcup b_{a_p}\) (by the definition of \(b_v\)). But since the CFL factorization of \(x\) is \((a_1, a_2, \ldots, a_p)\), we have

\[
\begin{align*}
b_x &= b_{a_1} \sqcup b_{a_2} \sqcup \cdots \sqcup b_{a_p} \quad \text{(by the definition of } b_x) \\
&= b_{a_1} \sqcup (b_{a_2} \sqcup b_{a_3} \sqcup \cdots \sqcup b_{a_p}) = b_{a_1} \sqcup b_v = b_u \sqcup b_v, \\
\end{align*}
\]

qed.
But \( u \uplus v \) is an element of the multiset \( u \uplus v \), and thus is \( \leq x \) (since the lexicographically highest element of the multiset \( u \uplus v \) is \( x \)). Thus, \( u \uplus v \leq x \), so that \( u \uplus z \leq u \uplus v \leq x \).

Now, let us forget that we fixed \( z \) and \( \sigma \). We thus have shown that

\[(11.1093) \quad \text{any } z \in \mathfrak{A}^{(v)} \text{ and } \sigma \in \text{Sh}_{(u),\ell(v)} \text{ satisfying } z \leq v \text{ satisfy } u \uplus z \in \mathfrak{A}^L \text{ and } u \uplus z \leq x.\]

A similar argument (but with some of the \( \leq \) signs replaced by < signs, and with a reference to Lemma 6.76(a) instead of a reference to Lemma 6.76(c)) shows that

\[(11.1094) \quad \text{any } z \in \mathfrak{A}^{(v)} \text{ and } \sigma \in \text{Sh}_{(u),\ell(v)} \text{ satisfying } z < v \text{ satisfy } u \uplus z \in \mathfrak{A}^L \text{ and } u \uplus z < x.\]

Now, (11.1092) becomes

\[
b_v = \sum_{z \in \mathfrak{A}^{(v)}; \ z \leq v} \eta_{v,z} \left( \sum_{\sigma \in \text{Sh}_{(u),\ell(v)}; \ u \uplus z \leq v} b_{u \uplus z} \right) = \sum_{z \in \mathfrak{A}^{(v)}; \ z \leq v} \eta_{v,z} \left( \sum_{y \in \mathfrak{A}^L; \ \sigma \in \text{Sh}_{(u),\ell(v)}; \ u \uplus z \leq y} b_{u \uplus z} \right) = \sum_{y \in \mathfrak{A}^L; \ y \leq x} \eta_{v,z} \left( \sum_{z \in \mathfrak{A}^{(v)}; \ u \uplus z = y} b_{u \uplus z} \right) \quad \text{(due to (11.1093))}.
\]

\[
= \sum_{y \in \mathfrak{A}^L; \ y \leq x} \sum_{z \in \mathfrak{A}^{(v)}; \ z \leq v} \eta_{v,z} \left| \left\{ \sigma \in \text{Sh}_{(u),\ell(v)} \mid u \uplus z = y \right\} \right| b_y.
\]

(11.1095)

Recall that we must prove that there is a family \((\eta_{x,y})_{y \in \mathfrak{A}^L} \in \mathbb{N}^{\mathfrak{A}^L}\) of elements of \( \mathbb{N} \) satisfying (11.1091) and \( \eta_{x,x} \neq 0 \) (in \( \mathbb{N} \)). In order to prove this, we define such a family \((\eta_{x,y})_{y \in \mathfrak{A}^L} \in \mathbb{N}^{\mathfrak{A}^L}\) by setting

\[
\eta_{x,y} = \sum_{z \in \mathfrak{A}^{(v)}; \ z \leq v} \eta_{v,z} \left| \left\{ \sigma \in \text{Sh}_{(u),\ell(v)} \mid u \uplus z = y \right\} \right| \quad \text{for every } y \in \mathfrak{A}^L.
\]

Then, (11.1095) shows that this family satisfies (11.1091). All that remains to be proven is now to show that \( \eta_{x,x} \neq 0 \) (in \( \mathbb{N} \)).
Indeed, the definition of $\eta_{x,z}$ yields
\[
\eta_{x,z} = \sum_{z \in \mathfrak{M}(\tau), \ z \leq v} \eta_{v,z} \left\{ \sigma \in \text{Sh}_{k(u), \ell(v)} \mid u \uplus \sigma z = x \right\} + \eta_{v,v} \left\{ \sigma \in \text{Sh}_{k(u), \ell(v)} \mid u \uplus \sigma v = x \right\}
\]
(since every $\sigma \in \text{Sh}_{k(u), \ell(v)}$ satisfies $u \uplus \sigma z = x$ (by (11.1094)), so that it cannot satisfy $u \uplus \sigma z = x$)

(here, we have split off the addend for $z = v$ from the sum)
\[
= \sum_{z \in \mathfrak{M}(\tau), \ z \leq v} \eta_{v,z} \left\{ \sigma \in \text{Sh}_{k(u), \ell(v)} \mid u \uplus \sigma v = x \right\} + \eta_{v,v} \left\{ \sigma \in \text{Sh}_{k(u), \ell(v)} \mid u \uplus \sigma v = x \right\}
\]
\[
= \sum_{z \in \mathfrak{M}(\tau), \ z \leq v} \eta_{v,z} 0 + \eta_{v,v} \left\{ \sigma \in \text{Sh}_{k(u), \ell(v)} \mid u \uplus \sigma v = x \right\}
\]
\[
= \eta_{v,v} \left\{ \sigma \in \text{Sh}_{k(u), \ell(v)} \mid u \uplus \sigma v = x \right\}
\]
\[
\neq 0 \text{ (in } \mathbb{N})
\]
(since there exists some $\sigma \in \text{Sh}_{k(u), \ell(v)}$ satisfying $u \uplus \sigma v = x$ (because $x = u \uplus \sigma v$))

\[
neq 0 \text{ (in } \mathbb{N}).
\]

This completes the proof that there is a family $(\eta_{x,y})_{y \in \mathfrak{M}} \in \mathbb{N}^{\mathfrak{M}}$ of elements of $\mathbb{N}$ satisfying (11.1091) and $\eta_{x,x} \neq 0$ (in $\mathbb{N}$). The induction step is thus complete, and Proposition 6.75 is proven by induction. \square


Proof of Theorem 6.71. It is clearly enough to prove that $(b_u)_{u \in \mathfrak{M}}$ is an algebraically independent generating set of the $k$-algebra Sh$(V)$.

For every word $u \in \mathfrak{M}^*$, define an element $b_u$ by $b_u = b_{u_1} \uplus b_{u_2} \uplus \cdots \uplus b_{u_p}$, where $(a_1, a_2, \ldots, a_p)$ is the CFL factorization of $u$. According to Lemma 6.73(a) (applied to $A = \text{Sh}(V)$) 836, the family $(b_u)_{u \in \mathfrak{M}}$ is an algebraically independent generating set of the $k$-algebra Sh$(V)$ if and only if the family $(b_u)_{u \in \mathfrak{M}^*}$ is a basis of the $k$-module Sh$(V)$. In order to prove the former statement (which is our goal), it is therefore enough to prove the latter statement.

So we must prove the family $(b_u)_{u \in \mathfrak{M}^*}$ is a basis of the $k$-module Sh$(V)$. We shall first prove a particular case of this statement:

Assertion A: If the set $\mathfrak{M}$ is finite, then the family $(b_u)_{u \in \mathfrak{M}}$ is a basis of the $k$-module Sh$(V)$.

Proof of Assertion A: Assume that the set $\mathfrak{M}$ is finite. Fix $\ell \in \mathbb{N}$. Regard $V^\otimes \ell$ as a $k$-submodule of $T(V)$. Then, $(b_{u_1} \otimes b_{u_2} \otimes \cdots \otimes b_{u_p})_{u \in \mathfrak{M}^*}$ is a basis of the $k$-module $V^\otimes \ell$ (since $(b_u)_{u \in \mathfrak{M}}$ is a basis of the $k$-module $V$). In other words,

(11.1096) $(b_u)_{u \in \mathfrak{M}^*}$ is a basis of the $k$-module $V^\otimes \ell$

(since $b_u = b_{u_1} b_{u_2} \cdots b_{u_p} = b_{u_1} \otimes b_{u_2} \otimes \cdots \otimes b_{u_p}$ for every $u \in \mathfrak{M}^*$). According to Proposition 6.75, the family $(b_u)_{u \in \mathfrak{M}^*}$ expands upper-triangularly (with respect to the total order $\leq$ on $\mathfrak{M}^*$) with respect to the basis

836 Don’t be confused by the fact that the multiplication of the $k$-algebra $\text{Sh}(V)$ is denoted by $\uplus$, but the multiplication of the $k$-algebra $A$ is denoted by $\cdot$ in Lemma 6.73(a).
$(b_u)_{u \in \mathcal{A}^\ell}$. The diagonal entries of this upper-triangular expansion are nonzero elements of $\mathbb{N}$ (indeed, these diagonal entries are precisely the elements $\eta_{x,x}$ defined in Proposition 6.75, and Proposition 6.75 shows that $\eta_{x,x} \neq 0$ in $\mathbb{N}$, and therefore are invertible elements of $k$ (since $Q$ is a subring of $k$). So the family $(b_u)_{u \in \mathcal{A}^\ell}$ expands upper-triangularly with respect to the basis $(b_u)_{u \in \mathcal{A}^\ell}$ of $V^{\mathcal{A}^\ell}$, and the diagonal entries of this upper-triangular expansion are invertible elements of $k$. As a consequence, this expansion is invertible \footnote{This is because the indexing set $\mathcal{A}^\ell$ is finite (since $\mathcal{A}$ is finite), and because any finite upper-triangular matrix over $k$ whose diagonal entries are invertible must be invertible.}, and thus the family $(b_u)_{u \in \mathcal{A}^\ell}$ is a basis of $V^{\mathcal{A}^\ell}$ as well.

Now, let us forget that we fixed $\ell$. We thus have shown that, for every $\ell \in \mathbb{N}$, the family $(b_u)_{u \in \mathcal{A}^\ell}$ is a basis of the $k$-module $V^{\mathcal{A}^\ell}$. Hence, the disjoint union of the families $(b_u)_{u \in \mathcal{A}^\ell}$ over all $\ell \in \mathbb{N}$ is a basis of the direct sum $\bigoplus_{\ell \in \mathbb{N}} V^{\mathcal{A}^\ell}$. Since the former disjoint union is the family $(b_u)_{u \in \mathcal{A}^\ast}$, while the latter direct sum is the $k$-module $\bigoplus_{\ell \in \mathbb{N}} V^{\mathcal{A}^\ell} = T(V) = Sh(V)$, this rewrites as follows: The family $(b_u)_{u \in \mathcal{A}^\ast}$ is a basis of the $k$-module $Sh(V)$. This proves Assertion A.

Now, we need to prove that $(b_u)_{u \in \mathcal{A}^\ast}$ is a basis of the $k$-module $Sh(V)$, without the assumption that $\mathcal{A}$ be finite. We will reduce this to Assertion A. First, we need some preparations:

Let $\mathcal{B}$ be any subset of $\mathcal{A}$. Then, $\mathcal{B}$ canonically becomes a totally ordered set (since $\mathcal{A}$ is totally ordered), so that the notion of a Lyndon word over $\mathcal{B}$ is well-defined. We view $\mathcal{B}^\ast$ as a subset of $\mathcal{A}^\ast$, and so the Lyndon words over $\mathcal{B}$ are precisely the Lyndon words over $\mathcal{A}$ which lie in $\mathcal{B}^\ast$.

We define a $k$-submodule $V_{\mathcal{B}}$ of $V$ as the $k$-linear span of the family $(b_u)_{u \in \mathcal{B}}$. Notice that the family $(b_u)_{u \in \mathcal{B}}$ is $k$-linearly independent (being a subfamily of the basis $(b_u)_{u \in \mathcal{A}}$ of $V$), and thus is a basis of the $k$-submodule $V_{\mathcal{B}}$.

The inclusion $V_{\mathcal{B}} \rightarrow V$ gives rise to an injective $k$-algebra homomorphism $Sh(V_{\mathcal{B}}) \rightarrow Sh(V)$, which sends every $b_u$ and every $b_v$ (for $u \in \mathcal{B}^\ast$ and $v \in \mathcal{B}^\ast$, respectively) to the corresponding elements $b_u$ and $b_v$ of $Sh(V)$, respectively. We regard this homomorphism as an inclusion, so that $Sh(V_{\mathcal{B}})$ is a $k$-subalgebra of $Sh(V)$.

If $\mathcal{B}$ is finite, then Assertion A (applied to $\mathcal{B}$ instead of $\mathcal{A}$) yields that the family $(b_u)_{u \in \mathcal{B}}$ is a basis of the $k$-module $Sh(V_{\mathcal{B}})$.

Now, let us forget that we fixed $\mathcal{B}$. We thus have shown that for every finite subset $\mathcal{B}$ of $\mathcal{A}$,

\begin{equation}
(11.1097) \quad \text{the family } (b_u)_{u \in \mathcal{B}} \text{ is a basis of the } k\text{-module } Sh(V_{\mathcal{B}}).
\end{equation}

Let us introduce one more notation: A family of elements of a $k$-module is said to be \textit{finitely supported} if all but finitely many elements of this family are 0.

Now, let us show that the family $(b_u)_{u \in \mathcal{A}^\ast}$ is $k$-linearly independent and spans the $k$-module $Sh(V)$.

**Proof that the family $(b_u)_{u \in \mathcal{A}^\ast}$ is $k$-linearly independent:** Let $(\lambda_u)_{u \in \mathcal{A}^\ast} \in k^{\mathcal{A}^\ast}$ be a finitely supported family of elements of $k$ satisfying $\sum_{u \in \mathcal{A}^\ast} \lambda_u b_u = 0$. We are going to prove that all $u \in \mathcal{A}^\ast$ satisfy $\lambda_u = 0$.

Indeed, the family $(\lambda_u)_{u \in \mathcal{A}^\ast}$ is finitely supported, so that there exists a finite subset $Z$ of $\mathcal{A}^\ast$ such that all $u \in \mathcal{A}^\ast \setminus Z$ satisfy $\lambda_u = 0$. Consider this $Z$. Since $Z$ is finite, there exists a finite subset $\mathcal{B}$ of $\mathcal{A}$ satisfying $Z \subseteq \mathcal{B}^\ast$ (in fact, we can take $\mathcal{B}$ to be the set of all letters occurring in the words lying in $Z$). Consider this $\mathcal{B}$. The family $(b_u)_{u \in \mathcal{B}}$, is a basis of the $k$-module $Sh(V_{\mathcal{B}})$ (by \footnote{This is because the indexing set $\mathcal{A}^\ell$ is finite (since $\mathcal{A}$ is finite), and because any finite upper-triangular matrix over $k$ whose diagonal entries are invertible must be invertible.}), and thus $k$-linearly independent. We have $\mathcal{A}^\ast \setminus \mathcal{B}^\ast \subseteq \mathcal{A}^\ast \setminus Z$ (since $Z \subseteq \mathcal{B}^\ast$), and therefore all $u \in \mathcal{A}^\ast \setminus \mathcal{B}^\ast$ satisfy $\lambda_u = 0$ (since all $u \in \mathcal{A}^\ast \setminus Z$ satisfy $\lambda_u = 0$). Hence, $\sum_{u \in \mathcal{A}^\ast \setminus \mathcal{B}^\ast} \lambda_u b_u = \sum_{u \in \mathcal{A}^\ast \setminus \mathcal{B}^\ast} 0 b_u = 0$. Now,

\[
0 = \sum_{u \in \mathcal{A}^\ast} \lambda_u b_u = \sum_{u \in \mathcal{B}^\ast} \lambda_u b_u + \sum_{u \in \mathcal{A}^\ast \setminus \mathcal{B}^\ast} \lambda_u b_u = \sum_{u \in \mathcal{A}^\ast \setminus \mathcal{B}^\ast} 0 b_u = 0.
\]

So we have $\sum_{u \in \mathcal{B}^\ast} \lambda_u b_u = 0$. Thus, all $u \in \mathcal{B}^\ast$ satisfy $\lambda_u = 0$ (since the family $(b_u)_{u \in \mathcal{B}^\ast}$ is $k$-linearly independent). Combining this with the fact that all $u \in \mathcal{A}^\ast \setminus \mathcal{B}^\ast$ satisfy $\lambda_u = 0$, we conclude that all $u \in \mathcal{A}^\ast$ satisfy $\lambda_u = 0$.\]
Now forget that we fixed \((\lambda_u)_{u \in \mathfrak{A}^*}\). We thus have shown that if \((\lambda_u)_{u \in \mathfrak{A}^*} \in k^{\mathfrak{A}^*}\) is a finitely supported family of elements of \(k\) satisfying \(\sum_{u \in \mathfrak{A}^*} \lambda_u b_u = 0\), then all \(u \in \mathfrak{A}^*\) satisfy \(\lambda_u = 0\). In other words, the family \((b_u)_{u \in \mathfrak{A}^*}\) is \(k\)-linearly independent.

Proof that the family \((b_u)_{u \in \mathfrak{A}^*}\) spans the \(k\)-module \(S_h(V)\): We are going to show that the family \((b_u)_{u \in \mathfrak{A}^*}\) spans the \(k\)-module \(S_h(V)\). In order to prove this, it is enough to show that \(b_w\) lies in the \(k\)-linear span of the family \((b_u)_{u \in \mathfrak{A}^*}\) for every \(w \in \mathfrak{A}^*\) (because the family \((b_w)_{w \in \mathfrak{A}^*}\) is a basis of the \(k\)-module \(S_h(V)\)).

So let us show this now.

Let \(w \in \mathfrak{A}^*\). Then, there exists a finite subset \(F\) of \(\mathfrak{A}\) such that \(w \in F^*\) (namely, we can take \(F\) to be the set of all letters of \(w\)). Consider this \(F\). The family \((b_u)_{u \in F}\), is a basis of the \(k\)-module \(S_h(V_F)\) (by \((11.1097)\)), and thus spans this \(k\)-module. But \(b_w \in S_h(V_F)\) (since \(w \in F^*\)), and thus \(b_w\) lies in the \(k\)-linear span of the family \((b_u)_{u \in F}\), since this family is a basis of the \(k\)-module \(S_h(V_F)\)). Hence, \(b_w\) lies in the \(k\)-linear span of the family \((b_u)_{u \in \mathfrak{A}^*}\) as well (since this family \((b_u)_{u \in \mathfrak{A}^*}\) includes the family \((b_u)_{u \in F}\), as a subfamily). Thus, we have proven that \(b_w\) lies in the \(k\)-linear span of the family \((b_u)_{u \in \mathfrak{A}^*}\) for every \(w \in \mathfrak{A}^*\). This completes the proof that the family \((b_u)_{u \in \mathfrak{A}^*}\) spans the \(k\)-module \(S_h(V)\).

Altogether, we now know that the family \((b_u)_{u \in \mathfrak{A}^*}\) is \(k\)-linearly independent and spans the \(k\)-module \(S_h(V)\). In other words, \((b_u)_{u \in \mathfrak{A}^*}\) is a basis of the \(k\)-module \(S_h(V)\). This completes our proof of Theorem 6.7.

11.135. Solution to Exercise 6.81. Solution to Exercise 6.81. We shall give two solutions to this exercise; but they both rest on the following lemmas:

Lemma 11.42. For every positive integer \(N\), we have

\[(11.1098) \sum_{d \mid N} \mu(d) \delta_{N,1} = \delta_{N,1}.\]

Lemma 11.42 is one of the most fundamental properties of the number-theoretic Möbius function; it will not be proven here.\(^{339}\)

Lemma 11.43. Every positive integer \(n\) satisfies

\[\frac{1}{n} \sum_{d \mid n} \mu(d) \left(2^{n/d} - 1\right) = \frac{1}{n} \sum_{d \mid n} \mu(d) 2^{n/d} - \delta_{n,1}.\]

Proof of Lemma 11.43. Let \(n\) be a positive integer. Then,

\[
\begin{align*}
\frac{1}{n} \sum_{d \mid n} \mu(d) \left(2^{n/d} - 1\right) &= \sum_{d \mid n} \mu(d) 2^{n/d} - \sum_{d \mid n} \mu(d) 1 \\
&= \frac{1}{n} \left( \sum_{d \mid n} \mu(d) 2^{n/d} - \sum_{d \mid n} \mu(d) \right) - \frac{1}{n} \left( \sum_{d \mid n} \mu(d) 2^{n/d} - \sum_{d \mid n} \right) \\
&= \frac{1}{n} \left( \sum_{d \mid n} \mu(d) 2^{n/d} - \frac{1}{n} \delta_{n,1} \right) = \frac{1}{n} \sum_{d \mid n} \mu(d) 2^{n/d} - \delta_{n,1}.
\end{align*}
\]

This proves Lemma 11.43. \(\square\)

\(^{338}\)Proof. For every \(\ell \in \mathbb{N}\), the family \((b_u)_{u \in \mathfrak{A}^*}\) of the \(k\)-module \(V^{\otimes \ell}\) (by \((11.1096)\)). Hence, the disjoint union of the families \((b_u)_{u \in \mathfrak{A}^*}\) over all \(\ell \in \mathbb{N}\) is a basis of the direct sum \(\bigoplus_{\ell \in \mathbb{N}} V^{\otimes \ell}\). Since the former disjoint union is the family \((b_u)_{u \in \mathfrak{A}^*}\), whereas the latter direct sum is the \(k\)-module \(\bigoplus_{\ell \in \mathbb{N}} V^{\otimes \ell} = T(V) = S_h(V)\), this rewrites as follows:

\[\text{The family } (b_u)_{u \in \mathfrak{A}^*} \text{ is a basis of the } k\text{-module } S_h(V), \text{ qed.}\]

\(^{339}\)For a proof of Lemma 11.42, see the solution of Exercise 2.82. (More precisely, Lemma 11.42 is obtained from \((11.337)\) by renaming \(n\) as \(N\).)
First solution to Exercise 6.81. [The following solution is similar to the solution of Exercise 6.29.]

For every positive integer \( n \) denote the number of Lyndon compositions of size \( n \). We need to prove that

\[
(11.1099) \quad \text{lync} \, n = \frac{1}{n} \sum_{d|n} \mu(d) \left( 2^{n/d} - 1 \right) = \frac{1}{n} \sum_{d|n} \mu(d) 2^{n/d} - \delta_{n,1}
\]

for every positive integer \( n \).

We recall that \( \mathfrak{A}^* = \text{Comp} \); that is, the elements of \( \mathfrak{A}^* \) are the compositions. Hence, the notation \(|w|\) makes sense for any \( w \in \mathfrak{A}^* \); it denotes the size of the composition \( w \).

For every \( n \in \mathbb{N} \), we have

\[
|\text{Comp}_n| = \begin{cases} 2^{n-1}, & \text{if } n \geq 1; \\ 1, & \text{if } n = 0 \end{cases}
\]

Thus, for every \( n \in \mathbb{N} \), we have

\[
\left| \left\{ w \in \mathfrak{A}^* \mid |w| = n \right\} \right| = \left| \left\{ w \in \text{Comp} \mid |w| = n \right\} \right| = |\text{Comp}_n|
\]

\[(11.1100) \quad = \begin{cases} 2^{n-1}, & \text{if } n \geq 1; \\ 1, & \text{if } n = 0 \end{cases} \]

Let \( \mathfrak{M} \) denote the set of all finite multisets of Lyndon compositions. We can then define a map \( \mathbf{m} : \mathfrak{M} \to \mathfrak{A}^* \) as follows: Given an \( M \in \mathfrak{M} \), we set \( \mathbf{m}(M) = a_1 a_2 \cdots a_k \), where \( a_1, a_2, \ldots, a_k \) denote the elements of \( M \) listed in decreasing order and each as often as it appears in \( M \). This map \( \mathbf{m} \) is bijective (its inverse is given by sending every word \( w \) to the multiset \( \{a_1, a_2, \ldots, a_k\}_{\text{multiset}} \), where \( (a_1, a_2, \ldots, a_k) \) is the CFL factorization of \( w \).

On the other hand, let \( \mathfrak{L} \) be the set of all Lyndon compositions. Thus, \( \mathfrak{M} \) is the set of all finite multisets of elements of \( \mathfrak{L} \). Also, the definition of \( \text{lync} \, n \) now rewrites as

\[
(11.1101) \quad \text{lync} \, n = |\{ w \in \mathfrak{L} \mid |w| = n \}| \quad \text{for every positive integer } n.
\]

Let \( \mathfrak{N} \) be the set of all finite families \( (k_w)_{w \in \mathfrak{L}} \in \mathbb{N}^\mathfrak{L} \) of nonnegative integers (indexed by the Lyndon compositions) such that all but finitely many \( w \in \mathfrak{L} \) satisfy \( k_w = 0 \). We can now define a bijection \( \text{mult} : \mathfrak{M} \to \mathfrak{N} \) by sending every \( M \in \mathfrak{M} \) to the family \((\text{multiplicity of } w \text{ in the multiset } M)_{w \in \mathfrak{L}} \in \mathfrak{N} \).

The composition \( \mathbf{m} \circ \text{mult}^{-1} : \mathfrak{N} \to \mathfrak{A}^* \) of the bijections \( \mathbf{m} \) and \( \text{mult}^{-1} \) is clearly a bijection. It can easily be seen to satisfy

\[
(11.1102) \quad |(\mathbf{m} \circ \text{mult}^{-1})( (k_w)_{w \in \mathfrak{L}} ) | = \sum_{w \in \mathfrak{L}} k_w \cdot |w| \quad \text{for every } (k_w)_{w \in \mathfrak{L}} \in \mathfrak{N}.
\]

\[\text{Proof.} \quad \text{Let } n \in \mathbb{N}. \text{ Recall that we have a bijection } \text{Comp}_n \to 2^{[n-1]}, \text{ where } [n - 1] = \{1, 2, \ldots, n - 1\}. \text{ Thus,}
\]

\[
|\text{Comp}_n| = 2^{[n-1]} = 2^{[n-1]} = \begin{cases} 2^{n-1}, & \text{if } n \geq 1; \\ 2^n, & \text{if } n = 0 \end{cases} \quad \text{since } |[n - 1]| = \begin{cases} n - 1, & \text{if } n \geq 1; \\ 0, & \text{if } n = 0 \end{cases}
\]

\[\text{qed.} \]
Now, in the ring \( \mathbb{Q}[[t]] \) of formal power series, we have

\[
\sum_{w \in \mathcal{A}^*} t^{\|w\|} = \sum_{n \in \mathbb{N}} \left\{ \begin{array}{ll}
2^{n-1}, & \text{if } n \geq 1; \\
1, & \text{if } n = 0
\end{array} \right. t^n = \sum_{n \in \mathbb{N}} \left\{ \begin{array}{ll}
2^{n-1}, & \text{if } n \geq 1; \\
1, & \text{if } n = 0
\end{array} \right. t^n
\]

(by (11.1100))

\[
= \frac{1}{1 - 2t} + \sum_{n \in \mathbb{N}; \ n \geq 1} 2^{n-1} t^n
\]

\[
= \frac{1}{1 - 2t} + \sum_{n \in \mathbb{N}; \ n \geq 1} 2^{n-1} t^n = 1 + \sum_{n \in \mathbb{N}} 2^n t^{n+1} = 1 + \sum_{n \in \mathbb{N}} (2t)^n t = 1 + \frac{1}{1 - 2t} = \frac{1 - t}{1 - 2t}.
\]

Proof of (11.1102): Let \( (k_w)_{w \in \mathcal{L}} \in \mathcal{L} \). Let \( M = \text{mult}^{-1} \left( (k_w)_{w \in \mathcal{L}} \right) \). Then, \( M \) is a multiset of elements of \( \mathcal{L} \) and satisfies \( (k_w)_{w \in \mathcal{L}} = \text{mult} M = \left( (\text{multiplicity of } w \text{ in the multiset } M) \right)_{w \in \mathcal{L}} \). In other words, every \( w \in \mathcal{L} \) satisfies

(11.1103)

\( k_w = \text{(multiplicity of } w \text{ in the multiset } M) \).

Let \( a_1, a_2, \ldots, a_k \) denote the elements of this multiset \( M \) listed in decreasing order. Then, the definition of \( \mathbf{m} \) yields \( \mathbf{m} (M) = a_1 a_2 \cdots a_k \), so that

\[
|\mathbf{m} (M)| = a_1 a_2 \cdots a_k = \sum_{i \in \{1,2,\ldots,k\}} |a_i| = \sum_{w \in \mathcal{L}} \sum_{i \in \{1,2,\ldots,k\}; \ a_i = w} |a_i| = \sum_{w \in \mathcal{L}} \sum_{i \in \{1,2,\ldots,k\}; \ a_i = w} \left( \begin{array}{c}
\text{number of } i \in \{1,2,\ldots,k\} \text{ satisfying } a_i = w \cdot |w| \\
\text{(multiplicity of } w \text{ in the multiset } M = k_w \text{) (by (11.1103))}
\end{array} \right)
\]

\[
= \sum_{w \in \mathcal{L}} k_w \cdot |w|.
\]

Now,

\[
|\left( \mathbf{m} \circ \text{mult}^{-1} \right) (k_w)_{w \in \mathcal{L}}| = \left| \mathbf{m} \left( \text{mult}^{-1} \left( (k_w)_{w \in \mathcal{L}} \right) \right) \right| = |\mathbf{m} (M)| = \sum_{w \in \mathcal{L}} k_w \cdot |w|,
\]

which proves (11.1102).
Hence,

\[
\frac{1 - t}{1 - 2t} = \sum_{w \in \mathfrak{A}} t^{|w|} = \sum_{(k_w)_{w \in \mathfrak{A}}} t^{|(m \circ \text{mult}^{-1})(k_w)_{w \in \mathfrak{A}}|} \\
\text{(here, we substituted } (m \circ \text{mult}^{-1})(k_w)_{w \in \mathfrak{A}} \text{ for } w \text{ in the sum, )}
\]

since the map \( m \circ \text{mult}^{-1} : \mathfrak{A} \to \mathfrak{A}^* \) is a bijection

\[
= \sum_{(k_w)_{w \in \mathfrak{A}}} t^{\sum_{w \in \mathfrak{A}} k_w |w|} \quad \text{(by (11.1102))}
\]

\[
= \prod_{w \in \mathfrak{A}} \prod_{n \geq 1 \atop |w| = n} \frac{1}{1 - t^n} = \prod_{n \geq 1} \prod_{w \in \mathfrak{A} \atop |w| = n} \frac{1}{1 - t^n} \quad \text{(by the product rule)}
\]

\[
= \prod_{n \geq 1} \prod_{w \in \mathfrak{A} \atop |w| = n} \frac{1}{1 - t^n} \quad \text{(since } |w| \geq 1 \text{ for every } w \in \mathfrak{A})
\]

Taking the logarithm of both sides of this identity, we obtain

\[
\log \frac{1 - t}{1 - 2t} = \log \left( \prod_{n \geq 1} \left( \frac{1}{1 - t^n} \right)^{\text{lync } n} \right) = \sum_{n \geq 1} (\text{lync } n) \cdot \log \left( \frac{1}{1 - t^n} \right) = -\log(1 - t^n) = \sum_{u \geq 1} \frac{1}{u} (t^n)^u
\]

(by the Mercator series for the logarithm)

\[
= \sum_{n \geq 1} (\text{lync } n) \cdot \sum_{u \geq 1} \frac{1}{u} (t^n)^u = \sum_{n \geq 1} \sum_{u \geq 1} (\text{lync } n) \frac{1}{u} (t^n)^u = \sum_{n \geq 1} \sum_{u \geq 1} (\text{lync } n) \frac{1}{u} t^{nu}
\]

(here, we substituted \( v/n \) for \( u \) in the second sum)

\[
= \sum_{n \geq 1} \sum_{v \geq 1 \atop n|v} (\text{lync } n) \frac{1}{v/n} t^v = \sum_{n \geq 1} \sum_{d|n} (\text{lync } d) \frac{1}{n} d^n t^n
\]

(here, we substituted \( v/n \) for \( u \) in the second sum)
DARIJ GRINBERG AND VICTOR REINER

(here, we renamed the summation indices \(v\) and \(n\) as \(n\) and \(d\)). Since

\[
\log \frac{1 - t}{1 - 2t} = \log (1 - t) - \log (1 - 2t) = -\frac{\log (1 - 2t)}{t} - \frac{\log (1 - t)}{2t}
\]

(by the Mercator series for the logarithm)

\[
= \sum_{n \geq 1} \frac{1}{n} (2t)^n = \sum_{n \geq 1} \frac{1}{n} (2t)^n - \sum_{n \geq 1} \frac{1}{n} (2t)^n = (2^n - 1) t^n
\]

this rewrites as

\[
\sum_{n \geq 1} \frac{1}{n} (2^n - 1) t^n = \sum_{n \geq 1} \sum_{d | n} (\text{lync } d) \frac{d}{n} t^n.
\]

Comparing coefficients, we conclude that every positive integer \(n\) satisfies

\[
\frac{1}{n} (2^n - 1) = \sum_{d | n} (\text{lync } d) \frac{d}{n}.
\]

Multiplying this with \(n\), we obtain

\[
(11.1104) \quad 2^n - 1 = \sum_{d | n} (\text{lync } d) d.
\]

Now, every positive integer \(n\) satisfies

\[
\sum_{d | n} \mu (d) \left( \frac{2^n}{d} - 1 \right) = \sum_{e | n} \mu (e) \left( \frac{2^n}{e} - 1 \right) = \sum_{e | n} \mu (e) \sum_{d | n/e} (\text{lync } d) d
\]

(by \(11.1094\), applied to \(n/e\) instead of \(n\))

\[
= \sum_{e | n} \sum_{d | n/e} \mu (e) (\text{lync } d) d = \sum_{d | n} \sum_{e | n/d} \mu (e) (\text{lync } d) d
\]

(by \(11.1098\), applied to \(N = n/d\))

\[
= \sum_{d | n} \delta_{n/d} (\text{lync } d) d = \sum_{d | n} \delta_{n/d} (\text{lync } d) d = (\text{lync } n) n.
\]

Dividing this by \(n\), we obtain

\[
\frac{1}{n} \sum_{d | n} \mu (d) \left( \frac{2^n}{d} - 1 \right) = \text{lync } n. \quad \text{Hence,}
\]

\[
\text{lync } n = \frac{1}{n} \sum_{d | n} \mu (d) \left( \frac{2^n}{d} - 1 \right) = \frac{1}{n} \sum_{d | n} \mu (d) 2^n - \delta_{n,1}
\]

(by Lemma 11.43). This proves \(11.1099\). Thus, Exercise 6.81 is solved.

Second solution to Exercise 6.81. Let \(\mathfrak{B}\) denote the two-element set \(\{0, 1\}\), where \(0\) and \(1\) are two new objects. We make \(\mathfrak{B}\) into a totally ordered set by setting \(0 < 1\). In the following, we will study not only words over the alphabet \(\mathfrak{A} = \{1, 2, 3, \ldots\}\), but also words over the alphabet \(\mathfrak{B}\). The latter words form the set \(\mathfrak{B}^*\). Let \(\mathfrak{L}_\mathfrak{B}\) denote the set of all Lyndon words over the alphabet \(\mathfrak{B}\).
For every $k \in \mathfrak{A}$, define an element $\tilde{k}$ of $\mathfrak{B}^*$ by $\tilde{k} = \left(0, 1, 1, \ldots, 1\right)$ \begin{equation}
abla \text{(This is well-defined, since } \mathfrak{A} = \{1, 2, 3, \ldots\}\text{.) We can rewrite the definition of } \tilde{k}\text{ as follows: We have} \end{equation}
\begin{equation}
(11.1105) \quad \tilde{k} = 01^{k-1} \quad \text{for every positive integer } k \quad \text{(where "0" and "1" are regarded as one-letter words). Thus,} \end{equation}
\begin{equation}
(11.1106) \quad \ell(\tilde{k}) = k \quad \text{for every positive integer } k \quad \text{\cite{844}}. \end{equation}
\begin{equation}
(11.1107) \quad \tilde{k} + 1 = \tilde{k}1 \quad \text{for every positive integer } k \quad \text{(where "\tilde{k}1" means the concatenation of } \tilde{k} \text{ with the one-letter word 1).} \end{equation}
\begin{equation}
(11.1108) \quad \ell(\tilde{k} + 1) = \ell(\tilde{k}1) = \ell(0) + \ell(1) = 1 + (k - 1) = k \quad \text{for every positive integer } k \quad \text{\cite{844}. In other words,} \end{equation}
\begin{equation}
\tilde{1} < \tilde{2} < \tilde{3} < \cdots \quad \text{in the lexicographic order on } \mathfrak{B}^*. \end{equation}
We notice a slightly stronger property: For any $a \in \mathfrak{A}$ and $b \in \mathfrak{A}$ satisfying $a < b$, we have
\begin{equation}
(11.1109) \quad \tilde{a}(c_1c_2 \cdots c_s) < \tilde{bw} \end{equation}
for any $s \in \mathbb{N}$, any $c_1, c_2, \ldots, c_s \in \mathfrak{A}$ and any $w \in \mathfrak{B}^*$. \cite{846}
Also,

\[ (11.1109) \quad \tilde{k} \text{ is a Lyndon word over the alphabet } \mathcal{B} \text{ for every positive integer } k. \]

847

We can now define a map \( \Phi : \mathfrak{A}^* \to \mathfrak{B}^* \) by setting

\[
(\Phi((w_1, w_2, \ldots, w_k)) = \tilde{w}_1 \tilde{w}_2 \cdots \tilde{w}_k \quad \text{for every word } (w_1, w_2, \ldots, w_k) \in \mathfrak{A}^*).
\]

In other words, \( \bar{0} \bar{1}^{k-a} \leq \bar{0}t \). Thus, Proposition 6.2(c) (applied to \( \mathcal{B}, \bar{a}, 1^{b-a} \) and \( 0t \) instead of \( \mathfrak{A}, a, c \) and \( d \)) yields \( 1^{b-a} \leq 0t \).

Now, \( 1^{(b-a)-1} = 1^{b-a} \leq 0t \). Therefore, Proposition 6.2(e) (applied to \( \mathcal{B}, 1, 1^{(b-a)-1} \) and \( 0t \) instead of \( \mathfrak{A}, a, b, c \) and \( d \)) yields that either we have \( 1 \leq 0 \) or the word \( 0t \) is a prefix of \( 1t \). Since the word \( 0t \) is not a prefix of \( 1t \), this shows that \( 1t \). But this contradicts \( 0t < 1t \). This contradiction proves that our assumption was wrong, qed.

847 Proof. Let \( k \) be a positive integer. From (11.1105), we obtain \( \tilde{k} = 01^{k-1} \).

From (11.1106), we have \( t(\tilde{k}) = k \geq 1 \). Thus, the word \( \tilde{k} \) is nonempty.

Let \( v \) be a nonempty proper suffix of \( \tilde{k} \). We will show that \( v > \tilde{k} \).

Indeed, there exists a nonempty \( u \in \mathfrak{B}^* \) satisfying \( \tilde{k} = uv \) (since \( v \) is a proper suffix of \( \tilde{k} \)). Consider this \( u \). Since \( u \) is nonempty, the first letter of \( u \) is well-defined. We have

\[ (\text{the first letter of } u) = \begin{cases} 
(\text{the first letter of } \tilde{k}) & \text{since } u \text{ is a prefix of } \tilde{k} \text{ (since } \tilde{k} = uv) \\
(\text{the first letter of } 01^{k-1}) & = 0.
\end{cases} \]

Thus, \( 0 \) is a prefix of \( u \). In other words, there exists a word \( u' \in \mathfrak{B}^* \) satisfying \( u = 0u' \). Consider this \( u' \). We have \( 0u' = uv = k = 01^{k-1} \). Cancelling \( 0 \) from this equality, we obtain \( u'v = 1^{k-1} \). Hence, \( v \) is a suffix of the word \( 1^{k-1} \). Thus, \( v \) has the form \( 1^p \) for some \( p \in \mathbb{N} \) (since every suffix of the word \( 1^{k-1} \) has this form). Consider this \( p \). The word \( v \) is nonempty; thus, \( v \neq \varnothing \). We have \( p \neq 0 \) (since otherwise, we would have

\[
v = 1^p = 1^0 \quad (\text{since } p = 0)
\]

\[ = \varnothing,
\]

contradicting \( v \neq \varnothing \). Hence, \( p \geq 1 \), so that \( p-1 \geq 0 \) and thus \( 1^p \leq 11^{p-1} \). Now, \( 0 < 1^{p-1} \). Hence, Proposition 6.2(d) (applied to \( \mathcal{B}, 0, 1^{k-1}, 1 \) and \( 1^{p-1} \) instead of \( \mathfrak{A}, a, b, c \) and \( d \)) yields that either we have \( 01^{k-1} \leq 11^{p-1} \) or the word \( 0 \) is a prefix of \( 1 \). Since the word \( 0 \) is not a prefix of \( 1 \), we thus obtain \( 01^{k-1} \leq 11^{p-1} \), so that \( \tilde{k} = 01^{k-1} \leq 11^{p-1} = 1^p = v \). Hence, \( v \geq \tilde{k} \). Since \( v \neq \tilde{k} \) (because \( v \) is a proper suffix of \( \tilde{k} \)), this yields \( v > \tilde{k} \).

Now, let us forget that we fixed \( c \). We thus have proven that every nonempty proper suffix \( v \) of \( \tilde{k} \) satisfies \( v > \tilde{k} \). Since the word \( \tilde{k} \) is nonempty, this shows that the word \( \tilde{k} \) is Lyndon (by the definition of a Lyndon word), qed.
This map $\Phi$ is clearly a monoid homomorphism. Hence, $\Phi(\emptyset) = \emptyset$. Also, the map $\Phi$ is strictly order-preserving (with respect to the lexicographical orders on $A^*$ and $B^*$). Consequently, the map $\Phi$ is injective.

---

\textit{449 Proof.} Let $u$ and $v$ be two words in $A^*$ satisfying $u < v$. We are going to prove that $\Phi(u) < \Phi(v)$ in $B^*$.

We have $u < v$. Thus, $u \leq v$. By the definition of the relation $\leq$, this means that

\begin{itemize}
  \item either there exists an $i \in \{1, 2, \ldots, \min \{\ell(u), \ell(v)\}\}$ such that $(u_i < v_i)$ and every $j \in \{1, 2, \ldots, i - 1\}$ satisfies $u_j = v_j$,
  \item or the word $u$ is a prefix of $v$.
\end{itemize}

We thus must be in one of the following two cases:

\textit{Case 1:} There exists an $i \in \{1, 2, \ldots, \min \{\ell(u), \ell(v)\}\}$ such that $(u_i < v_i)$ and every $j \in \{1, 2, \ldots, i - 1\}$ satisfies $u_j = v_j$.

\textit{Case 2:} The word $u$ is a prefix of $v$.

Let us first consider Case 1. In this case, there exists an $i \in \{1, 2, \ldots, \min \{\ell(u), \ell(v)\}\}$ such that $(u_i < v_i)$ and every $j \in \{1, 2, \ldots, i - 1\}$ satisfies $u_j = v_j$. Consider this $i$. We have $\hat{a} < \hat{b}$ for any two elements $a$ and $b$ of $A$ satisfying $a < b$ (because $1 < 2 < 3 < \cdots$). Applying this to $a = u_i$ and $b = v_i$, we obtain $u_i < v_i$.

But let $g = u_i^1v_i^2 \cdots \hat{u_i} \hat{v_i}$. Every $j \in \{1, 2, \ldots, i - 1\}$ satisfies $u_j = v_j$. Thus, every $j \in \{1, 2, \ldots, i - 1\}$ satisfies $\hat{u_i} = \hat{v_i}$. Taking the product of this equality over all $j \in \{1, 2, \ldots, i - 1\}$, we obtain $\hat{u_i}u_i^1 \hat{v_i}v_i^2 \cdots \hat{u_i} \hat{v_i}^{-1}$, so that $g = u_i^1v_i^2 \cdots \hat{u_i} \hat{v_i}$. We have $u_i < v_i$. Thus, $u_i u_i^1u_i^2 \cdots \hat{u_i} u_i = v_i \hat{v_i}v_i^1v_i^2 \cdots \hat{v_i}$ (by (11.1108), applied to $a = u_i$, $b = v_i$, $s = \ell(u)-i$, $c_k = u_i+k$ and $w = v_i^1v_i^2 \cdots \hat{v_i}$). Thus, $\Phi(u) = \Phi(u_i u_i^1u_i^2 \cdots \hat{u_i} u_i) = g \Phi(u_i^1v_i^2 \cdots \hat{u_i} \hat{v_i}) = v_i^1v_i^2 \cdots \hat{v_i}$.

Hence, Proposition 6.2(b) (applied to $\Phi$, $g$, $u_i u_i^1 \cdots \hat{u_i}$ and $v_i \hat{v_i} v_i^1 \cdots \hat{v_i}$ instead of $A$, $a$, $c$ and $d$) yields

\begin{align*}
g(u_i u_i^1 \cdots \hat{u_i}) &\leq g(u_i^1v_i^2 \cdots \hat{v_i})
\end{align*}

Moreover, $g(u_i u_i^1 \cdots \hat{u_i}) \neq g(v_i \hat{v_i} v_i^1 \cdots \hat{v_i})$ (because otherwise, we would have $g(u_i u_i^1 \cdots \hat{u_i}) = g(v_i \hat{v_i} v_i^1 \cdots \hat{v_i})$, and thus (by cancelling $g$ from the equality $g(u_i u_i^1 \cdots \hat{u_i}) = g(v_i \hat{v_i} v_i^1 \cdots \hat{v_i})$) we would obtain $\hat{u_i} u_i^1 \cdots \hat{u_i} = \hat{v_i} v_i^1 \cdots \hat{v_i}$, which would contradict $\hat{u_i} u_i^1 \cdots \hat{u_i} < \hat{v_i} v_i^1 \cdots \hat{v_i}$). Combining this with $g(u_i u_i^1 \cdots \hat{u_i}) \leq g(v_i \hat{v_i} v_i^1 \cdots \hat{v_i})$, we obtain $g(u_i u_i^1 \cdots \hat{u_i}) < g(v_i \hat{v_i} v_i^1 \cdots \hat{v_i})$.

But $u = (u_1, u_2, \ldots, u_{\ell(u)})$, so that

\begin{align*}
\Phi(u) &= \Phi(u_1, u_2, \ldots, u_{\ell(u)}) \\
&= u_1 u_2 \cdots u_{\ell(u)} = (u_1 u_2 \cdots \hat{u}_{i-1}) (u_i u_i^1 \cdots \hat{u}_i) = g(u_i u_i^1 \cdots \hat{u}_i) < \Phi(v)
\end{align*}

(since $\Phi(v) = v_i^1v_i^2 \cdots \hat{v}_i$ (by the definition of $\Phi(v)$)). Thus, $\Phi(u) < \Phi(v)$ is proven in Case 1.

Let us now consider Case 2. In this case, the word $u$ is a prefix of $v$. That is, there exists a word $r \in A^*$ such that $v = ur$.

Consider this $r$. We have $\Phi(v) = \Phi(ur) = \Phi(u) \Phi(r)$ (since $\Phi$ is a monoid homomorphism). If we had $r = \emptyset$, then we would have $v = u$. But $u < v$, which is absurd. Hence, we cannot have $r = \emptyset$. In other words, $r$ is nonempty, so that $\ell(r) \geq 1$.

Now, let us check that $\Phi(r)$ is nonempty. In fact, $r = (r_1, r_2, \ldots, r_{\ell(r)})$, so that $\Phi(r) = \hat{r}_1 \hat{r}_2 \cdots \hat{r}_{\ell(r)}$ (by the definition of $\Phi(r)$). Notice that $r_1$ is well-defined, since $\ell(r) \geq 1$. From (11.1105) (applied to $k = r_1$), we obtain $\hat{r}_1 = 0^1 \hat{1}^{-1}$, so that $0$ is a prefix of the word $\hat{r}_1$. Since $\hat{r}_1$ (in turn) is a prefix of the word $\hat{r}_1 \hat{r}_2 \cdots \hat{r}_{\ell(r)}$, this yields that $0$ is a prefix of the word $\hat{r}_1 \hat{r}_2 \cdots \hat{r}_{\ell(r)}$. In other words, $0$ is a prefix of the word $\Phi(r)$ (since $\Phi(r) = \hat{r}_1 \hat{r}_2 \cdots \hat{r}_{\ell(r)}$). Hence, the word $\Phi(r)$ has a nonempty prefix (namely, $0$). Thus, the word $\Phi(r)$ is nonempty. In other words, $\Phi(r) \neq \emptyset$. That is, $\emptyset \neq \Phi(r)$.

Since $\Phi(v) = \Phi(u) \Phi(r)$, we now see that the word $\Phi(u)$ is a prefix of $\Phi(v)$ (since $\Phi(r)$ is nonempty). Hence, $\Phi(u) \leq \Phi(v)$. On the other hand, if we had $\Phi(u) = \Phi(v)$, then we would have $\Phi(u) \neq \Phi(u) = \Phi(v) = \Phi(u) \Phi(r)$, and therefore we would have $\emptyset = \Phi(r)$ (as a result of cancelling $\Phi(u)$ from the equality $\Phi(u) = \Phi(u) \Phi(r)$), which would contradict $\emptyset \neq \Phi(r)$. Hence, we cannot have $\Phi(u) = \Phi(v)$. We therefore must have $\Phi(u) \neq \Phi(v)$. Hence, $\Phi(u) \leq \Phi(v)$ becomes $\Phi(u) < \Phi(v)$. We thus have proven that $\Phi(u) < \Phi(v)$ in Case 2.

Now, we have proven that $\Phi(u) < \Phi(v)$ in each of the two cases 1 and 2. Since these two Cases cover all possibilities, this shows that $\Phi(u) < \Phi(v)$ always holds.
For every \( w \in \mathfrak{B}^* \), we denote by \( w\mathfrak{B}^* \) the set \( \{wu \mid u \in \mathfrak{B}^*\} \) (that is, the set of all words in \( \mathfrak{B}^* \) which have \( w \) as a prefix). Every nonempty word in \( \mathfrak{B}^* \) starts with either the letter 0 or the letter 1, but not both. In other words, every nonempty word in \( \mathfrak{B}^* \) belongs to either \( 0\mathfrak{B}^* \) or \( 1\mathfrak{B}^* \) (where 0 and 1 are regarded as one-letter words), but not both. In other words, \( \mathfrak{B}^* \setminus \{\emptyset\} = 0\mathfrak{B}^* \cup 1\mathfrak{B}^* \) and \( 0\mathfrak{B}^* \cap 1\mathfrak{B}^* = \emptyset \). Thus, \( \mathfrak{B}^* \setminus 1\mathfrak{B}^* = \{\emptyset\} \cup 0\mathfrak{B}^* \).

Now, let us forget that we fixed \( u \) and \( v \). We thus have proven that if \( u \) and \( v \) are two words in \( \mathfrak{A}^* \) satisfying \( u < v \), then \( \Phi(u) < \Phi(v) \) in \( \mathfrak{B}^* \). In other words, the map \( \Phi \) is strictly order-preserving, qed.

\[ \mathfrak{B}^* \setminus (0\mathfrak{B}^* \cup 1\mathfrak{B}^*) = \{\emptyset\} \cup 0\mathfrak{B}^* \]
Furthermore, $\Phi(\mathfrak{A}^*) \subset \mathfrak{B}^* \setminus 1\mathfrak{B}^*$ and $\mathfrak{B}^* \setminus 1\mathfrak{B}^* \subset \Phi(\mathfrak{A}^*)$. Combining these two relations, we obtain

$$\tag{11.1112} \Phi(\mathfrak{A}^*) = \mathfrak{B}^* \setminus 1\mathfrak{B}^*.$$ 

\textit{Proof.} Let $u \in \Phi(\mathfrak{A}^*)$. We will prove that $u \in \mathfrak{B}^* \setminus 1\mathfrak{B}^*$.

Indeed, there exists some $w \in \mathfrak{A}^*$ satisfying $u = \Phi(w)$ (since $u \in \Phi(\mathfrak{A}^*)$). Consider this $w$. If $w = \emptyset$, then $u = \Phi(\overline{w}) = \Phi(\emptyset) = \emptyset \in \mathfrak{B}^* \setminus 1\mathfrak{B}^*$ (since clearly, $\emptyset \notin 1\mathfrak{B}^*$). Hence, if $w = \emptyset$, then $u \in \mathfrak{B}^* \setminus 1\mathfrak{B}^*$ is proven. Thus, for the rest of this proof, we can WLOG assume that we don’t have $w = \emptyset$. Assume this.

The word $w$ is nonempty (since $w \neq \emptyset$), so that $\ell(w) \geq 1$. Hence, the letter $w_1$ is well-defined. We have $w = \overline{(w_1, w_2, \ldots, w_{\ell(w)})}$ and thus

$$\Phi(w) = \overline{w_1 w_2 \cdots w_{\ell(w)}} \quad \text{(by the definition of $\Phi$)}$$

$$= \overline{w_1 w_2 w_3 \cdots w_{\ell(w)}} = 0 \bigwedge_{i=1}^{\ell(w)} \bigwedge_{i=0}^{w_i - 1} \bigwedge_{i=0}^{1} \bigwedge_{i=0}^{w_i - 1} \in \mathfrak{B}^* \setminus 1\mathfrak{B}^*.$$ 

Hence, $u = \Phi(w) \in \mathfrak{B}^* \setminus 1\mathfrak{B}^*$.

Now, let us forget that we fixed $u$. We thus have proven that $u \in \mathfrak{B}^* \setminus 1\mathfrak{B}^*$ for every $u \in \Phi(\mathfrak{A}^*)$. In other words, $\Phi(\mathfrak{A}^*) \subset \mathfrak{B}^* \setminus 1\mathfrak{B}^*$, qed.

\textit{Proof of (11.1110):} We will prove (11.1110) by strong induction over $\ell(u)$:

\textit{Induction step:} Let $N \in \mathbb{N}$. Assume that (11.1110) holds whenever $\ell(u) < N$. We will now prove that (11.1110) holds whenever $\ell(u) = N$.

We know that (11.1110) holds whenever $\ell(u) < N$. In other words,

$$\tag{11.1111} \text{every } u \in \mathfrak{B}^* \setminus 1\mathfrak{B}^* \text{ satisfying } \ell(u) < N \text{ satisfies } u \in \Phi(\mathfrak{A}^*).$$

Now, fix an $u \in \mathfrak{B}^* \setminus 1\mathfrak{B}^*$ satisfying $\ell(u) = N$. We shall show that $u \in \Phi(\mathfrak{A}^*)$.

If $u = \emptyset$, then $u = \emptyset = \Phi(\emptyset) = \Phi(\mathfrak{A}^*)$. Hence, for the rest of our proof of $u \in \Phi(\mathfrak{A}^*)$, we can WLOG assume that we don’t have $u = \emptyset$. Assume this.

We have $u \neq \emptyset$ (since we don’t have $u = \emptyset$), thus $u \notin \{\emptyset\}$. We have $u \in \mathfrak{B}^* \setminus 1\mathfrak{B}^* = \{\emptyset\} \cup 0\mathfrak{B}^*$ but $u \notin \{\emptyset\}$. Hence, $u \in \{\emptyset\} \cup 0\mathfrak{B}^* \setminus \{\emptyset\} = 0\mathfrak{B}^*$. In other words, there exists a $p \in 0\mathfrak{B}^*$ such that $u = 0p$. Consider this $p$. The one-letter word 0 is a prefix of $u$ (since $0p$ is a prefix of $u$). Hence, there exists a $k \in \mathfrak{A}$ such that $k$ is a prefix of $u$ (namely, $k = 1$).

On the other hand, if $k \in \mathfrak{A}$ is such that $k$ is a prefix of $u$, then $\ell(k) \leq \ell(u)$. Hence, if $k \in \mathfrak{A}$ is such that $k$ is a prefix of $u$, then $\ell(u) \geq \ell(k) = k$ (by (11.1106)). Hence, if $k \in \mathfrak{A}$ is such that $k$ is a prefix of $u$, then $k \leq \ell(u)$. Hence, only finitely many $k \in \mathfrak{A}$ have the property that $k$ is a prefix of $u$ (because only finitely many $k \in \mathfrak{A}$ have the property that $k \leq \ell(u)$).

We now have made the following two observations:

- There exists a $k \in \mathfrak{A}$ such that $k$ is a prefix of $u$.
- Only finitely many $k \in \mathfrak{A}$ have the property that $k$ is a prefix of $u$.

Combining these two observations, we conclude that there exists a \textbf{largest} $k \in \mathfrak{A}$ such that $k$ is a prefix of $u$. Consider this largest $k$. Then, $k$ is a prefix of $u$, but $k + 1$ is not a prefix of $u$.

There exists a $v \in \mathfrak{B}^*$ such that $u = kv$ (since $k$ is a prefix of $u$). Consider this $v$. We have $\ell\left(\overline{u_{=k}}\right) = \ell\left(\overline{kv}\right) = \ell(k) + \ell(v) = k + \ell(v) \geq 1 + \ell(v) > \ell(v)$, so that $\ell(v) < \ell(u) = N$.

Let us now assume that $v \in \mathfrak{B}^* \setminus 1\mathfrak{B}^*$. Then, $v \in \Phi(\mathfrak{A}^*)$ (by (11.1111), applied to $v$ instead of $u$). In other words, there exists a $v' \in \mathfrak{A}^*$ such that $v = \Phi(v')$. Consider this $v'$. The definition of $\Phi(k)$ (where $k$ stands for the one-letter word $(k) \in \mathfrak{A}^*$)
Next, we recall that $\mathcal{L}_B$ is the set of all Lyndon words over the alphabet $B$. We notice that $\Phi(\mathcal{L}) \subset \mathcal{L}_B \setminus \{1\}$ and $\mathcal{L}_B \setminus \{1\} \subset \Phi(\mathcal{L})$. Combining the latter two relations, we obtain
\[
\Phi(\mathcal{L}) = \mathcal{L}_B \setminus \{1\}.
\]
Hence, the map $\Phi$ restricts to a bijection $\mathcal{L} \rightarrow \mathcal{L}_B \setminus \{1\}$.

\[
\Phi(k) = \tilde{k}, \quad \text{Hence,} \quad u = \begin{pmatrix} k \end{pmatrix} \frac{v}{\Phi(k) = \Phi(v')} = \Phi(kv') = \Phi(kv') \quad \text{(since $\Phi$ is a monoid homomorphism)}
\]
\[
\in \Phi(\mathcal{S}^*). \quad (by \ (11.1107))
\]

Now, let us forget that we have assumed that $v \in B^* \setminus 1B^*$. We have thus proven that $u \in \Phi(\mathcal{S}^*)$ under the assumption that $v \in \mathcal{S}^* \setminus 1\mathcal{S}^*$. Thus, for the rest of the proof of $u \in \Phi(\mathcal{S}^*)$, we can WLOG assume that we don’t have $v \in \mathcal{S}^* \setminus 1\mathcal{S}^*$. Assume this.

We have $v \in 1\mathcal{S}^*$ (since we don’t have $v \in \mathcal{S}^* \setminus 1\mathcal{S}^*$). Thus, there exists a $v' \in \mathcal{S}^*$ such that $v = 1v'$. Consider this $v'$.

We have
\[
u = \begin{pmatrix} k \end{pmatrix} \frac{v}{\Phi(\mathcal{L}) = \Phi(u')} = \begin{pmatrix} k \end{pmatrix} \frac{1}{\Phi(\mathcal{L}) = \Phi(u')} \left( \begin{pmatrix} k \end{pmatrix} + 1 \right) v'.
\]

Hence, $\tilde{k} + 1$ is a prefix of $u$. This contradicts the fact that $\tilde{k} + 1$ is not a prefix of $u$. From this contradiction, we conclude that $u \in \Phi(\mathcal{S}^*)$ (since ex falso quodlibet). Thus, $u \in \Phi(\mathcal{S}^*)$ is proven.

Proof. Let $u \in \Phi(\mathcal{L})$. We are going to prove that $u \in \mathcal{L}_B \setminus \{1\}$.

There exists a $w \in \mathcal{L}$ such that $u = \Phi(w)$ (since $u \in \Phi(\mathcal{L})$). Consider this $w$. The word $w$ is Lyndon (since $w \in \mathcal{L}$), and thus nonempty. Hence, $\ell(w) \geq 1$.

Let $n = \ell(w)$. Thus, $n = \ell(w) \geq 1$. We have $w = (w_1, w_2, \ldots, w_{\ell(w)}) = (w_1, w_2, \ldots, w_n)$ (since $\ell(w) = n$).

Thus, $w = (w_1, w_2, \ldots, w_n)$, so that $\Phi(w) = \tilde{w_1} \tilde{w_2} \cdots \tilde{w_n}$ (by the definition of $\Phi(w)$). For every $i \in \{1, 2, \ldots, n\}$, the word $\tilde{w_i}$ is a Lyndon word over the alphabet $B$ (by (11.1109), applied to $k = w_i$). Thus, $w_1, w_2, \ldots, w_n$ are Lyndon words over the alphabet $B$.

Let now $i \in \{1, 2, \ldots, n\}$. Then, $(w_i, w_{i+1}, \ldots, w_n)$ is a nonempty suffix of $w$ (since $w = (w_1, w_2, \ldots, w_n)$).

Recall that $w$ is Lyndon. Hence, Corollary 6.15 (applied to $v = (w_1, w_{i+1}, \ldots, w_n)$) yields that $(w_i, w_{i+1}, \ldots, w_n) \geq v$ (since $(w_i, w_{i+1}, \ldots, w_n)$ is a nonempty suffix of $w$). Hence, $\Phi((w_i, w_{i+1}, \ldots, w_n)) \geq \Phi(v)$ (since the map $\Phi$ is strictly order-preserving). But the definition of the map $\Phi$ yields $\Phi((w_i, w_{i+1}, \ldots, w_n)) = \tilde{w_i} \tilde{w_{i+1}} \cdots \tilde{w_n}$. Thus, $\tilde{w_i} \tilde{w_{i+1}} \cdots \tilde{w_n} = \Phi((w_i, w_{i+1}, \ldots, w_n)) \geq \Phi(v) = \tilde{w_1} \tilde{w_2} \cdots \tilde{w_n}$.

Now, let us forget that we fixed $i$. We thus have shown that $\tilde{w_i} \tilde{w_{i+1}} \cdots \tilde{w_n} \geq \tilde{w_1} \tilde{w_2} \cdots \tilde{w_n}$ for every $i \in \{1, 2, \ldots, n\}$. Hence, Exercise 6.24 (applied to $B$ and $\tilde{w_i}$ instead of $B$ and $w_i$) yields that $\tilde{w_1} \tilde{w_2} \cdots \tilde{w_n}$ is a Lyndon word. In other words, we have a Lyndon word (since $u = \Phi(w) = \tilde{w_1} \tilde{w_2} \cdots \tilde{w_n}$). More precisely, $u$ is a Lyndon word over the alphabet $B$. In other words, $u \in \mathcal{L}_B$ (since $\mathcal{L}_B$ is the set of all Lyndon words over the alphabet $B$).

On the other hand, $u \in \Phi \left( \begin{pmatrix} \mathcal{L} \end{pmatrix} \setminus \{1\} \right) \subset \Phi(\mathcal{S}^*) = B^* \setminus 1B^*$ (by (11.1112)), so that $u \notin 1B^*$. Hence, we cannot have $u = 1$ (because if we had $u = 1$, we would have $u = 1 = 1 \in 1B^*$, which would contradict $u \notin 1B^*$). Thus, we have $u \neq 1$, so that $u \notin \{1\}$. Combined with $u \in \mathcal{L}_B$, this yields $u \in \mathcal{L}_B \setminus \{1\}$, which we are trying to prove.

Proof. Let $u \in \mathcal{L}_B \setminus \{1\}$. We are going to prove that $u \in \Phi(\mathcal{L})$.

We have $u \in \mathcal{L}_B \setminus \{1\} \subset \mathcal{L}_B$. Thus, $u$ is a Lyndon word over the alphabet $B$ (since $\mathcal{L}_B$ is the set of all Lyndon words over the alphabet $B$). Also, $u \notin \{1\}$ (since $u \in \mathcal{L}_B \setminus \{1\}$), so that $u \neq 1$.

Let us now assume (for the sake of contradition) that $u \in 1B^*$. Then, there exists a $u' \in \mathcal{S}^*$ such that $u = 1u'$. Consider this $u'$.

The word $u$ is nonempty (since it is Lyndon), and thus the last letter of $u$ is well-defined. Let $g$ be this last letter of $u$. Then, $g$ is a suffix of $u$. Clearly, $g$ is nonempty (when regarded as a word). Thus, Corollary 6.15 (applied to $B$, $u$ and $g$ instead of $B$, $w$ and $v$) yields $g \geq 1 = u'$, so that $1u' \geq g = g\circ$. Also, $\ell(g) \geq 1$ (since $g$ is nonempty), whence $\ell(g) \geq 1 = \ell(1)$ and thus $\ell(1) \leq \ell(g)$. Now, Proposition 6.2(f) (applied to $B$, $1$, $u'$, $g$ and $\circ$ instead of $B$, $a$, $b$ and $d$) yields $1 \leq g$. Thus, $g \geq 1$.

Since $g$ is a single letter, this yields that $g = 1$ (because the only letter of $B\circ$ which is $\geq 1$ is $1\circ$ itself). Thus, $1u' \leq g = 1 = 1\circ$.

Now, Proposition 6.2(f) (applied to $B$, $1$, $u'$ and $\circ$ instead of $B$, $a$, $c$ and $d$) yields $u' \leq \circ$. This yields $u' = \circ$ (since the only word which is $\leq \circ$ is $\circ$ itself). Hence, $u = 1u' = 1\circ = 1$, which contradicts $u \neq 1$. This contradiction shows that our assumption (that $u \in 1B^*$) was wrong. Hence, we have $u \notin 1B^*$.
Next, we notice that
\[(11.1114) \quad \ell (\Phi (w)) = |w| \quad \text{for every } w \in \mathfrak{A}^* \]
(where the notation \(|w|\) makes sense because every \(w \in \mathfrak{A}^*\) is a composition and thus has a size)\(^{856}\).

Incidentally, here is a similar identity which will not use:
\[(11.1115) \quad (\text{the number of letters } 0 \text{ in } \Phi (w)) = \ell (w) \quad \text{for every } w \in \mathfrak{A}^* \]
\(^{857}\]

Now, fix a positive integer \(n\). The alphabet \(\mathfrak{B}\) is finite and satisfies \(2 = |\mathfrak{B}|\). Hence, Exercise 6.29 (applied to \(\mathfrak{B}\) and 2 instead of \(\mathfrak{A}\) and \(q\)) yields that the number of Lyndon words of length \(n\) over the alphabet \(\mathfrak{B}\) equals \(\frac{1}{n} \sum \mu (d) 2^{n/d}\). That is,
\[(11.1117) \quad (\text{the number of all Lyndon words of length } n \text{ over the alphabet } \mathfrak{B}) = \frac{1}{n} \sum \mu (d) 2^{n/d}.\]

Since \(u \in \mathfrak{B}^*\) and \(u \notin \mathfrak{A}^*\), we must have \(u \in \mathfrak{B}^* \setminus \mathfrak{B}^* = \Phi (\mathfrak{A}^*)\) (by (11.1112)). In other words, there exists a \(w \in \mathfrak{A}^*\) such that \(u = \Phi (w)\). Consider this \(w\).

We have \(u \neq \emptyset\) (since \(u\) is nonempty) and thus \(w \neq \emptyset\) (because otherwise, we would have \(w = \emptyset\) and thus \(u = \Phi (w) = \emptyset\)). Hence, the word \(w\) is nonempty.

Let now \(v\) be a nonempty proper suffix of \(w\). We assume (for the sake of contradiction) that \(v \leq w\). Since \(v \neq w\) (because \(v\) is a proper suffix of \(w\)) and \(v \leq w\), we have \(v < w\). Thus, \(\Phi (v) < \Phi (w)\) (since the map \(\Phi\) is strictly order-preserving). Hence, \(\Phi (v) \neq \Phi (w) = u\).

Also, \(v \neq \emptyset\) (since \(v\) is nonempty). Thus, \(\Phi (v) \neq \Phi (\emptyset)\) (because otherwise, we would have \(\Phi (v) = \Phi (\emptyset)\), so that \(v = \emptyset\) (since the map \(\Phi\) is injective), which would contradict \(v \neq \emptyset\)). Hence, \(\Phi (v) \neq \Phi (\emptyset) = \emptyset\), so that the word \(\Phi (v)\) is nonempty.

But there exists a \(p \in \mathfrak{A}^*\) such that \(w = pv\) (since \(v\) is a suffix of \(w\)). Consider this \(p\). We have \(u = \Phi \left( \begin{array}{c} w \\ v \end{array} \right) = \Phi (pv) = \Phi (p) \Phi (v)\) (since \(\Phi\) is a monoid homomorphism). Thus, \(\Phi (v)\) is a suffix of \(u\). Now, Corollary 6.15 (applied to \(\mathfrak{B}\), \(u\) and \(\Phi (v)\) instead of \(\mathfrak{A}\), \(w\) and \(v\)) yields \(\Phi (v) \geq u\). This contradicts \(\Phi (v) < u\). This contradiction shows that our assumption (that \(v \leq w\)) was wrong. Hence, we cannot have \(v \leq w\). We thus have \(v > w\).

Now, let us forget that we fixed \(v\). We thus have proven that every nonempty proper suffix \(v\) of \(w\) satisfies \(v > w\). Since the word \(w\) is nonempty, this yields that the word \(w\) is Lyndon (by the definition of a Lyndon word). In other words, \(w \in \mathcal{L}\) (since \(\mathcal{L}\) is the set of all Lyndon words over the alphabet \(\mathfrak{A}\)). Now, \(u = \Phi \left( \begin{array}{c} w \\ v \end{array} \right) \in \Phi (\mathcal{L})\).

Now, let us forget that we fixed \(u\). We thus have shown that \(u \in \Phi (\mathcal{L})\) for every \(u \in \mathfrak{L}_\mathfrak{B} \setminus \{1\}\). In other words, \(\mathfrak{L}_\mathfrak{B} \setminus \{1\} \subset \Phi (\mathcal{L})\), qed.

\(^{859}\) \textit{Proof}. Let \(w \in \mathfrak{A}^*\). Then, \(w = (w_1, w_2, \ldots, w_{\ell(w)})\), so that \(\Phi (w) = \overline{w_1 w_2 \ldots w_{\ell(w)}}\) (by the definition of \(\Phi (w)\)). Hence,
\[
\ell \left( \Phi (w) \right) = \ell (\overline{w_1 w_2 \ldots w_{\ell(w)}}) = \ell (\overline{w_1}) + \ell (\overline{w_2}) + \cdots + \ell (\overline{w_{\ell(w)}})
\]
\[
= \sum_{i=1}^{\ell(w)} w_i = w_1 + w_2 + \cdots + w_{\ell(w)} = |w|
\]
(since \(|w| = w_1 + w_2 + \cdots + w_{\ell(w)}\) (since \(w = (w_1, w_2, \ldots, w_{\ell(w)})\)), qed.

\(^{857}\) \textit{Proof}. Let \(w \in \mathfrak{A}^*\).

We notice that
\[(11.1116) \quad \left( \text{the number of letters } 0 \text{ in } \underbrace{\overline{w}}_{=01^{k-1}} \right) = \left( \text{the number of letters } 0 \text{ in } 01^{k-1} \right) = 1\]
for every \(k \in \mathfrak{A}\).
Now, let \( \text{lync}_n \) denote the number of Lyndon compositions of size \( n \). We need to prove that

\[
\text{lync}_n = \frac{1}{n} \sum_{d|n} \mu(d) \left( 2^{n/d} - 1 \right) = \frac{1}{n} \sum_{d|n} \mu(d) 2^{n/d} - \delta_{n,1}.
\]

By the definition of \( \text{lync}_n \), we have

\[
\text{lync}_n = (\text{the number of all } w \in \mathcal{L} \text{ such that } |w| = n)
\]

(since the set of all Lyndon compositions is \( \mathcal{L} \))

\[
= \left\{ w \in \mathcal{L} \mid |w| = n \right\}
\]

(by \( \Phi(w) \))

\[
= \left\{ w \in \mathcal{L} \mid \ell \left( \Phi \left( w \right) \right) = n \right\} = \left\{ w \in \mathcal{L}_{\mathcal{B}_n} \setminus \{1\} \mid \ell(w) = n \right\}
\]

\[
= \left\{ w \in \mathcal{L}_{\mathcal{B}_n} \setminus \{1\} \mid \ell(w) = n \right\} - \left\{ w \in \{1\} \mid \ell(w) = n \right\}
\]

\[
= \left\{ w \in \mathcal{L}_{\mathcal{B}_n} \setminus \{1\} \mid \ell(w) = n \right\} - \left\{ w \in \{1\} \mid \ell(w) = n \right\}
\]

\[
= \left\{ w \in \{1\} \mid \ell(w) = n \right\} \subset \mathcal{L}_{\mathcal{B}_n}
\]

\[
= (\text{the number of all Lyndon words of length } n \text{ over the alphabet } \mathcal{B}) - \delta_{n,1}
\]

\[
= \frac{1}{n} \sum_{d|n} \mu(d) 2^{n/d} - \delta_{n,1} = \frac{1}{n} \sum_{d|n} \mu(d) \left( 2^{n/d} - 1 \right)
\]

(by Lemma 11.43). This solves Exercise 6.81 again.

We have \( w = (w_1, w_2, \ldots, w_{\ell(w)}) \), so that \( \Phi(w) = \overline{w_1 w_2 \cdots w_{\ell(w)}} \) (by the definition of \( \Phi(w) \)). Hence,

\[
\text{the number of letters } 0 \text{ in } \Phi(w) = \overline{w_1 w_2 \cdots w_{\ell(w)}}
\]

= (the number of letters 0 in \( w_1 w_2 \cdots w_{\ell(w)} \))

= (the number of letters 0 in \( w_1 \)) + (the number of letters 0 in \( w_2 \))

+ \cdots + (the number of letters 0 in \( w_{\ell(w)} \))

\[
= \sum_{i=1}^{\ell(w)} \text{(the number of letters 0 in } \overline{w_i}) = \sum_{i=1}^{\ell(w)} 1 = \ell(w).
\]

(by \( (11.1116) \), applied to \( k=w_1 \))

qed.

Proof of Proposition 6.84. Write $\alpha$ and $\beta$ as $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ and $\beta = (\beta_1, \beta_2, \ldots, \beta_m)$, respectively; then $\ell(\alpha) = \ell$ and $\ell(\beta) = m$.

Fix three pairwise disjoint chain posets $(i_1 < i_2 < \cdots < i_\ell)$, $(j_1 < j_2 < \cdots < j_m)$ and $(k_1 < k_2 < k_3 < \cdots)$. For any $p \in \mathbb{N}$, we define a $p$-shuffling map to mean a map $f$ from the disjoint union of two chains to a chain

$$(i_1 < i_2 < \cdots < i_\ell) \sqcup (j_1 < j_2 < \cdots < j_m) \xrightarrow{f} (k_1 < k_2 < \cdots < k_p)$$

which is both surjective and strictly order-preserving (that is, $x < y$ implies $f(x) < f(y)$). For every $p \in \mathbb{N}$ and every $p$-shuffling map $f$, we define a composition $w(f) := (w_1(f), w_2(f), \ldots, w_p(f))$ by

$$w_s(f) := \sum_{i_u \in f^{-1}(k_s)} \alpha_u + \sum_{j_v \in f^{-1}(k_s)} \beta_v.$$ 

Then, Proposition 5.3 yields

$$M_\alpha M_\beta = \sum_f M_{w(f)},$$

where the sum is over all $p \in \mathbb{N}$ and all $p$-shuffling maps $f$. In other words,

\[
M_\alpha M_\beta = \sum_{(p,f): \quad p \in \mathbb{N}; f \text{ is a } p\text{-shuffling map}} M_{w(f)} = \sum_{p \in \mathbb{N}; \quad f \text{ is a } p\text{-shuffling map}} M_{w(f)} = \sum_{p \in \mathbb{N}; \quad f \text{ is a } p\text{-shuffling map}} M_{w(f)} = \sum_{p \in \mathbb{N}; \quad f \text{ is a } p\text{-shuffling map}} M_{w(f)}
\]

(11.119)

\[
= \sum_{p \in \mathbb{N}; \quad f \text{ is a } p\text{-shuffling map}} M_{w(f)} + \sum_{p \in \mathbb{N}; \quad f \text{ is a } p\text{-shuffling map}} \sum_{p \in \mathbb{N}; \quad f \text{ is a } p\text{-shuffling map}} M_{w(f)} = (\text{empty sum}) = 0
\]

(11.1120)

\[
= \sum_{p \in \mathbb{N}; \quad f \text{ is an } (\ell+m)\text{-shuffling map}} M_{w(f)} + \sum_{p \in \mathbb{N}; \quad f \text{ is a } p\text{-shuffling map}} M_{w(f)}
\]

(11.1121)

\[
= (\text{a sum of terms of the form } M_\delta \text{ with } \delta \in \mathfrak{X}^* \text{ satisfying } \ell(\delta) < \ell(\alpha) + \ell(\beta))
\]

Our next goal is to prove

(11.1122)

\[
\sum_{f \text{ is an } (\ell+m)\text{-shuffling map}} M_{w(f)} = \sum_{\sigma \in \mathrm{Sh}_{\ell,m}} M_{\alpha \uparrow \sigma \downarrow \beta}.
\]

\[\text{Proof of (11.1121): If } p \in \mathbb{N} \text{ is such that } p < \ell + m, \text{ and if } f \text{ is a } p\text{-shuffling map, then } w(f) = (w_1(f), \ldots, w_p(f)) \text{ is a composition of length } p < \ell + m \text{, hence, if } p \in \mathbb{N} \text{ is such that } p < \ell + m, \text{ and if } f \text{ is a } p\text{-shuffling map, then } M_{w(f)} \text{ is a term of the form } M_\delta \text{ with } \delta \in \mathfrak{X}^* \text{ satisfying } \ell(\delta) < \ell(\alpha) + \ell(\beta). \text{ Therefore, } \sum_{p \in \mathbb{N}; \quad p < \ell + m} \sum_{f \text{ is a } p\text{-shuffling map}} M_{w(f)} \text{ is a sum of terms of the form } M_\delta \text{ with } \delta \in \mathfrak{X}^* \text{ satisfying } \ell(\delta) < \ell(\alpha) + \ell(\beta). \text{ This proves (11.1121).} \]
Once (11.1122) is proven, we will immediately conclude that

\[ M_{\text{wt} f} = \sum_{f \text{ is an } (\ell+m)\text{-shuffling map}} M_{\alpha \uplus \beta} = \sum_{\gamma \in \alpha \uplus \beta} M_{\gamma} \]

(since the multiset \(\alpha \uplus \beta\) is defined as \(\{\alpha \uplus \beta : \sigma \in \text{Sh}_\ell\}\)), and therefore we will obtain

\[ M_{\alpha M_{\beta}} = \sum_{f \text{ is an } (\ell+m)\text{-shuffling map}} M_{\text{wt } f} + \sum_{p \in \mathbb{N}; f \text{ is a } p\text{-shuffling map}} \sum_{\gamma \in \alpha \uplus \beta} M_{\text{wt } f} \]

\[ = \sum_{\gamma \in \alpha \uplus \beta} M_{\gamma} + \text{(a sum of terms of the form } M_{\delta} \text{ with } \delta \in A^* \text{ satisfying } \ell(\delta) < \ell(\alpha) + \ell(\beta) \text{) (by (11.1121))} \]

which will complete the proof of Proposition 6.84. Hence, it only remains to prove (11.1122).

We let \(D\) be the poset \((i_1 < i_2 < \cdots < i_{\ell}) \cup (j_1 < j_2 < \cdots < j_{m})\), and we let \(R\) be the poset \((k_1 < k_2 < \cdots < k_{\ell+m})\). Note that \(|D| = \ell + m = |R|\). Hence, \(D\) and \(R\) are two finite sets of the same cardinality. Consequently, a given map from \(D\) to \(R\) is surjective if and only if it is bijective.859

Recall that an \((\ell+m)\)-shuffling map means a map \(f\) from \((i_1 < i_2 < \cdots < i_{\ell}) \cup (j_1 < j_2 < \cdots < j_{m})\) to \((k_1 < k_2 < \cdots < k_{\ell+m})\) which is both surjective and strictly order-preserving. In other words, an \((\ell+m)\)-shuffling map means a map \(f\) from \(D\) to \(R\) which is both surjective and strictly order-preserving.860 In other words, an \((\ell+m)\)-shuffling map means a map \(f\) from \(D\) to \(R\) which is both bijective and strictly order-preserving.861

Let us now define a bijection \(d : \{1, 2, \ldots, \ell + m\} \to D\) by

\[ d(u) = \begin{cases} 
  i_u, & \text{if } u \leq \ell; \\
  j_{u-\ell}, & \text{if } u > \ell
\end{cases}\]

for every \(u \in \{1, 2, \ldots, \ell + m\}\).

Let us further define a bijection \(r : \{1, 2, \ldots, \ell + m\} \to R\) by

\[ r(u) = k_u \]

for every \(u \in \{1, 2, \ldots, \ell + m\}\).

Notice that \(r\) is an isomorphism of posets, where we endow the set \(\{1, 2, \ldots, \ell + m\}\) with its natural total order (i.e., the order \(1 < 2 < \cdots < \ell + m\)).

We can then define a bijection

\[ \Phi : (\text{the set of all bijective maps from } \{1, 2, \ldots, \ell + m\} \text{ to } \{1, 2, \ldots, \ell + m\}) \to (\text{the set of all bijective maps from } D \text{ to } R) \]

by setting

\[ \Phi(\sigma) = r \circ \sigma \circ d^{-1} \]

for every bijective map \(\sigma : \{1, 2, \ldots, \ell + m\} \to \{1, 2, \ldots, \ell + m\}\).

862 Consider this bijection \(\Phi\). Then, \(\Phi\) is a bijection from \(S_{\ell+m}\) to (the set of all bijective maps from \(D\) to \(R\)) (because (the set of all bijective maps from \(\{1, 2, \ldots, \ell + m\}\) to \(\{1, 2, \ldots, \ell + m\}\)) = \(S_{\ell+m}\)). It is now easy to see that if \(\sigma \in S_{\ell+m}\), then we have the following equivalence of assertions:

\[ (\sigma \in \text{Sh}_{\ell,m}) \iff (\text{the map } \Phi(\sigma^{-1}) : D \to R \text{ is strictly order-preserving}) \]

863 Moreover, every \(\sigma \in \text{Sh}_{\ell,m}\) satisfies

\[ \text{wt} \left(\Phi(\sigma^{-1})\right) = \alpha \uplus \beta. \]

864 Of course, the letters \(D\) and \(R\) have been chosen to remind of “domain” and “range”.

865 Since \(D = (i_1 < i_2 < \cdots < i_{\ell}) \cup (j_1 < j_2 < \cdots < j_{m})\) and \(R = (k_1 < k_2 < \cdots < k_{\ell+m})\).

866 This is because a given map from \(D\) to \(R\) is surjective if and only if it is bijective.

867 This is a bijection since both \(d\) and \(r\) are bijections.

868 Proof of (11.1123): Let \(\sigma \in S_{\ell+m}\). The poset \(D\) is the disjoint union of the totally ordered posets \((i_1 < i_2 < \cdots < i_{\ell})\) and \((j_1 < j_2 < \cdots < j_{m})\). Hence, if \(P\) is any other poset, and \(f : D \to P\) is any map, then the map \(f : D \to P\) is strictly order-preserving if and only if it satisfies

\[ (f(i_1) < f(i_2) < \cdots < f(i_{\ell}) \text{ and } f(j_1) < f(j_2) < \cdots < f(j_{m})). \]
Now, recall that an \((\ell + m)\)-shuffling map means a map \(f\) from \(D\) to \(R\) which is both bijective and strictly order-preserving. In other words, an \((\ell + m)\)-shuffling map means a bijective map from \(D\) to \(R\) which is

\[
\sum_{i \in D} \gamma_i \sigma = (\text{the set of all bijective maps from } D \text{ to } R),
\]

Similarly, we have the following equivalence of assertions:

\[
((\Phi (\sigma^{-1})) (i_1) < (\Phi (\sigma^{-1})) (i_2) < \cdots < (\Phi (\sigma^{-1})) (i_{\ell}) \quad \text{and} \quad (\Phi (\sigma^{-1})) (j_1) < (\Phi (\sigma^{-1})) (j_2) < \cdots < (\Phi (\sigma^{-1})) (j_m)).
\]

In other words, we have the following equivalence of assertions:

\[(\text{the map } \Phi (\sigma^{-1}) : D \to R \text{ is strictly order-preserving}) \iff ((\Phi (\sigma^{-1})) (i_1) < (\Phi (\sigma^{-1})) (i_2) < \cdots < (\Phi (\sigma^{-1})) (i_{\ell}) \quad \text{and} \quad (\Phi (\sigma^{-1})) (j_1) < (\Phi (\sigma^{-1})) (j_2) < \cdots < (\Phi (\sigma^{-1})) (j_m)).\]

Now, recall that \(\Phi (\sigma^{-1}) = r \circ \sigma^{-1} \circ d^{-1}\) (by the definition of \(\Phi\)). Hence, every \(u \in \{1, 2, \ldots, \ell\}\) satisfies

\[
(\Phi (\sigma^{-1})) (i_u) = (r \circ \sigma^{-1} \circ d^{-1}) (i_u) = r \left( \sigma^{-1} \left( \frac{d^{-1} (i_u)}{u} \right) \right) = r (\sigma^{-1} (u)).
\]

Hence, we have the following equivalence of assertions:

\[
((\Phi (\sigma^{-1})) (i_1) < (\Phi (\sigma^{-1})) (i_2) < \cdots < (\Phi (\sigma^{-1})) (i_{\ell}) \quad \iff \quad (r \sigma^{-1} (1)) < (r \sigma^{-1} (2)) < \cdots < (r \sigma^{-1} (\ell)));
\]

\[
(\Phi (\sigma^{-1})) (j_1) < (\Phi (\sigma^{-1})) (j_2) < \cdots < (\Phi (\sigma^{-1})) (j_m) \quad \iff \quad (\sigma^{-1} (1) < \sigma^{-1} (2) < \cdots < \sigma^{-1} (\ell) \quad \text{and} \quad (\sigma^{-1} (\ell + 1) < \sigma^{-1} (\ell + 2) < \cdots < \sigma^{-1} (\ell + m))\).
\]

Now, the equivalence \((11.1124)\) becomes

\[
(\text{the map } \Phi (\sigma^{-1}) : D \to R \text{ is strictly order-preserving}) \iff \begin{align*}
((\Phi (\sigma^{-1})) (i_1) < (\Phi (\sigma^{-1})) (i_2) < \cdots < (\Phi (\sigma^{-1})) (i_{\ell}) & \quad \text{is equivalent to} \quad (\sigma^{-1} (1) < \sigma^{-1} (2) < \cdots < (r \sigma^{-1} (\ell)) \\
\text{(by \((11.1125)\))} & \quad \iff \quad (r \sigma^{-1} (1)) < (r \sigma^{-1} (2)) < \cdots < (r \sigma^{-1} (\ell))\), \end{align*}
\]

and

\[
(\Phi (\sigma^{-1})) (j_1) < (\Phi (\sigma^{-1})) (j_2) < \cdots < (\Phi (\sigma^{-1})) (j_m) \quad \text{is equivalent to} \quad (\sigma^{-1} (\ell + 1) < \sigma^{-1} (\ell + 2) < \cdots < (r \sigma^{-1} (\ell + m)) \\
\text{(by \((11.1126)\))} & \quad \iff \quad (\sigma^{-1} (1) < \sigma^{-1} (2) < \cdots < \sigma^{-1} (\ell)) \quad \text{and} \quad (\sigma^{-1} (\ell + 1) < \sigma^{-1} (\ell + 2) < \cdots < \sigma^{-1} (\ell + m))\).
\]

This proves \((11.1123)\).

\section*{Proof of \((11.1127)\):}

Let \(\sigma \in Sh_{\ell, m}\). Set \(f = \Phi (\sigma^{-1})\). Then, \(f\) is a bijective map from \(D\) to \(R\) (because the domain of \(\Phi\) is \((\text{the set of all bijective maps from } D \text{ to } R)\)).

By the definition of \(wt (f)\), we have \(wt (f) = (wt_1 (f), wt_2 (f), \ldots, wt_{\ell + m} (f))\), where we set \(wt_s (f) := \sum_{u \in f^{-1} (k_s)} \alpha_u + \sum_{u \in f^{-1} (k_s)} \beta_u \) for every \(s \in \{1, 2, \ldots, \ell + m\}\).

On the other hand, let \((\gamma_1, \gamma_2, \ldots, c_{\ell, m})\) be the concatenation \(\alpha \cdot \beta = (\alpha_1, \alpha_2, \ldots, \alpha_\ell, \beta_1, \beta_2, \ldots, \beta_m)\). Then, \(\alpha \circ \beta = \left(\gamma_{\alpha(1)}, \gamma_{\alpha(2)}, \ldots, \gamma_{\alpha(\ell + m)}\right)\). Let \(s \in \{1, 2, \ldots, \ell + m\}\). We are going to prove that \(\gamma_{\alpha(s)} = wt_s (f)\).

We must be in one of the following two cases:
strictly order-preserving. Hence,
\[
\sum f \text{ is an } (\ell + m)\text{-shuffling map } M_{\text{wt } f}
\]
\[
= \sum f \text{ is a bijective map from } D \to R; \quad M_{\text{wt } f} = \sum_{\sigma \in \mathcal{S}_{\ell + m}; \Phi(\sigma) \text{ is strictly order-preserving}} M_{\text{wt } (\Phi(\sigma))}
\]
\[
\begin{pmatrix}
\text{here, we have substituted } \Phi(\sigma) \text{ for } f \text{ in the sum,}
\end{pmatrix}
\]
\[
\text{since } \Phi : \mathcal{S}_{\ell + m} \to \text{(the set of all bijective maps from } D \to R) \text{ is a bijection}
\]
\[
= \sum_{\sigma \in \mathcal{S}_{\ell + m}; \Phi(\sigma^{-1}) \text{ is strictly order-preserving}} M_{\text{wt } (\Phi(\sigma^{-1}))}
\]
(because \(\Phi(\sigma^{-1})\) is strictly order-preserving if and only if \(\sigma \in \text{Sh}_{\ell, m}\) (according to \((11.1121))\))
\[
\begin{pmatrix}
\text{here, we have substituted } \sigma^{-1} \text{ for } \sigma \text{ in the sum,}
\end{pmatrix}
\]
\[
\text{since the map } \mathcal{S}_{\ell + m} \to \mathcal{S}_{\ell + m}, \sigma \mapsto \sigma^{-1} \text{ is a bijection}
\]
\[
= \sum_{\sigma \in \mathcal{S}_{\ell + m}; \Phi(\sigma) \to \mathcal{S}_{\ell + m}, \sigma \mapsto \sigma^{-1}} M_{\text{wt } \sigma^{-1}}.
\]
\[
\text{(by} \quad (\text{11.1127))}
\]

This proves \((11.1122)\). The proof of Proposition 6.84 is thus complete. \hfill \square

**Case 1:** We have \(\sigma(s) \leq \ell\).

**Case 2:** We have \(\sigma(s) > \ell\).

Let us consider Case 1 first. In this case, we have \(\sigma(s) \leq \ell\). Thus, \(\sigma(s) \in \{1, 2, \ldots, \ell\}\), so that \(i_{\sigma(s)}\) is well-defined. We have \(d(\sigma(s)) = i_{\sigma(s)}\) (by the definition of \(d\), since \(\sigma(s) \leq \ell\)), so that \(d^{-1}(i_{\sigma(s)}) = \sigma(s)\). Now, \(f = \Phi(\sigma^{-1}) = r \circ \sigma^{-1} \circ d^{-1}\) (by the definition of \(\Phi\)), so that

\[
f(i_{\sigma(s)}) = (r \circ \sigma^{-1} \circ d^{-1})(i_{\sigma(s)}) = r\left(\sigma^{-1}\left(d^{-1}(i_{\sigma(s)})\right)\right) = r\left(\sigma^{-1}(\sigma(s))\right) = r(s) = k_s
\]

(by the definition of \(r\)). Since the map \(f\) is a bijection, this yields that the set \(f^{-1}(k_s)\) equals \(\{i_{\sigma(s)}\}\). Hence, the sum \(\sum_{i_{\sigma(s)} \in f^{-1}(k_s)} \alpha_u\) contains precisely one addend, namely the one for \(u = \sigma(s)\); as a consequence, this sum simplifies to \(\sum_{i_{\sigma(s)} \in f^{-1}(k_s)} \alpha_u = \alpha_{\sigma(s)}\). On the other hand, the sum \(\sum_{j_u \in f^{-1}(k_s)} \beta_u\) is empty (since the set \(f^{-1}(k_s)\) equals \(\{i_{\sigma(s)}\}\), and thus contains no elements of the form \(j_u\)), and thus vanishes, i.e., we have \(\sum_{j_u \in f^{-1}(k_s)} \beta_u = 0\). Now,

\[
\text{wt}_s(f) = \sum_{i_{\sigma(s)} \in f^{-1}(k_s)} \alpha_u + \sum_{j_u \in f^{-1}(k_s)} \beta_u = \alpha_{\sigma(s)} = \gamma_{\sigma(s)}
\]

(because \(\gamma_{\sigma(s)} = \alpha_{\sigma(s)}\) (since \(\gamma_1, \gamma_2, \ldots, \gamma_{\ell + m} = (\alpha_1, \alpha_2, \ldots, \alpha_\ell, \beta_1, \beta_2, \ldots, \beta_m)\) and \(\sigma(s) \leq \ell\)). In other words, \(\gamma_{\sigma(s)} = \text{wt}_s(f)\).

We have thus shown that \(\gamma_{\sigma(s)} = \text{wt}_s(f)\) holds in Case 1. A similar argument (but relying on \(d(\sigma(s)) = j_{\sigma(s) - \ell}\) instead of \(d(\sigma(s)) = i_{\sigma(s)}\)) shows that \(\gamma_{\sigma(s)} = \text{wt}_s(f)\) holds in Case 2.

Thus, \(\gamma_{\sigma(s)} = \text{wt}_s(f)\) holds in both Cases 1 and 2. Since these two Cases cover all possibilities, this yields that \(\gamma_{\sigma(s)} = \text{wt}_s(f)\) always holds.

Hence, \(\gamma_{\sigma(s)} = \text{wt}_s(f)\) for every \(s \in \{1, 2, \ldots, \ell + m\}\). Thus, \((\gamma_{\sigma(1)}, \gamma_{\sigma(2)}, \ldots, \gamma_{\sigma(\ell + m)}) = (\text{wt}_1(f), \text{wt}_2(f), \ldots, \text{wt}_{\ell + m}(f))\).

Hence,

\[
\alpha \cup_{\sigma} \beta = (\gamma_{\sigma(1)}, \gamma_{\sigma(2)}, \ldots, \gamma_{\sigma(\ell + m)}) = (\text{wt}_1(f), \text{wt}_2(f), \ldots, \text{wt}_{\ell + m}(f)) = \text{wt}\left(\begin{array}{c}
f \\ = \Phi(\sigma^{-1})
\end{array}\right).
\]

This proves \((11.1127)\).
11.137. **Solution to Exercise 6.87.** 

**Solution to Exercise 6.87.**

**Proof of Corollary 6.86.** Write \( \alpha \) and \( \beta \) as \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_t) \) and \( \beta = (\beta_1, \beta_2, \ldots, \beta_m) \), respectively; then \( \ell(\alpha) = \ell \) and \( \ell(\beta) = m \).

Fix three pairwise disjoint chain posets \((i_1 < i_2 < \cdots < i_t)\), \((j_1 < j_2 < \cdots < j_m)\) and \((k_1 < k_2 < k_3 < \cdots)\). We have

\[
\sum_{p \in \mathbb{N}; \quad f \text{ is a p-shuffling map}} \sum_{p \leq \ell + m} M_{\text{wt } f} = (\text{a sum of terms of the form } M_\delta \text{ satisfying } \ell(\delta) \leq \ell(\alpha) + \ell(\beta))
\]

(by (11.1128)). This proves Corollary 6.86.

\[\Box\]

11.138. **Solution to Exercise 6.91.** 

**Solution to Exercise 6.91.**

**Proof of Lemma 6.90.** (a) Let \( \gamma \in u \sqcup v \) be arbitrary. Then, the multiset of all letters of \( \gamma \) is the disjoint union of the multiset of all letters of \( u \) with the multiset of all letters of \( v \). Hence, the sum of all letters of \( \gamma \) equals the sum of all letters of \( u \) plus the sum of all letters of \( v \). In other words, \( |\gamma| = |u| + |v| \) (because \( |\gamma| \) is the sum of all letters of \( \gamma \), because \( |u| \) is the sum of all letters of \( u \), and because \( |v| \) is the sum of all letters of \( v \)). Since \( u \in \text{Comp}_n \), we have \( |u| = n \). Similarly, \( |v| = m \). Thus, \( |\gamma| = \underbrace{|u| + |v|}_{=n+m} = n + m \), so that \( \gamma \in \text{Comp}_{n+m} \).

Now, let us forget that we fixed \( \gamma \). We thus have proven that

\[
\gamma \in \text{Comp}_{n+m} \quad \text{for every } \gamma \in u \sqcup v.
\]

Applying (11.1129) to \( \gamma = z \), we obtain \( z \in \text{Comp}_{n+m} \). This proves Lemma 6.90(a).

(b) Since \( z \in u \sqcup v \), we have \( \ell(z) = \ell(u) + \ell(v) \).

We have \( u \in \text{Comp}_n \subset \text{Comp} = \mathbb{A}^* \) and similarly \( v \in \mathbb{A}^* \). Thus, Proposition 6.84 (applied to \( \alpha = u \) and \( \beta = v \)) yields

\[
M_u M_v = \sum_{\gamma \in u \sqcup v} M_\gamma + \text{(a sum of terms of the form } M_\delta \text{ satisfying } \ell(\delta) < \ell(u) + \ell(v))
\]

so that

\[
M_u M_v - \sum_{\gamma \in u \sqcup v} M_\gamma
= \left( \text{a sum of terms of the form } M_\delta \text{ satisfying } \ell(\delta) < \ell(u) + \ell(v) \right)_{\text{Comp} \leq \ell(z)}
\]

(11.1130) \( = \text{(a sum of terms of the form } M_\delta \text{ with } \delta \in \text{Comp satisfying } \ell(\delta) < \ell(z)) \).

\[\Box\]

---

Provisional proofs are included for (11.1128): The proof of this is analogous to the proof of (11.1121), with the only difference that some \( < \) signs are replaced by \( \leq \) signs.
Now, let \( \pi \) denote the projection from the direct sum \( \text{QSym} = \bigoplus_{k \in \mathbb{N}} \text{QSym}_k \) onto its \((n + m)\)-th homogeneous component \( \text{QSym}_{n+m} \). Notice that 
\[
M_u \quad M_v \in \text{QSym}_n \cdot \text{QSym}_m \subset \text{QSym}_{n+m},
\]
so that \( \pi(M_uM_v) = M_uM_v \). Also, for every \( \gamma \in u \cup v \), we have \( M_\gamma \in \text{QSym}_{n+m} \) (because (11.1129) shows that \( \gamma \in \text{Comp}_{n+m} \)) and therefore \( \pi(M_\gamma) = M_\gamma \). Hence, the \( k \)-linearity of \( \pi \) yields
\[
\pi \left( M_uM_v - \sum_{\gamma \in u \cup v} M_\gamma \right) = \pi (M_uM_v) - \sum_{\gamma \in u \cup v} \pi(M_\gamma) = M_uM_v - \sum_{\gamma \in u \cup v} M_\gamma.
\]

(11.1131)

On the other hand,
\[
\pi (M_\delta) = 0 \quad \text{for every } \delta \in \text{Comp} \setminus \text{Comp}_{n+m}.
\]

But applying the projection \( \pi \) to the equality (11.1130), we obtain
\[
\pi \left( M_uM_v - \sum_{\gamma \in u \cup v} M_\gamma \right) = \pi \left( M_uM_v \right) - \sum_{\gamma \in u \cup v} \pi(M_\gamma)
\]
\[
= (a \text{ sum of terms of the form } M_\delta \text{ with } \delta \in \text{Comp} \text{ satisfying } \ell(\delta) < \ell(z))
\]
\[
= (a \text{ sum of terms of the form } \pi(M_\delta) \text{ with } \delta \in \text{Comp} \text{ satisfying } \ell(\delta) < \ell(z))
\]
\[
= \left( \text{a sum of terms of the form } \pi(M_\delta) \quad \text{with } \delta \in \text{Comp}_{n+m} \text{ satisfying } \ell(\delta) < \ell(z) \right)
\]
\[
\quad \left( \text{since } M_\delta \in \text{QSym}_{n+m} \text{ (since } \delta \in \text{Comp}_{n+m}) \right)
\]
\[
\quad + \left( \text{a sum of terms of the form } \pi(M_\delta) \quad \text{with } \delta \in \text{Comp} \setminus \text{Comp}_{n+m} \text{ satisfying } \ell(\delta) < \ell(z) \right)
\]
\[
\quad \text{(by (11.1132))}
\]
\[
= (a \text{ sum of terms of the form } M_\delta \text{ with } \delta \in \text{Comp}_{n+m} \text{ satisfying } \ell(\delta) < \ell(z))
\]
\[
\quad + (a \text{ sum of terms of the form } 0 \text{ with } \delta \in \text{Comp} \setminus \text{Comp}_{n+m} \text{ satisfying } \ell(\delta) < \ell(z))
\]
\[
= (a \text{ sum of terms of the form } M_\delta \text{ with } \delta \in \text{Comp}_{n+m} \text{ satisfying } \ell(\delta) < \ell(z))
\]
\[
= (a \text{ sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } \ell(w) < \ell(z))
\]
\[
\quad \text{(here, we renamed the index } \delta \text{ as } w)
\]
\[
= \left( \text{a sum of terms of the form } M_W \text{ with } W \in \text{Comp}_{n+m} \text{ satisfying } W < z \right)
\]

866 Proof of (11.1132): Let \( \delta \in \text{Comp} \setminus \text{Comp}_{n+m} \). Then, \( \delta \in \text{Comp} \) but \( \delta \notin \text{Comp}_{n+m} \). In other words, \( \delta \) is a composition with size \(|\delta| \neq n + m \). As a consequence, \( M_\delta \) is a homogeneous element of \( \text{QSym} \) of degree \(|\delta| \neq n + m \). Therefore, \( \pi(M_\delta) = 0 \) (since \( \pi \) is the projection from the direct sum \( \text{QSym} = \bigoplus_{k \in \mathbb{N}} \text{QSym}_k \) onto its \((n + m)\)-th homogeneous component \( \text{QSym}_{n+m} \)). This proves (11.1132).
(since every \( w \in \text{Comp}_{n+m} \) satisfying \( \ell (w) < \ell (z) \) must satisfy \( w < z \)). Compared with (11.1131), this yields

\[
M_u M_v - \sum_{\gamma \in u \uplus v} M_\gamma = \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < z \right).
\]

Adding \( \sum_{\gamma \in u \uplus v} M_\gamma \) to both sides of this equality, we obtain

\[
M_u M_v = \sum_{\gamma \in u \uplus v} M_\gamma + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < z \right).
\]

But every \( \gamma \in u \uplus v \) satisfying \( \gamma \neq z \) must satisfy \( \gamma \in \text{Comp}_{n+m} \) (by (11.1129)) and \( \gamma < z \) \( \text{will} \) \( 867 \). Hence,

\[
\sum_{\gamma \in u \uplus v; \gamma \neq z} M_\gamma = \left( \text{a sum of terms of the form } M_\gamma \text{ with } \gamma \in \text{Comp}_{n+m} \text{ satisfying } \gamma < z \right).
\]

But let \( h \) be the multiplicity with which the word \( z \) appears in the multiset \( u \uplus v \). Then, \( h \) is a positive integer (since \( z \) is an element of the multiset \( u \uplus v \)), and satisfies \( \sum_{\gamma \in u \uplus v; \gamma = z} M_\gamma = h M_z \). Now,

\[
\sum_{\gamma \in u \uplus v} M_\gamma = \sum_{\gamma \in u \uplus v; \gamma \neq z} M_\gamma + \sum_{\gamma \in u \uplus v; \gamma = z} M_\gamma
\]

\[
= \sum_{\gamma \in u \uplus v; \gamma = z} M_\gamma + \sum_{\gamma \in u \uplus v; \gamma \neq z} M_\gamma \quad \text{(since } \gamma = z)\]

\[
= h M_z + \sum_{\gamma \in u \uplus v; \gamma \neq z} M_\gamma \quad \text{(a sum of terms of the form } M_\gamma \text{ with } \gamma \in \text{Comp}_{n+m} \text{ satisfying } \gamma < z \text{ will }) \quad \text{(by (11.1134))}
\]

\[
= h M_z + \left( \text{a sum of terms of the form } M_\gamma \text{ with } \gamma \in \text{Comp}_{n+m} \text{ satisfying } \gamma < z \right) \quad \text{(here, we have renamed the index } \gamma \text{ as } w)\]

\[
= h M_z + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < z \right).
\]

\(\textbf{867 Proof.} \) Let \( \gamma \in u \uplus v \) be such that \( \gamma \neq z \). Since \( \gamma \in u \uplus v \), we must have \( \ell (\gamma) = \ell (u) + \ell (v) = \ell (z) \). Also, \( \gamma \) and \( z \) both belong to \( \text{Comp}_{n+m} \). Now, \( \gamma \) is an element of the multiset \( u \uplus v \), whereas \( z \) is the lexicographically highest element of this multiset. Hence, \( \gamma \leq z \) with respect to the lexicographic order. Since the elements \( \gamma \) and \( z \) of \( \text{Comp}_{n+m} \) satisfy \( \ell (\gamma) = \ell (z) \) and \( \gamma \leq z \) with respect to the lexicographic order, we must have \( \gamma \leq z \) will and thus \( \gamma < z \) will (since \( \gamma \neq z \)), qed.
Hence, (11.1133) becomes
\[
M_u M_v = \sum_{\gamma \in \text{Comp}_{n+m}} M_{\gamma} + (\text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < \ell)
\]
\[
\quad + (\text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < \ell)
\]
\[
\quad = h M_z + (\text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < \ell)
\]
\[
\quad + (\text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < \ell)
\]
This proves Lemma 6.90(b).

(c) Since \(z \in u \uplus v\), we have \(\ell(z) = \ell(u) + \ell(v)\). Since \(v' < v\), we have \(v' \neq v\).

Let \(z'\) be the lexicographically highest element of the multiset \(u \uplus v'\). Lemma 6.90(a) (applied to \(v'\) and \(z'\) instead of \(v\) and \(z\)) yields \(z' \in \text{Comp}_{n+m}\). Since \(z'\) is an element of the multiset \(u \uplus v'\), we have \(\ell(z') = \ell(u) + \ell(v')\).

Now, it is easy to see (using Lemma 6.76) that \(wll\). Hence, every \(w \in \text{Comp}_{n+m}\) satisfying \(w < z'\) also satisfies \(w < z\) (because it satisfies \(w < z' < z\)). Applying Lemma 6.90(b) to \(v'\) and \(z'\) instead of \(v\) and \(z\), we conclude that there exists a positive integer \(h\) such that
\[
M_u M_v = h M_z + (\text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < z')
\]

Proof. We have \(v' < v\), thus \(v' \leq v\). According to the definition of the relation \(\leq\), this means that we are in one of the following two cases:

Case 1: We have \(\ell(v') < \ell(v)\).

Case 2: We have \(\ell(v') = \ell(v)\) and \(v' < v\) in lexicographic order.

Let us first consider Case 1. In this case, we have \(\ell(v') < \ell(v)\). Now, \(\ell(z') = \ell(u) + \ell(v') < \ell(u) + \ell(v) = \ell(z)\). But any two elements \(\alpha\) and \(\beta\) of \(\text{Comp}_{n+m}\) satisfying \(\ell(\alpha) < \ell(\beta)\) must satisfy \(\alpha \leq \beta\) (by the definition of \(\leq\)). Applying this to \(\alpha = z'\) and \(\beta = z\), we obtain \(z' \leq z\). Combined with \(z' \neq z\) (since \(\ell(z') < \ell(z)\)), this yields \(z' < z\). Thus, \(z' < z\) is proven in Case 1.

Let us now consider Case 2. In this case, we have \(\ell(v') = \ell(v)\) and \(v' < v\) in lexicographic order. Now, \(\ell(z') = \ell(u) + \ell(v') = \ell(u) + \ell(v) = \ell(z)\).

We know that \(z'\) is an element of the multiset \(u \uplus v'\); in other words, \(z'\) can be written in the form \(z' = u \uplus v'\) for some \(\sigma \in \text{Sh}(u), \ell(v')\) (because every element of the multiset \(u \uplus v'\) can be written in this form). Consider this \(\sigma\). We have \(\sigma \in \text{Sh}(u), \ell(v') = \text{Sh}(u), \ell(v)\) (since \(\ell(v') = \ell(v)\)), and thus \(u \uplus v\) is a well-defined element of the multiset \(u \uplus v\). Therefore, \(u \uplus v \leq z\) (because \(z\) is the lexicographically highest element of this multiset \(u \uplus v\)).

But we have \(v' \leq v\) with respect to the relation \(\leq\) on \(W^*\) defined in Definition 6.1 (since \(v' \leq v\) in lexicographic order). Thus, \(v' < v\) with respect to this relation (since \(v' \neq v\)). Hence, Lemma 6.76(a) (applied to \(\ell(u)\) and \(\ell(v)\)) instead of \(n\) and \(m\) yields \(u \uplus v' < u \uplus v \leq z\). Thus, \(z' = u \uplus v' < z\) with respect to the relation \(\leq\) on \(W^*\). In other words, \(z' < z\) in lexicographic order (since \(\ell(z') = \ell(z)\)), thus \(z' < z\).

Now, the two elements \(z'\) and \(z\) of \(\text{Comp}_{n+m}\) satisfy \(\ell(z') = \ell(z)\) and \(z' \leq z\) in lexicographic order. But any two elements \(\alpha\) and \(\beta\) of \(\text{Comp}_{n+m}\) satisfying \(\ell(\alpha) = \ell(\beta)\) and \(\alpha \leq \beta\) in lexicographic order) must satisfy \(\alpha \leq \beta\) (by the definition of \(\leq\)). Applying this to \(\alpha = z'\) and \(\beta = z\), we obtain \(z' \leq z\). Since \(z' \neq z\), this yields \(z' < z\). Hence, \(z' < z\) is proven in Case 2.

Now, we have proved \(z' < z\) in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that \(z' < z\) always holds, qed.
Consider this \( h \). We have
\[
M_{h}M'_{w} = \left( \sum_{i=1}^{h} M_{z'} \right) = \sum_{i=1}^{h} M_{z'} = \left( \text{a sum of terms of the form } M_{w} \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < z' \right)
\]
\[
\text{since } z' \in \text{Comp}_{n+m} \text{ satisfies } z' < z
\]
\[
+ \left( \text{a sum of terms of the form } M_{w} \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < z' \right)
\]
\[
\text{will also satisfy } w < z
\]
\[
= \left( \text{a sum of terms of the form } M_{w} \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < z \right)
\]
\[
\text{will}
\]
\[
\text{will satisfy } w < z
\]
\[
M_{x} = \sum_{y \in \text{Comp}_{N}; y \leq x} \eta_{x,y} M_{y}
\]
This proves Lemma 6.90(c).


Proof of Proposition 6.89. We shall prove Proposition 6.89 by strong induction over \( n \). So we fix some \( N \in \mathbb{N} \), and we assume that Proposition 6.89 holds whenever \( n < N \). We now need to prove that Proposition 6.89 holds for \( n = N \). In other words, we need to prove that for every \( x \in \text{Comp}_{N} \), there is a family \( (\eta_{x,y})_{y \in \text{Comp}_{N}} \in \mathbb{N}^{\text{Comp}_{N}} \) of elements of \( N \) satisfying

\[
(11.1135) \quad M_{x} = \sum_{y \in \text{Comp}_{N}; y \leq x} \eta_{x,y} M_{y}
\]

and \( \eta_{x,x} \neq 0 \) (in \( \mathbb{N} \)).

Let \( x \in \text{Comp}_{N} \) be arbitrary. We need to prove that there is a family \( (\eta_{x,y})_{y \in \text{Comp}_{N}} \in \mathbb{N}^{\text{Comp}_{N}} \) of elements of \( N \) satisfying (11.1135) and \( \eta_{x,x} \neq 0 \).

Let \( (a_{1}, a_{2}, \ldots, a_{p}) \) be the CFL factorization of \( x \). Then, \( a_{1}, a_{2}, \ldots, a_{p} \) are Lyndon words satisfying \( x = a_{1}a_{2}\cdots a_{p} \) and \( a_{1} \geq a_{2} \geq \cdots \geq a_{p} \).

If \( x \) is the empty word, then our claim is trivial (in fact, we can just set \( \eta_{x,x} = 1 \) in this case, and (11.1135) holds obviously). Hence, for the rest of this proof, we WLOG assume that \( x \) is not the empty word. Thus, \( p \neq 0 \) (because otherwise, \( x = a_{1}a_{2}\cdots a_{p} \) would be an empty product and thus the empty word, contradicting the assumption that \( x \) is not the empty word). Hence, we can define two words \( u \in \mathbb{A}^{*} \) and \( v \in \mathbb{A}^{*} \) by \( u = a_{1} \) and \( v = a_{2}a_{3}\cdots a_{p} \). The word \( u \) is Lyndon (since \( u = a_{1} \) and since \( a_{1} \) is Lyndon), and thus has \( (u) \) as its CFL factorization. The CFL factorization of the word \( v \) is \( (a_{2}, a_{3}, \ldots, a_{p}) \) (since the words \( a_{2}, a_{3}, \ldots, a_{p} \) are Lyndon and satisfy \( v = a_{2}a_{3}\cdots a_{p} \) and \( a_{2} \geq a_{3} \geq \cdots \geq a_{p} \) (because \( a_{1} \geq a_{2} \geq \cdots \geq a_{p} \)). Also, \( u = a_{1} \geq a_{j+1} \) for every \( i \in \{1, 2, \ldots, 1\} \) and \( j \in \{1, 2, \ldots, p-1\} \) (because \( a_{1} \geq a_{2} \geq \cdots \geq a_{p} \)). Hence, we can apply Theorem 6.44(c) to \( 1, (u), p-1 \) and \( (a_{2}, a_{3}, \ldots, a_{p}) \) instead of \( p, (a_{1}, a_{2}, \ldots, a_{p}), q \) and \( (b_{1}, b_{2}, \ldots, b_{q}) \). As a result, we conclude that the lexicographically highest element of the multiset \( u \uplus v \) is
\[
\sum_{a_{1}}^{u} \sum_{a_{2}a_{3}\cdots a_{p}}^{v} = a_{1} \left( a_{2}a_{3}\cdots a_{p} \right) = a_{1}a_{2}\cdots a_{p} = x.
\]
But $M_u = M_u$ (by the definition of $M_u$, since $u$ has CFL factorization $(u)$), and $M_x = M_u M_v$.

The word $u$ is Lyndon and thus nonempty, so that $|u| > 0$. Now, $|x| = |uv| = |u| + |v|$, so that $|u| + |v| = |x| = N$ (since $x \in \text{Comp}_N$). Thus, $N = |u| + |v| > |v|$, so that $|v| < N$. Hence, the induction hypothesis tells us that we can apply Proposition 6.89 to $|v|$ and $v$ instead of $u$ and $x$ (since $v \in \text{Comp}_{|v|}$).

As a result, we see that there is a family $(\eta_{v,y})_{y \in \text{Comp}_{|v|}} \in \mathbb{N}^{\text{Comp}_{|v|}}$ of elements of $\mathbb{N}$ satisfying

$$M_v = \sum_{y \in \text{Comp}_{|v|}} \eta_{v,y} M_y$$

and $\eta_{v,v} \neq 0$ (in $\mathbb{N}$). Consider this family $(\eta_{v,y})_{y \in \text{Comp}_{|v|}}$.

We have $u \in \text{Comp}_{|u|}$ and $v \in \text{Comp}_{|v|}$. Hence, Lemma 6.90(b) (applied to $x$, $|u|$ and $|v|$ instead of $z$, $u$ and $v$) yields that there exists a positive integer $h$ such that

$$M_u M_v = h M_x + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{|u|+|v|} \text{ satisfying } w < x \right)$$

(11.1136) $$= h M_x + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_N \text{ satisfying } w < x \right).$$

Consider this $h$. Then, $h \in \mathbb{N}$ and $h \neq 0$ (since $h$ is a positive integer).

---

\textit{Proof.} Since the CFL factorization of $v$ is $(a_2, a_3, \ldots, a_p)$, we have $M_v = M_{a_2} M_{a_3} \cdots M_{a_p}$ (by the definition of $M_v$). But since the CFL factorization of $x$ is $(a_1, a_2, \ldots, a_p)$, we have

$$M_x = M_{a_1} M_{a_2} \cdots M_{a_p} = M_u \left( M_{a_2} M_{a_3} \cdots M_{a_p} \right) = M_v$$

$$= M_u \left( M_{a_2} \cdots M_{a_p} \right) = M_v.$$

\text{qed.}
In other words, we can write
\[ M_x = M_u \sum_{y \in \text{Comp}_v|\cdot|; \ y \leq v} \eta_{v,y}M_y = M_u \sum_{y \in \text{Comp}_v|\cdot|; \ y \leq v} \eta_{v,y}M_y = \sum \eta_{v,y}M_uM_y \]

\[ = \eta_{v,v} M_x + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_N \text{ satisfying } w < x \right) \]
(by (11.1136))

\[ + \sum_{y \in \text{Comp}_v|\cdot|; \ y \leq v} \eta_{v,y} \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_v|\cdot| \text{ satisfying } w < x \right) \]
(by Lemma 6.90(c), applied to \( n = |u|, m = |v|, v' = y \) and \( z = x \))

(here, we have split off the addend for \( y = v \) from the sum)

\[ = \eta_{v,v} \left( hM_x + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_N \text{ satisfying } w < x \right) \right) \]

\[ = \eta_{v,v} hM_x + \eta_{v,v} \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_N \text{ satisfying } w < x \right) \]

\[ + \sum_{y \in \text{Comp}_v|\cdot|; \ y \leq v} \eta_{v,y} \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_v|\cdot| \text{ satisfying } w < x \right) \]
(since \( \eta_{v,y} \in \mathbb{N} \) for every \( y \in \text{Comp}_v|\cdot| \) satisfying \( y < v \))

\[ = \eta_{v,v} \left( hM_x + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_N \text{ satisfying } w < x \right) \right) \]

\[ + \left( \text{an } \mathbb{N}\text{-linear combination of terms of the form } M_w \text{ with } w \in \text{Comp}_N \text{ satisfying } w < x \right) \]
(since \( \eta_{v,y} \in \mathbb{N} \))

\[ + \left( \text{an } \mathbb{N}\text{-linear combination of terms of the form } M_w \text{ with } w \in \text{Comp}_N \text{ satisfying } w < x \right) \]

\[ = \eta_{v,v} hM_x + \left( \text{an } \mathbb{N}\text{-linear combination of terms of the form } M_w \text{ with } w \in \text{Comp}_N \text{ satisfying } w < x \right) \]

\[ + \left( \text{an } \mathbb{N}\text{-linear combination of terms of the form } M_w \text{ with } w \in \text{Comp}_N \text{ satisfying } w < x \right) \]

\[ = \eta_{v,v} hM_x + \left( \text{an } \mathbb{N}\text{-linear combination of terms of the form } M_w \text{ with } w \in \text{Comp}_N \text{ satisfying } w < x \right) \].

In other words, we can write \( M_x \) in the form
\[ (11.1137) \quad M_x = \eta_{v,v} hM_x + c, \]
where \( c \) is an \( \mathbb{N} \)-linear combination of terms of the form \( M_w \) with \( w \in \text{Comp}_N \) satisfying \( w < x \). Consider this \( c \).

Write \( c \) in the form
\[
(11.1138) \quad c = \sum_{y \in \text{Comp}_N; \ y \leq x; w \in \text{Comp}} \eta_{x,y} M_y,
\]
where \( \eta_{x,y} \) is an element of \( \mathbb{N} \) for every \( y \in \text{Comp}_N \) satisfying \( y < x \). (This is possible since \( c \) is an \( \mathbb{N} \)-linear combination of terms of the form \( M_w \) with \( w \in \text{Comp}_N \) satisfying \( w < x \).) Thus, we have defined a family \( (\eta_{x,y})_{y \in \text{Comp}_N; \ y \leq x; w \in \text{Comp}} \) of elements of \( \mathbb{N} \). Extend this family to a family \( (\eta_{x,y})_{y \in \text{Comp}_N} \in \mathbb{N}^{\text{Comp}_N} \) of elements of \( \mathbb{N} \) by defining
\[
\eta_{x,y} = \delta_{x,y} \eta_{v,v} h \quad \text{for every} \quad y \in \text{Comp}_N \text{ which does not satisfy} \ y < x.
\]
Notice that the definition of \( \eta_{x,y} \) yields \( \eta_{x,x} = \delta_{x,x} \eta_{v,v} h \) (since \( x \) does not satisfy \( x < x \)), and thus \( \eta_{x,x} = \delta_{x,x} \eta_{v,v} h = \frac{\eta_{v,v}}{v \neq 0} \neq 0 \).

Now,
\[
\sum_{y \in \text{Comp}_N; \ y \leq x; w \in \text{Comp}} \eta_{x,y} M_y = \sum_{y \in \text{Comp}_N; \ y < x; w \in \text{Comp}} \eta_{x,y} M_y + \sum_{y \in \text{Comp}_N; \ y = x; w \in \text{Comp}} \eta_{x,y} M_y = \eta_{v,v} h M_x + c = M_x \quad \text{(by (11.1137))},
\]
and thus \( M_x = \sum_{y \in \text{Comp}_N; \ y \leq x; w \in \text{Comp}} \eta_{x,y} M_y \). In other words, (11.1135) is satisfied.

Hence, we have constructed a family \( (\eta_{x,y})_{y \in \text{Comp}_N} \in \mathbb{N}^{\text{Comp}_N} \) of elements of \( \mathbb{N} \) satisfying (11.1135) and \( \eta_{x,x} \neq 0 \). Thus, we have shown the existence of such a family. The induction step is thus complete, and Proposition 6.89 is proven by induction. \( \square \)

11.140. Solution to Exercise 6.94. Solution to Exercise 6.94.

Proof of Proposition 6.93. For every composition \( u \in \text{Comp} = \mathbb{N}^* \), define an element \( M_u \in \text{QSym} \) by \( M_u = M_{a_1} M_{a_2} \cdots M_{a_p} \), where \( (a_1, a_2, \ldots, a_p) \) is the CFL factorization of the word \( u \).

Fix \( n \in \mathbb{N} \). It is easy to see that \( (M_u)_{u \in \text{Comp}_n} \) is a basis of the \( k \)-module \( \text{QSym}_n \). According to Proposition 6.89, the family \( (M_u)_{u \in \text{Comp}_n} \) expands upper-triangularly (with respect to the total order \( \leq \) on \( \text{Comp}_n \)) with respect to the basis \( (M_u)_{u \in \text{Comp}_n} \). The diagonal entries of this upper-triangular expansion are nonzero elements of \( \mathbb{N} \) (indeed, these diagonal entries are precisely the elements \( \eta_{x,x} \) defined in Proposition 6.89, and Proposition 6.89 shows that \( \eta_{x,x} \neq 0 \) in \( \mathbb{N} \)), and therefore are invertible elements of \( k \) (since \( \mathbb{Q} \) is a subring of \( k \)). So the family \( (M_u)_{u \in \text{Comp}_n} \) expands upper-triangularly with respect to the basis \( (M_u)_{u \in \text{Comp}_n} \) of \( \text{QSym}_n \), and the diagonal entries of this upper-triangular expansion are invertible elements of \( k \). As a consequence, this expansion is invertible\(^{870}\), and thus the family \( (M_u)_{u \in \text{Comp}_n} \) is a basis of \( \text{QSym}_n \) as well.

Now, let us forget that we fixed \( n \). We thus have proven that the family \( (M_u)_{u \in \text{Comp}_n} \) is a basis of \( \text{QSym}_n \) for every \( n \in \mathbb{N} \). Hence, the disjoint union of the families \( (M_u)_{u \in \text{Comp}_n} \) over all \( n \in \mathbb{N} \) is a basis of the direct sum \( \bigoplus_{n \in \mathbb{N}} \text{QSym}_n \). Since the former disjoint union is the family \( (M_u)_{u \in \text{Comp}_n} \), while the latter direct sum

\(^{870}\) This is because the indexing set \( \text{Comp}_n \) is finite, and because any finite upper-triangular matrix over \( k \) whose diagonal entries are invertible must be invertible.
is the $k$-module $\bigoplus_{n \in \mathbb{N}} \text{QSym}_n = \text{QSym}$, this rewrites as follows: The family $(M_w)_{w \in \text{Comp}}$ is a basis of the $k$-module $\text{QSym}$. In other words, the family $(M_u)_{u \in \mathfrak{A}}$ is a basis of the $k$-module $\text{QSym}$ (since $\mathfrak{A}^* = \text{Comp}$).

But Lemma 6.73(a) (applied to $A = \text{QSym}$, $b_u = M_w$ and $b_u = M_u$) yields that the family $(M_u)_{u \in \mathfrak{A}}$ is an algebraically independent generating set of the $k$-algebra $\text{QSym}$ if and only if the family $(M_u)_{u \in \mathfrak{A}}$ is a basis of the $k$-module $\text{QSym}$. Hence, the family $(M_w)_{w \in \mathfrak{A}}$ is an algebraically independent generating set of the $k$-algebra $\text{QSym}$ (since we know that the family $(M_u)_{u \in \mathfrak{A}}$ is a basis of the $k$-module $\text{QSym}$). This proves Proposition 6.93. \qed

11.141. **Solution to Exercise 6.98.** Solution to Exercise 6.98. We first notice that every $s \in \{1, 2, 3, \ldots\}$ and $i \in \text{SIS}(\ell)$ satisfy

$$\text{(11.1139)} \quad (x^s_i)^s = x_i^{s(s)}$$

(a) Let $s$ be a positive integer.

We begin with the following general observation:

If $R$ is a topological commutative $k$-algebra, if $I$ is a countable (finite or not) set, and if $(s_i)_{i \in I} \in R^I$ is a power-summable family of elements of $R$, then

$$\text{(11.1140)} \quad p_s\left((s_i)_{i \in I}\right) = \sum_{i \in I} s_i^s.$$ 

Applying this to $R = k[[x]]$, $I = \text{SIS}(\ell)$ and $s_i = x_i^s$, we obtain

$$p_s\left((x^s_i)_{i \in \text{SIS}(\ell)}\right) = \sum_{i \in \text{SIS}(\ell)} (x^s_i)^s = \sum_{i \in \text{SIS}(\ell)} x_i^{s(s)}.$$ 

(by (11.1139))

---

Proof of (11.1139): Let $s \in \{1, 2, 3, \ldots\}$ and $i \in \text{SIS}(\ell)$. Then, $i$ is an $\ell$-tuple of positive integers. Write $i$ in the form $i = (i_1, i_2, \ldots, i_\ell)$. Then, $x^s_i = x_{i_1}^{s_1}x_{i_2}^{s_2}\cdots x_{i_\ell}^{s_\ell}$ (by the definition of $x^s_i$). Also, by the definition of $x_i^{s(s)}$, we have $x_i^{s(s)} = x_{i_1}^{s_1}x_{i_2}^{s_2}\cdots x_{i_\ell}^{s_\ell}$ (since $s = (s_1, s_2, \ldots, s_\ell)$). Now, taking both sides of the equality $x^s_i = x_{i_1}^{s_1}x_{i_2}^{s_2}\cdots x_{i_\ell}^{s_\ell}$ to the $s$-th power, we obtain

$$(x^s_i)^s = (x_{i_1}^{s_1}x_{i_2}^{s_2}\cdots x_{i_\ell}^{s_\ell})^s = (x_{i_1}^{s_1})^s(x_{i_2}^{s_2})^s\cdots (x_{i_\ell}^{s_\ell})^s = x_{i_1}^{s_1s}s_2^{s_2}\cdots x_{i_\ell}^{s_\ell} = x_i^{s(s)}.$$ 

This proves (11.1139).

Proof of (11.1140): Let $R$ be a topological commutative $k$-algebra. Let $I$ be a countable set. Let $(s_i)_{i \in I} \in R^I$ be a power-summable family of elements of $R$. We need to prove that (11.1140) holds.

We must be in one of the following two cases:

**Case 1:** The set $I$ is infinite.

**Case 2:** The set $I$ is finite.

Let us first consider Case 1. In this case, the set $I$ is infinite, and therefore countably infinite (since it is countable). Fix a bijection $j : \{1, 2, 3, \ldots\} \to I$ (such a bijection clearly exists). Then, $p_s\left((s_i)_{i \in I}\right)$ is the result of substituting $s_{j(1)}$, $s_{j(2)}$, $s_{j(3)}$, ... for the variables $x_1$, $x_2$, $x_3$, ... in $p_s$ (by the definition of $p_s\left((s_i)_{i \in I}\right)$). Thus,

$$p_s\left((s_i)_{i \in I}\right)$$

= \left(\text{the result of substituting } s_{j(1)}, s_{j(2)}, s_{j(3)}, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } p_s\left((s_i)_{i \in I}\right)\right)$$

= \left(\text{the result of substituting } s_{j(1)}, s_{j(2)}, s_{j(3)}, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } p_s\left((s_i)_{i \in I}\right)\right)$$

= \sum_{i \in \{1,2,3,\ldots\}} s_i^s = \sum_{i \in I} s_i^s$$

(here, we have substituted $i$ for $j(i)$ in the sum, since the map $j : \{1, 2, 3, \ldots\} \to I$ is a bijection). Thus, (11.1140) is proven in Case 1.

Let us now consider Case 2. In this case, the set $I$ is finite. Fix a bijection $j : \{1, 2, 3, \ldots, |I|\} \to I$ (such a bijection clearly exists). Then, $p_s\left((s_i)_{i \in I}\right)$ is the result of substituting $s_{j(1)}$, $s_{j(2)}$, ..., $s_{j(|I|)}$, 0, 0, 0, ... for the variables $x_1$, $x_2$, $x_3$, ... in $p_s$
Compared with $M_{\alpha(s)} = \sum_{i \in \text{SIS}(t)} x^s_{i}$ (by (6.26), applied to $\alpha \{s\}$ instead of $\alpha$), this yields $p_s \left( (x^s_i)_{i \in \text{SIS}(t)} \right) = M_{\alpha(s)}$. Exercise 6.98(a) is now solved.

(b) Let $s \in \mathbb{N}$. We shall first make some general statements.

- If $R$ is a topological commutative $k$-algebra, if $I$ is a countable (finite or not) totally ordered set, and if $(s_i)_{i \in I} \in R^I$ is a power-summable family of elements of $R$, then

$$\sum_{(i_1, i_2, \ldots, i_s) \in I^s; \ i_1 < i_2 < \cdots < i_s} s_{i_1} s_{i_2} \cdots s_{i_s} = \sum_{K \subseteq I; |K| = s} \prod_{i \in K} s_i. \tag{11.1141}$$

- If $R$ is a commutative $k$-algebra, if $I$ and $J$ are two sets, if $j : J \to I$ is a bijection, and if $(s_i)_{i \in I} \in R^I$ is a family of elements of $R$, then

$$\prod_{i \in |K|} s_i = \prod_{i \in K} s_{j(i)} \quad \text{for every finite subset } K \text{ of } J. \tag{11.1142}$$

(by the definition of $p_s \left( (s_i)_{i \in I} \right)$). Thus, $p_s \left( (s_i)_{i \in I} \right)$

$$= \left( \text{the result of substituting } s_{j(1)}, s_{j(2)}, \ldots, s_{j(|I|)}, 0, 0, 0, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \right) \sum_{i=1}^{s_i} x_i^{s_i} + \sum_{i=|I|+1}^{\infty} x_i^{s_i}$$

$$= \left( \text{the result of substituting } s_{j(1)}, s_{j(2)}, \ldots, s_{j(|I|)}, 0, 0, 0, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \right) \sum_{i=1}^{s_i} x_i^{s_i} + \sum_{i=|I|+1}^{\infty} x_i^{s_i}$$

$$= \sum_{i \in I} s_i$$

(here, we have substituted $i$ for $j(i)$ in the sum, since the map $j : \{1, 2, \ldots, |I|\} \to I$ is a bijection). Thus, (11.1140) is proven in Case 2.

We have thus proven (11.1140) in each of the Cases 1 and 2. Since these two Cases are the only cases that can occur, we thus conclude that (11.1140) holds.

Proof of (11.1141): Let $R$ be a topological commutative $k$-algebra. Let $I$ be a countable totally ordered set. Let $(s_i)_{i \in I} \in R^I$ be a power-summable family of elements of $R$. We need to prove that (11.1141) holds.

The order on $I$ is total. Hence, the map

$$\{(i_1, i_2, \ldots, i_s) \in I^s \mid i_1 < i_2 < \cdots < i_s\} \to \{K \subseteq I \mid |K| = s\}, \quad (i_1, i_2, \ldots, i_s) \mapsto \{i_1, i_2, \ldots, i_s\}$$

is a bijection (because every subset $K$ of $I$ satisfying $|K| = s$ can be written in the form $K = \{i_1, i_2, \ldots, i_s\}$ for a unique $s$-tuple $(i_1, i_2, \ldots, i_s) \in I^s$ satisfying $i_1 < i_2 < \cdots < i_s$).

Now, let $(i_1, i_2, \ldots, i_s) \in I^s$ be such that $i_1 < i_2 < \cdots < i_s$. Then, the $s$-tuple $(i_1, i_2, \ldots, i_s)$ is a list of all elements of the set $\{i_1, i_2, \ldots, i_s\}$, each occurring exactly once (since the elements $i_1, i_2, \ldots, i_s$ are pairwise distinct (since $i_1 < i_2 < \cdots < i_s$)). Hence, $\prod_{i \in (i_1, i_2, \ldots, i_s)} s_i = s_i s_{i_2} \cdots s_{i_s}$, so that $s_1 s_2 \cdots s_s = \prod_{i \in (i_1, i_2, \ldots, i_s)} s_i$.

Now, let us forget that we fixed $(i_1, i_2, \ldots, i_s) \in I^s$. We thus have shown that $s_1 s_2 \cdots s_s = \prod_{i \in (i_1, i_2, \ldots, i_s)} s_i$ for every $(i_1, i_2, \ldots, i_s) \in I^s$ such that $i_1 < i_2 < \cdots < i_s$. Hence,

$$\sum_{(i_1, i_2, \ldots, i_s) \in I^s; \ i_1 < i_2 < \cdots < i_s} s_1 s_2 \cdots s_s = \sum_{(i_1, i_2, \ldots, i_s) \in I^s; \ i \in (i_1, i_2, \ldots, i_s)} \prod_{i \in (i_1, i_2, \ldots, i_s)} s_i = \sum_{K \subseteq I; |K| = s} \prod_{i \in K} s_i$$

(here, we have substituted $K$ for $(i_1, i_2, \ldots, i_s)$ in the sum, since the map

$$\{(i_1, i_2, \ldots, i_s) \in I^s \mid i_1 < i_2 < \cdots < i_s\} \to \{K \subseteq I \mid |K| = s\}, \quad (i_1, i_2, \ldots, i_s) \mapsto \{i_1, i_2, \ldots, i_s\}$$

is a bijection). This proves (11.1141).
If $R$ is a topological commutative $k$-algebra, if $I$ and $J$ are two countable totally ordered sets, if $j : J \to I$ is a (not necessarily order-preserving!) bijection, and if $(s_i)_{i \in I} \in R^I$ is a power-summable family of elements of $R$, then

$$(11.1143) \quad \sum_{(i_1,i_2,\ldots,i_s) \in \Gamma'} s_{i_1}s_{i_2}\cdots s_{i_s} = \sum_{(j_1,j_2,\ldots,j_s) \in \Gamma'} s_{j_1}s_{j_2}\cdots s_{j_s}.$$  

If $R$ is a topological commutative $k$-algebra, if $I$ is a countable (finite or not) totally ordered set, and if $(s_i)_{i \in I} \in R^I$ is a power-summable family of elements of $R$, then

$$(11.1145) \quad e_s\left((s_i)_{i \in I}\right) = \sum_{(i_1,i_2,\ldots,i_s) \in \Gamma'} s_{i_1}s_{i_2}\cdots s_{i_s}.$$  

Proof of (11.1142): Let $R$ be a commutative $k$-algebra. Let $I$ and $J$ be two sets. Let $j : J \to I$ be a bijection. Let $(s_i)_{i \in I} \in R^I$ be a family of elements of $R$. Let $K$ be a subset of $J$. The map $j$ is a bijection, and thus injective. Hence, the map $K \to j(K)$, $i \mapsto j(i)$ is a bijection. Hence, we can substitute $j(i)$ for $i$ in the product $\prod_{i \in j(K)} s_i$. As a result, we obtain

$$(11.1142) \quad \prod_{i \in K} s_i = \prod_{i \in j(K)} s_{j(i)}. \quad \text{This proves (11.1142).}$$  

Proof of (11.1143): Let $R$ be a topological commutative $k$-algebra. Let $I$ and $J$ be two countable totally ordered sets. Let $j : J \to I$ be a bijection. Let $(s_i)_{i \in I} \in R^I$ be a power-summable family of elements of $R$.

Notice that $(s_i)_{i \in I} \in R^I$ is a reindexing of the family $(\alpha_i)_{i \in I} \in R^I$ (since $j : J \to I$ is a bijection), and thus a power-summable family of elements of $R$ (since $(s_i)_{i \in I}$ is a power-summable family of elements of $R$).

The map

$$\{K \subseteq J \mid |K| = s\} \to \{K \subseteq I \mid |K| = s\},$$

$$K \mapsto j(K)$$

is a bijection (since $j$ is a bijection from $J$ to $I$). Hence, we can substitute $j(K)$ for $K$ in the sum $\sum_{|K| = s} \prod_{i \in K} s_i$, and as a result we obtain

$$\sum_{|K| = s} \prod_{i \in K} s_i = \sum_{|K| = s} \prod_{i \in j(K)} s_{j(i)}.$$

Thus, (11.1141) becomes

$$(11.1144) \quad \sum_{(i_1,i_2,\ldots,i_s) \in \Gamma'} s_{i_1}s_{i_2}\cdots s_{i_s} = \sum_{|K| = s} \prod_{i \in K} s_i = \sum_{|K| = s} \prod_{i \in j(K)} s_{j(i)}.$$  

But

$$\sum_{(j_1,j_2,\ldots,j_s) \in \Gamma'} s_{j_1}s_{j_2}\cdots s_{j_s} = \sum_{(i_1,i_2,\ldots,i_s) \in \Gamma'} s_{i_1}s_{i_2}\cdots s_{i_s}.$$  

(here, we renamed the summation index $(j_1,j_2,\ldots,j_s)$ as $(i_1,i_2,\ldots,i_s)$)

$$= \sum_{|K| = s} \prod_{i \in K} s_{j(i)}$$

(by (11.1141), applied to $J$ and $s_{j(i)}$ instead of $I$ and $s_i$). Compared with (11.1144), this yields

$$\sum_{(i_1,i_2,\ldots,i_s) \in \Gamma'} s_{i_1}s_{i_2}\cdots s_{i_s} = \sum_{(j_1,j_2,\ldots,j_s) \in \Gamma'} s_{j_1}s_{j_2}\cdots s_{j_s}.$$  

This proves (11.1143).

Proof of (11.1145): Let $R$ be a topological commutative $k$-algebra. Let $I$ be a countable totally ordered set. Let $(s_i)_{i \in I} \in R^I$ be a power-summable family of elements of $R$. We need to prove that (11.1145) holds.

The definition of $e_s$ yields

$$(11.1146) \quad e_s = \sum_{(i_1,i_2,\ldots,i_n) \in \{1,2,3,\ldots\}^s} x_{i_1}x_{i_2}\cdots x_{i_n} = \sum_{(j_1,j_2,\ldots,j_n) \in \{1,2,3,\ldots\}^s} x_{j_1}x_{j_2}\cdots x_{j_n}.$$  

(here, we renamed the summation index $(i_1,i_2,\ldots,i_n)$ as $(j_1,j_2,\ldots,j_n)$).
Now, let us return to our proof. Applying (11.1145) to $R = k[[x]]$, $I = \text{SIS} (\ell)$ and $s_i = x_i^\alpha$, we obtain

$$e_s \left( \left( x_i^\alpha \right)_{i \in \text{SIS} (\ell)} \right) = \sum_{(i_1, i_2, \ldots, i_s) \in (\text{SIS}(\ell))^s; \ i_1 < i_2 < \cdots < i_s} x_{i_1}^\alpha x_{i_2}^\alpha \cdots x_{i_s}^\alpha.$$  

We must be in one of the following two cases:

**Case 1:** The set $I$ is infinite.

**Case 2:** The set $I$ is finite.

Let us first consider Case 1. In this case, the set $I$ is infinite, and therefore countably infinite (since it is countable). Let $J$ denote the totally ordered set $\{1, 2, 3, \ldots \}$. Then, (11.1146) rewrites as

$$(11.1147)\quad e_s = \sum_{(j_1, j_2, \ldots, j_s) \in \{1, 2, 3, \ldots \}^s; \ j_1 < j_2 < \cdots < j_s} x_{j_1} x_{j_2} \cdots x_{j_s}$$

(since $\{1, 2, 3, \ldots \} = J$).

Fix a bijection $j : \{1, 2, 3, \ldots \} \rightarrow I$ (such a bijection clearly exists, since $I$ is countably infinite). Then, $e_s \left( (s_i)_{i \in I} \right)$ is the result of substituting $s_{j(1)}$, $s_{j(2)}$, $s_{j(3)}$, ... for the variables $x_1$, $x_2$, $x_3$, ... in $e_s$ (by the definition of $e_s \left( (s_i)_{i \in I} \right)$). Thus,

$$e_s \left( (s_i)_{i \in I} \right) = \left( \text{the result of substituting } s_{j(1)}, s_{j(2)}, s_{j(3)}, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \right)$$

(by (11.1147))

$$(11.1148)\quad = \sum_{(j_1, j_2, \ldots, j_s) \in \{1, 2, 3, \ldots \}^s; \ j_1 < j_2 < \cdots < j_s} s_{j(j_1)} s_{j(j_2)} \cdots s_{j(j_s)}.$$  

But $J$ is a countable totally ordered set (since $J = \{1, 2, 3, \ldots \}$). Also, $j$ is a bijection $\{1, 2, 3, \ldots \} \rightarrow I$. In other words, $j$ is a bijection $J \rightarrow I$ (since $J = \{1, 2, 3, \ldots \}$). Now, (11.1148) becomes

$$e_s \left( (s_i)_{i \in I} \right) = \sum_{(j_1, j_2, \ldots, j_s) \in \{1, 2, 3, \ldots \}^s; \ j_1 < j_2 < \cdots < j_s} s_{j(j_1)} s_{j(j_2)} \cdots s_{j(j_s)}$$

(by (11.1143)). Thus, (11.1145) is proven in Case 1.

Let us now consider Case 2. In this case, the set $I$ is finite. Let $J$ denote the totally ordered set $\{1, 2, \ldots, |I| \}$. Then, $J = \{1, 2, \ldots, |I| \} \subset \{1, 2, 3, \ldots \}$ and $\bigcup J = \{1, 2, 3, \ldots \} \setminus \{1, 2, 3, \ldots, |I| \} = \{1, 2, 3, \ldots \} \setminus \{1, 2, \ldots, |I| \} = \{|I| + 1, |I| + 2, |I| + 3, \ldots \}$.  

Fix a bijection $j : \{1, 2, \ldots, |I| \} \rightarrow I$ (such a bijection clearly exists, since $I$ is finite). Then, $e_s \left( (s_i)_{i \in I} \right)$ is the result of substituting $s_{j(1)}$, $s_{j(2)}$, ..., $s_{j(|I|)}$, 0, 0, 0, ... for the variables $x_1$, $x_2$, $x_3$, ... in $e_s$ (by the definition of $e_s \left( (s_i)_{i \in I} \right)$). Thus,

$$e_s \left( (s_i)_{i \in I} \right) = \left( \text{the result of substituting } s_{j(1)}, s_{j(2)}, \ldots, s_{j(|I|)}, 0, 0, 0, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \right)$$

(by (11.1147))

$$(11.1149)\quad = \sum_{(j_1, j_2, \ldots, j_s) \in \{1, 2, 3, \ldots \}^s; \ j_1 < j_2 < \cdots < j_s} \left( \text{the result of substituting } s_{j(1)}, s_{j(2)}, \ldots, s_{j(|I|)}, 0, 0, 0, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \right) x_{j_1} x_{j_2} \cdots x_{j_s}.$$
Thus,

\[ M_\alpha^{(s)} = e_s \left( \left( x_i^\alpha \right)_{i \in \text{SIS}(t)} \right) = \sum_{(i_1, i_2, \ldots, i_t) \in (\text{SIS}(t))^t; \ i_1 < i_2 < \cdots < i_t} x_i^\alpha x_i^\alpha \cdots x_i^\alpha. \]

This solves Exercise 6.98(b).

(c) Every \( i \in \text{SIS}(t) \) satisfies \( (x_i^\alpha)^n = x_i^{\alpha^{(n)}} \) (by (11.1139)), applied to \( n \) instead of \( s \). In other words, every \( i \in \text{SIS}(t) \) satisfies \( x_i^{\alpha^{(n)}} = (x_i^\alpha)^n \). Now, by the definition of \( M_\alpha^{(n)}(n) \), we have

\[ M_\alpha^{(n)} = e_s \left( \left( x_i^{\alpha^{(n)}} \right)_{i \in \text{SIS}(t)} \right) = e_s \left( \left( (x_i^\alpha)^n \right)_{i \in \text{SIS}(t)} \right). \]

We now notice that if \( R \) is a topological commutative \( k \)-algebra, if \( I \) is a countable set, and if \( (s_i)_{i \in I} \in R^I \) is a power-summable family of elements of \( R \), then

\[ e_s^{(n)} \left( (s_i)_{i \in I} \right) = e_s \left( (s_i^n)_{i \in I} \right). \]

Now, let us fix some \((j_1, j_2, \ldots, j_s) \in J^s\). Then, every \( k \in \{1, 2, \ldots, s\} \) satisfies \( j_k \in J = \{1, 2, \ldots, |I|\} \). Hence, for every \( k \in \{1, 2, \ldots, s\} \), the substitution of \( s_{(1)}(j_1), s_{(2)}(j_2), \ldots, s_{(|I|)}(j_s) \) for the variables \( x_1, x_2, x_3, \ldots \) transforms the indeterminate \( x_{j_k} \) into \( s_{(j_k)} \). Consequently, the substitution of \( s_{(1)}(j_1), s_{(2)}(j_2), \ldots, s_{(|I|)}(j_s) \) for the variables \( x_1, x_2, x_3, \ldots \) transforms the product \( x_{j_1}x_{j_2} \cdots x_{j_s} \) into \( s_{(j_1)}s_{(j_2)} \cdots s_{(j_s)} \). In other words,

\[ \begin{align*}
&\quad \left( \text{the result of substituting } s_{(1)}(j_1), s_{(2)}(j_2), \ldots, s_{(|I|)}(j_s) \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } x_{j_1}x_{j_2} \cdots x_{j_s} \right) \\
&= s_{(j_1)}s_{(j_2)} \cdots s_{(j_s)} \\
&= s_{(j_k)}. 
\end{align*} \]

Now, let us forget that we fixed \((j_1, j_2, \ldots, j_s) \in J^s\). We thus have shown that (11.1150) holds for every \((j_1, j_2, \ldots, j_s) \in J^s\).

On the other hand, let us fix some \((j_1, j_2, \ldots, j_s) \in \{1, 2, 3, \ldots\}^s \setminus J^s\). Since \((j_1, j_2, \ldots, j_s) \in \{1, 2, 3, \ldots\}^s \setminus J^s\), we must have \((j_1, j_2, \ldots, j_s) \in \{1, 2, 3, \ldots\}^s \setminus J^s\) but \((j_1, j_2, \ldots, j_s) \notin J^s\). Since \((j_1, j_2, \ldots, j_s) \notin J^s\), there exists some \( k \in \{1, 2, \ldots, s\} \) such that \( j_k \notin J \). Fix this \( k \). We have \( j_k \in \{1, 2, 3, \ldots\} \) but \( j_k \notin J \); therefore, \( j_k \in \{1, 2, 3, \ldots\} \setminus J = \{1\} + 1, |I| + 2, |I| + 3, \ldots. \) In other words, \( j_k > |I| \). Hence, the substitution of \( s_{(1)}(j_1), s_{(2)}(j_2), \ldots, s_{(|I|)}(j_s) \) for the variables \( x_1, x_2, x_3, \ldots \) transforms the indeterminate \( x_{j_k} \) into 0. In other words,

\[ \begin{align*}
&\quad \left( \text{the result of substituting } s_{(1)}(j_1), s_{(2)}(j_2), \ldots, s_{(|I|)}(j_s) \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } x_{j_k} \right) \\
&= 0.
\end{align*} \]

Now,

\[ \begin{align*}
&\quad \left( \text{the result of substituting } s_{(1)}(j_1), s_{(2)}(j_2), \ldots, s_{(|I|)}(j_s) \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } x_{j_1}x_{j_2} \cdots x_{j_s} \text{ instead of } s_{(j_1)}s_{(j_2)} \cdots s_{(j_s)} \right) \\
&= \prod_{z \in \{1, 2, \ldots, s\}} \left( \text{the result of substituting } s_{(1)}(j_1), s_{(2)}(j_2), \ldots, s_{(|I|)}(j_s) \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } x_{j_z} \right) \\
&= \left( \prod_{z \in \{1, 2, \ldots, s\}} \left( \text{the result of substituting } s_{(1)}(j_1), s_{(2)}(j_2), \ldots, s_{(|I|)}(j_s) \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } x_{j_z} \right) \\
&\quad \text{for } z \neq k \right) \\
&\quad \text{for } z = k \text{ from the product} \right) \\
&\quad \text{for } z = k \text{ from the product} \right) \\
&= 0.
\end{align*} \]

Now, let us forget that we fixed \((j_1, j_2, \ldots, j_s) \in \{1, 2, 3, \ldots\}^s \setminus J^s\). We thus have shown that (11.1151) holds for every \((j_1, j_2, \ldots, j_s) \in \{1, 2, 3, \ldots\}^s \setminus J^s\).

But \( J^s \subset \{1, 2, 3, \ldots\}^s \) (since \( J \subset \{1, 2, 3, \ldots\} \)). Hence, the set \( \{1, 2, 3, \ldots\}^s \) is the union of its two disjoint subsets \( J^s \) and \( \{1, 2, 3, \ldots\}^s \setminus J^s \). Now, (11.1149) becomes
Applying this to $R = k[[x]]$, $I = \text{SIS}(\ell)$ and $s_i = x_i^n$, we obtain

$$e_s(\delta^{(n)}_\ell I_{\text{SIS}(\ell)}) = e_s(\delta^{(n)}_\ell I_{\text{SIS}(\ell)}) = \left(\sum_{i_1 < i_2 < \cdots < i_s} x_{i_1}^{n_1} x_{i_2}^{n_2} \cdots x_{i_s}^{n_s}\right)^n.$$

Compared with (11.1154), this yields $M^{(s)}_{\alpha | \ell I} = e_s^{(n)}(\delta^{(n)}_\ell I_{\text{SIS}(\ell)})$. This solves Exercise 6.98(c).

---

We must be in one of the following two cases:

**Case 1.** In this case, the set $J$ is finite.

**Case 2.** But $J$ is a countable totally ordered set (since $J = \{1, 2, 3, \ldots, \ell\}$). Also, $j$ is a bijection $\{1, 2, 3, \ldots, \ell\} \to I$. In other words, $j$ is a bijection $J \to I$ (since $J = \{1, 2, 3, \ldots, \ell\}$). Now, (11.1153) becomes

$$e_s(\delta^{(n)}_\ell I_{\text{SIS}(\ell)}) = \sum_{i_1 < i_2 < \cdots < i_s} s_{i_1} s_{i_2} \cdots s_{i_s} \left(\sum_{i_1 < i_2 < \cdots < i_s} x_{i_1}^{n_1} x_{i_2}^{n_2} \cdots x_{i_s}^{n_s}\right)^n.$$

(by (11.1143)). Thus, (11.1145) is proven in Case 2.

We have thus proven (11.1145) in each of the Cases 1 and 2. Since these two Cases are the only cases that can occur, we thus conclude that (11.1154) holds.

---

**Proof of (11.1155):** Let $R$ be a topological commutative $k$-algebra. Let $I$ be a countable set. Let $(s_i)_{i \in I} \in R^I$ be a power-summable family of elements of $R$.

We must be in one of the following two cases:

**Case 1.** The set $I$ is infinite.

**Case 2.** The set $I$ is finite.

Let us first consider Case 1. In this case, the set $I$ is infinite, and therefore countably infinite (since it is countable). Fix a bijection $j : \{1, 2, 3, \ldots\} \to I$ (such a bijection clearly exists, since $I$ is countably infinite). Then, $e_s^{(n)}((s_i)_{i \in I})$ is the result of substituting $s_{j(1)}$, $s_{j(2)}$, $s_{j(3)}$, ... for the variables $x_1$, $x_2$, $x_3$, ... in $e_s^{(n)}$ (by the definition of $e_s^{(n)}((s_i)_{i \in I})$). Thus,

$$e_s^{(n)}((s_i)_{i \in I}) = \left(\sum_{i_1 < i_2 < \cdots < i_s} x_{i_1}^{n_1} x_{i_2}^{n_2} \cdots x_{i_s}^{n_s}\right)^n.$$

(by (11.1143)). Thus, (11.1145) is proven in Case 2.
(d) The first sentence of Proposition 2.28 yields that the family \((e_1, e_2, e_3, \ldots)\) generates the \(k\)-algebra \(\Lambda\). Thus, every element of \(\Lambda\) can be written as a polynomial in the elements \(e_1, e_2, e_3, \ldots\) Applying this to the element \(e_s^{(n)}\) of \(\Lambda\), we conclude that \(e_s^{(n)}\) can be written as a polynomial in the elements \(e_1, e_2, e_3, \ldots\) In other words, there exists a polynomial \(Q \in k[z_1, z_2, z_3, \ldots]\) such that \(e_s^{(n)} = Q(e_1, e_2, e_3, \ldots)\). Consider this polynomial \(Q\).

Let us now define a map \(\Phi : \Lambda \rightarrow k[[x]]\) by

\[
\Phi(f) = f \left( (x_i^n)_{i \in \text{SIS}(\ell)} \right)
\]

for every \(f \in \Lambda\)

(where \(f \left( (x_i^n)_{i \in \text{SIS}(\ell)} \right)\) is defined as in Definition 6.95(b)). This map \(\Phi\) is a \(k\)-algebra homomorphism (since it amounts to a substitution of certain elements for the variables in a power series). Therefore, it commutes with polynomials, i.e., it satisfies

\[
\Phi \left( R(f_1, f_2, f_3, \ldots) \right) = R(\Phi(f_1), \Phi(f_2), \Phi(f_3), \ldots)
\]

On the other hand, \(e_s \left( (s_i^n)_{i \in I} \right)\) is the result of substituting \(s_i^{(1)}, s_i^{(2)}, s_i^{(3)}, \ldots\) for the variables \(x_1, x_2, x_3, \ldots\) in \(e_s\) (by the definition of \(e_s \left( (s_i^n)_{i \in I} \right)\)). Thus, \(e_s \left( (s_i^n)_{i \in I} \right)\)

\[
= \left( \text{the result of substituting } s_i^{(1)}, s_i^{(2)}, s_i^{(3)}, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } e_s \right)
\]

\[
= e_s \left( \sum_{i_1 < i_2 < \cdots < i_s} s_i^{(1)} s_i^{(2)} \cdots s_i^{(n)} \right)
\]

Compared with (11.1156), this yields \(e_s^{(n)} \left( (s_i^n)_{i \in I} \right) = e_s \left( (s_i^n)_{i \in I} \right)\). Thus, (11.1155) is proven in Case 1.

Let us now consider Case 2. In this case, the set \(I\) is finite. Fix a bijection \(\iota: \{1, 2, \ldots, |I|\} \rightarrow I\) (such a bijection clearly exists, since \(I\) is finite).

Define an infinite sequence \((t_1, t_2, t_3, \ldots)\) of elements of \(R\) by

\[
(11.1157) \quad (t_1, t_2, t_3, \ldots) = (s_{\iota(1)}, s_{\iota(2)}, \ldots, s_{\iota(|I|)}, 0, 0, 0, \ldots)
\]

Then,

\[
(11.1158) \quad (t_1^n, t_2^n, t_3^n, \ldots) = \left( s_{\iota(1)}^{n}, s_{\iota(2)}^{n}, \ldots, s_{\iota(|I|)}^{n} \right)^T = \left( s_{\iota(1)}^{n}, s_{\iota(2)}^{n}, \ldots, s_{\iota(|I|)}^{n}, 0, 0, 0, \ldots \right)
\]

Recall that \(e_s^{(n)} \left( (s_i^n)_{i \in I} \right)\) is the result of substituting \(s_{\iota(1)}, s_{\iota(2)}, \ldots, s_{\iota(|I|)}, 0, 0, 0, \ldots\) for the variables \(x_1, x_2, x_3, \ldots\) in \(e_s^{(n)}\) (by the definition of \(e_s^{(n)} \left( (s_i^n)_{i \in I} \right)\)). Thus,

\[
e_s^{(n)} \left( (s_i^n)_{i \in I} \right) = \left( \text{the result of substituting } s_{\iota(1)}, s_{\iota(2)}, \ldots, s_{\iota(|I|)}, 0, 0, 0, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } e_s^{(n)} \right)
\]

\[
= \left( \text{the result of substituting } t_1, t_2, t_3, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } e_s^{(n)} \right)
\]

\[
= \sum_{i_1 < i_2 < \cdots < i_s} t_{i_1}^{n} t_{i_2}^{n} \cdots t_{i_s}^{n}
\]

(11.1159)
for every $f_1, f_2, f_3, \ldots \in \Lambda$ and every polynomial $R \in k[z_1, z_2, z_3, \ldots]$. Applying this to $f_i = e_i$ and $R = Q$, we obtain

\[(11.1160) \quad \Phi(Q(e_1, e_2, e_3, \ldots)) = Q(\Phi(e_1), \Phi(e_2), \Phi(e_3), \ldots).\]

But for every $j \in \{1, 2, 3, \ldots\}$, we have

\[\Phi(e_j) = e_j \left( (x_i^\alpha)_{i \in \text{SIS}(\ell)} \right) \quad \text{(by the definition of } \Phi(e_j) \text{)} \]
\[= M_{\alpha}^{(j)} \quad \text{ (since } M_{\alpha}^{(j)} = e_j \left( (x_i^\alpha)_{i \in \text{SIS}(\ell)} \right) \text{ (by the definition of } M_{\alpha}^{(j)} \text{)} \).

Thus, $(\Phi(e_1), \Phi(e_2), \Phi(e_3), \ldots) = (M_{\alpha}^{(1)}, M_{\alpha}^{(2)}, M_{\alpha}^{(3)}, \ldots)$, so that

\[Q(\Phi(e_1), \Phi(e_2), \Phi(e_3), \ldots) = Q(M_{\alpha}^{(1)}, M_{\alpha}^{(2)}, M_{\alpha}^{(3)}, \ldots).\]

Hence, $(11.1160)$ becomes

\[\Phi(Q(e_1, e_2, e_3, \ldots)) = Q(\Phi(e_1), \Phi(e_2), \Phi(e_3), \ldots) = Q(M_{\alpha}^{(1)}, M_{\alpha}^{(2)}, M_{\alpha}^{(3)}, \ldots).\]

Compared with

\[\Phi \left( \Phi \left( \frac{Q(e_1, e_2, e_3, \ldots)}{= e_3^{(n)} = \left( x_i^\alpha \right)_{i \in \text{SIS}(\ell)}} \right) \right) = \Phi \left( e_3^{(n)} \left( x_i^\alpha \right)_{i \in \text{SIS}(\ell)} \right) \quad \text{(by the definition of } \Phi(e_3^{(n)}) \text{)} \]
\[= M_{\alpha}^{(n)} \quad \text{(by Exercise 6.98(c))},\]

this yields $M_{\alpha}^{(n)} = Q(M_{\alpha}^{(1)}, M_{\alpha}^{(2)}, M_{\alpha}^{(3)}, \ldots)$. Thus, there exists a polynomial $P \in k[z_1, z_2, z_3, \ldots]$ such that $M_{\alpha}^{(n)} = P(M_{\alpha}^{(1)}, M_{\alpha}^{(2)}, M_{\alpha}^{(3)}, \ldots)$ (namely, $P = Q$). This solves Exercise 6.98(d).

On the other hand, $e_s \left( \left( s_i^n \right)_{i \in \mathbb{I}} \right)$ is the result of substituting $s_1^n, s_2^n, \ldots, s_{|\mathbb{I}|}^n$, 0, 0, 0, \ldots for the variables $x_1, x_2, x_3, \ldots$ in $e_s$ (by the definition of $e_s \left( \left( s_i^n \right)_{i \in \mathbb{I}} \right)$). Thus,

\[e_s \left( \left( s_i^n \right)_{i \in \mathbb{I}} \right) = \begin{cases} \text{the result of substituting } s_1^n, s_2^n, \ldots, s_{|\mathbb{I}|}^n, 0, 0, 0, \ldots \text{ for the variables } x_1, x_2, x_3, \ldots \text{ in } e_s \\ = \sum_{t_1 < t_2 < \cdots < t_s} x_{t_1} x_{t_2} \cdots x_{t_s} \text{ (by (11.1159))} \end{cases} = \sum_{t_1 < t_2 < \cdots < t_s} t_1^n t_2^n \cdots t_s^n.

Compared with (11.1159), this yields $e_s^{(n)} \left( \left( s_i^n \right)_{i \in \mathbb{I}} \right) = e_s \left( \left( s_i^n \right)_{i \in \mathbb{I}} \right)$. Thus, (11.1155) is proven in Case 2.

We have thus proven (11.1155) in each of the Cases 1 and 2. Since these two Cases are the only cases that can occur, we thus conclude that (11.1155) holds, qed.
11.142. **Solution to Exercise 6.99.** Solution to Exercise 6.99. The definition of \( M^{(s)} \) yields \( M^{(s)} = e_s \left( (x_i^{(1)})_{i \in \text{SIS}(1)} \right) \).

But \( \text{SIS}(1) \) is defined as the set of all strictly increasing 1-tuples of positive integers. Clearly, such 1-tuples are in bijection with positive integers; the bijection sends a positive integer \( i \) to the strictly increasing 1-tuple \( (i) \). This bijection shows that the family \( \left( (x_i^{(1)})_{i \in \text{SIS}(1)} \right) \) is a reparametrization of the family \( \left( (x_{i(i)})_{i \in \{1,2,3,\ldots\}} \right) \).

Thus, \( e_s \left( (x_i^{(1)})_{i \in \text{SIS}(1)} \right) = e_s \left( (x_{i(i)})_{i \in \{1,2,3,\ldots\}} \right) \). Thus,

\[
M^{(s)}_{(1)} = e_s \left( (x_i^{(1)})_{i \in \text{SIS}(1)} \right) = e_s \left( \begin{pmatrix} x_{1}^{(1)} \\ \vdots \\ x_{i}^{(1)} \end{pmatrix} \right) \quad \text{(by the definition of } x_{i}^{(1)} \text{)} \quad i \in \{1,2,3,\ldots\} \\
= e_s \left( (x_i)_{i \in \{1,2,3,\ldots\}} \right) = e_s.
\]

This solves Exercise 6.99.


**Proof of Proposition 6.100.** We first recall a general fact from algebra. Namely, if \( \mathcal{C} \) and \( \mathcal{D} \) are two commutative rings, if \( \varphi : \mathcal{C} \to \mathcal{D} \) is a ring homomorphism, and if \( u \) and \( v \) are two nonnegative integers, then we can define a homomorphism \( \varphi^{u \times v} : \mathcal{C}^{u \times v} \to \mathcal{D}^{u \times v} \) of additive groups by sending every matrix \( (c_{i,j})_{i=1,2,\ldots,u; \ j=1,2,\ldots,v} \in \mathcal{C}^{u \times v} \) to the matrix \( (\varphi(c_{i,j}))_{i=1,2,\ldots,u; \ j=1,2,\ldots,v} \in \mathcal{D}^{u \times v} \). This homomorphism \( \varphi^{u \times v} \) is the map from \( \mathcal{C}^{u \times v} \) to \( \mathcal{D}^{u \times v} \) canonically induced by \( \varphi \), and it has many structure-preserving properties (for instance, it respects the multiplication of matrices, in the sense that we have \( \varphi^{u \times v} (X) \cdot \varphi^{u \times v} (Y) = \varphi^{u \times v} (XY) \) whenever \( X \in \mathcal{C}^{u \times v} \) and \( Y \in \mathcal{C}^{v \times w} \)). We furthermore have

\[
(11.1161) \quad \det \left( \varphi^{u \times u} (X) \right) \cdot \det \left( \varphi^{v \times v} (Y) \right) = \det \left( \varphi^{u \times v} (XY) \right)
\]

for any two commutative rings \( \mathcal{C} \) and \( \mathcal{D} \), any ring homomorphism \( \varphi : \mathcal{C} \to \mathcal{D} \), any nonnegative integer \( u \) and any matrix \( X \in \mathcal{C}^{u \times u} \). (This is because the determinant of a matrix is a polynomial in its entries, and polynomials commute with ring homomorphisms.)

We shall now construct a particular \( \mathbf{k} \)-algebra homomorphism \( \Lambda \to \mathbf{k}[[x]] \) which will help us in our proof. Namely, we define a map \( \Phi : \Lambda \to \mathbf{k}[[x]] \) by

\[
\Phi(f) = f \left( (x_i^{(1)})_{i \in \text{SIS}(1)} \right) \quad \text{for every } f \in \Lambda
\]

(where \( f \left( (x_i^{(1)})_{i \in \text{SIS}(1)} \right) \) is defined as in Definition 6.95(b)). This map \( \Phi \) is a \( \mathbf{k} \)-algebra homomorphism (since it amounts to a substitution of certain elements for the variables in a power series). Every \( s \in \mathbb{N} \) satisfies

\[
(11.1162) \quad \Phi(e_s) = e_s \left( (x_i^{(1)})_{i \in \text{SIS}(1)} \right) \quad \text{(by the definition of } \Phi(e_s) \text{)}
\]

\[
(11.1163) \quad \Phi(p_s) = p_s \left( (x_i^{(1)})_{i \in \text{SIS}(1)} \right) \quad \text{(by the definition of } \Phi(p_s) \text{)}
\]

Every positive integer \( s \) satisfies

\[
\Phi(p_s) = p_s \left( (x_i^{(1)})_{i \in \text{SIS}(1)} \right) \quad \text{(by Exercise 6.98(a))}
\]
(a) Define a matrix $A_n = (a_{i,j})_{i,j=1,2,...,n}$ as in Exercise 2.89(a). Then, Exercise 2.89(a) yields $\det(A_n) = n!e_n$. Applying the map $\Phi$ to both sides of this equality, we obtain

$$\Phi(\det(A_n)) = \Phi(n!e_n) = n!$$

(by (11.1162), applied to $s=n$)

But (11.1161) (applied to $C=\Lambda$, $D=k[[x]]$, $\varphi=\Phi$, $u=n$ and $X=A_n$) yields

$$\det(\Phi^{n\times n}(A_n)) = \Phi(\det(A_n)) = n!M^{(n)}_\alpha.$$  

But

$$\Phi^{n\times n}(A_n) = \Phi^{n\times n}(a_{i,j})_{i,j=1,2,...,n} = \Phi(\Phi(\Phi(0))) = \Phi(\Phi(\Phi(0))) = \Phi(\Phi(\Phi(0)))$$

(by the definition of $\Phi^{n\times n}(a_{i,j})_{i,j=1,2,...,n}$). But every $(i,j) \in \{1,2,\ldots,n\}^2$ satisfies

$$\Phi(a_{i,j}) = \Phi\left( \begin{cases} p_{i-j+1}, & \text{if } i \geq j; \\ i, & \text{if } i = j - 1; \\ 0, & \text{if } i < j - 1 \end{cases} \right)$$

$$= \left( \begin{array}{c} \Phi(p_{i-j+1}), \\
\Phi(i), \\
\Phi(0), \end{array} \right)$$

$$= \left( \begin{array}{c} M_{\alpha(i-j+1)}, \\
on, \\
0, \end{array} \right)$$

$$= a^{(\alpha)}_{i,j}$$

Hence, $\Phi(a_{i,j}) = a^{(\alpha)}_{i,j}$. Thus, (11.1165) becomes $\Phi^{n\times n}(A_n) = (\Phi(a_{i,j}))_{i,j=1,2,...,n} = A_n^{(\alpha)}$. Thus, (11.1164) becomes $\det\left(A_n^{(\alpha)}\right) = n!M^{(n)}_\alpha$. This proves Proposition 6.100(a).

(b) The proof of Proposition 6.100(b) proceeds similarly to our proof of Proposition 6.100(a) above, as long as the obvious changes are made (one needs to consider $B_n$, $b_{i,j}$, $B_n^{(\alpha)}$ and $b^{(\alpha)}_{i,j}$ instead of $A_n$, $a_{i,j}$, $A_n^{(\alpha)}$ and $a^{(\alpha)}_{i,j}$).

\[\square\]
11.144. Solution to Exercise 6.103. Solution to Exercise 6.103.

Proof of Corollary 6.102. (a) We WLOG assume that \( s \in \mathbb{N} \) (because otherwise, \( s \) is negative and thus satisfies \( M_\alpha^{(s)} = 0 \in \text{QSym} \)).

Define a matrix \( A_\alpha^{(s)} = \left( a_{i,j}^{(\alpha)} \right)_{i,j=1,2,\ldots,s} \) as in Proposition 6.100(a) (but for \( s \) instead of \( n \)). Then, \( A_\alpha^{(s)} \in \text{QSym}^{s \times s} \) (since every \( (i,j) \in \{1,2,\ldots,s\}^2 \) satisfies \( a_{i,j}^{(\alpha)} \in \text{QSym} \) (as follows from the definition of \( a_{i,j}^{(\alpha)} \))). Hence, \( \det (A_\alpha^{(s)}) \in \text{QSym} \). But Proposition 6.100(a) (applied to \( n = s \)) yields \( \det (A_\alpha^{(s)}) = s!M_\alpha^{(s)} \).

Hence, \( s!M_\alpha^{(s)} = \det (A_\alpha^{(s)}) \in \text{QSym} \).

Now, let us recall that there is a canonical ring homomorphism \( \varphi : \mathbb{Z} \to k \). This homomorphism gives rise to a ring homomorphism \( \varphi [[x]] : \mathbb{Z}[[x]] \to k[[x]] \), and this latter homomorphism \( \varphi [[x]] \) sends \( \text{QSym}_\mathbb{Z} \) to \( \text{QSym}_k \); that is, we have \( (\varphi [[x]])(\text{QSym}_\mathbb{Z}) \subset \text{QSym}_k \). Moreover, it is clear that the ring homomorphism \( \varphi [[x]] \) sends the element \( M_\alpha^{(s)} \) of \( \mathbb{Z}[[x]] \) to the element \( M_\alpha^{(s)} \) of \( k[[x]] \) (because the definition of \( M_\alpha^{(s)} \) is functorial in the base ring \( k \)). Therefore, if we can prove that the element \( M_\alpha^{(s)} \) of \( \mathbb{Z}[[x]] \) belongs to \( \text{QSym}_\mathbb{Z} \), then the element \( M_\alpha^{(s)} \) of \( k[[x]] \) belongs to \( (\varphi [[x]])(\text{QSym}_\mathbb{Z}) \subset \text{QSym}_k \); this will complete the proof of Corollary 6.102(a). Hence, in order to prove Corollary 6.102(a), it only remains to prove that the element \( M_\alpha^{(s)} \) of \( \mathbb{Z}[[x]] \) belongs to \( \text{QSym}_\mathbb{Z} \). In other words, it only remains to prove Corollary 6.102(a) in the case of \( k = \mathbb{Z} \). Hence, in proving Corollary 6.102(a), we can WLOG assume that \( k = \mathbb{Z} \). Assume this. Since \( k = \mathbb{Z} \), we have \( \text{QSym} = \text{QSym}_\mathbb{Z} \).

If a positive integer \( N \) and an element \( f \) of \( \mathbb{Z}[[x]] \) satisfy \( Nf \in \text{QSym}_\mathbb{Z} \), then \( f \) also lies in \( \text{QSym}_\mathbb{Z} \) (because \( N \) is not a zero-divisor in \( k = \mathbb{Z} \), and therefore \( f \) is obtained from the power series \( Nf \) by dividing all coefficients by \( N \)). Applying this to \( N = s! \) and \( f = M_\alpha^{(s)} \), we obtain \( M_\alpha^{(s)} \in \text{QSym}_\mathbb{Z} \) (since \( s!M_\alpha^{(s)} = M_\alpha^{(s)} \in \text{QSym} = \text{QSym}_\mathbb{Z} \)). In other words, \( M_\alpha^{(s)} \in \text{QSym} \) (since \( k = \mathbb{Z} \)). This completes the proof of Corollary 6.102(a).

(b) Recall that \( M_\alpha^{(s)} = e_s(\binom{x^\alpha}{\lambda})_{\ell \in \text{SIS}(\ell)} \). Thus, \( M_\alpha^{(s)} \) is a homogeneous power series of degree \( s|\alpha| \) (since each \( x_i^\alpha \) is a monomial of degree \( |\alpha| \), and since \( e_s \) is a power series of degree \( s \)). Combined with \( M_\alpha^{(s)} \in \text{QSym} \) (which follows from Corollary 6.102(a)), this yields \( M_\alpha^{(s)} \in \text{QSym}_{s|\alpha|} \). This proves Corollary 6.102(b).


Proof of Remark 6.105. (a) Write the composition \( \alpha \) in the form \( (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \). Then,

\[
\text{red} \alpha = \left( \frac{\alpha_1}{\gcd \alpha}, \frac{\alpha_2}{\gcd \alpha}, \ldots, \frac{\alpha_\ell}{\gcd \alpha} \right)
\]

(by the definition of \( \text{red} \alpha \)). Hence, the definition of \( \text{red} \alpha \{\gcd \alpha\} \)

\[
\text{red} \alpha \{\gcd \alpha\} = \left( \gcd \alpha, \frac{\alpha_1}{\gcd \alpha}, \frac{\alpha_2}{\gcd \alpha}, \ldots, \frac{\alpha_\ell}{\gcd \alpha} \right) = (\alpha_1, \alpha_2, \ldots, \alpha_\ell) = \alpha,
\]

so that Remark 6.105(a) is proven.

(c) Write the composition \( \alpha \) in the form \( (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \). Then,

\[
\text{red} \alpha = \left( \frac{\alpha_1}{\gcd \alpha}, \frac{\alpha_2}{\gcd \alpha}, \ldots, \frac{\alpha_\ell}{\gcd \alpha} \right)
\]

(by the definition of \( \text{red} \alpha \)). Hence, the definition of \( \gcd (\text{red} \alpha) \) yields

\[
\gcd (\text{red} \alpha) = \gcd \left( \frac{\alpha_1}{\gcd \alpha}, \frac{\alpha_2}{\gcd \alpha}, \ldots, \frac{\alpha_\ell}{\gcd \alpha} \right) = \gcd \left( \frac{\alpha_1, \alpha_2, \ldots, \alpha_\ell}{\gcd \alpha} \right) = \gcd \left( \frac{\alpha_1, \alpha_2, \ldots, \alpha_\ell}{\gcd \alpha} \right)
\]

(since \( \gcd \alpha = \gcd (\alpha_1, \alpha_2, \ldots, \alpha_\ell) \) (by the definition of \( \gcd \alpha \))

\[
= 1.
\]

In other words, the composition \( \text{red} \alpha \) is reduced. This proves Remark 6.105(c).
\[(d)\] Assume that \(\alpha\) is reduced. Write the composition \(\alpha\) in the form \((\alpha_1, \alpha_2, \ldots, \alpha_\ell)\). Then, \(\text{red } \alpha = \left(\frac{\alpha_1}{\gcd \alpha}, \frac{\alpha_2}{\gcd \alpha}, \ldots, \frac{\alpha_\ell}{\gcd \alpha}\right)\) (by the definition of \(\text{red } \alpha\)). But \(\gcd \alpha = 1\) (since \(\alpha\) is reduced). Now,

\[
\text{red } \alpha = \left(\frac{\alpha_1}{\gcd \alpha}, \frac{\alpha_2}{\gcd \alpha}, \ldots, \frac{\alpha_\ell}{\gcd \alpha}\right) = \left(\frac{\alpha_1}{1}, \frac{\alpha_2}{1}, \ldots, \frac{\alpha_\ell}{1}\right) \quad \text{(since } \gcd \alpha = 1)\]

\((\alpha_1, \alpha_2, \ldots, \alpha_\ell) = \alpha\).

This proves Remark 6.105(d).

\[(c)\] Let \(s \in \{1, 2, 3, \ldots\}\). Write the composition \(\alpha\) in the form \((\alpha_1, \alpha_2, \ldots, \alpha_\ell)\). Then,

\[
\text{red } \alpha = \left(\frac{\alpha_1}{\gcd \alpha}, \frac{\alpha_2}{\gcd \alpha}, \ldots, \frac{\alpha_\ell}{\gcd \alpha}\right) \quad \text{(by the definition of \(\text{red } \alpha\)). Also, } \alpha \{s\} = (s\alpha_1, s\alpha_2, \ldots, s\alpha_\ell) \quad \text{(by the definition of } \alpha \{s\})\).

Now, \(\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell)\), so that \(\ell(\alpha) = \ell\). Hence, \(\ell = \ell(\alpha) > 0\) (since \(\alpha\) is nonempty). Since \(\alpha \{s\} = (s\alpha_1, s\alpha_2, \ldots, s\alpha_\ell)\), we have \(\ell(\alpha \{s\}) = \ell > 0\), so that the composition \(\alpha \{s\}\) is nonempty.

By the definition of \(\gcd \alpha\), we have \(\gcd \alpha = \gcd (\alpha_1, \alpha_2, \ldots, \alpha_\ell)\). By the definition of \(\gcd (\alpha \{s\})\), we have

\[
\gcd (\alpha \{s\}) = \gcd (s\alpha_1, s\alpha_2, \ldots, s\alpha_\ell) \quad \text{(since } \alpha \{s\} = (s\alpha_1, s\alpha_2, \ldots, s\alpha_\ell))
\]

\[
= s(\gcd (\alpha_1, \alpha_2, \ldots, \alpha_\ell)) = s\gcd \alpha.
\]

Now, recall that \(\alpha \{s\} = (s\alpha_1, s\alpha_2, \ldots, s\alpha_\ell)\). Hence, the definition of \(\text{red } (\alpha \{s\})\) yields

\[
\text{red } (\alpha \{s\}) = \left(\frac{s\alpha_1}{\gcd (\alpha \{s\})}, \frac{s\alpha_2}{\gcd (\alpha \{s\})}, \ldots, \frac{s\alpha_\ell}{\gcd (\alpha \{s\})}\right)
\]

\[
= \left(\frac{s\alpha_1}{s\gcd \alpha}, \frac{s\alpha_2}{s\gcd \alpha}, \ldots, \frac{s\alpha_\ell}{s\gcd \alpha}\right) \quad \text{(since } \gcd (\alpha \{s\}) = s\gcd \alpha)\)

\[
= \left(\frac{\alpha_1}{\gcd \alpha}, \frac{\alpha_2}{\gcd \alpha}, \ldots, \frac{\alpha_\ell}{\gcd \alpha}\right) = \text{red } \alpha.
\]

This proves Remark 6.105(e).

\[(f)\] Write the composition \(\alpha\) in the form \((\alpha_1, \alpha_2, \ldots, \alpha_\ell)\). Then, \(\text{red } \alpha = \left(\frac{\alpha_1}{\gcd \alpha}, \frac{\alpha_2}{\gcd \alpha}, \ldots, \frac{\alpha_\ell}{\gcd \alpha}\right) \quad \text{(by the definition of } \text{red } \alpha)\).

Thus,

\[
|\text{red } \alpha| = \frac{\alpha_1}{\gcd \alpha} + \frac{\alpha_2}{\gcd \alpha} + \cdots + \frac{\alpha_\ell}{\gcd \alpha} = \frac{\alpha_1 + \alpha_2 + \cdots + \alpha_\ell}{\gcd \alpha},
\]

so that \((\gcd \alpha)|\text{red } \alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_\ell = |\alpha|\) (since \(|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_\ell\) (because \(\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell))\)).

This proves Remark 6.105(f).

\[(b)\] We have \((\text{red } \alpha) \{\gcd \alpha\} = \alpha\) (by Remark 6.105(a)).

The property of a composition to be Lyndon does not change if all entries of the composition are multiplied by a fixed positive integer \(m\) (because this property only depends on the relative order of the parts of the composition). In other words, if \(\beta\) is a composition and \(m\) is a positive integer, then \(\beta\) is Lyndon if and only if \(\beta \{m\}\) is Lyndon. Applying this to \(\beta = \text{red } \alpha\) and \(m = \gcd \alpha\), we conclude that \(\text{red } \alpha\) is Lyndon if and only if \((\text{red } \alpha) \{\gcd \alpha\}\) is Lyndon. In other words, \(\text{red } \alpha\) is Lyndon if and only if \(\alpha\) is Lyndon (because \((\text{red } \alpha) \{\gcd \alpha\} = \alpha\)). This proves Remark 6.105(b). \(\Box\)
Proof of Lemma 6.108. For every $\alpha \in \mathcal{L}$, the pair $(\text{red } \alpha, \gcd \alpha)$ is a well-defined element of $\mathcal{RL} \times \{1, 2, 3, \ldots\}$ $^878$. Hence, we can define a map $\mathcal{R} : \mathcal{L} \to \mathcal{RL} \times \{1, 2, 3, \ldots\}$ by

$$(\mathcal{R} (\alpha) = (\text{red } \alpha, \gcd \alpha) \quad \text{for every } \alpha \in \mathcal{L}).$$

Consider this $\mathcal{R}$. For every $(w, s) \in \mathcal{RL} \times \{1, 2, 3, \ldots\}$, the element $w\{s\}$ is a well-defined element of $\mathcal{L}$ $^879$. Hence, we can define a map $\mathcal{M} : \mathcal{RL} \times \{1, 2, 3, \ldots\} \to \mathcal{L}$ by

$$(\mathcal{M} (w, s) = w\{s\} \quad \text{for every } (w, s) \in \mathcal{RL} \times \{1, 2, 3, \ldots\}).$$

Consider this $\mathcal{M}$. It is now easy to see that $\mathcal{R} \circ \mathcal{M} = \text{id}$ $^880$ and $\mathcal{M} \circ \mathcal{R} = \text{id}$ $^881$. Therefore, the maps $\mathcal{M}$ and $\mathcal{R}$ are mutually inverse. Hence, the map $\mathcal{M}$ is a bijection. Hence, the family $\left(\mathcal{M}^{(\gcd (\mathcal{M}(w, s)))}_{\text{red } \alpha} \right)_{(w, s) \in \mathcal{RL} \times \{1, 2, 3, \ldots\}}$ is a reindexing of the family $\left(\mathcal{M}^{(\gcd (\alpha))}_{\text{red } \alpha} \right)_{\alpha \in \mathcal{L}}$. Since every $(w, s) \in \mathcal{RL} \times \{1, 2, 3, \ldots\}$ satisfies

$$\gcd (\mathcal{M}(w, s)) = s \quad \text{and} \quad \text{red} (\mathcal{M}(w, s)) = w$$

Proof. Let $\alpha \in \mathcal{L}$. Then, $\alpha$ is a Lyndon word (since $\mathcal{L}$ is the set of all Lyndon words), and thus nonempty. Hence, $\alpha$ is a nonempty composition, so that $\text{red } \alpha$ is a nonempty composition, and $\gcd \alpha$ is a well-defined positive integer. Remark 6.105(b) yields that the composition $\alpha$ is Lyndon if and only if the composition $\text{red } \alpha$ is Lyndon. Since $\alpha$ is Lyndon, this yields that $\text{red } \alpha$ is Lyndon. Also, $\alpha$ is reduced (by Remark 6.105(c)). Thus, $\alpha$ is a reduced Lyndon composition. In other words, $\text{red } \alpha \in \mathcal{RL}$ (since $\mathcal{RL}$ is the set of all reduced Lyndon compositions). Combined with $\gcd \alpha \in \{1, 2, 3, \ldots\}$ (since $\gcd \alpha$ is a well-defined positive integer), this yields (red $\alpha, \gcd \alpha) \in \mathcal{RL} \times \{1, 2, 3, \ldots\}$, qed.

Proof. Let $(w, s) \in \mathcal{RL} \times \{1, 2, 3, \ldots\}$. Then, $w \in \mathcal{RL}$ and $s \in \{1, 2, 3, \ldots\}$. Since $\mathcal{RL}$ is the set of all reduced Lyndon compositions, we see that $w$ is a reduced Lyndon composition (since $w \in \mathcal{RL}$). Thus, $w$ is nonempty (since $w$ is Lyndon).

We shall now show that $w\{s\} \in \mathcal{L}$.

Remark 6.105(e) (applied to $\alpha = w$) yields that the composition $w\{s\}$ is nonempty and satisfies $\text{red} (w\{s\}) = \text{red } w$ and $\gcd (w\{s\}) = s \gcd w$. We have $\text{red} (w\{s\}) = \text{red } w = w$ (by Remark 6.105(d), applied to $\alpha = w$). Now, recall that the composition $w\{s\}$ is Lyndon. Hence, the composition $\text{red} (w\{s\})$ is Lyndon (since $\text{red} (w\{s\}) = w$).

Remark 6.105(b) (applied to $\alpha = w\{s\}$) yields that the composition $w\{s\}$ is Lyndon if and only if the composition $\text{red} (w\{s\})$ is Lyndon. Since the composition $\text{red} (w\{s\})$ is Lyndon, this yields that the composition $w\{s\}$ is Lyndon. In other words, $w\{s\} \in \mathcal{L}$ (since $\mathcal{L}$ is the set of all Lyndon words), qed.

Proof. Let $(w, s) \in \mathcal{RL} \times \{1, 2, 3, \ldots\}$. Thus, $w \in \mathcal{RL}$ and $s \in \{1, 2, 3, \ldots\}$. Since $\mathcal{RL}$ is the set of all reduced Lyndon compositions, we see that $w$ is a reduced Lyndon composition (since $w \in \mathcal{RL}$). Thus, $w$ is nonempty (since $w$ is Lyndon).

Remark 6.105(e) (applied to $\alpha = w$) yields that the composition $w\{s\}$ is nonempty and satisfies $\text{red} (w\{s\}) = \text{red } w$ and $\gcd (w\{s\}) = s \gcd w$. But $\gcd w = 1$ (since $w$ is reduced), so that $\gcd (w\{s\}) = s \gcd w = s$. Also, $\text{red} (w\{s\}) = \text{red } w = w$ (by Remark 6.105(d), applied to $\alpha = w$). Now,

$$(\mathcal{R} \circ \mathcal{M})(w, s) = \mathcal{R} \left( \frac{\mathcal{M}(w, s)}{w\{s\}} \right) = \mathcal{R} \left( \frac{w\{s\}}{w\{s\}} \right) = \mathcal{R} \left( \frac{w\{s\}}{w\{s\}} \right) = (w, s).$$

Now, let us forget that we fixed $(w, s)$. We thus have shown that $(\mathcal{R} \circ \mathcal{M})(w, s) = \text{id} (w, s)$ for every $(w, s) \in \mathcal{RL} \times \{1, 2, 3, \ldots\}$. In other words, $\mathcal{R} \circ \mathcal{M} = \text{id}$, qed.

Proof. Let $\alpha \in \mathcal{L}$. Then, $\alpha$ is a Lyndon word (since $\mathcal{L}$ is the set of all Lyndon words), and thus nonempty. Hence, $\alpha$ is a nonempty composition. Now,

$$(\mathcal{M} \circ \mathcal{R})(\alpha) = \mathcal{M} \left( \frac{\mathcal{R}(\alpha)}{\text{red } \alpha, \gcd \alpha} \right) = \mathcal{M}(\text{red } \alpha, \gcd \alpha) = (\text{red } \alpha) \{\gcd \alpha\} \quad \text{by the definition of } \mathcal{M}(\text{red } \alpha, \gcd \alpha)$$

(by the definition of $\mathcal{M}(\alpha)$)

$$(\text{by Remark 6.105(a)})$$

$$(\text{by the definition of } \mathcal{M}(\alpha))$$

$$(\text{id } (\alpha)).$$

Now, let us forget that we fixed $\alpha$. We thus have proven that $(\mathcal{M} \circ \mathcal{R})(\alpha) = \text{id } (\alpha)$ for every $\alpha \in \mathcal{L}$. Thus, $\mathcal{M} \circ \mathcal{R} = \text{id}$, qed.
of this. Similarly, assume that there exists a \( \psi \) satisfying \( \ell(\psi) \leq \delta \). In other words, \((11.1167)\) holds for \( \alpha \in \mathbb{C} \).

This proves Lemma 6.108.

11.147. **Solution to Exercise 6.111.**  **Solution to Exercise 6.111.** We shall give two proofs of Lemma 6.110.

**First proof of Lemma 6.110.** We first introduce some notation.

For every \( m \in \mathbb{Z} \), let \( \mathcal{F}_m \) denote the \( k \)-submodule of \( \text{QSym} \) spanned by \((M_\beta)_{\beta \in \text{Comp}; \; \ell(\beta) \leq m}\). Then, \( 1 \in \mathcal{F}_0 \).

Furthermore,

\[
\mathcal{F}_u \mathcal{F}_v \subset \mathcal{F}_{u+v} \quad \text{for every } u \in \mathbb{Z} \text{ and } v \in \mathbb{Z}
\]

It is now easy to see that

\[
(\mathcal{F}_m)^k \subset \mathcal{F}_{km} \quad \text{for every } m \in \mathbb{Z} \text{ and } k \in \mathbb{N}
\]

(where \((\mathcal{F}_m)^k\) means \(\underbrace{\mathcal{F}_m \mathcal{F}_m \cdots \mathcal{F}_m}_{\text{k times}}\)).

Write the composition \( \alpha \) in the form \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_t) \); then, \( \ell = \ell(\alpha) \).

**Proof.** We have \( R \circ M = \text{id} \), thus \( R \circ M \equiv (w, s) = \text{id}(w, s) = (w, s) \) and therefore

\[
(w, s) = (R \circ M \equiv (w, s) - \text{id}) \equiv (R \circ M \equiv (w, s)) = \text{id}(M(w, s)) = (\text{id}(M(w, s)) \circ \text{id}(M(w, s))) \quad \text{(by the definition of } R \circ M \text{)}.
\]

Hence, \( (\text{red}(M(w, s)), \text{gcd}(M(w, s))) = (w, s) \). In other words, \( \text{red}(M(w, s)) = w \) and \( \text{gcd}(M(w, s)) = s \), qed.

**Proof of (11.1166):** Let \( u \in \mathbb{Z} \) and \( v \in \mathbb{Z} \). Let \( p \in \mathcal{F}_u \) and \( q \in \mathcal{F}_v \). We are going to prove that \( pq \in \mathcal{F}_{u+v} \).

We know that \( p \) belongs to \( \mathcal{F}_u \). In other words, \( p \) is a \( k \)-linear combination of the elements \((M_\beta)_{\beta \in \text{Comp}; \; \ell(\beta) \leq u}\) (since \( \mathcal{F}_u \) is the submodule of \( \text{QSym} \) spanned by \((M_\beta)_{\beta \in \text{Comp}; \; \ell(\beta) \leq u}\)). Hence, we can WLOG assume that \( p \) is one of the elements \((M_\beta)_{\beta \in \text{Comp}; \; \ell(\beta) \leq u}\) (because the claim that we are proving – namely, the claim that \( pq \in \mathcal{F}_{u+v} \) – is \( k \)-linear in \( p \)). Assume this. Similarly, assume that \( q \) is one of the elements \((M_\beta)_{\beta \in \text{Comp}; \; \ell(\beta) \leq v}\).

There exists a \( \varphi \in \text{Comp} \) satisfying \( \ell(\varphi) \leq u \) and \( p = M_\varphi \) (since \( p \) is one of the elements \((M_\beta)_{\beta \in \text{Comp}; \; \ell(\beta) \leq u}\)). Similarly, there exists a \( \psi \in \text{Comp} \) satisfying \( \ell(\psi) \leq v \) and \( q = M_\psi \). Consider these \( \varphi \) and \( \psi \). Corollary 6.86 (applied to \( \varphi \) and \( \psi \) instead of \( \alpha \) and \( \beta \)) shows that \( M_\varphi M_\psi \) is a sum of terms of the form \( M_\delta \) with \( \delta \in 2^* \) satisfying \( \ell(\delta) \leq u + v \). In particular, \( M_\varphi M_\psi \) is a \( k \)-linear combination of \((M_\delta)_{\delta \in 2^*; \; \ell(\delta) \leq u+v}\).

In other words, \( M_\varphi M_\psi \) is a \( k \)-linear combination of \((M_\delta)_{\delta \in \text{Comp}; \; \ell(\delta) \leq u+v}\) (since \( 2^* = \text{Comp} \)). In other words, \( M_\varphi M_\psi \) is a \( k \)-linear combination of \((M_\delta)_{\delta \in \text{Comp}; \; \ell(\delta) \leq u+v}\) (here, we renamed the index \( \delta \) as \( \beta \)). In other words, \( M_\varphi M_\psi \) belongs to the \text{k-submodule of QSym spanned by} \((M_\beta)_{\beta \in \text{Comp}; \; \ell(\beta) \leq u+v}\). In other words, \( M_\varphi M_\psi \) belongs to \( \mathcal{F}_{u+v} \) (since \( \mathcal{F}_{u+v} \) was defined as the \text{k-submodule of QSym spanned by} \((M_\beta)_{\beta \in \text{Comp}; \; \ell(\beta) \leq u+v}\)). Hence, \( M_\varphi M_\psi \in \mathcal{F}_{u+v} \), so that \( pq = M_\varphi M_\psi \in \mathcal{F}_{u+v} \).

**Proof of (11.1167):** Let \( m \in \mathbb{Z} \). We shall prove that \((11.1167)\) holds for every \( k \in \mathbb{N} \). Indeed, we shall prove this by induction over \( k \).

**Induction base:** We have

\[
(\mathcal{F}_m)^0 = k \cdot \mathcal{F}_0 \subset \mathcal{F}_0 \quad \text{(since } \mathcal{F}_0 \text{ is a } k\text{-module)}
\]

In other words, \((11.1167)\) holds for \( k = 0 \). This completes the induction base.

**Induction step:** Let \( k \in \mathbb{N} \). Assume that \((11.1167)\) holds for \( k = K \). We must then show that \((11.1167)\) holds for \( k = K+1 \).
Define $A_s^{(\alpha)}$ and $a_{i,j}^{(\alpha)}$ as in Proposition 6.100(a) (but with $s$ instead of $n$). Then,

$$a_{i,j}^{(\alpha)} = \begin{cases} M_{\alpha(i-j+1)}, & \text{if } i \geq j; \\ i, & \text{if } i = j - 1; \\ 0, & \text{if } i < j - 1 \end{cases}$$

for all $(i,j) \in \{1,2,\ldots,s\}^2$, and we have $A_s^{(\alpha)} = \left(a_{i,j}^{(\alpha)}\right)_{i,j=1,2,\ldots,s}$. Proposition 6.100(a) (applied to $s$ instead of $n$) yields $\det(A_s^{(\alpha)}) = s!M_{\alpha}^{(s)}$. Since $A_s^{(\alpha)} = \left(a_{i,j}^{(\alpha)}\right)_{i,j=1,2,\ldots,s}$, this rewrites as

$$(11.1168) \quad \det\left(a_{i,j}^{(\alpha)}\right)_{i,j=1,2,\ldots,s} = s!M_{\alpha}^{(s)}.$$

On the other hand, it is easy to see that

$$(11.1169) \quad a_{i,j}^{(\alpha)} \in k \cdot 1 \quad \text{for every } (i,j) \in \{1,2,\ldots,s\}^2 \text{ satisfying } i < j.$$

Furthermore,

$$(11.1170) \quad a_{i,j}^{(\alpha)} \in \mathcal{F}_\ell \quad \text{for every } (i,j) \in \{1,2,\ldots,s\}^2.$$

We have $(\mathcal{F}_m)^K \subset \mathcal{F}_{Km}$ (since (11.1167) holds for $k = K$). Thus,

$$(\mathcal{F}_m)^{K+1} = (\mathcal{F}_m)^K \cdot \mathcal{F}_m \subset \mathcal{F}_{Km} \cdot \mathcal{F}_m \subset \mathcal{F}_{Km+m} \quad \text{(by (11.1166), applied to } u = Km \text{ and } v = m)$$

$$= \mathcal{F}_{(K+1)m} \quad \text{(since } Km + m = (K+1)m).$$

In other words, (11.1167) holds for $k = K + 1$. This completes the induction step. The proof of (11.1167) is thus complete.

Proof of (11.1169): Let $(i,j) \in \{1,2,\ldots,s\}^2$ be such that $i < j$. We need to prove that $a_{i,j}^{(\alpha)} \in k \cdot 1$. This is clear in the case when $i = j - 1$ (because in this case, we have

$$a_{i,j}^{(\alpha)} = \begin{cases} M_{\alpha(i-j+1)}, & \text{if } i \geq j; \\ i, & \text{if } i = j - 1; = i \quad \text{(since } i = j - 1) \\ 0, & \text{if } i < j - 1 \end{cases} \in k \cdot 1$$

). Hence, for the rest of this proof, we can WLOG assume that we don’t have $i = j - 1$. Assume this.

We have $i < j$. Thus, $i \leq j - 1$ (since $i$ and $j$ are integers), so that $i < j - 1$ (since $i \neq j - 1$). Now,

$$a_{i,j}^{(\alpha)} = \begin{cases} M_{\alpha(i-j+1)}, & \text{if } i \geq j; \\ i, & \text{if } i = j - 1; = 0 \quad \text{(since } i < j - 1) \\ 0, & \text{if } i < j - 1 \end{cases} \in k \cdot 1,$$

and thus (11.1169) is proven.
Finally, \( a_{i,j}^{(\alpha)} = M_\alpha \) for every \( i \in \{1, 2, \ldots, s\} \).

But recall that every commutative ring \( A \), every \( m \in \mathbb{N} \) and every matrix \( (u_{i,j})_{i,j=1,2,\ldots,m} \in A^{m \times m} \) satisfy
\[
\det \begin{pmatrix} (u_{i,j})_{i,j=1,2,\ldots,m} \end{pmatrix} = \sum_{\sigma \in \mathfrak{S}_m} (-1)^\sigma \prod_{i=1}^m u_{i,\sigma(i)}.
\]

Applying this equality to \( A = \text{QSym}, m = s \) and \( u_{i,j} = a_{i,j}^{(\alpha)} \), we obtain
\[
\det \begin{pmatrix} (a_{i,j}^{(\alpha)})_{i,j=1,2,\ldots,s} \end{pmatrix} = \sum_{\sigma \in \mathfrak{S}_s} (-1)^\sigma \prod_{i=1}^s a_{i,\sigma(i)}^{(\alpha)}
= \left( -1 \right)^{\#(\sigma)} \sum_{i=1}^s a_{i,\sigma(i)}^{(\alpha)} + \sum_{\sigma \in \mathfrak{S}_s; \sigma \neq \text{id}} (-1)^\sigma \prod_{i=1}^s a_{i,\sigma(i)}^{(\alpha)}.
\]
(here, we have split off the addend for \( \sigma = \text{id} \) from the sum)
\[
= \sum_{\sigma \in \mathfrak{S}_s; \sigma \neq \text{id}} (-1)^\sigma \prod_{i=1}^s a_{i,\sigma(i)}^{(\alpha)} = M_\alpha^s + \sum_{\sigma \in \mathfrak{S}_s; \sigma \neq \text{id}} (-1)^\sigma \prod_{i=1}^s a_{i,\sigma(i)}^{(\alpha)}.
\]

Compared with (11.1168), this yields
\[
s! M_\alpha^s = M_\alpha^s + \sum_{\sigma \in \mathfrak{S}_s; \sigma \neq \text{id}} (-1)^\sigma \prod_{i=1}^s a_{i,\sigma(i)}^{(\alpha)}.
\]

887 Proof of (11.1170): Let \((i, j) \in \{1, 2, \ldots, s\}^2\). We need to show that \( a_{i,j}^{(\alpha)} \in \mathcal{F}_\ell \).

First, recall that \( \mathcal{F}_\ell \) is the \( k \)-submodule of \( \text{QSym} \) spanned by \( (M_\beta)_{\beta \in \text{Comp}, \ell(\beta) \leq \ell} \) (by the definition of \( \mathcal{F}_\ell \)). In particular, \( M_\beta \in \mathcal{F}_\ell \) for every \( \beta \in \text{Comp} \) satisfying \( \ell(\beta) \leq \ell \). Applied to \( \beta = \emptyset \), this yields that \( M_\emptyset \in \mathcal{F}_\ell \) (since \( \emptyset \in \text{Comp} \) satisfies \( \ell(\emptyset) = 0 \leq \ell(\alpha) = \ell \)). Since \( M_\emptyset = 1 \), this rewrites as \( 1 \in \mathcal{F}_\ell \), whence \( k \cdot 1 \subseteq k \cdot \mathcal{F}_\ell \subseteq \mathcal{F}_\ell \) (since \( \mathcal{F}_\ell \) is a \( k \)-module).

Now, if \( i < j \), then (11.1169) yields \( a_{i,j}^{(\alpha)} \in k \cdot 1 \subseteq \mathcal{F}_\ell \). Hence, (11.1170) is proven in the case when \( i < j \). For the rest of our proof of (11.1170), we can thus WLOG assume that we don’t have \( i < j \). Assume this.

We have \( i \geq j \) (since we don’t have \( i < j \)) and \( \ell(\{i - j + 1\}) = \ell(\alpha) \) (since \( \ell(\{k\}) = \ell(\alpha) \) for every positive integer \( k \)). Now, recall that \( M_\beta \in \mathcal{F}_\ell \) for every \( \beta \in \text{Comp} \) satisfying \( \ell(\beta) \leq \ell \). Applied to \( \beta = \alpha \{i - j + 1\} \), this yields that \( M_{\alpha \{i - j + 1\}} \in \mathcal{F}_\ell \) (since \( \ell(\alpha \{i - j + 1\}) = \ell(\alpha) = \ell \)). Now,
\[
a_{i,j}^{(\alpha)} = \begin{cases} M_{\alpha \{i - j + 1\}}, & \text{if } i \geq j; \\ i, & \text{if } i = j - 1; \\ M_{\alpha \{i - j + 1\}} & \text{if } i < j - 1. \\
\end{cases}
\]
and this proves (11.1170).

888 Proof of (11.1171): Let \( i \in \{1, 2, \ldots, s\} \). Then, the definition of \( a_{i,i}^{(\alpha)} \) yields
\[
a_{i,i}^{(\alpha)} = \begin{cases} M_{\alpha \{i - i + 1\}}, & \text{if } i \geq i; \\ i, & \text{if } i = i - 1; \\ M_{\alpha \{i - i + 1\}} & \text{if } i < i - 1. \\
\end{cases}
\]
(qed)

This is simply the explicit formula for the determinant of a matrix as a sum over permutations.
Subtracting $M_\alpha^s$ from both sides of this equality, we obtain

\[(11.1172)\quad s!M_\alpha^{(s)} - M_\alpha^s = \sum_{\sigma \in \mathfrak{S}_s; \sigma \neq \text{id}} (-1)^\sigma \prod_{i=1}^{s} a_{i,\sigma(i)}^{(\alpha)}.
\]

But it is easy to see that

\[(11.1173)\quad \prod_{i=1}^{s} a_{i,\sigma(i)}^{(\alpha)} \in F_{(s-1)\ell} \quad \text{for every } \sigma \in \mathfrak{S}_s \text{ satisfying } \sigma \neq \text{id}
\]

\[890.\text{ Hence, (11.1172) becomes (11.1174) becomes}
\]

\[(11.1174)\quad s!M_\alpha^{(s)} - M_\alpha^s = \sum_{\sigma \in \mathfrak{S}_s; \sigma \neq \text{id}} (-1)^\sigma \prod_{i=1}^{s} a_{i,\sigma(i)}^{(\alpha)} \in \sum_{\sigma \in \mathfrak{S}_s; \sigma \neq \text{id}} (-1)^\sigma F_{(s-1)\ell} \subset F_{(s-1)\ell}
\]

\[\text{(since } F_{(s-1)\ell} \text{ is a } k\text{-module}).
\]

But $F_{(s-1)\ell}$ is the $k$-submodule of QSym spanned by $(M_\beta)_{\beta \in \text{Comp}; \ell(\beta) \leq (s-1)\ell}$ (because this is how $F_{(s-1)\ell}$ was defined). In other words,

\[
F_{(s-1)\ell} = \sum_{\beta \in \text{Comp}; \ell(\beta) \leq (s-1)\ell} kM_\beta.
\]

Hence, (11.1174) becomes

\[(11.1175)\quad s!M_\alpha^{(s)} - M_\alpha^s \in F_{(s-1)\ell} = \sum_{\beta \in \text{Comp}; \ell(\beta) \leq (s-1)\ell} kM_\beta.
\]

Also, $s!M_\alpha^{(s)} - M_\alpha^s \in F_{(s-1)\ell} \subset \text{QSym}$, so that $s!M_\alpha^{(s)} \in M_\alpha^s + \text{QSym} \subset \text{QSym} + \text{QSym} \subset \text{QSym}.

Recall that $M_\alpha^{(s)} = e_s \left( \sum_{i=1}^{s} x_i^\alpha \right)$. Thus, $M_\alpha^{(s)}$ is a homogeneous power series of degree $s \mid \alpha \mid$ (since each $x_i^\alpha$ is a monomial of degree $\mid \alpha \mid$, and since $e_s$ is a power series of degree $s$). Hence, $s!M_\alpha^{(s)}$ is a homogeneous power series of degree $s \mid \alpha \mid$ as well. Thus, $s!M_\alpha^{(s)} \in \text{QSym}_{s|\alpha|}$ (since $s!M_\alpha^{(s)} \in \text{QSym}$). Also, $M_\alpha^s \in \text{QSym}_{s|\alpha|}$ (since $M_\alpha^s \in \text{QSym}_{s|\alpha|}$). Thus, $s!M_\alpha^{(s)} - M_\alpha^s \in \text{QSym}_{s|\alpha|} - \text{QSym}_{s|\alpha|} \subset \text{QSym}_{s|\alpha|}$ (since $\text{QSym}_{s|\alpha|}$ is a $k$-module).

Now, let $\pi$ denote the projection from the direct sum $\text{QSym} = \bigoplus_{k \in \mathbb{N}} \text{QSym}_k$ onto its $(s \mid \alpha \mid)$-th homogeneous component $\text{QSym}_{s|\alpha|}$. Notice that $\pi \left( s!M_\alpha^{(s)} - M_\alpha^s \right) = s!M_\alpha^{(s)} - M_\alpha^s$ (since $s!M_\alpha^{(s)} - M_\alpha^s \in \text{QSym}_{s|\alpha|}$).

\[890.\text{Proof of (11.1173): Let } \sigma \in \mathfrak{S}_s \text{ be such that } \sigma \neq \text{id}. \text{ Then, there exists a } k \in \{1,2,\ldots,s\} \text{ such that } \sigma(k) > k. \text{ Consider this } k.
\]

We have $k < \sigma(k)$ (since $\sigma(k) > k$) and thus $a_{k,\sigma(k)}^{(\alpha)} \in k \cdot 1$ (by (11.1169), applied to $i = k$ and $j = \sigma(k)$). But

\[
\prod_{i=1}^{s} a_{i,\sigma(i)}^{(\alpha)} = \prod_{i \in \{1,2,\ldots,s\}} a_{i,\sigma(i)}^{(\alpha)} = \prod_{i \in \{1,2,\ldots,s\}} a_{i,\sigma(i)}^{(\alpha)} \quad \text{(by (11.1170), applied to } j = \sigma(i))
\]

\[
\text{(here, we have split off the factor for } i = k \text{ from the product)}
\]

\[
\prod_{i \in \{1,2,\ldots,s\}} a_{i,\sigma(i)}^{(\alpha)} \in k \cdot F_{\ell} \subset k \cdot F_{(s-1)\ell} = k \cdot F_{(s-1)\ell} \subset F_{(s-1)\ell}
\]

\[\text{(since } F_{(s-1)\ell} \text{ is a } k\text{-module). This proves (11.1173).}
We have

\[(11.1176) \quad \pi(M_\beta) = 0 \quad \text{for every } \beta \in \text{Comp} \setminus \text{Comp}_s[\alpha].\]

Now, applying the map \(\pi\) to both sides of the relation \[(11.1175),\] we obtain

\[
\pi\left(s! M^{(s)}_\alpha - M^s_\alpha \right) \in \pi \left(\sum_{\beta \in \text{Comp}_s[\alpha]: \ell(\beta) \leq (s-1)\ell} k M_\beta\right) = \sum_{\beta \in \text{Comp}: \ell(\beta) \leq (s-1)\ell} k \pi(M_\beta) = \sum_{\beta \in \text{Comp} \setminus \text{Comp}_s[\alpha]: \ell(\beta) \leq (s-1)\ell} k \pi(M_\beta),
\]

(since the map \(\pi\) is \(k\)-linear)

\[
= \sum_{\beta \in \text{Comp}_s[\alpha]: \ell(\beta) \leq (s-1)\ell} k M_\beta + \sum_{\beta \in \text{Comp} \setminus \text{Comp}_s[\alpha]: \ell(\beta) \leq (s-1)\ell} k \pi(M_\beta) = \sum_{\beta \in \text{Comp}_s[\alpha]: \ell(\beta) \leq (s-1)\ell} k M_\beta
\]

(since \(\ell = \ell(\alpha)\)). Since \(\pi\left(s! M^{(s)}_\alpha - M^s_\alpha \right) = s! M^{(s)}_\alpha - M^s_\alpha\), this rewrites as \(s! M^{(s)}_\alpha - M^s_\alpha \in \sum_{\beta \in \text{Comp}_s[\alpha]: \ell(\beta) \leq (s-1)\ell(\alpha)} k M_\beta\).

This proves Lemma 6.110.

Second proof of Lemma 6.110. Let us first notice that if \(\gamma\) and \(\beta\) are two compositions, then

\[(11.1177) \quad \text{(the coefficient of the monomial } x^\gamma \text{ in } M_\beta \text{)} = \delta_{\gamma,\beta}.\]

(This follows from the definition of \(M_\beta\).)

Write the composition \(\alpha\) in the form \(\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_\ell);\) then, \(\ell = \ell(\alpha)\). Let us now fix a total order on the set SIS(\(\ell\)) (for example, the lexicographic order). Exercise 6.98(b) yields

\[(11.1178) \quad M^{(s)}_\alpha = \sum_{(i_1, i_2, \ldots, i_s) \in (\text{SIS}(\ell))^s; \ i_1 < i_2 < \cdots < i_s} x^{i_1}_1 x^{i_2}_2 \cdots x^{i_s}_s.\]

Now, for every \((i_1, i_2, \ldots, i_s) \in (\text{SIS}(\ell))^s\) such that \(i_1, i_2, \ldots, i_s\) are pairwise distinct, there exists a unique \(\sigma \in \mathcal{S}_s\) satisfying \(i_{\sigma(1)} < i_{\sigma(2)} < \cdots < i_{\sigma(s)}\) (because there is exactly one way to sort the \(\ell\)-tuple \((i_1, i_2, \ldots, i_s)\) into increasing order\footnote{Here, we are using that the order on SIS(\(\ell\)) is total.}). Hence, we can split the sum \(\sum_{(i_1, i_2, \ldots, i_s) \in (\text{SIS}(\ell))^s; \ i_1 < i_2 < \cdots < i_s} x^{i_1}_1 x^{i_2}_2 \cdots x^{i_s}_s\) into several

\footnote{Proof of \((11.1176)\): Let \(\beta \in \text{Comp} \setminus \text{Comp}_s[\alpha].\) Then, \(\beta \in \text{Comp}\) but \(\beta \notin \text{Comp}_s[\alpha].\) In other words, \(\beta\) is a composition with size \(|\beta| \neq s|\alpha|\). As a consequence, \(M_\beta\) is a homogeneous element of QSym of degree \(|\beta| \neq s|\alpha|\). Therefore, \(\pi(M_\beta) = 0\) (since \(\pi\) is the projection from the direct sum \(QSym = \bigoplus_{k \in \mathbb{N}} QSym_k\) onto its \((s|\alpha|)\)-th homogeneous component \(QSym_{s|\alpha|}\)). This proves \((11.1176)\).}
subsums, one for each $\sigma \in \mathfrak{S}_s$, as follows:

$$\sum_{(i_1, i_2, \ldots, i_s) \in (\text{SIS}(t))^s; \ i_1, i_2, \ldots, i_s \text{ are pairwise distinct}} x_{i_1}^{\alpha} x_{i_2}^{\alpha} \cdots x_{i_s}^{\alpha}$$

$$= \sum_{\sigma \in \mathfrak{S}_s} \sum_{(i_1, i_2, \ldots, i_s) \in (\text{SIS}(t))^s; \ i_1, i_2, \ldots, i_s \text{ are pairwise distinct; } i_\sigma(1) < i_\sigma(2) < \cdots < i_\sigma(s)} x_{i_1}^{\alpha} x_{i_2}^{\alpha} \cdots x_{i_s}^{\alpha}$$

$$= \sum_{(i_1, i_2, \ldots, i_s) \in (\text{SIS}(t))^s; \ i_\sigma(1) < i_\sigma(2) < \cdots < i_\sigma(s)} x_{i_1}^{\alpha} x_{i_2}^{\alpha} \cdots x_{i_s}^{\alpha}$$

(11.1179)

But every $\sigma \in \mathfrak{S}_s$ and every $(i_1, i_2, \ldots, i_s) \in (\text{SIS}(t))^s$ satisfy

(11.1180)

$$x_{i_1}^{\alpha} x_{i_2}^{\alpha} \cdots x_{i_s}^{\alpha} = x_{i_{\sigma(1)}}^{\alpha} x_{i_{\sigma(2)}}^{\alpha} \cdots x_{i_{\sigma(s)}}^{\alpha}$$

893. Now, (11.1179) becomes

$$\sum_{(i_1, i_2, \ldots, i_s) \in (\text{SIS}(t))^s; \ i_1, i_2, \ldots, i_s \text{ are pairwise distinct}} x_{i_1}^{\alpha} x_{i_2}^{\alpha} \cdots x_{i_s}^{\alpha}$$

$$= \sum_{\sigma \in \mathfrak{S}_s} \sum_{(i_1, i_2, \ldots, i_s) \in (\text{SIS}(t))^s; \ i_\sigma(1) < i_\sigma(2) < \cdots < i_\sigma(s)} x_{i_{\sigma(1)}}^{\alpha} x_{i_{\sigma(2)}}^{\alpha} \cdots x_{i_{\sigma(s)}}^{\alpha}$$

(11.1181)

$$= \sum_{\sigma \in \mathfrak{S}_s; \ i_\sigma(1) < i_\sigma(2) < \cdots < i_\sigma(s)} x_{i_{\sigma(1)}}^{\alpha} x_{i_{\sigma(2)}}^{\alpha} \cdots x_{i_{\sigma(s)}}^{\alpha} = \sum_{\sigma \in \mathfrak{S}_s} M_\alpha^{(s)} = |\mathfrak{S}_s| M_\alpha^{(s)} = s! M_\alpha^{(s)}.$$
On the other hand, taking both sides of the identity (6.26) to the $s$-th power, we obtain

$$M_s^α = \left( \sum_{i \in \text{SIS}(ℓ)} x_i^α \right)^s = \sum_{(i_1, i_2, \ldots, i_s) \in \text{(SIS}(ℓ))^s; \atop i_1, i_2, \ldots, i_s \text{ are pairwise distinct}} x_{i_1}^α x_{i_2}^α \cdots x_{i_s}^α \quad \text{(by the product rule)}$$

$$= \sum_{(i_1, i_2, \ldots, i_s) \in \text{(SIS}(ℓ))^s; \atop i_1, i_2, \ldots, i_s \text{ are pairwise distinct}} x_{i_1}^α x_{i_2}^α \cdots x_{i_s}^α + \sum_{(i_1, i_2, \ldots, i_s) \in \text{(SIS}(ℓ))^s; \atop i_1, i_2, \ldots, i_s \text{ are not pairwise distinct}} x_{i_1}^α x_{i_2}^α \cdots x_{i_s}^α$$

$$= s! M_s^{(s)} + \sum_{(i_1, i_2, \ldots, i_s) \in \text{(SIS}(ℓ))^s; \atop i_1, i_2, \ldots, i_s \text{ are not pairwise distinct}} x_{i_1}^α x_{i_2}^α \cdots x_{i_s}^α.$$

Hence,

(11.1182) \[ M_s^α - s! M_s^{(s)} = \sum_{(i_1, i_2, \ldots, i_s) \in \text{(SIS}(ℓ))^s; \atop i_1, i_2, \ldots, i_s \text{ are not pairwise distinct}} x_{i_1}^α x_{i_2}^α \cdots x_{i_s}^α. \]

But $M_s^{(s)} \in \text{QSym}$ (by Corollary 6.102(a)), so that $M_s^α - s! M_s^{(s)} \in \text{QSym} - s! \text{QSym} \subset \text{QSym}$. Hence, $M_s^α - s! M_s^{(s)}$ is a $\mathbf{k}$-linear combination of the family $(M_β)_{β \in \text{Comp}}$ (since this family $(M_β)_{β \in \text{Comp}}$ is a basis of the $\mathbf{k}$-module QSym). In other words, there exists a family $(c_β)_{β \in \text{Comp}} \in \mathbf{k}^{\text{Comp}}$ of elements of $\mathbf{k}$ such that (all but finitely many $β \in \text{Comp}$ satisfy $c_β = 0$) and

(11.1183) \[ s! M_s^{(s)} - M_s^α = \sum_{β \in \text{Comp}} c_β M_β. \]

Consider this family $(c_β)_{β \in \text{Comp}} \in \mathbf{k}^{\text{Comp}}$. Every composition $γ$ satisfies

$$
\left( \begin{array}{c}
\text{the coefficient of the monomial } x_γ^n \text{ in } s! M_s^{(s)} - M_s^α \\
= \sum_{β \in \text{Comp}} c_β M_β \\
(\text{by (11.1183)})
\end{array} \right)
= \left( \begin{array}{c}
\text{the coefficient of the monomial } x_γ^n \text{ in } \sum_{β \in \text{Comp}} c_β M_β \\
= \sum_{β \in \text{Comp}} c_β \left( \text{the coefficient of the monomial } x_γ^n \text{ in } M_β \right) \\
= \sum_{β \in \text{Comp}} c_β δ_{γ, β} \quad (\text{by (11.1177)})
\end{array} \right)
= \sum_{β \in \text{Comp}} c_β δ_{γ, β} = c_γ.
$$

Hence, every composition $γ$ satisfies

$$c_γ = \left( \text{the coefficient of the monomial } x_γ^n \text{ in } s! M_s^{(s)} - M_s^α \right).$$
Thus, every composition \( \gamma \) satisfies

\[
-c_{\gamma} = - \left( \text{the coefficient of the monomial } x^{\gamma} \text{ in } s!M^{(s)}_{\alpha} - M^s_{\alpha} \right)
\]

\[
= \left( \text{the coefficient of the monomial } x^{\gamma} \text{ in } \sum_{(i_1, i_2, \ldots, i_s) \in (\text{SIS}(\ell))^s} x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \right) \quad \text{(by (11.1182))}
\]

\[
= \sum_{(i_1, i_2, \ldots, i_s) \in (\text{SIS}(\ell))^s; \ i_1, i_2, \ldots, i_s \text{ are not pairwise distinct} } \left( \text{the coefficient of the monomial } x^{\gamma} \text{ in } x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \right).
\]

(11.1184)

We now notice that if \( \gamma \) is a composition satisfying \( |\gamma| \neq s \ |\alpha| \), then

\[
(11.1185) \quad \left( \text{the coefficient of the monomial } x^{\gamma} \text{ in } x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \right) = 0
\]

for every \((i_1, i_2, \ldots, i_s) \in (\text{SIS}(\ell))^s\). Also, if \( \gamma \) is a composition satisfying \( \ell(\gamma) > (s - 1) \ell(\alpha) \), then

\[
(11.1187) \quad \left( \text{the coefficient of the monomial } x^{\gamma} \text{ in } x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \right) = 0
\]

\[\textit{Proof of (11.1185):} \] Let \( \gamma \) be a composition satisfying \( |\gamma| \neq s \ |\alpha| \). Let \((i_1, i_2, \ldots, i_s) \in (\text{SIS}(\ell))^s\). Then, the monomial \( x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \) has degree

\[
\deg \left( x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \right) = \deg \left( x_{i_1}^{\alpha_1} \right) + \deg \left( x_{i_2}^{\alpha_2} \right) + \cdots + \deg \left( x_{i_s}^{\alpha_s} \right).
\]

But for every \( i \in \text{SIS}(\ell) \), the monomial \( x_{i}^{\alpha} \) is a monomial of degree \(|\alpha|\). Thus, for every \( i \in \text{SIS}(\ell) \), we have

\[
\deg \left( x_{i}^{\alpha} \right) = |\alpha|.
\]

(11.1186)

Hence, for every \( k \in \{1, 2, \ldots, s\} \), we have \( \deg \left( x_{i_k}^{\alpha_k} \right) = |\alpha| \) (by (11.1186), applied to \( i = i_k \)). Adding up these equalities over all \( k \in \{1, 2, \ldots, s\} \), we obtain \( \deg \left( x_{i_1}^{\alpha_1} \right) + \deg \left( x_{i_2}^{\alpha_2} \right) + \cdots + \deg \left( x_{i_s}^{\alpha_s} \right) = s |\alpha| \). Hence,

\[
\deg \left( x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \right) = \deg \left( x_{i_1}^{\alpha_1} \right) + \deg \left( x_{i_2}^{\alpha_2} \right) + \cdots + \deg \left( x_{i_s}^{\alpha_s} \right) = s |\alpha|,
\]

so that \( s |\alpha| = \deg \left( x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \right) \). But \( \deg(x^{\gamma}) = |\gamma| \neq s |\alpha| = \deg \left( x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \right) \). In other words, the monomials \( x^{\gamma} \) and \( x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \) have different degrees; thus, these monomials are distinct. Therefore, \( \left( \text{the coefficient of the monomial } x^{\gamma} \text{ in } x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \right) = 0 \). This proves (11.1185).
for every \((i_1, i_2, \ldots, i_s) \in (\text{SIS}(\ell))^s\) satisfying \((i_1, i_2, \ldots, i_s)\) are not pairwise distinct, then

\[-c_\gamma = \sum_{(i_1, i_2, \ldots, i_s) \in (\text{SIS}(\ell))^s; \quad i_1, i_2, \ldots, i_s \text{ are not pairwise distinct}} \left( \text{the coefficient of the monomial } x^\gamma \text{ in } x_{i_1}^{\alpha_1}x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \right) \quad \text{(by (11.1184))} \]

\[= 0. \quad \text{(by (11.1185))} \]

**Proof of (11.1187):** Let \(\gamma\) be a composition satisfying \(\ell(\gamma) > (s - 1)\ell(\alpha)\). Let \((i_1, i_2, \ldots, i_s) \in (\text{SIS}(\ell))^s\) be such that \((i_1, i_2, \ldots, i_s)\) are not pairwise distinct.

For every monomial \(m\) in the variables \(x_1, x_2, x_3, \ldots\), we denote by \(\text{Supp}(m)\) the set of all variables which occur in the monomial \(m\) (where we say that a variable occurs in \(m\) if and only if its exponent in \(m\) is positive). For instance, \(\text{Supp}(1) = \emptyset\) and \(\text{Supp}(x_2 x_3 x_5) = \{x_2, x_3, x_5\}\) and \(\text{Supp}(x_3 x_5^2) = \text{Supp}(x_3) = \{x_3\}\). It is easy to see that \(\text{Supp}(mn) = (\text{Supp}(m)) \cup (\text{Supp}(n))\) for any two monomials \(m\) and \(n\). More generally, if \(k \in \mathbb{N}\), and if \(m_1, m_2, \ldots, m_k\) are \(k\) monomials, then

(11.1188) \(\text{Supp}(m_1 m_2 \cdots m_k) = (\text{Supp}(m_1)) \cup (\text{Supp}(m_2)) \cup \cdots \cup (\text{Supp}(m_k))\).

On the other hand,

(11.1189) \(|\text{Supp}(x^\beta)| = \ell(\beta)\) for every composition \(\beta\).

**Proof of (11.1189):** Let \(\beta\) be a composition. Write \(\beta\) in the form \(\beta = (\beta_1, \beta_2, \ldots, \beta_{\ell(\beta)})\). Then, \(x^\beta = x_1^{\beta_1} x_2^{\beta_2} \cdots x_{\ell(\beta)}^{\beta_{\ell(\beta)}}\). Hence, the variables \(x_1, x_2, \ldots, x_{\ell(\beta)}\) all occur in the monomial \(x^\beta\) (since their exponents \(\beta_1, \beta_2, \ldots, \beta_{\ell(\beta)}\) are positive), and no other variables do. In other words, the set of all variables which occur in the monomial \(x^\beta\) is \(\{x_1, x_2, \ldots, x_{\ell(\beta)}\}\). In other words, \(\text{Supp}(x^\beta) = \{x_1, x_2, \ldots, x_{\ell(\beta)}\}\) (since Supp \((x^\beta)\) is the set of all variables which occur in the monomial \(x^\beta\) (by the definition of \(\text{Supp}(x^\beta)\))). In other words, \(\text{Supp}(x^\beta) = \{x_1, x_2, \ldots, x_{\ell(\beta)}\}\), so that \(|\text{Supp}(x^\beta)| = \ell(\beta)\). This proves (11.1189).

Also,

(11.1190) \(|\text{Supp}(x^\alpha_i)| = \ell(\beta)\) for every \(i \in \text{SIS}(\ell)\).

**Proof of (11.1190):** Let \(i \in \text{SIS}(\ell)\). Then, \(i\) is a strictly increasing \(\ell\)-tuple of positive integers. In other words, we can write \(i\) in the form \(i = (i_1, i_2, \ldots, i_\ell)\) for some positive integers \(i_1, i_2, \ldots, i_\ell\) satisfying \(i_1 < i_2 < \cdots < i_\ell\). Consider these \(i_1, i_2, \ldots, i_\ell\). Notice that \(i_1, i_2, \ldots, i_\ell\) are pairwise distinct (since \(i_1 < i_2 < \cdots < i_\ell\)), so that the variables \(x_{i_1}, x_{i_2}, \ldots, x_{i_\ell}\) are pairwise distinct.

The definition of \(x^\alpha_i\) yields \(x^\alpha_i = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_{i_\ell}^{\alpha_{i_\ell}}\). Thus, the variables \(x_{i_1}, x_{i_2}, \ldots, x_{i_\ell}\) all occur in the monomial \(x^\alpha_i\) (since their exponents \(\alpha_1, \alpha_2, \ldots, \alpha_{i_\ell}\) are positive), and no other variables do. In other words, the set of all variables which occur in the monomial \(x^\alpha_i\) is \(\{x_1, x_2, \ldots, x_{i_\ell}\}\). In other words, \(\text{Supp}(x^\alpha_i) = \{x_1, x_2, \ldots, x_{i_\ell}\}\) (since Supp \((x^\alpha_i)\) is the set of all variables which occur in the monomial \(x^\alpha_i\) (by the definition of \(\text{Supp}(x^\alpha_i)\))). In other words, \(\text{Supp}(x^\alpha_i) = \{x_1, x_2, \ldots, x_{i_\ell}\}\), so that \(|\text{Supp}(x^\alpha_i)| = \ell(\beta)\) (since the variables \(x_{i_1}, x_{i_2}, \ldots, x_{i_\ell}\) are pairwise distinct). This proves (11.1190).

We know that \(i_1, i_2, \ldots, i_\ell\) are not pairwise distinct. In other words, (at least) two of the elements \(i_1, i_2, \ldots, i_\ell\) are equal. In other words, there exist two distinct elements \(u\) and \(v\) of \(\{1, 2, \ldots, s\}\) such that \(i_u = i_v\). Consider these \(u\) and \(v\). Notice that \(u \in \{1, 2, \ldots, s\}\), so that \(1 \leq u \leq s\) and thus \(s \geq 1\).

Since \(u\) and \(v\) are distinct, we have \(u \in \{1, 2, \ldots, s\} \setminus \{v\}\); thus, \(\text{Supp}(x^\alpha_i)\) is an addend of the union

\(\bigcup_{j \in \{1, 2, \ldots, s\} \setminus \{v\}} \text{Supp}(x^\alpha_i)\). Consequently, \(\text{Supp}(x^\alpha_i) \subset \bigcup_{j \in \{1, 2, \ldots, s\} \setminus \{v\}} \text{Supp}(x^\alpha_i)\). Since \(i_u = i_v\), this rewrites as

(11.1191) \(\text{Supp}(x^\alpha_i) \subset \bigcup_{j \in \{1, 2, \ldots, s\} \setminus \{v\}} \text{Supp}(x^\alpha_i)\).

Now, (11.1188) (applied to \(k = s\) and \(m_j = x^\alpha_i\)) yields

\[\text{Supp}(x^\alpha_i x^\alpha_{i_2} \cdots x^\alpha_{i_j}) = \left(\text{Supp}(x^\alpha_i)\right) \cup \left(\text{Supp}(x^\alpha_{i_2})\right) \cup \cdots \cup \left(\text{Supp}(x^\alpha_{i_j})\right)\]

\[= \bigcup_{j \in \{1, 2, \ldots, s\}} \text{Supp}(x^\alpha_i) = \left(\bigcup_{j \in \{1, 2, \ldots, s\} \setminus \{v\}} \text{Supp}(x^\alpha_i)\right) \cup \text{Supp}(x^\alpha_i)\]

\[\subset \left(\bigcup_{j \in \{1, 2, \ldots, s\} \setminus \{v\}} \text{Supp}(x^\alpha_i)\right) \cup \text{Supp}(x^\alpha_i)\]

\[= \bigcup_{j \in \{1, 2, \ldots, s\} \setminus \{v\}} \text{Supp}(x^\alpha_i)\].
Hence, if $\gamma$ is a composition satisfying $|\gamma| \neq s|\alpha|$, then

(11.1193)  
$c_\gamma = 0$.

Also, if $\gamma$ is a composition satisfying $\ell(\gamma) > (s - 1) \ell(\alpha)$, then

$$- c_\gamma = \sum_{(i_1, i_2, \ldots, i_s) \in \text{SIS}(\ell)^s;\ i_1, i_2, \ldots, i_s \text{ are not pairwise distinct}} \text{(the coefficient of the monomial } x^\gamma \text{ in } x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \text{)} (\text{by (11.1184)})$$  

(by (11.1187))

$$= \sum_{(i_1, i_2, \ldots, i_s) \in \text{SIS}(\ell)^s;\ i_1, i_2, \ldots, i_s \text{ are not pairwise distinct}} 0 = 0.$$  

Hence, if $\gamma$ is a composition satisfying $\ell(\gamma) > (s - 1) \ell(\alpha)$, then

(11.1194)  
$c_\gamma = 0$.

Now, (11.1183) becomes

$$s!M_\alpha^{(s)} - M_\alpha^s = \sum_{\beta \in \text{Comp}} c_\beta M_\beta = \sum_{\beta \in \text{Comp};\ |\beta| = s|\alpha|} c_\beta M_\beta + \sum_{\beta \in \text{Comp};\ |\beta| \neq s|\alpha|} c_\beta M_\beta$$  

(by (11.1193), applied to $\gamma = \beta$)

$$= \sum_{\beta \in \text{Comp}_{s|\alpha|}: \ell(\beta) \leq (s - 1) \ell(\alpha)} c_\beta M_\beta + \sum_{\beta \in \text{Comp}_{s|\alpha|}: \ell(\beta) > (s - 1) \ell(\alpha)} c_\beta M_\beta$$  

(by (11.1194), applied to $\gamma = \beta$)

$$= \sum_{\beta \in \text{Comp}_{s|\alpha|}: \ell(\beta) \leq (s - 1) \ell(\alpha)} c_\beta M_\beta + \sum_{\beta \in \text{Comp}_{s|\alpha|}: \ell(\beta) > (s - 1) \ell(\alpha)} 0 M_\beta$$  

(by (11.1193), applied to $\gamma = \beta$)

This proves Lemma 6.110 once again.

Thus,

$$\left| \text{Supp} \left( x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \right) \right| \leq \left| \bigcup_{j \in \{1, 2, \ldots, s\} \setminus \{v\}} \text{Supp} \left( x_{i_j}^{\alpha_j} \right) \right| \leq \sum_{j \in \{1, 2, \ldots, s\} \setminus \{v\}} \left| \text{Supp} \left( x_{i_j}^{\alpha_j} \right) \right|$$  

(by (11.1190), applied to $1 = l_1$)

$$= \sum_{j \in \{1, 2, \ldots, s\} \setminus \{v\}} \ell = \ell = (s - 1) \ell(\alpha).$$  

Thus,

(11.1192)  
$$(s - 1) \ell(\alpha) \geq \left| \text{Supp} \left( x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} \right) \right|.$$

But (11.1189) (applied to $\beta = \gamma$) yields $|\text{Supp}(x^\gamma)| = \ell(\gamma) > (s - 1) \ell(\alpha) \geq |\text{Supp}(x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s})| (\text{by (11.1192)}).$ Hence, $|\text{Supp}(x^\gamma)| \neq |\text{Supp}(x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s})|$, so that $x^\gamma \neq x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s}$. Hence, $x^\gamma$ and $x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s}$ are distinct monomials.

Therefore, the coefficient of the monomial $x^\gamma$ in $x_{i_1}^{\alpha_1} x_{i_2}^{\alpha_2} \cdots x_{i_s}^{\alpha_s} = 0$. This proves (11.1187).

**Proof of Corollary 6.113.** We know that $u$ is a Lyndon word and satisfies $u = u$. Hence, $(u)$ is the CFL factorization of the word $u$.

Define the notion $\text{mult}_{u} z$ (for any Lyndon word $w$ and any word $z$) as in Theorem 6.44(b).

Theorem 6.44(e) yields that the lexicographically highest element of the multiset $u \sqcup v$ is $uv$, and the multiplicity with which this word $uv$ appears in the multiset $u \sqcup v$ is $\text{mult}_{u} v + 1$.

Recall that $\text{mult}_{u} v$ is the number of terms in the CFL factorization of $v$ which are equal to $u$ (by the definition of $\text{mult}_{u} v$). In other words,

$$\text{mult}_{u} v = \left(\text{the number of terms in the CFL factorization of } v \text{ which are equal to } u\right)_{= (b_1, b_2, \ldots, b_q)}$$

$$= (\text{the number of terms in } (b_1, b_2, \ldots, b_q) \text{ which are equal to } u)$$

$$= (\text{the number of } j \in \{1, 2, \ldots, q\} \text{ satisfying } b_j = u) = |\{ j \in \{1, 2, \ldots, q\} \mid b_j = u\}|.$$  

Now, the multiplicity with which the word $uv$ appears in the multiset $u \sqcup v$ is

$$\text{mult}_{u} v \quad + 1 = |\{ j \in \{1, 2, \ldots, q\} \mid b_j = u\}| + 1$$

$$= 1 + |\{ j \in \{1, 2, \ldots, q\} \mid b_j = u\}| = h.$$  

Hence, we can apply Lemma 6.112 to $z = uv$. As a result, we conclude that

$$M_u M_v = h M_{uv} + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < uv\right).$$

This proves Corollary 6.113.

\[\square\]


**Proof of Corollary 6.115.** Notice that $|x| = k$ (since $x \in \text{Comp}_k$), thus $|x^s| = s |x| = sk$ and thus $x^s \in \text{Comp}_{sk}$.

Clearly, $(x)$ is the CFL factorization of the word $x$ (since $x$ is Lyndon). On the other hand, $(x, x, \ldots, x)_{s \text{ times}}$ is the CFL factorization of the word $x^s$ (since $(x, x, \ldots, x)_{s \text{ times}}$ is a tuple of Lyndon words (since $x$ is Lyndon) satisfying $x^s = xx \ldots x$ and $x \geq x \geq \cdots \geq x$). Hence, Theorem 6.44(c) (applied to $u = x$, $v = x^s$, $p = 1$, $q = s$, $(a_1, a_2, \ldots, a_p) = (x)$ and $(b_1, b_2, \ldots, b_q) = (x, x, \ldots, x)_{s \text{ times}}$) yields that the lexicographically highest element of the multiset $x \sqcup x^s$ is $xx^s$ (since $x \geq x$ for every $i \in \{1, 2, \ldots, 1\}$ and $j \in \{1, 2, \ldots, s\}$). In other words, the lexicographically highest element of the multiset $x \sqcup x^s$ is $xs^s + 1$ (since $xx^s = x^s + 1$). This proves Corollary 6.115(a).
(c) Let \( t \in \text{Comp}_{sk} \) be such that \( t < x^s \). Then, Lemma 6.90(c) (applied to \( n = k, m = sk, u = x, v = x^s \), \( z = x^{s+1} \) and \( v' = t \)) yields

\[
M_x M_t = \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{k+sk} \text{ satisfying } w < x^{s+1} \right)
\]

\[
= \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{(s+1)k} \text{ satisfying } w < x^{s+1} \right).
\]

This proves Corollary 6.115(c).

(b) Let \( h = 1 + \{ j \in \{1, 2, \ldots, s \} \mid x = x \} \). Of course, \( h = s + 1 \). Recall that \( x \) is a Lyndon word, and that \( (x, x, \ldots, x) \) is the CFL factorization of the word \( x^s \). Notice also that \( x \geq x \) for every \( j \in \{1, 2, \ldots, s\} \).

Thus, Corollary 6.113 (applied to \( n = k, m = sk, u = x, v = x^s, q = s \) and \( (b_1, b_2, \ldots, b_q) = (x, x, \ldots, x) \) \( s \) times) yields that

\[
M_x M_{x^s} = \left( \frac{h}{s+1} \right) \left( \frac{M_{x^s}}{M_{x^{s+1}}} \right) + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{k+sk} \text{ satisfying } w < x^{s+1} \right)
\]

\[
= (s + 1) M_{x^{s+1}} + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{(s+1)k} \text{ satisfying } w < x^{s+1} \right).
\]

This proves Corollary 6.115(b). □


Proof of Corollary 6.117. We assumed that \( a_i > b_j \) for every \( i \in \{1, 2, \ldots, p\} \) and \( j \in \{1, 2, \ldots, q\} \). Thus, \( a_i \geq b_j \) for every \( i \in \{1, 2, \ldots, p\} \) and \( j \in \{1, 2, \ldots, q\} \). Hence, the lexicographically highest element of the multiset \( u \sqcup v \) is \( uv \) (by Theorem 6.44(c)). Also, the multiplicity with which the word \( uv \) appears in the multiset \( u \sqcup v \) is 1 (by Theorem 6.44(d)). Hence, Lemma 6.112 (applied to \( z = uv \) and \( h = 1 \)) yields

\[
M_u M_v = 1M_{uv} + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < uv \right)
\]

\[
= M_{uv} + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{n+m} \text{ satisfying } w < uv \right).
\]

This proves Corollary 6.117. □

896Proof. Every \( j \in \{1, 2, \ldots, s\} \) satisfies \( x = x \). Hence, the set \( \{ j \in \{1, 2, \ldots, s\} \mid x = x \} \) equals the whole \( \{1, 2, \ldots, s\} \).

In other words, \( \{ j \in \{1, 2, \ldots, s\} \mid x = x \} = \{1, 2, \ldots, s\} \). Now, \( h = 1 + \{ j \in \{1, 2, \ldots, s\} \mid x = x \} = 1 + \{1, 2, \ldots, s\} = 1 + s = s + 1 \), qed.
11.151. Solution to Exercise 6.120. Solution to Exercise 6.120.

Proof of Corollary 6.119. The tuple \((a_1, a_2, \ldots, a_p)\) is the CFL factorization of \(u\). Thus, \((a_1, a_2, \ldots, a_p)\) is a tuple of Lyndon words satisfying \(u = a_1 a_2 \cdots a_p\) and \(a_1 \geq a_2 \geq \cdots \geq a_p\).

Now, \((a_1, a_2, \ldots, a_k)\) is a tuple of Lyndon words (since \((a_1, a_2, \ldots, a_p)\) is a tuple of Lyndon words) satisfying
\[
x = a_1 a_2 \cdots a_k
\]
and \(a_1 \geq a_2 \geq \cdots \geq a_k\) (since \(a_1 \geq a_2 \geq \cdots \geq a_p\)). In other words, \((a_1, a_2, \ldots, a_k)\) is the CFL factorization of \(x\) (by the definition of a CFL factorization).

Also, \((a_{k+1}, a_{k+2}, \ldots, a_p)\) is a tuple of Lyndon words (since \((a_1, a_2, \ldots, a_p)\) is a tuple of Lyndon words) satisfying
\[
y = a_{k+1} a_{k+2} \cdots a_p
\]
and \(a_{k+1} \geq a_{k+2} \geq \cdots \geq a_p\) (since \(a_1 \geq a_2 \geq \cdots \geq a_p\)). In other words, \((a_{k+1}, a_{k+2}, \ldots, a_p)\) is the CFL factorization of \(y\) (by the definition of a CFL factorization).

We have \(x \in \text{Comp}_{|x|}\) and \(y \in \text{Comp}_{|y|}\). Also, multiplying the equalities \(x = a_1 a_2 \cdots a_k\) and \(y = a_{k+1} a_{k+2} \cdots a_p\), we obtain
\[
xy = (a_1 a_2 \cdots a_k) (a_{k+1} a_{k+2} \cdots a_p) = a_1 a_2 \cdots a_p = u.
\]
Thus, \(|x| + |y| = xy\).

We have \(a_i > a_{i+j}\) for every \(i \in \{1, 2, \ldots, k\}\) and \(j \in \{1, 2, \ldots, p-k\}\) \(897\). Hence, Corollary 6.117 (applied to \(x, y, |x|, |y|, k, p-k, (a_1, a_2, \ldots, a_k)\) and \((a_{k+1}, a_{k+2}, \ldots, a_p)\) instead of \(u, v, n, m, p, q, (a_1, a_2, \ldots, a_p)\) and \((b_1, b_2, \ldots, b_q)\)) yields
\[
M_x M_y = M_u + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{|x|+|y|} \text{ satisfying } w < xy\right)
\]
\[
= M_u + \left(\text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_n \text{ satisfying } w < u\right).
\]
Thus,
\[
M_u = M_x M_y - \left(\text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_n \text{ satisfying } w < u\right).
\]
This proves Corollary 6.119. \(\square\)


Proof of Corollary 6.121. We shall prove Corollary 6.121 by induction over \(s\):

Induction base: We have \(x^0 = \emptyset\) and therefore \(M_{x^0} = M_{\emptyset} = 1\). Now, \(M_x^0 - 0! M_{x^0} = 1 - 1 = 0\) in
\[
\sum_{w \in \text{Comp}_{|x^0|}^k} k M_w.
\]
In other words, Corollary 6.121 holds for \(s = 0\). This completes the induction base.

Induction step: Let \(S\) be a nonnegative integer. Assume that Corollary 6.121 holds for \(s = S\). We need to prove that Corollary 6.121 holds for \(s = S + 1\).

Notice that \(x \in \text{Comp}_k\), thus \(|x| = k\) and thus \(|x^S| = S |x| = Sk\), so that \(x^S \in \text{Comp}_{Sk}\).

\(897\)Proof: Recall that \(a_1 \geq a_2 \geq \cdots \geq a_p\). Thus,
\[
a_s \geq a_t \quad \text{for any } s \in \{1, 2, \ldots, p\} \text{ and } t \in \{1, 2, \ldots, p\} \text{ satisfying } s \leq t.
\]

Let \(i \in \{1, 2, \ldots, k\}\) and \(j \in \{1, 2, \ldots, p-k\}\). Then, \(i \leq k\) (since \(i \in \{1, 2, \ldots, k\}\)) and thus \(a_i \geq a_k\) (by (11.1195), applied to \(s = i\) and \(t = k\)). Also, \(j \geq 1\) (since \(j \in \{1, 2, \ldots, p-k\}\)), and thus \(k \geq j \geq k+1\), so that \(k+1 \leq k+j\). Hence, \(a_{k+1} \geq a_{k+j}\)
\[
\text{(by (11.1195), applied to } s = k+1\text{ and } t = k+j\text{). Now, } a_i \geq a_k \geq a_{k+1} \geq a_{k+j}, \text{ qed.}
\]
We know that Corollary 6.121 holds for \( s = S \). In other words,

\[
M_x^S - S! M_x S \in \sum_{w \in \text{Comp}_{S,k}; \ w < x^S} kM_w = \sum_{t \in \text{Comp}_{S,S}; \ t < x^S} kM_t
\]

(here, we renamed the summation index \( w \) as \( t \)).

Using Corollary 6.115(c), it is easy to see that

\[
M_x M_t \in \sum_{w \in \text{Comp}_{(S+1),k}; \ w < x^{S+1}} kM_w
\]

for every \( t \in \text{Comp}_{S,k} \) satisfying \( t < x^S \).

Now,

\[
M_x \left( M_x^S - S! M_x S \right) = M_x M_x^S - S! M_x M_x S = M_x^{S+1} - S! M_x M_x S,
\]

so that

\[
M_x^{S+1} - S! M_x M_x S = M_x \left( M_x^S - S! M_x S \right) \in M_x \left( \sum_{t \in \text{Comp}_{S,k}; \ t < x^S} kM_t \right)
\]

(by (11.1196))

\[
= \sum_{t \in \text{Comp}_{S,k}; \ t < x^S} \left( \sum_{w \in \text{Comp}_{(S+1),k}; \ w < x^{S+1}} kM_w \right)
\]

(by (11.1197))

\[
\subset \sum_{w \in \text{Comp}_{(S+1),k}; \ w < x^{S+1}} kM_w
\]

since \( \sum_{w \in \text{Comp}_{(S+1),k}; \ w < x^{S+1}} kM_w \) is a \( k \)-module.

(11.1198)

Now, Corollary 6.115(b) (applied to \( s = S \)) yields

\[
M_x M_x S = (S + 1) M_x S + \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{(S+1),k} \text{ satisfying } w < x^{S+1} \right).
\]

\[\text{Proof of (11.1197): Let } t \in \text{Comp}_{S,k} \text{ be such that } t < x^S. \text{ Then, Corollary 6.115(c) (applied to } s = S) \text{ yields}

\[
M_x M_t = \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_{(S+1),k} \text{ satisfying } w < x^{S+1} \right)
\]

\[
\in \sum_{w \in \text{Comp}_{(S+1),k}; \ w < x^{S+1}} kM_w.
\]

This proves (11.1197).
Thus,

\[
M_{x}M_{x^{s}} - (S + 1) M_{x^{s+1}} = \left( \text{a sum of terms of the form } M_{w} \text{ with } w \in \text{Comp}(S+1)k \text{ satisfying } w < x^{S+1} \right)
\]

(11.1199)

\[
\sum_{w \in \text{Comp}(S+1)k; \ w < x^{S+1}} kM_{w}.
\]

Now,

\[
M_{x}^{S+1} - (S + 1)! M_{x^{S+1}} = (M_{x}^{S+1} - S!M_{x^{s}}) + (S!M_{x}M_{x^{s}} - (S + 1)! M_{x^{S+1}})
\]

\[
= (M_{x}^{S+1} - S!M_{x^{s}}) + (S!M_{x}M_{x^{s}} - S! \cdot (S + 1)! M_{x^{S+1}})
\]

\[
= (M_{x}^{S+1} - S!M_{x^{s}}) + !S! \cdot (M_{x}M_{x^{s}} - (S + 1)! M_{x^{S+1}})
\]

\[
\sum_{w \in \text{Comp}(S+1)k; \ w < x^{S+1}} kM_{w}
\]

(by (11.1198))

\[
\sum_{w \in \text{Comp}(S+1)k; \ w < x^{S+1}} kM_{w}
\]

(by (11.1199))

\[
\sum_{w \in \text{Comp}(S+1)k; \ w < x^{S+1}} kM_{w} + \sum_{w \in \text{Comp}(S+1)k; \ w < x^{S+1}} kM_{w}
\]

(\text{since } \sum_{w \in \text{Comp}(S+1)k; \ w < x^{S+1}} kM_{w} \text{ is a } k\text{-module). In other words, Corollary 6.121 holds for } s = S + 1. \text{ This completes the induction step. Thus, Corollary 6.121 is proven by induction.}
\]

\[\square\]


Proof of Corollary 6.123. We first assume that \(k = \mathbb{Z}\).

Notice that \(|x| = k\) (since \(x \in \text{Comp}_k\)). Also, \(M_{x}^{(s)} \in \text{QSym} \) (by Corollary 6.102(a), applied to \(\alpha = x\)) and thus \(M_{x}^{(s)} - M_{x^{s}} \in \text{QSym} - \text{QSym} \subset \text{QSym} \).

Corollary 6.121 yields

(11.1200)

\[
M_{x}^{s} - s!M_{x^{s}} \in \sum_{w \in \text{Comp}_{x^{k}}; \ w < x^{s}} kM_{w}.
\]

On the other hand, the composition \(x\) is Lyndon and therefore nonempty. Hence, (6.27) (applied to \(\alpha = x\)) yields

\[
s!M_{x}^{(s)} - M_{x}^{s} \in \sum_{\beta \in \text{Comp}_{x^{k}}; \ \ell(\beta) \leq (s-1)\ell(x)} kM_{\beta} = \sum_{\beta \in \text{Comp}_{x^{k}}; \ \ell(\beta) \leq (s-1)\ell(x)} kM_{\beta} = \sum_{\beta \in \text{Comp}_{x^{k}}; \ \ell(\beta) \leq (s-1)\ell(x)} kM_{w}
\]

\[
(\text{since } |x| = k)
\]

\[
= \sum_{\beta \in \text{Comp}_{x^{k}}; \ \ell(\beta) \leq (s-1)\ell(x)} kM_{w}.
\]
(here, we renamed the summation index \( \beta \) as \( w \)). But every \( w \in \text{Comp}_{sk} \) satisfying \( \ell(w) \leq (s - 1) \ell(x) \) must also satisfy \( w < x^s \). Hence, the sum \( \sum_{w \in \text{Comp}_{sk}; \ell(w) \leq (s - 1) \ell(x)} kM_w \) is a subsum of the sum \( \sum_{w \in \text{Comp}_{sk}; w < x^s} kM_w \).

Thus, \( \sum_{w \in \text{Comp}_{sk}; \ell(w) \leq (s - 1) \ell(x)} kM_w \subset \sum_{w \in \text{Comp}_{sk}; w < x^s} kM_w \). Thus,

\[
\text{(11.1201)}
\]

\[
s!M_x^s = M_x \in \sum_{w \in \text{Comp}_{sk}; \ell(w) \leq (s - 1) \ell(x)} kM_w \subset \sum_{w \in \text{Comp}_{sk}; w < x^s} kM_w.
\]

Now,

\[
s! \left( M_x^s - M_x \right) = s!M_x^s - s!M_x = \left( s!M_x^s - M_x^s \right) + \left( M_x^s - s!M_x \right)
\]

\[
\in \sum_{w \in \text{Comp}_{sk}; w < x^s} kM_w + \sum_{w \in \text{Comp}_{sk}; w < x^s} kM_w \subset \sum_{w \in \text{Comp}_{sk}; w < x^s} kM_w
\]

(by (11.1201)).

(since \( \sum_{w \in \text{Comp}_{sk}; w < x^s} kM_w \) is a \( k \)-module).

But recall that we assumed that \( k = \mathbb{Z} \). Hence, if \( N \) is a positive integer and if \( f \) is an element of \( \text{QSym} \) satisfying \( NF \in \sum_{w \in \text{Comp}_{sk}; w < x^s} kM_w \), then \( f \in \sum_{w \in \text{Comp}_{sk}; w < x^s} kM_w \). Applying this to \( N = s! \) and \( f = M_x^s - M_x \), we obtain \( M_x^s - M_x \in \sum_{w \in \text{Comp}_{sk}; w < x^s} kM_w \) (because of (11.1202)). In other words, Corollary 6.123 holds under the assumption that \( k = \mathbb{Z} \).

---

**Proof.** Let \( w \in \text{Comp}_{sk} \). Notice that \( |x^s| = s|x| = sk \), so that \( x^s \in \text{Comp}_{sk} \). Now, we have \( \ell(x) > 0 \) (since \( x \) is nonempty), and

\[
\ell(x^s) = s\ell(x) = (s - 1)\ell(x) + \ell(x) > (s - 1)\ell(x).
\]

Thus, \( (s - 1)\ell(x) < \ell(x^s) \), so that \( \ell(w) \leq (s - 1)\ell(x) < \ell(x^s) \).

But the definition of the \( \sigma \)-order shows that if \( n \in \mathbb{N} \), and if \( \alpha \) and \( \beta \) are two elements of \( \text{Comp}_n \) satisfying \( \ell(\alpha) < \ell(\beta) \), then \( \alpha < \beta \). Applying this to \( \alpha = w \) and \( \beta = x^s \), we obtain \( w < x^s \) (since \( \ell(w) < \ell(x^s) \)), qed.

**Proof.** Let \( N \) be a positive integer. Let \( f \) be an element of \( \text{QSym} \) satisfying \( NF \in \sum_{w \in \text{Comp}_{sk}; w < x^s} kM_w \).

We know that \( (M_x)_{\alpha \in \text{Comp}} \) is a basis of the \( k \)-module \( \text{QSym} \). In the following, whenever \( g \in \text{QSym} \) and \( \beta \in \text{Comp} \), we will let \( \text{coord}_{M_\beta} g \) denote the \( M_\beta \)-coordinate of \( g \) with respect to the basis \( (M_x)_{\alpha \in \text{Comp}} \) of \( \text{QSym} \). Then, every \( g \in \text{QSym} \) satisfies

\[
\text{(11.1203)}
\]

(by the definition of coordinates). Notice also that

\[
\text{(11.1204)}
\]

\[
\text{coord}_{M_\beta} (M_x) = \delta_{\beta, \gamma} \quad \text{for every compositions } \beta \text{ and } \gamma.
\]

Now, \( NF \in \sum_{w \in \text{Comp}_{sk}; w < x^s} kM_w \). Thus, there exists a family \( (\lambda_w)_{w \in \text{Comp}_{sk}; w < x^s} \subset k \{ w \in \text{Comp}_{sk}; w < x^s \} \) of elements of \( k \) satisfying \( NF = \sum_{w \in \text{Comp}_{sk}; w < x^s} \lambda_w M_w \). Consider this family \( (\lambda_w)_{w \in \text{Comp}_{sk}; w < x^s} \).

Let \( \beta \) be a composition such that we don’t have \( \beta \in \text{Comp}_{sk} \) and \( \beta < x^s \). Then,

\[
\text{(11.1205)}
\]

every composition \( w \in \text{Comp}_{sk} \) satisfying \( w < x^s \) satisfies \( \beta \neq w \).
Now, let us forget that we have assumed that \( k = \mathbb{Z} \). We thus have shown that Corollary 6.123 holds under the assumption that \( k = \mathbb{Z} \). In other words, we have shown that the relation

\[
M^{(s)}_x - M_{x^s} \in \sum_{w \in \text{Comp}, \ w < x^s \ \text{will}} \mathbb{Z} M_w
\]

holds in \( \mathbb{Z}[x] \).

Now, recall that there is a canonical ring homomorphism \( \varphi : \mathbb{Z} \to k \). This homomorphism gives rise to a ring homomorphism \( \varphi[[x]] : \mathbb{Z}[[x]] \to k[[x]] \), and this latter homomorphism \( \varphi[[x]] \) has the properties that:

- it is \( \mathbb{Z} \)-linear;
- it sends the element \( M_\alpha \) of \( \mathbb{Z}[[x]] \) to the element \( M_\alpha \) of \( k[[x]] \) for every composition \( \alpha \);
- it sends the element \( M^{(s)}_\alpha \) of \( \mathbb{Z}[[x]] \) to the element \( M^{(s)}_\alpha \) of \( k[[x]] \) for every composition \( \alpha \) and every nonnegative integer \( s \).

(since every such \( w \) satisfies \( (w \in \text{Comp}, \ w < x^s \) will), whereas \( \beta \) does not satisfy \( (\beta \in \text{Comp}, \ \beta < x^s \) will \)). Now,

\[
\text{coord}_{M_{\beta}} \left( \sum_{w \in \text{Comp}, \ w < x^s \ \text{will}} N f \right) = \text{coord}_{M_{\beta}} \left( \sum_{w \in \text{Comp}, \ w < x^s \ \text{will}} \lambda_w M_w \right) = \sum_{w \in \text{Comp}, \ w < x^s \ \text{will}} \lambda_w \text{coord}_{M_{\beta}}(M_w) = \sum_{w \in \text{Comp}, \ w < x^s \ \text{will}} \lambda_w \delta_{\beta,w} = \sum_{w \in \text{Comp}, \ w < x^s \ \text{will}} \lambda_w 0 = 0.
\]

Compared with \( \text{coord}_{M_{\beta}} (N f) = N \text{coord}_{M_{\beta}} f \), this yields \( N \text{coord}_{M_{\beta}} f = 0 \).

But \( N \) is a positive integer. Thus, if an element \( \rho \) of \( \mathbb{Z} \) satisfies \( N\rho = 0 \), then \( \rho = 0 \). Applying this to \( \rho = \text{coord}_{M_{\beta}} f \), we obtain \( \text{coord}_{M_{\beta}} f = 0 \) (since \( \text{coord}_{M_{\beta}} f \in k = \mathbb{Z} \) and \( N \text{coord}_{M_{\beta}} f = 0 \)).

Now, let us forget that we fixed \( \beta \). We thus have shown that if \( \beta \) is a composition such that we don’t have \( (\beta \in \text{Comp}, \ \beta < x^s \) will \), then

\[
\text{coord}_{M_{\beta}} f = 0.
\]

Now, (11.1205) (applied to \( g = f \)) yields

\[
f = \sum_{\beta \in \text{Comp}} \left( \text{coord}_{M_{\beta}} f \right) M_{\beta} = \sum_{\beta \in \text{Comp}, \ \beta < x^s \ \text{will}} \left( \text{coord}_{M_{\beta}} f \right) M_{\beta} + \sum_{\beta \in \text{Comp}, \ \beta < x^s \ \text{will}} 0 M_{\beta} \text{ we don’t have (by (11.1206))}
\]

\[
= \sum_{\beta \in \text{Comp}, \ \beta < x^s \ \text{will}} \left( \text{coord}_{M_{\beta}} f \right) M_{\beta} + \sum_{\beta \in \text{Comp}, \ \beta < x^s \ \text{will}} 0 M_{\beta} \text{ we don’t have (by (11.1206))}
\]

\[
= \sum_{w \in \text{Comp}, \ w < x^s \ \text{will}} \left( \text{coord}_{M_{w}} f \right) M_w \text{ (here, we renamed the summation index \( \beta \) as \( w \))}
\]

\[
\in \sum_{w \in \text{Comp}, \ w < x^s \ \text{will}} k M_w,
\]

qed.
Hence, by applying the homomorphism \( \varphi [\mathfrak{A}] \) to both sides of (11.1207), we obtain
\[
M^{(s)}_x - M^{(s)}_x \in \sum_{w \in \text{Comp}, k} \mathbb{Z}M_w \quad \text{(due to the canonical ring homomorphism } \mathbb{Z} \rightarrow k) \quad \subseteq \sum_{w \in \text{Comp}, k} kM_w.
\]
This proves Corollary 6.123. \( \square \)


Proof of Theorem 6.107. The family \( \left( M_w^{(s)} \right)_{(w,s) \in M \times \{1,2,3,\ldots\}} \) is a reindexing of the family \( \left( M_{\text{red } \alpha}^{(\gcd \alpha)} \right)_{\alpha \in \mathfrak{A}} \) (according to Lemma 6.108). In other words, the family \( \left( M_w^{(s)} \right)_{(w,s) \in M \times \{1,2,3,\ldots\}} \) is a reindexing of the family \( \left( M_{\text{red } w}^{(\gcd w)} \right)_{w \in \mathfrak{A}} \) (here, we renamed the index \( \alpha \) as \( w \)).

We need to show that the family \( \left( M_w^{(s)} \right)_{(w,s) \in M \times \{1,2,3,\ldots\}} \) is an algebraically independent generating set of the \( k \)-algebra \( Q\text{Sym} \). It is enough to prove that the family \( \left( M_{\text{red } w}^{(\gcd w)} \right)_{w \in \mathfrak{A}} \) is an algebraically independent generating set of the \( k \)-algebra \( Q\text{Sym} \) (since the family \( \left( M_w^{(s)} \right)_{(w,s) \in M \times \{1,2,3,\ldots\}} \) is a reindexing of the family \( \left( M_{\text{red } w}^{(\gcd w)} \right)_{w \in \mathfrak{A}} \). We shall prove the latter claim.

Indeed, the main difficulty is to show that
\[
(11.1208) \quad \text{the family } \left( M_{\text{red } w}^{(\gcd w)} \right)_{w \in \mathfrak{A}} \text{ generates the } k \text{-algebra } Q\text{Sym}.
\]

Once (11.1208) is proven, we will be able to complete the proof of Theorem 6.107 as follows:

For every \( w \in \mathfrak{A} \), we have \( M_{\text{red } w}^{(\gcd w)} \in Q\text{Sym} \).

Let \( wt : \mathfrak{A} \rightarrow \{1,2,3,\ldots\} \) be the identity map (this is well-defined since \( \mathfrak{A} = \{1,2,3,\ldots\} \)). Obviously, for every \( N \in \{1,2,3,\ldots\} \), the set \( wt^{-1}(N) \) is finite.

For every word \( w \in \mathfrak{A}^* \), define an element \( Wt(w) \in N \) by \( Wt(w) = wt(w_1) + wt(w_2) + \cdots + wt(w_k) \), where \( k \) is the length of \( w \). Then,
\[
(11.1209) \quad \text{every } w \in \mathfrak{A}^* \text{ satisfies } Wt(w) = |w|
\]

For every \( w \in \mathfrak{A} \), the element \( M_{\text{red } w}^{(\gcd w)} \) of \( Q\text{Sym} \) is homogeneous of degree \( Wt(w) \).

The \( k \)-module \( Q\text{Sym} \) has a basis \( \left( g_u \right)_{u \in \mathfrak{A}^*} \) having the property that for every \( u \in \mathfrak{A}^* \), the element \( g_u \) of \( Q\text{Sym} \) is homogeneous of degree \( Wt(u) \).

\[\text{Proof.}\] Let \( w \in \mathfrak{A} \). Then, Corollary 6.102(a) (applied to \( \gcd w \) and \( \text{red } w \) instead of \( s \) and \( w \)) yields \( M_{\text{red } w}^{(\gcd w)} \in Q\text{Sym} \), qed.

\[\text{Proof of (11.1209):}\] Let \( w \in \mathfrak{A}^* \). Let \( k \) be the length of \( w \). Then, \( w = (w_1, w_2, \ldots, w_k) \), so that \( |w| = w_1 + w_2 + \cdots + w_k \).

Recall that \( wt \) is the identity map. In other words, \( wt = \text{id} \). Now, the definition of \( Wt(w) \) yields
\[
Wt(w) = wt(w_1) + wt(w_2) + \cdots + wt(w_k) = \text{id}(w_1) + \text{id}(w_2) + \cdots + \text{id}(w_k) \quad (\text{since } wt = \text{id})
\]
and thus (11.1209) is proven.

\[\text{Proof.}\] Let \( w \in \mathfrak{A} \). Then, \( w \) is a Lyndon word (since \( \mathfrak{A} \) is the set of all Lyndon words), hence nonempty. Remark 6.105(f) (applied to \( \alpha = w \)) now yields \( (\gcd w)(\text{red } w) = |w| = Wt(w) \) (by (11.1209)). But Corollary 6.102(b) (applied to \( \gcd w \) and \( \text{red } w \) instead of \( s \) and \( w \)) yields \( M_{\text{red } w}^{(\gcd w)} \in Q\text{Sym} \) (since \( \gcd w \) \text{red } w \) \( Wt(w) \)). In other words, the element \( M_{\text{red } w}^{(\gcd w)} \) of \( Q\text{Sym} \) is homogeneous of degree \( Wt(w) \), qed.

\[\text{Proof.}\] We shall show that \( \left( M_u \right)_{u \in \mathfrak{A}^*} \) is a basis.

Indeed, \( \mathfrak{A}^* = \text{Comp} \), so that \( \left( M_u \right)_{u \in \mathfrak{A}^*} = \left( M_u \right)_{u \in \text{Comp}} \) is clearly a basis of the \( k \)-module \( Q\text{Sym} \). Furthermore, for every \( u \in \mathfrak{A}^* \), the element \( M_u \) of \( Q\text{Sym} \) is homogeneous of degree \( |u| \). Since \( Wt(u) = |u| \) for every \( u \in \mathfrak{A}^* \) (by (11.1209), applied to \( w = u \)), this rewrites as follows: For every \( u \in \mathfrak{A}^* \), the element \( M_u \) of \( Q\text{Sym} \) is homogeneous of degree \( Wt(u) \).
Once (11.1208) is proven, it thus follows that we can apply Lemma 6.73(c) to $A = \text{QSym}$ and $b_w = M_{\text{red } w}^{(\gcd w)}$. From this, we can conclude that the family $\left(M_{\text{red } w}^{(\gcd w)}\right)_{w \in \Sigma}$ is an algebraically independent generating set of the $k$-algebra QSym (provided that (11.1208) is proven); this is precisely what we need to prove.

Hence, in order to complete the proof of Theorem 6.107, it is sufficient to prove (11.1208). So we shall now prove (11.1208).

Let $U$ denote the $k$-subalgebra of QSym generated by the family $\left(M_{\text{red } w}^{(\gcd w)}\right)_{w \in \Sigma}$. Then, $U$ is a $k$-submodule of QSym. It is clear that $M_{\text{red } w}^{(\gcd w)}$ is nonempty and satisfies red ($\beta$) for every reduced Lyndon composition $\beta$ and every $s \in \{1, 2, 3, \ldots\}$.

Using this and applying Exercise 6.98(d), it is now easy to see that

(11.1211) $M_{\beta}^{(s)} \in U$ for every Lyndon composition $\beta$ and every $s \in \{1, 2, 3, \ldots\}$.

We shall now prove that

(11.1212) $M_{\beta} \in U$ for every composition $\beta$.

**Proof of (11.1212):** We will prove (11.1212) by strong induction over $|\beta|$.

**Induction step:** Let $N \in \mathbb{N}$. We assume that (11.1212) holds for all compositions $\beta$ satisfying $|\beta| < N$. We need to show that (11.1212) also holds for all compositions $\beta$ satisfying $|\beta| = N$. In other words, we

Thus, $(M_{\alpha})_{\alpha \in \Sigma^*}$ is a basis of the $k$-module QSym having the property that for every $u \in \Sigma^*$, the element $M_u$ of QSym is homogeneous of degree $\text{wt}(u)$. Hence, the $k$-module QSym has a basis $(g_u)_{u \in \Sigma^*}$, having the property that for every $u \in \Sigma^*$, the element $g_u$ of QSym is homogeneous of degree $\text{wt}(u)$ (namely, $(M_u)_{u \in \Sigma^*}$ is such a basis), qed.

**Proof of (11.1210):** Let $\beta$ be a reduced Lyndon composition. Let $s \in \{1, 2, 3, \ldots\}$.

The composition $\beta$ is Lyndon and thus nonempty. Remark 6.105(e) (applied to $\alpha = \beta$) yields that the composition $\beta \{s\}$ is nonempty and satisfies red $(\beta \{s\}) = \text{red } \beta$ and $\gcd (\beta \{s\}) = s \gcd \beta$. But $\beta$ is reduced; thus, gcd $\beta = 1$ (by the definition of “reduced”). Hence, $\gcd (\beta \{s\}) = s \gcd \beta = s$. Also, Remark 6.105(d) (applied to $\alpha = \beta$) yields $\text{red } \beta = \beta$, so that

\[
\text{red } (\beta \{s\}) = \text{red } \beta = \beta.
\]

Remark 6.105(b) (applied to $\alpha = \beta \{s\}$) shows that the composition $\beta \{s\}$ is Lyndon if and only if the composition red $(\beta \{s\})$ is Lyndon. Since the composition red $(\beta \{s\}) = \beta$ is Lyndon, this yields that the composition $\beta \{s\}$ is Lyndon. In other words, $\beta \{s\} \in \Sigma$ (since $\Sigma$ is the set of all Lyndon compositions). Hence, $M_{\text{red } \beta \{s\}}$ is an element of the family $\left(M_{\text{red } w}^{(\gcd w)}\right)_{w \in \Sigma}$ (namely, the element for $w = \beta \{s\}$). In other words, $M_{\beta}^{(s)}$ is an element of the family $\left(M_{\text{red } w}^{(\gcd w)}\right)_{w \in \Sigma}$ (since $\text{red } (\beta \{s\}) = \beta$ and $\gcd (\beta \{s\}) = s$). Hence, $M_{\beta}^{(s)}$ belongs to the $k$-subalgebra of QSym generated by this family $\left(M_{\text{red } w}^{(\gcd w)}\right)_{w \in \Sigma}$. Since the $k$-subalgebra of QSym generated by the family $\left(M_{\text{red } w}^{(\gcd w)}\right)_{w \in \Sigma}$ is $U$, this rewrites as follows: $M_{\beta}^{(s)} \in U$.

**Proof of (11.1210).** Let $\beta$ be a Lyndon composition, and let $s \in \{1, 2, 3, \ldots\}$. The composition $\beta$ is Lyndon and thus nonempty; hence, gcd $\beta$ and red $\beta$ are well-defined.

Remark 6.105(b) (applied to $\alpha = \beta$) yields that the composition $\beta$ is Lyndon if and only if the composition red $\beta$ is Lyndon. Since the composition $\beta$ is Lyndon, we therefore conclude that the composition red $\beta$ is Lyndon. Also, the composition red $\beta$ is reduced (by Remark 6.105(c)). Thus, red $\beta$ is a reduced Lyndon composition. Remark 6.105(a) (applied to $\alpha = \beta$) yields red $\beta = \beta$, so that (red $\beta$) $\gcd \beta = \beta$.

Let $\alpha = \text{red } \beta$. Recall that red $\beta$ is a reduced Lyndon composition. In other words, $\alpha$ is a reduced Lyndon composition (since $\alpha = \text{red } \beta$).

Exercise 6.98(d) (applied to $n = \gcd \beta$) yields that there exists a polynomial $P \in k[z_1, z_2, z_3, \ldots]$ such that $M_{\alpha}^{(s)} = P \left(M_{\alpha}^{(s)}(1), M_{\alpha}^{(s)}(2), M_{\alpha}^{(s)}(3), \ldots\right)$. Consider this $P$. We have $\alpha = \text{red } \beta$, so that $\alpha$ $\gcd \beta = \beta$. Hence, $M_{\alpha}^{(s)} = P \left(M_{\alpha}^{(s)}(1), M_{\alpha}^{(s)}(2), M_{\alpha}^{(s)}(3), \ldots\right)$ rewrites as $M_{\beta}^{(s)} = P \left(M_{\beta}^{(s)}(1), M_{\beta}^{(s)}(2), M_{\beta}^{(s)}(3), \ldots\right)$.

But for every $j \in \{1, 2, 3, \ldots\}$, we have $M_{\alpha}^{(j)} \in U$ (by (11.1210), applied to $\alpha$ and $j$ instead of $\beta$ and $s$). In other words, $M_{\alpha}^{(1)}$, $M_{\alpha}^{(2)}$, $M_{\alpha}^{(3)}$, $\ldots$ are elements of $U$. Therefore, $Q \left(M_{\alpha}^{(1)}, M_{\alpha}^{(2)}, M_{\alpha}^{(3)}, \ldots\right) \in U$ for every $Q \in k[z_1, z_2, z_3, \ldots]$ (since $U$ is a $k$-subalgebra of QSym). Applied to $Q = P$, this yields $P \left(M_{\alpha}^{(1)}, M_{\alpha}^{(2)}, M_{\alpha}^{(3)}, \ldots\right) \in U$. Thus, $M_{\beta}^{(s)} = P \left(M_{\beta}^{(1)}, M_{\beta}^{(2)}, M_{\beta}^{(3)}, \ldots\right) \in U$. This proves (11.1211).
need to prove that $M_\beta \in U$ for every composition $\beta$ satisfying $|\beta| = N$. In other words, we need to prove that
\begin{equation}
(11.1213) \quad M_\beta \in U \quad \text{for every } \beta \in \text{Comp}_N.
\end{equation}

**Proof of (11.1213):** We will prove (11.1213) by strong induction over $\beta$ with respect to the wll-order on $\text{Comp}_N$. In other words, we fix some $\alpha \in \text{Comp}_N$, and we assume that (11.1213) holds for all $\beta < \alpha$. We now need to prove that (11.1213) holds for $\beta = \alpha$. In other words, we need to prove that $M_\alpha \in U$.

If $\alpha = \emptyset$, then $M_\alpha \in U$ is obvious.\(^{907}\) Hence, for the rest of this proof of $M_\alpha \in U$, we can WLOG assume that $\alpha \neq \emptyset$. Assume this. The composition $\alpha$ is nonempty (since $\alpha \neq \emptyset$), so that it satisfies $|\alpha| \neq 0$. Since $|\alpha| = N$ (because $\alpha \in \text{Comp}_N$), we have $N = |\alpha| \neq 0$. Thus, $N$ is a positive integer.

We have assumed that (11.1213) holds for all $\beta < \alpha$. We now need to prove that (11.1213) holds for all $\beta \in \text{Comp}_N$ satisfying $\beta < \alpha$. In other words,
\begin{equation}
(11.1214) \quad M_\beta \in U \quad \text{for all } \beta \in \text{Comp}_N \text{ satisfying } \beta < \alpha.
\end{equation}

Also, we have assumed that (11.1212) holds for all compositions $\beta$ satisfying $|\beta| < N$. In other words,
\begin{equation}
(11.1215) \quad M_\beta \in U \quad \text{for all compositions } \beta \text{ satisfying } |\beta| < N.
\end{equation}

Let $(a_1, a_2, \ldots, a_p)$ be the CFL factorization of the word $\alpha$. Then, $(a_1, a_2, \ldots, a_p)$ is a tuple of Lyndon words satisfying $\alpha = a_1a_2 \cdots a_p$ and $a_1 \geq a_2 \geq \cdots \geq a_p$ (according to the definition of a CFL factorization). We have $p \neq 0$ (since otherwise, we would have $p = 0$ and thus $\alpha = a_1a_2 \cdots a_p = (\text{empty product}) = \emptyset$, contradicting $\alpha \neq \emptyset$). Thus, $p \in \{1, 2, 3, \ldots\}$. Hence, the word $a_1$ is well-defined. Clearly, $a_1$ is a Lyndon word (since $(a_1, a_2, \ldots, a_p)$ is a tuple of Lyndon words).

We distinguish between two cases:

**Case 1:** All of the words $a_1, a_2, \ldots, a_p$ are equal.

**Case 2:** Not all of the words $a_1, a_2, \ldots, a_p$ are equal.

Let us consider Case 1 first. In this case, all of the words $a_1, a_2, \ldots, a_p$ are equal. In other words, $a_1 = a_2 = \cdots = a_p$. Thus, $a_1 = a_i$ for every $i \in \{1, 2, \ldots, p\}$.

Let $x = a_1$. Then, $x = x_1 = a_i$ for every $i \in \{1, 2, \ldots, p\}$. Multiplying these identities for all $i \in \{1, 2, \ldots, p\}$, we obtain $x \underbrace{\cdots x}_p = a_1a_2 \cdots a_p = \alpha$, so that $\alpha = x \underbrace{\cdots x}_p = x^p$, thus $x^p = \alpha$. Also, $x = a_1$ is a Lyndon word. Let $k = |x|$. Then, $x \in \text{Comp}_k$. Also, $N = \left\lfloor \frac{\alpha}{x^p} \right\rfloor = |x^p| = p |x|$, so that $p |x| = N$. Thus, $p \frac{k}{|x|} = p |x| = N$.

Now, Corollary 6.123 (applied to $s = p$) yields
\[
M_\alpha - M_{x^p} = \sum_{w \in \text{Comp}_\alpha, \ w \ll \alpha} k M_w = \sum_{\underbrace{\text{Comp}_\alpha, \ w \ll \alpha}} k M_w \quad \text{(since } pk = N \text{ and } x^p = \alpha) \quad \text{(by (11.1214), applied to } \beta = \alpha) \quad \text{(since } U \text{ is a } k\text{-submodule of } Q\text{Sym})}.
\]

Hence,
\[
M_{x^p} - M_{x^p} = - \underbrace{\left(M_\alpha - M_{x^p}\right)}_{\in U} \in - U \subset U \quad \text{(since } U \text{ is a } k\text{-submodule of } Q\text{Sym}), \quad \text{(by (11.1211), applied to } x \text{ and } p \text{ instead of } \beta \text{ and } s)
\]
so that
\[
M_{x^p} \in \underbrace{M_{x^p}}_{\in U} + U \subset U
\]
(since $U$ is a $k$-submodule of $QSym$). Since $x^p = \alpha$, this rewrites as $M_\alpha \in U$. Thus, we have proven $M_\alpha \in U$ in Case 1.

Let us now consider Case 2. In this case, not all of the words $a_1, a_2, \ldots, a_p$ are equal. Hence, there exists some $k \in \{1, 2, \ldots, p - 1\}$ such that $a_k \neq a_{k+1}$. Consider this $k$.

We have $a_k \geq a_{k+1}$ (since $a_1 \geq a_2 \geq \cdots \geq a_p$). Combined with $a_k \neq a_{k+1}$, this yields $a_k > a_{k+1}$. Let $x$ be the word $a_1 a_2 \cdots a_k$, and let $y$ be the word $a_{k+1} a_{k+2} \cdots a_p$. Then, Corollary 6.119 (applied to $u = \alpha$ and $n = N$) yields

$$M_\alpha = M_x M_y - \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_N \text{ satisfying } w < \alpha \right).$$

Thus,

$$M_x M_y - M_\alpha = \left( \text{a sum of terms of the form } M_w \text{ with } w \in \text{Comp}_N \text{ satisfying } w < \alpha \right) \in \sum_{\substack{w \in \text{Comp}_N; \ w \text{ will} \ w < \alpha}} e_{M_w} \subset \sum_{\substack{w \in \text{Comp}_N; \ w \text{ will} \ w < \alpha}} kU \subset U$$

(since $U$ is a $k$-submodule of $QSym$). Hence,

(11.1216) $M_\alpha \in M_x M_y - U.$

Now, it is easy to see that $N = |x| + |y|$ \footnote{Proof. Multiplying the equalities $x = a_1 a_2 \cdots a_k$ and $y = a_{k+1} a_{k+2} \cdots a_p$, we obtain $xy = (a_1 a_2 \cdots a_k)(a_{k+1} a_{k+2} \cdots a_p) = a_1 a_2 \cdots a_p = \alpha,$ so that $x = \alpha$ and thus $|\alpha| = |xy| = |x| + |y|$, so that $|x| + |y| = |\alpha| = N$, qed.} \footnote{Proof. Notice that $a_k$ is a Lyndon word (since $(a_1, a_2, \ldots, a_p)$ is a tuple of Lyndon words), and thus nonempty.

But $a_k$ is a suffix of the word $a_1 a_2 \cdots a_k$. Hence, $\ell(a_k) \leq \ell(a_1 a_2 \cdots a_k)$, so that $\ell(a_1 a_2 \cdots a_k) \geq \ell(a_k) > 0$ (since the word $a_k$ is nonempty). Now, $\ell \left( \frac{x}{a_1 a_2 \cdots a_k} \right) = \ell(a_1 a_2 \cdots a_k) > 0$, so that the word $x$ is nonempty, and therefore $|x| > 0$, qed.} and $|y| > 0$ \footnote{Proof. Notice that $a_{k+1}$ is a Lyndon word (since $(a_1, a_2, \ldots, a_p)$ is a tuple of Lyndon words), and thus nonempty.

But $a_{k+1}$ is a prefix of the word $a_{k+1} a_{k+2} \cdots a_p$. Hence, $\ell(a_{k+1}) \leq \ell(a_{k+1} a_{k+2} \cdots a_p)$, so that $\ell(a_{k+1} a_{k+2} \cdots a_p) \geq \ell(a_{k+1}) > 0$ (since the word $a_{k+1}$ is nonempty). Now, $\ell \left( \frac{y}{a_{k+1} a_{k+2} \cdots a_p} \right) = \ell(a_{k+1} a_{k+2} \cdots a_p) > 0$, so that the word $y$ is nonempty, and therefore $|y| > 0$, qed.} Hence, $N = |x| + |y| > |x|$, which yields $|x| < N$. Hence, (11.1215) (applied to $\beta = x$) yields $M_x \in U$. Also, $N = |x| + |y| > |y|$, and thus $|y| < N$, so that (11.1215) (applied to $\beta = y$) yields $M_y \in U$. Now, (11.1216) becomes $M_\alpha \in M_x M_y - U \subset UU - U \subset U$ (since $U$ is a $k$-subalgebra of $QSym$). Thus, we have proven $M_\alpha \in U$ in Case 2.

Now, we have proven $M_\alpha \in U$ in each of the two Cases 1 and 2. Since these two Cases cover all possibilities, this yields that we always have $M_\alpha \in U$. In other words, (11.1213) holds for $\beta = \alpha$.

Thus, we have completed the induction step of our induction over $\beta$. Therefore, we have proven (11.1213) by induction. So we now know that $M_\beta \in U$ for every $\beta \in \text{Comp}_N$. In other words, $M_\beta \in U$ for every composition $\beta$ satisfying $|\beta| = N$. In other words, (11.1208) holds for all compositions $\beta$ satisfying $|\beta| = N$.

Thus, we have completed the induction step of our induction over $N$. Hence, (11.1212) is proven by induction over $N$.
Now, recall that the family \((M_\beta)_{\beta \in \text{Comp}}\) is a basis of the \(k\)-module \(Q\text{Sym}\), and thus generates this \(k\)-module. Hence,

\[
Q\text{Sym} = \sum_{\beta \in \text{Comp}} k M_\beta \supseteq \sum_{\beta \in \text{Comp}} kU \subseteq U
\]

(by (11.1212)) (since \(U\) is a \(k\)-submodule of \(Q\text{Sym}\)). Combined with \(U \subseteq Q\text{Sym}\), this yields \(U = Q\text{Sym}\). Since \(U\) is the \(k\)-subalgebra of \(Q\text{Sym}\) generated by the family \(\left(M_{\text{red}\, w}^{(\text{gcd}\, w)}\right)_{w \in \mathcal{L}}\), this shows that the \(k\)-subalgebra of \(Q\text{Sym}\) generated by the family \(\left(M_{\text{red}\, w}^{(\text{gcd}\, w)}\right)_{w \in \mathcal{L}}\) is \(Q\text{Sym}\) itself. In other words, the family \(\left(M_{\text{red}\, w}^{(\text{gcd}\, w)}\right)_{w \in \mathcal{L}}\) generates the \(k\)-algebra \(Q\text{Sym}\). Thus, (11.1208) is proven. As we know, this completes the proof of Theorem 6.107. \(\square\)


Proof of Theorem 6.82. We know (from the proof of Theorem 6.107) that the family \(\left(M_{\text{red}\, w}^{(\text{gcd}\, w)}\right)_{w \in \mathcal{L}}\) is an algebraically independent generating set of the \(k\)-algebra \(Q\text{Sym}\).

Now, define a grading on the \(k\)-algebra \(k[x_w \mid w \in \mathcal{L}]\) by setting \(\deg(x_w) = \sum_{i=1}^{(w)} w_i\) for every \(w \in \mathcal{L}\). By the universal property of the polynomial algebra \(k[x_w \mid w \in \mathcal{L}]\), we can define a \(k\)-algebra homomorphism \(\Phi : k[x_w \mid w \in \mathcal{L}] \to Q\text{Sym}\) by setting

\[
\Phi(x_w) = M_{\text{red}\, w}^{(\text{gcd}\, w)} \quad \text{for every} \ w \in \mathcal{L}.
\]

This homomorphism \(\Phi\) is a \(k\)-algebra isomorphism (since \(\left(M_{\text{red}\, w}^{(\text{gcd}\, w)}\right)_{w \in \mathcal{L}}\) is an algebraically independent generating set of the \(k\)-algebra \(Q\text{Sym}\)) and is graded (because for every \(w \in \mathcal{L}\), the element \(M_{\text{red}\, w}^{(\text{gcd}\, w)}\) of \(Q\text{Sym}\) is homogeneous of degree \(\deg(x_w)\)). Thus, \(\Phi\) is an isomorphism of graded \(k\)-algebras. Hence, \(Q\text{Sym} \cong k[x_w \mid w \in \mathcal{L}]\) as graded \(k\)-algebras. Thus, \(Q\text{Sym}\) is a polynomial algebra. This proves Theorem 6.82. \(\square\)


Proof of Corollary 6.127. Theorem 6.107 yields that the family \(\left(M_{w}^{(s)}\right)_{(w,s) \in \mathcal{RL} \times \{1,2,3,\ldots\}}\) is an algebraically independent generating set of the \(k\)-algebra \(Q\text{Sym}\).

Notice that (1) is a reduced Lyndon composition; that is, \((1) \in \mathcal{RL}\) (since \(\mathcal{RL}\) is the set of all reduced Lyndon compositions). Hence, \(\{(1)\} \subseteq \mathcal{RL}\), whence \(\{(1)\} \times \{1,2,3,\ldots\} \subseteq \mathcal{RL} \times \{1,2,3,\ldots\}\).

The following fact is straightforward to check: If \(A\) is a commutative \(k\)-algebra, and if \((a_i)_{i \in I}\) is an algebraically independent generating set of the \(k\)-algebra \(A\), and if \(J\) is a subset of \(I\), then \((a_i)_{i \in J}\) is an algebraically independent generating set of the \(k\)-algebra \(A\). \(\square\)

We can apply this fact to \(A = Q\text{Sym}\), \(I = \mathcal{RL} \times \{1,2,3,\ldots\}\), \((a_i)_{i \in I} = \left(M_{w}^{(s)}\right)_{(w,s) \in \mathcal{RL} \times \{1,2,3,\ldots\}}\) and \(J = \{(1)\} \times \{1,2,3,\ldots\}\). As a result, we conclude that \(\left(M_{w}^{(s)}\right)_{(w,s) \in \{(1) \times \{1,2,3,\ldots\}\}}\) is an algebraically independent generating set of the \(k\left[M_{w}^{(s)} \mid (w,s) \in \{1\} \times \{1,2,3,\ldots\}\right]\)-algebra \(Q\text{Sym}\). Hence, the \(k\left[M_{w}^{(s)} \mid (w,s) \in \{1\} \times \{1,2,3,\ldots\}\right]\)-algebra \(Q\text{Sym}\) has an algebraically independent generating set; it is therefore a polynomial algebra over \(Q\text{Sym}\). That is, the polynomial ring in the indeterminates \((x_i)_{i \in I}\) can be regarded as a polynomial ring in the indeterminates \((x_i)_{i \in J}\) over the polynomial ring in the indeterminates \((x_i)_{i \in J}\).
$\mathbf{k} \left[ M_w^{(s)} \mid (w, s) \in \{1\} \times \{1, 2, 3, \ldots\} \right]$. Since $\mathbf{k} \left[ M_w^{(s)} \mid (w, s) \in \{1\} \times \{1, 2, 3, \ldots\} \right] = \Lambda^{914}$, this rewrites as follows: The $\Lambda$-algebra $Q\text{Sym}$ is a polynomial algebra over $\Lambda$. This proves Corollary 6.127. □

11.157. **Solution to Exercise 7.9.** Solution to Exercise 7.9. Fix $m \in \{0, 1, 2, \ldots\}$. In order to check that $\zeta^{*m} (f) = \text{ps}^1 (f) (m)$, it is clearly enough to show that $\zeta^{*m} (M_\alpha) = \text{ps}^1 (M_\alpha) (m)$ for all compositions $\alpha$. So let $\alpha$ be a composition. Iterated application of Proposition 5.7 yields

\[
\Delta^{(m-1)} M_\alpha = \sum_{(\beta_1, \beta_2, \ldots, \beta_m) : \beta_1 \beta_2 \cdots \beta_m = \alpha} M_{\beta_1} \otimes M_{\beta_2} \otimes \cdots \otimes M_{\beta_m}.
\]

An addend on the right hand side of this equality is annihilated by the map $\zeta^{*m} : \text{QSym}^{*m} \rightarrow \mathbf{k}^{*m} \cong \mathbf{k}$ unless each of the compositions $\beta_1, \beta_2, \ldots, \beta_m$ has length $\leq 1$; all remaining terms are mapped to $1 \cdot 1 \cdots 1 = 1$. Hence,

\[
\zeta^{*m} \left( \Delta^{(m-1)} M_\alpha \right) = \sum_{(\beta_1, \beta_2, \ldots, \beta_m) : \beta_1 \beta_2 \cdots \beta_m = \alpha; \text{ each of the compositions } \beta_1, \beta_2, \ldots, \beta_m \text{ has length } \leq 1} 1 = \binom{m}{\ell},
\]

where $\ell$ is the length of $\alpha$. Hence,

\[
\zeta^{*m} (M_\alpha) = \zeta^{*m} \left( \Delta^{(m-1)} M_\alpha \right) = \binom{m}{\ell} = \text{ps}^1 (M_\alpha) (m)
\]

(by Proposition 7.7(i)). This completes the proof.

11.158. **Solution to Exercise 7.33.** Solution to Exercise 7.33.

**Proof of Proposition 7.28.** We recall the following fundamental fact (a version of inclusion-exclusion or a very simple special case of Möbius inversion): If $R$ is a finite set, then

\[
\sum_{T \subset R} (-1)^{|T|} = \delta_{R, \emptyset}.
\]

\[914\]Proof. The first sentence of Proposition 2.28 yields that the family $(e_1, e_2, e_3, \ldots)$ generates the $\mathbf{k}$-algebra $\Lambda$. In other words, $\Lambda = \mathbf{k} [e_1, e_2, e_3, \ldots] = \mathbf{k} [e_s \mid s \in \{1, 2, 3, \ldots\}]$.

But

\[
\mathbf{k} \left[ M_w^{(s)} \mid (w, s) \in \{1\} \times \{1, 2, 3, \ldots\} \right] = \mathbf{k} \left[ M_w^{(1)} \mid s \in \{1, 2, 3, \ldots\} \right] = \mathbf{k} [e_s \mid s \in \{1, 2, 3, \ldots\}] = \Lambda
\]

(since $\Lambda = \mathbf{k} [e_s \mid s \in \{1, 2, 3, \ldots\}]$), qed.
We can use this to show a slightly more complicated fact: If $P$ and $R$ are two finite sets, then

\[(11.1218) \quad \sum_{F \subseteq R, \; F \supsetneq P} (-1)^{|F \setminus P|} = \delta_{P,R}.\]

---

For the sake of completeness, here is a short proof of (11.1217): Let $R$ be a finite set. We have

\[
\sum_{T \subset R} (-1)^{|T|} = \sum_{k \in \mathbb{N}} \sum_{T \subset R, \; |T| = k} (-1)^{|T|} = \sum_{k \in \mathbb{N}} \sum_{T \subset R, \; |T| = k} (-1)^k = \sum_{k \in \mathbb{N}} \binom{|T \subset R | \cap \{T \mid \text{|T| = k}\}}{(-1)^k} = \binom{|T \subset R | \cap \{T \mid \text{|T| = k}\}}{|R|} (-1)^k
\]

(by the combinatorial interpretation of binomial coefficients)

\[
= \sum_{k \in \mathbb{N}} \binom{|R|}{k} (-1)^k = \left(1 + (-1)^{|R|}\right)^{|R|} \quad \text{since } (1 + (-1))^n = \sum_{k \in \mathbb{N}} \binom{n}{k} (-1)^k \quad \text{(by the binomial formula)}
\]

\[
= 0^{|R|} = \delta_{|R|,0} = \delta_{R,\emptyset} \quad \text{(since } |R| = 0 \text{ holds if and only if } R = \emptyset).\]

This proves (11.1217).

Proof of (11.1218): Let $P$ and $R$ be two finite sets.

Let us first assume that $P \not\subseteq R$. Then, there exists no $F \subset R$ such that $F \supsetneq P$ (because if such a $F$ would exist, then we would have $P \subset F \subset R$, which would contradict $P \not\subseteq R$). Hence, the sum $\sum_{F \supsetneq P} (-1)^{|F \setminus P|}$ is empty, and thus it simplifies to $\sum_{F \supsetneq P} (-1)^{|F \setminus P|} = 0$. On the other hand, $P \not\subset R$, so that $P \neq R$ and thus $\delta_{P,R} = 0$. Hence, $\sum_{F \supsetneq P} (-1)^{|F \setminus P|} = 0 = \delta_{P,R}$. Thus, (11.1218) is proven under the assumption that $P \not\subseteq R$.

Now, let us forget that we have assumed $P \not\subseteq R$. We thus have shown that (11.1218) holds if $P \not\subseteq R$. Hence, for the rest of this proof, we can WLOG assume that we don’t have $P \not\subseteq R$. Assume this.

We have $P \subset R$ (since we don’t have $P \not\subset R$). Hence, there exists a bijection from the set of all $F \subset R$ satisfying $F \supsetneq P$ to the set of all $T \subset R \setminus P$; this bijection sends every $F$ to $F \setminus P$. Hence, we can substitute $T$ for $F \setminus P$ in the sum $\sum_{F \supsetneq P} (-1)^{|F \setminus P|}$.

We thus obtain

\[
\sum_{F \subset R, \; F \supsetneq P} (-1)^{|F \setminus P|} = \sum_{T \subset R \setminus P} (-1)^{|T|} = \delta_{R,P,\emptyset} \quad \text{(by (11.1217), applied to } R \setminus P \text{ instead of } R)
\]

\[
= \delta_{P,R} \quad \text{(since } R \setminus P = \emptyset \text{ if and only if } P = R \text{ (because } P \subset R)).
\]

This proves (11.1218).
(a) Let \( G = (V, E) \) be a finite graph. We have
\[
\sum_{H = (V, E') \subseteq E' \cap E = \emptyset} [H]^f = \sum_{K = (V, F); F' \subseteq E} \left( (-1)^{|E' \setminus E^c|} \sum_{H = (V, E'); E' \supseteq F^c} [K]^f \right).
\]
(here, we renamed \( H \) and \( E' \) as \( K \) and \( F \) in the sum)
\[
= \sum_{K = (V, F); H = (V, E') \subseteq E' \supseteq F^c} (-1)^{|E' \setminus E^c|} [H]
\]
\[
= \sum_{H = (V, E')} \sum_{K = (V, F); F' \subseteq E^c; E' \supseteq F^c} (-1)^{|F' \setminus (E')^c|} [H]
\]
\[
= \sum_{H = (V, E')} \sum_{F \subseteq E^c; F' \subseteq (E')^c} (-1)^{|F' \setminus (E')^c|} \delta_{F' \subseteq F^c}
\]
\[
= \sum_{H = (V, E')} \sum_{E' \subseteq E^c} \delta_{E' \subseteq E^c} [H] = \sum_{H = (V, E')} \delta_{E' \subseteq E^c} [H] = [(V, E)] = [G].
\]
This proves Proposition 7.28(a).

(b) Let \( \varphi : \mathcal{G} \to \mathcal{G} \) be the \( k \)-linear map which sends every \([G]\) to the corresponding \([G]^f\). For every \( n \in \mathbb{N} \), let \( \mathcal{G}_n \) denote the \( n \)-th homogeneous component of the graded \( k \)-module \( \mathcal{G} \). (The grading is by the number of vertices of a graph.)

The map \( \varphi \) is graded. Hence, for every \( n \in \mathbb{N} \), the map \( \varphi \) restricts to a \( k \)-linear map \( \varphi_n : \mathcal{G}_n \to \mathcal{G}_n \). This latter map \( \varphi_n \) is unitriangular — or, more precisely, it is representable by a unitriangular matrix with respect to two appropriate bases of \( \mathcal{G}_n \).\(^{917}\) Hence, \( \varphi_n \) is invertible (since every unitriangular matrix is invertible). Therefore, \( \varphi \) is invertible. Hence, the elements \([G]^f\) form a basis of the \( k \)-module \( \mathcal{G} \) (since they are the images of the elements \([G]\) under \( \varphi \), and those latter elements form a basis). Proposition 7.28(b) is proven.

\(^{917}\)These appropriate bases are the following: For the basis of the domain, take a list \([(V, E_1)], [(V, E_2)], \ldots, [(V, E_N)]\) of all isomorphism classes of graphs with \( n \) vertices for which \(|E_1| \leq |E_2| \leq \cdots \leq |E_N|\). For the basis of the codomain, take the list \([(V, (E_1)^c)], [(V, (E_2)^c)], \ldots, [(V, (E_N)^c)]\). The matrix which represents the map \( \varphi_n \) with respect to these two bases is unitriangular, because every \( G \) satisfies
\[
\varphi_n ([G]) = \varphi ([G]) = [G]^f = \sum_{H = (V, E')} (-1)^{|E' \setminus E^c|} [H].
\]
(Apart from the graph \((V, E^c)\), which appears exactly once as an addend in the right hand side of this equality, every other graph \( H = (V, E') \) appearing on this right hand side has more edges than \((V, E')\) and thus comes before \((V, E')\) in the basis of the codomain.)
(c) We define a $k$-linear map $\Delta': \mathcal{G} \to \mathcal{G} \otimes \mathcal{G}$ by

$$\Delta'[H]^g = \sum_{(V_1,V_2); \ V = V_1 \sqcup V_2; \ H = H_{|V_1 \sqcup H_{|V_2}}} [H_{|V_1}]^g \otimes [H_{|V_2}]^g.$$ 

918 In order to prove Proposition 7.28(c), it clearly suffices to show that $\Delta' = \Delta$.

Let $G = (V, E)$ be a finite graph. If $T$ is a set of two-element subsets of $V$, and if $R$ is a subset of $V$, then $T \mid R$ shall denote the subset $\{X \in T \mid X \subset R\}$ of $T$. (Thus, if $R$ is a subset of $V$, then $E \mid R$ is the set of edges of the graph $G \mid R$.)

918 This is well-defined, since the $[G]^g$ form a basis of the $k$-module $\mathcal{G}$. 
Now, Proposition 7.28(a) yields \( |G| = \sum_{H=(V,E')} [H]^2 \). Hence,

\[
\Delta'[G] = \Delta' \sum_{H=(V,E'); \ E' \cap E = \emptyset} [H]^2 = \sum_{H=(V,E'); \ E' \cap E = \emptyset} \Delta'[H]^2 = \sum_{H=(V,E'); \ E' \cap E = \emptyset} [H|V_1]^2 \otimes [H|V_2]^2
\]

\[
= \sum_{H=(V,E'); \ E' \cap E = \emptyset} \sum_{\substack{(V_1,V_2); \ V=V_1 \cup V_2; \ H=H|V_1 \cup H|V_2}} [H|V_1]^2 \otimes [H|V_2]^2
\]

\[
= \sum_{\substack{(V_1,V_2); \ V=V_1 \cup V_2}} \sum_{\substack{H=(V,E'); \ E' \cap E = \emptyset; \ H=H|V_1 \cup H|V_2}} [H|V_1]^2 \otimes [H|V_2]^2
\]

\[
= \sum_{H_1=(V_1,E'_1); \ E'_1 \cap (E|V_1) = \emptyset} \sum_{H_2=(V_2,E'_2); \ E'_2 \cap (E|V_2) = \emptyset} \sum_{\substack{H=(V,E'); \ E' \cap E = \emptyset; \ H=H|V_1 \cup H|V_2}} [H|V_1|^2 \otimes [H|V_2]^2
\]

(by every \( H=(V,E') \) satisfying \( E' \cap E = \emptyset \),
we can write the subgraph \( H|V_1 \) in the form \( H_1=(V_1,E'_1) \)
with some \( E'_1 \) satisfying \( E'_1 \cap (E|V_1) = \emptyset \), and we can
write the subgraph \( H|V_2 \) in the form \( H_2=(V_2,E'_2) \)
with some \( E'_2 \) satisfying \( E'_2 \cap (E|V_2) = \emptyset \)

\[
= \sum_{\substack{(V_1,V_2); \ V=V_1 \cup V_2; \ V_1 \cup V_2 = V}} \sum_{\substack{H_1=(V_1,E'_1); \ E'_1 \cap (E|V_1) = \emptyset \ E'_1 \cap (E|V_2) = \emptyset \}} \sum_{H=(V,E'); \ E' \cap E = \emptyset; \ H=H|V_1 \cup H|V_2} [H|V_1]^2 \otimes [H|V_2]^2
\]

\[
= \sum_{\substack{(V_1,V_2); \ V=V_1 \cup V_2; \ V_1 \cup V_2 = V}} \sum_{\substack{H_1=(V_1,E'_1); \ E'_1 \cap (E|V_1) = \emptyset \ E'_1 \cap (E|V_2) = \emptyset \}} \sum_{\substack{H=(V,E'); \ E' \cap E = \emptyset; \ H=H|V_1 \cup H|V_2}} [H|V_1]^2 \otimes [H|V_2]^2
\]

\[
= \sum_{\substack{(V_1,V_2); \ V=V_1 \cup V_2; \ V_1 \cup V_2 = V}} \sum_{\substack{H_1=(V_1,E'_1); \ E'_1 \cap (E|V_1) = \emptyset \ E'_1 \cap (E|V_2) = \emptyset \}} \sum_{\substack{H=(V,E'); \ E' \cap E = \emptyset; \ H=H|V_1 \cup H|V_2}} [H_1]^2 \otimes [H_2]^2
\]

\[
= \sum_{\substack{(V_1,V_2); \ V=V_1 \cup V_2; \ V_1 \cup V_2 = V}} \sum_{\substack{H_1=(V_1,E'_1); \ E'_1 \cap (E|V_1) = \emptyset \ E'_1 \cap (E|V_2) = \emptyset \}} \sum_{\substack{H=(V,E'); \ E' \cap E = \emptyset; \ H=H|V_1 \cup H|V_2}} [H_1]^2 \otimes [H_2]^2
\]

\[
= \sum_{\substack{(V_1,V_2); \ V=V_1 \cup V_2; \ V_1 \cup V_2 = V}} \sum_{\substack{H_1=(V_1,E'_1); \ E'_1 \cap (E|V_1) = \emptyset \ E'_1 \cap (E|V_2) = \emptyset \}} \sum_{\substack{H=(V,E'); \ E' \cap E = \emptyset; \ H=H|V_1 \cup H|V_2}} [H_1]^2 \otimes [H_2]^2
\]
Comparing this with

\[
\Delta [G] = \sum_{V_1 \sqcup V_2 = V} \frac{[G | V_1]}{1} \otimes \frac{[G | V_2]}{1} = \sum_{H_1 = (V_1, E_1') \cap E | V_1 = \emptyset} [H_1]^g \otimes [H_2]^g \\
= \sum_{H_1 = (V_1, E_1') \cap E | V_1 = \emptyset} [H_1]^g \otimes [H_2]^g,
\]

we obtain \( \Delta'[G] = \Delta [G] \). Since this holds for every graph \( G \), we thus obtain \( \Delta' = \Delta \). As we know, this proves Proposition 7.28(c).

(d) We define a \( k \)-bilinear operation \( \sharp : \mathcal{G} \times \mathcal{G} \to \mathcal{G} \) (written infix\(^{919}\)) by

\[
[H_1]^g \sharp [H_2]^g = \sum_{H = (V_1 \cup V_2, E)} [H]^g.
\]

\(^{920}\) In order to prove Proposition 7.28(d), it clearly suffices to show that this operation \( \sharp \) is identical with the usual multiplication on \( \mathcal{G} \).

Let \( G_1 \) and \( G_2 \) be any two finite graphs. We shall show that \( [G_1] \sharp [G_2] = [G_1] [G_2] \).

\(^{919}\) This means that we write \( a \sharp b \) to denote the image of a pair \( (a, b) \in \mathcal{G} \times \mathcal{G} \) under this operation \( \sharp \).

\(^{920}\) This is well-defined, since the \([G]^g\) form a basis of the \( k \)-module \( \mathcal{G} \).
Indeed, write \( G_1 \) and \( G_2 \) in the forms \( G_1 = (V_1, E_1) \) and \( G_2 = (V_2, E_2) \). Then,

\[
\begin{align*}
\mathbb{G}[G_1] &\equiv \sum_{H_1 = (V_1, E_1')} [H_1]^3 \quad \mathbb{G}[G_2] = \sum_{H_2 = (V_2, E_2')} [H_2]^3 \\
&\text{(by Proposition 7.28(a))} \\
&\text{(by Proposition 7.28(a))} \\
= \left( \sum_{H_1 = (V_1, E_1') \mid E_1' \cap E_1 = \emptyset} [H_1]^3 \right) \# \left( \sum_{H_2 = (V_2, E_2') \mid E_2' \cap E_2 = \emptyset} [H_2]^3 \right) \\
= \sum_{H_1 = (V_1, E_1')} \sum_{H_2 = (V_2, E_2')} [H] \sum_{H = (V_1 \cup V_2, E)} [H]^3 \\
&\text{(because for a graph } H = (V_1 \cup V_2, E) \text{ satisfying } H|_{V_1} = H_1 \text{ and } H|_{V_2} = H_2, \text{ we have } E_1' = E|_{V_1} \text{ and } E_2' = E|_{V_2}) \\
= \sum_{H = (V_1 \cup V_2, E)} [H]^3 \\
\text{(since the statement } (E|_{V_1}) \cap E_1 = \emptyset \text{ and } (E|_{V_2}) \cap E_2 = \emptyset \text{ is equivalent to } E \cap (E_1 \cup E_2) = \emptyset) \\
\end{align*}
\]

Compared with

\[
\mathbb{G}[G_1] \mathbb{G}[G_2] = [G_1 \cup G_2] = [(V_1 \cup V_2, E_1 \cup E_2)] \\
= \sum_{H = (V_1 \cup V_2, E) \mid E \cap (E_1 \cup E_2) = \emptyset} [H]^3 \\
\text{(by Proposition 7.28(a))} \\
\]

this yields \( \mathbb{G}[G_1] \mathbb{G}[G_2] = \mathbb{G}[G_1] \mathbb{G}[G_2] \).
We now forget that we fixed $G_1$ and $G_2$. We have thus shown that $[G_1] \ast [G_2] = [G_1] \cdot [G_2]$ for any two finite graphs $G_1$ and $G_2$. Thus, the operation $\ast$ is identical with the usual multiplication on $\mathcal{G}$. This completes the proof of Proposition 7.28(d). \qed

\textit{Proof of Proposition 7.30.} Proposition 7.28(a) shows that every finite graph $G = (V, E)$ satisfies
\begin{equation}
[G] = \sum_{H = (V', E') \mid E' \cap E = \emptyset} [H]^2 = \sum_{K = (V, E') \mid E' \cap E = \emptyset} [K]^2
\end{equation}
(here, we renamed the summation index $H$ as $K$). Let us now show that
\begin{equation}
([G], [H]) = ([H], [G])
\end{equation}
for any two finite graphs $G$ and $H$.

\textit{Proof of (11.1220):} We shall use the following notation: If $V$ and $W$ are two sets, and if $\varphi : V \to W$ is a map, then $\varphi_*$ will denote the map from the powerset of $V$ to the powerset of $W$ which sends every $T \subset V$ to $\varphi(T) \subset W$. This map $\varphi_*$ is a bijection if $\varphi$ is a bijection.

Let $G$ and $H$ be two finite graphs. Let us write $G$ and $H$ in the forms $G = (V, E)$ and $H = (W, F)$. Now,
\begin{equation}
\left(\sum_{K = (V, E') \mid E' \cap E = \emptyset} [K]^2, [H]\right) = \sum_{K = (V, E') \mid E' \cap E = \emptyset} \left(\sum_{K = (V', E') \mid E' \cap E = \emptyset} [K]^2, [H]\right) = \sum_{K = (V, E') \mid E' \cap E = \emptyset} [K]^2, [H]
\end{equation}
(since the form $(\cdot, \cdot)$ is $k$-bilinear)
\begin{equation}
= \sum_{K = (V, E') \mid E' \cap E = \emptyset} \left[\text{Iso}(K, H)\right].
\end{equation}

But for every graph $K = (V, E')$, we have
\[\text{Iso}(K, H) = \{\varphi : V \to W \mid \varphi \text{ is a bijection and satisfies } \varphi_*(E') = F\}\]
(because the isomorphisms from $K$ to $H$ are defined to be the bijections $\varphi : V \to W$ such that $\varphi_*(E') = F$).

Hence,
\begin{equation}
\sum_{K = (V, E') \mid E' \cap E = \emptyset} \left[\text{Iso}(K, H)\right] = \sum_{K = (V, E') \mid E' \cap E = \emptyset} \left[\varphi : V \to W \mid \varphi \text{ is a bijection and satisfies } \varphi_*(E') = F\right]
\end{equation}
\begin{align*}
&= \sum_{E' \text{ is a set of two-element subsets of } V; \ E' \cap E = \emptyset} \left[\varphi : V \to W \mid \varphi \text{ is a bijection and satisfies } E' = (\varphi_*)^{-1}(F)\right] \left[\text{this is equivalent to } E' = (\varphi_*)^{-1}(F) \text{ (since } \varphi_* \text{ is a bijection (and } \varphi \text{ is a bijection))}\right] \\
&= \sum_{E' \text{ is a set of two-element subsets of } V; \ E' \cap E = \emptyset} \left[\varphi : V \to W \mid \varphi \text{ is a bijection and satisfies } E' = (\varphi_*)^{-1}(F)\right] \\
&= \left[\varphi : V \to W \mid \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(F) \cap E = \emptyset\right].
\end{align*}
Now, (11.1221) becomes

\[ ([G], [H]) = \sum_{K=(V,E) : E \cap E = \emptyset} |\text{Iso}(K, H)| \]

(11.1222)
\[ = \left| \left\{ \varphi : V \to W \mid \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(F) \cap E = \emptyset \right\} \right| \]

The same argument (applied to \(H, W, F, G, V \) and \(E\) instead of \(G, V, E, H, W \) and \(F\)) yields

(11.1223)
\[ ([H], [G]) = \left| \left\{ \varphi : W \to V \mid \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(E) \cap F = \emptyset \right\} \right| \]

Now, we define two sets \(\mathfrak{P}\) and \(\mathfrak{Q}\) by

\[ \mathfrak{P} = \left\{ \varphi : V \to W \mid \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(F) \cap E = \emptyset \right\} \]

and

\[ \mathfrak{Q} = \left\{ \varphi : W \to V \mid \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(E) \cap F = \emptyset \right\} \]

Then, (11.1222) becomes

(11.1224)
\[ ([G], [H]) = \left| \left\{ \varphi : V \to W \mid \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(F) \cap E = \emptyset \right\} \right| = |\mathfrak{P}|. \]

Also, (11.1223) becomes

(11.1225)
\[ ([H], [G]) = \left| \left\{ \varphi : W \to V \mid \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(E) \cap F = \emptyset \right\} \right| = |\mathfrak{Q}|. \]

But every \(\psi \in \mathfrak{P}\) satisfies \(\psi^{-1} \in \mathfrak{Q}\) \(^{921}\). Hence, we can define a map

\[ A : \mathfrak{P} \to \mathfrak{Q}, \]

\[ \psi \mapsto \psi^{-1}. \]

\(^{921}\)Proof.\ Let \(\psi \in \mathfrak{P}\). Then, \(\psi \in \mathfrak{P} = \left\{ \varphi : V \to W \mid \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(F) \cap E = \emptyset \right\} \). Hence, \(\psi\) is a bijection \(V \to W\) and satisfies \((\psi_*)^{-1}(F) \cap E = \emptyset\).

Let \(\rho\) be the map \(\psi^{-1} : W \to V\). This map \(\rho\) is well-defined (since \(\psi\) is a bijection) and is a bijection itself (since \(\rho = \psi^{-1}\)).

But \(\psi\) is a bijection. Thus, \(\psi_*\) is a bijection as well, and satisfies \((\psi_*)^{-1} = \left( \begin{array}{c} \psi^{-1} \\ \mu \end{array} \right) \). Of course, \(\rho_*\) is also a bijection (since \(\rho\) is a bijection).

Now,

\[ \rho_* \left( (\rho_*)^{-1}(E) \cap F \right) = \left( \rho_* \left( (\rho_*)^{-1}(F) \right) \right) \cap (\psi_*)^{-1}(E) \]

\[ = E \cap (\psi_*)^{-1}(F) = (\psi_*)^{-1}(F) \cap E = \emptyset. \]

Thus,

\[ (\rho_*)^{-1}(E) \cap F = \rho_* \emptyset \]

\[ = \emptyset. \]

Hence, \(\rho : W \to V\) is a bijection and satisfies \((\rho_*)^{-1}(E) \cap F = 0\). In other words,

\[ \rho \in \left\{ \varphi : W \to V \mid \varphi \text{ is a bijection and satisfies } (\varphi_*)^{-1}(E) \cap F = \emptyset \right\} = \mathfrak{Q}. \]

Thus, \(\psi^{-1} = \rho \in \mathfrak{Q}\), qed.
Similarly, we can define a map

\[ B : \Omega \to \mathcal{Q}, \]

\[ \psi \mapsto \psi^{-1}. \]

Consider these two maps \( A \) and \( B \). Clearly, these maps \( A \) and \( B \) are mutually inverse. Hence, \( A \) is a bijection. Thus, there exists a bijection \( \mathcal{Q} \to \Omega \) (namely, \( A \)). Consequently, \(|\mathcal{Q}| = |\Omega|\). Now, (11.1224) becomes

\[ ([G], [H]) = |\mathcal{Q}| = ([H], [G]) \]

(by (11.1225)). This proves (11.1220).

Now, let \( a \) and \( b \) be two elements of \( \mathcal{G} \). We want to show that \((a, b) = (b, a)\). This equality is \( k \)-linear in each of \( a \) and \( b \). Therefore, we can WLOG assume that \( a \) and \( b \) belong to the family \(([G], [G])\) is an isomorphism class of finite graphs (because this family is a basis of the \( k \)-module \( \mathcal{G} \)). Assume this. Then, \( a = [G] \) and \( b = [H] \) for two finite graphs \( G \) and \( H \). Consider these \( G \) and \( H \). Now,

\[
\left( \begin{array}{c}
[a] \\
[b]
\end{array} \right)
= ([G], [H]) = \left( \begin{array}{c}
[H] \\
[a]
\end{array} \right)
= (b, a).
\]

We thus have shown \((a, b) = (b, a)\).

Let us now forget that we fixed \( a \) and \( b \). We thus have proven that \((a, b) = (b, a)\) for any \( a \in \mathcal{G} \) and \( b \in \mathcal{G} \). In other words, the form \((\cdot, \cdot) : \mathcal{G} \times \mathcal{G} \to k\) is symmetric. This proves Proposition 7.30.

**Proof of Proposition 7.32.** In this proof, we shall use all the notations that appeared in Definition 7.31. We shall also use the fact that the family \(([G], [G])\) is an isomorphism class of finite graphs \( \mathcal{G} \). (This fact is Proposition 7.28(b).)

(a) We first notice that any two finite graphs \( G \) and \( H \) satisfy

\[
\left( \psi \left( [G] \right) \right) [H] = |\text{Iso}(G, H)|
\]

923. Now, let \( a \in \mathcal{G} \) and \( b \in \mathcal{G} \). We want to prove the equality \((\psi(a))(b) = (a, b)\). This equality is \( k \)-linear in each of \( a \) and \( b \). Thus, we can WLOG assume that \( a \) belongs to the family \(([G], [G])\) is an isomorphism class of finite graphs and that \( b \) belongs to the family \(([G], [G])\) is an isomorphism class of finite graphs (because both of these families are

\[ (B \circ A) (\psi) = B \left( \begin{array}{c}
A (\psi) \\
\psi^{-1}
\end{array} \right) = B (\psi^{-1}) = (\psi^{-1})^{-1}
\]

(by the definition of \( B \))

\[ = \psi = \text{id}(\psi). \]

Hence, \( B \circ A = \text{id} \). Similarly, \( A \circ B = \text{id} \). Combining this with \( B \circ A = \text{id} \), we conclude that the maps \( A \) and \( B \) are mutually inverse. Qed.

\[ 923 \text{ Proof of (11.1226):} \]

Let \( G \) and \( H \) be two finite graphs. Then,

\[
\psi \left( [G] \right) [H] = (\text{aut}(G) \cdot [G]^*) [H] = \text{aut}(G) \cdot [G]^* [H] = \text{aut}(G) \cdot \delta_{[G]} [H].
\]

We are in one of the following two cases:

*Case 1:* The graphs \( G \) and \( H \) are isomorphic.

*Case 2:* The graphs \( G \) and \( H \) are not isomorphic.

Let us first consider Case 1. In this case, the graphs \( G \) and \( H \) are isomorphic. In other words, there exists a graph isomorphism \( \alpha \) from \( G \) to \( H \). Consider this \( \alpha \). Then,

\[ \text{Iso}(G, G) \to \text{Iso}(G, H), \]

\[ \psi \mapsto \alpha \circ \psi \]
bases of the $k$-module $G$). Assume this. Then, $a = [G]^\sharp$ and $b = [H]$ for two finite graphs $G$ and $H$. Consider these $G$ and $H$. Now,

$$
\left(\psi\left(\begin{array}{c} a \\
\end{array}\right)\right) \left(\begin{array}{c} b \\
\end{array}\right) = \left(\psi\left([G]^\sharp\right)\right) [H] = |\text{Iso}(G, H)| \quad \text{(by (11.1226))}
$$

Hence, (11.1226) is proven in Case 2.

Let us now consider Case 2. In this case, the graphs $G$ and $H$ are not isomorphic. Thus, $[G] \neq [H]$, so that $\delta_{[G],[H]} = 0$.

Now, (11.1226) is proven in both Cases 1 and 2. Hence, (11.1226) always holds.

(a) We have $\psi(1_G) = 1_{G^\circ}$.

(b) We have $\psi(1_G) = 1_{G^\circ}$.

Hence, $\psi : G \rightarrow G^\circ$ is a $k$-algebra map (since $\psi(1_G) = 1_{G^\circ}$).

is a bijection (this is very easy to check). Thus, there exists a bijection $\text{Iso}(G, G) \rightarrow \text{Iso}(G, H)$. Hence, $|\text{Iso}(G, G)| = |\text{Iso}(G, H)|$.

But the definition of $\text{aut}(G)$ yields $\text{aut}(G) = |\text{Iso}(G, G)|$. Now,

$$
\left(\psi\left([G]^\sharp\right)\right) [H] = \text{aut}(G) \cdot \delta_{[G],[H]} = \text{aut}(G) = |\text{Iso}(G, G)| = |\text{Iso}(G, H)|.
$$

Hence, (11.1226) is proven in Case 1.

Let us now consider Case 2. In this case, the graphs $G$ and $H$ are not isomorphic. Thus, $[G] \neq [H]$, so that $\delta_{[G],[H]} = 0$.

But the graphs $G$ and $H$ are not isomorphic. Hence, the set of all isomorphisms from $G$ to $H$ is empty. In other words, $|\text{Iso}(G, H)| = 2$ (since $|\text{Iso}(G, H)|$ is the set of all isomorphisms from $G$ to $H$). Hence, $|\text{Iso}(G, H)| = |\emptyset| = 0$.

Now, (11.1226) is proven in both Cases 1 and 2. Hence, (11.1226) always holds.

924 Proof. Let $\emptyset$ denote the empty graph.

We have $1_{G^\circ} = \epsilon_G$ (by the definition of the $k$-algebra $G^\circ$). Thus, for every finite graph $G$, we have

$$
\epsilon_G(\emptyset) = \epsilon_G([G]) = \delta_{[G],[\emptyset]} \quad \text{(by the definition of $\epsilon_G$)}
$$

Hence, the two maps $1_{G^\circ}$ and $[0]^\ast$ are equal to each other on every element of the basis $\text{Iso}(G)$ of $G$. Since these two maps are $k$-linear, this yields that these two maps $1_{G^\circ}$ and $[0]^\ast$ are identical. In other words, $1_{G^\circ} = [0]^\ast$.

On the other hand, the definition of $[0]^2$ readily yields $[0]^2 = [0] = 1_G$. But the definition of $\psi$ yields $\psi([0]^2) = \psi(0) \cdot [0]^\ast = [0]^\ast = 1_{G^\circ}$ (since $1_{G^\circ} = [0]^\ast$), so that $1_{G^\circ} = \psi\left(\begin{array}{c} [0]^2 \\
\end{array}\right) = \psi(1_G)$, qed.

925 Proof. Let $a \in G$ and $b \in G$. We need to prove the equality $\psi(ab) = \psi(a) \cdot \psi(b)$. Since this equality is $k$-linear in each of $a$ and $b$, we can WLOG assume that $a$ and $b$ are elements of the family $\text{Iso}(G)$ (since this family is a basis of the $k$-module $G$). Assume this.

There exist two finite graphs $H_1$ and $H_2$ such that $a = [H_1]^\sharp$ and $b = [H_2]^\sharp$ (since $a$ and $b$ are elements of the family $\text{Iso}(G)$).

Multiplying the equalities $a = [H_1]^\sharp$ and $b = [H_2]^\sharp$, we obtain

$$
ab = [H_1]^\sharp [H_2]^\sharp = \sum_{H = (V_1 \cup V_2, E), \begin{array}{c} H_{V_1} = H_1, \\
H_{V_2} = H_2 \\
\end{array}} [H]^\sharp \quad \text{(by Proposition 7.28(d))}.
$$

Now, fix a finite graph $G$. We shall show that $\psi(ab) [G] = (\psi(a) \cdot \psi(b)) [G]$. 
Next, we notice that the \( k \)-linear map \( \psi \) is graded (because for any finite graph \( G \), both \( [G]^2 \) and \( [G]^* \) are homogeneous elements having degree \( |G| \)). Thus, the \( k \)-linear map \( \psi : G \to G^o \) gives rise to an adjoint \( k \)-linear map \( \psi^* : (G^o)^o \to G^o \).

Now, the canonical \( k \)-module homomorphism \( G \to (G^o)^o \) (which sends every \( a \in G \) to the map \( G^o \to k, f \mapsto f(a) \)) is a \( k \)-module isomorphism (since \( G \) is of finite type). We thus identify \( G \) with \( (G^o)^o \) along

Write the graph \( G \) in the form \( G = (V, F) \). Then,

\[
\begin{pmatrix}
\psi \\
(ab)
\end{pmatrix}
= \sum_{H = (V_1 \cup V_2, E); H|_{V_1} = H_1; H|_{V_2} = H_2} \left( \psi \left( \sum_{H = (V_1 \cup V_2, E); H|_{V_1} = H_1; H|_{V_2} = H_2} [H]^2 \right) \right)
\]

\[= \sum_{H = (V_1 \cup V_2, E); H|_{V_1} = H_1; H|_{V_2} = H_2} \left( \sum_{H = (V_1 \cup V_2, E); H|_{V_1} = H_1; H|_{V_2} = H_2} \psi \left( [H]^2 \right) \right)
\]

\[= \sum_{H = (V_1 \cup V_2, E); H|_{V_1} = H_1; H|_{V_2} = H_2} \left( \sum_{H = (V_1 \cup V_2, E); H|_{V_1} = H_1; H|_{V_2} = H_2} \psi \left( [H]^2 \right) \right)
\]

On the other hand, Exercise 1.67(b) shows that the \( k \)-algebra structure defined on \( G^* \) is precisely the one defined on \( \text{Hom}(G, k) = G^* \) according to Definition 1.27. Thus, the product of two elements of \( G^* \) is the convolution of these elements (viewed as maps from \( G \) to \( k \)). Applying this to the two elements \( \psi(a) \) and \( \psi(b) \), we obtain \( \psi(a) \cdot \psi(b) = \psi(a) \ast \psi(b) = m_k (\psi(a) \otimes \psi(b)) \circ \Delta_G \), so that

\[(\psi(a) \cdot \psi(b)) [G] = (m_k \circ (\psi(a) \otimes \psi(b)) \circ \Delta_G) ([G])
\]

\[= \sum_{V_1 \cup V_2 = V} \psi \left( [H_1]^2 \right) \otimes \psi \left( [H_2]^2 \right) \left( \sum_{V_1 \cup V_2 = V} \right) [G|_{V_1}] \otimes [G|_{V_2}]
\]

\[= \sum_{V_1 \cup V_2 = V} \psi \left( [H_1]^2 \right) \otimes \psi \left( [H_2]^2 \right) \left( \sum_{V_1 \cup V_2 = V} \right) [G|_{V_1}] \otimes [G|_{V_2}]
\]

\[= \sum_{V_1 \cup V_2 = V} \psi \left( [H_1]^2 \right) \left( \sum_{V_1 \cup V_2 = V} \right) [G|_{V_1}] \otimes \psi \left( [H_2]^2 \right) \left( \sum_{V_1 \cup V_2 = V} \right) [G|_{V_2}]
\]

\[= \sum_{H = (V_1 \cup V_2, E); H|_{V_1} = H_1; H|_{V_2} = H_2} \left( \psi \left( [H_1]^2 \right) \right) [G|_{V_1}] \cdot \left( \psi \left( [H_2]^2 \right) \right) [G|_{V_2}]
\]

\[= \sum_{H = (V_1 \cup V_2, E); H|_{V_1} = H_1; H|_{V_2} = H_2} \left( \psi \left( [H_1]^2 \right) \right) [G|_{V_1}] \cdot \left( \psi \left( [H_2]^2 \right) \right) [G|_{V_2}]
\]

\[= \sum_{H = (V_1 \cup V_2, E); H|_{V_1} = H_1; H|_{V_2} = H_2} \left( \psi \left( [H_1]^2 \right) \right) [G|_{V_1}] \cdot \left( \psi \left( [H_2]^2 \right) \right) [G|_{V_2}]
\]
this isomorphism. Then, \((G^\circ)^\circ = G\) as Hopf algebras\(^{926}\). Therefore, \(\psi : (G^\circ)^\circ \rightarrow G^\circ\) is a \(k\)-algebra map (since \(\psi : G \rightarrow G^\circ\) is a \(k\)-algebra map). Also, \(G^\circ\) is of finite type (since \(G\) is of finite type). Thus, Exercise 1.67(f) \((\text{applied to} C = G, D = G^\circ \text{and} f = \psi)\) yields that \(\psi : G \rightarrow G^\circ\) is a \(k\)-coalgebra map if and only if \(\psi^* : (G^\circ)^\circ \rightarrow G^\circ\) is a \(k\)-algebra map.

Now, we are going to construct a bijection between the sets

\[
\bigcup_{H=(V_1 \sqcup V_2, E); \ H|_{V_1}=H_1; \ H|_{V_2}=H_2} \text{Iso}(H, G) \quad \text{and} \quad \bigcup_{(W_1, W_2); \ W_1 \sqcup W_2 = V} \text{Iso}(H_1, G|_{W_1}) \times \text{Iso}(H_2, G|_{W_2}).
\]

First, let us agree to encode the elements of the set

\[
\bigcup_{H=(V_1 \sqcup V_2, E); \ H|_{V_1}=H_1; \ H|_{V_2}=H_2} \text{Iso}(H, G)
\]

as triples \((H, E, \varphi)\) consisting of a graph \(H = (V_1 \sqcup V_2, E)\), its set of edges \(E\), and an isomorphism \(\varphi \in \text{Iso}(H, G)\). Let us also agree to encode the elements of the set

\[
\bigcup_{(W_1, W_2); \ W_1 \sqcup W_2 = V} \text{Iso}(H_1, G|_{W_1}) \times \text{Iso}(H_2, G|_{W_2})
\]

as triples \(((W_1, W_2), \varphi_1, \varphi_2)\) consisting of a pair \((W_1, W_2)\) of subsets of \(V\) (which satisfies \(W_1 \sqcup W_2 = V\)), an isomorphism \(\varphi_1 \in \text{Iso}(H_1, G|_{W_1})\), and an isomorphism \(\varphi_2 \in \text{Iso}(H_2, G|_{W_2})\).

Now, we claim that there exists a bijection from the set

\[
\bigcup_{H=(V_1 \sqcup V_2, E); \ H|_{V_1}=H_1; \ H|_{V_2}=H_2} \text{Iso}(H, G)
\]

to the set

\[
\bigcup_{(W_1, W_2); \ W_1 \sqcup W_2 = V} \text{Iso}(H_1, G|_{W_1}) \times \text{Iso}(H_2, G|_{W_2}).
\]

Namely, this bijection sends every \((H, E, \varphi)\) (where \(H = (V_1 \sqcup V_2, E)\) and \(\varphi \in \text{Iso}(H, G)\)) to \(((\varphi(V_1), \varphi(V_2), \varphi_1, \varphi_2))\), where \(\varphi_1 : V_1 \rightarrow \varphi(V_1)\) is the isomorphism from \(H_1\) to \(G|_{\varphi(V_1)}\) which is obtained by restricting \(\varphi\) to \(V_1\), and where \(\varphi_2\) is the isomorphism from \(H_2\) to \(G|_{\varphi(V_2)}\) which is obtained by restricting \(\varphi\) to \(V_2\). (The inverse map of this bijection sends every \(((W_1, W_2), \varphi_1, \varphi_2)\) (with \(W_1 \sqcup W_2 = V\) and \(\varphi_1 \in \text{Iso}(H_1, G|_{W_1})\) and \(\varphi_2 \in \text{Iso}(H_2, G|_{W_2})\)) to \(((V_1 \sqcup V_2, E), \varphi, \varphi')\), where \(\varphi : V_1 \sqcup V_2 \rightarrow V\) is the map glued together from the maps \(\varphi_1 : V_1 \rightarrow W_1\) and \(\varphi_2 : V_2 \rightarrow W_2\), and where \(E = (\varphi^{-1})_*=F\).) The existence of this bijection yields

\[
\bigcup_{H=(V_1 \sqcup V_2, E); \ H|_{V_1}=H_1; \ H|_{V_2}=H_2} \text{Iso}(H, G) = \bigcup_{(W_1, W_2); \ W_1 \sqcup W_2 = V} \text{Iso}(H_1, G|_{W_1}) \times \text{Iso}(H_2, G|_{W_2}).
\]

Hence,

\[
(\psi(ab))[G] = \bigcup_{H=(V_1 \sqcup V_2, E); \ H|_{V_1}=H_1; \ H|_{V_2}=H_2} \text{Iso}(H, G) = \bigcup_{(W_1, W_2); \ W_1 \sqcup W_2 = V} \text{Iso}(H_1, G|_{W_1}) \times \text{Iso}(H_2, G|_{W_2}) = (\psi(a) \cdot \psi(b))[G].
\]

Let us now forget that we fixed \(G\). We thus have shown that \((\psi(ab))[G] = (\psi(a) \cdot \psi(b))[G]\) for any finite graph \(G\). In other words, the two maps \(\psi(ab)\) and \(\psi(a) \cdot \psi(b)\) are equal to each other on every element of the basis \(\{[G]\}|G\) is an isomorphism class of finite graphs of \(G\). Since these two maps are \(k\)-linear, this yields that these two maps \(\psi(ab)\) and \(\psi(a) \cdot \psi(b)\) are identical. In other words, \(\psi(ab) = \psi(a) \cdot \psi(b)\). Q.e.d.

\(^{926}\)This is because the Hopf algebra structure on \(G^\circ\) was defined by taking adjoints of the structure maps of the Hopf algebra structure on \(G\) (for example, the comultiplication \(\Delta_g^\circ\) on \(G^\circ\) is the adjoint of the multipication \(m_g\) on \(G\), if \((G \otimes G)^\circ\) is identified with \(G^\circ \otimes (G^\circ)^\circ\), and the Hopf algebra structure on \((G^\circ)^\circ\) was defined by taking adjoints of the structure maps of the Hopf algebra structure on \(G^\circ\); but the adjoint of the adjoint of a linear map \(F\) between two graded \(k\)-modules of finite type is the map \(F\) again.
Let us now prove that \( \psi^* = \psi \) (as maps from \( \mathcal{G} \) to \( \mathcal{G}^o \)). Indeed, let \( a \in \mathcal{G} \). Then, every \( b \in \mathcal{G} \) satisfies

\[
(\psi^* (a)) (b) = a (\psi (b)) \quad \text{(by the definition of the adjoint map } \psi^* : (\mathcal{G}^o)^o \rightarrow \mathcal{G}^o, \text{ where } a \in \mathcal{G} \text{ is regarded as an element of } (\mathcal{G}^o)^o,)
\]

\[
= (\psi (b)) (a) \quad \text{(because the } \mathbb{k} \text{-module homomorphism } \mathcal{G} \rightarrow (\mathcal{G}^o)^o \text{ which we use to identify } \mathcal{G} \text{ with } (\mathcal{G}^o)^o \text{ sends } a \in \mathcal{G} \text{ to the map } \mathcal{G}^o \rightarrow \mathbb{k}, \ f \mapsto f (a))
\]

\[
= (b, a) \quad \text{(by Proposition 7.32(a), applied to } b \text{ and } a \text{ instead of } a \text{ and } b)
\]

\[
= (a, b) \quad \text{(since the form } (\cdot, \cdot) : \mathcal{G} \times \mathcal{G} \rightarrow \mathbb{k} \text{ is symmetric)}
\]

\[
= (\psi (a)) (b) \quad \text{(by Proposition 7.32(a)).}
\]

Hence, \( \psi^* (a) = \psi (a) \). Now, let us forget that we fixed \( a \). We thus have shown that \( \psi^* (a) = \psi (a) \) for every \( a \in \mathcal{G} \). In other words, \( \psi^* = \psi \). Hence, \( \psi^* : (\mathcal{G}^o)^o \rightarrow \mathcal{G}^o \) is a \( \mathbb{k} \)-algebra map (since \( \psi : (\mathcal{G}^o)^o \rightarrow \mathcal{G}^o \) is a \( \mathbb{k} \)-algebra map). Thus, \( \psi : \mathcal{G} \rightarrow \mathcal{G}^o \) is a \( \mathbb{k} \)-coalgebra map (because we know that \( \psi : \mathcal{G} \rightarrow \mathcal{G}^o \) is a \( \mathbb{k} \)-coalgebra map if and only if \( \psi^* : (\mathcal{G}^o)^o \rightarrow \mathcal{G}^o \) is a \( \mathbb{k} \)-algebra map). This yields that \( \psi : \mathcal{G} \rightarrow \mathcal{G}^o \) is a \( \mathbb{k} \)-bialgebra map (since we already know that \( \psi : \mathcal{G} \rightarrow \mathcal{G}^o \) is a \( \mathbb{k} \)-algebra map). Thus, \( \psi : \mathcal{G} \rightarrow \mathcal{G}^o \) is a Hopf algebra map (by Proposition 1.46(c), applied to \( H_1 = \mathbb{G} \), \( H_2 = \mathbb{G}^o \), \( S_1 = \mathbb{S} \), \( S_2 = \mathbb{S}^o \) and \( \beta = \psi \)). This proves Proposition 7.32(b).

(c) Assume that \( Q \) is a subring of \( \mathbb{k} \). Then, \( \text{aut} (\mathbb{G}) \) is invertible for every finite graph \( G \) (because \( \text{aut} (\mathbb{G}) \) is a positive integer). Hence, \( \left( \text{aut} (\mathbb{G}) \cdot [\mathbb{G}]^* \right)_{[\mathbb{G}]} \) is an isomorphism class of finite graphs is a basis of the \( \mathbb{k} \)-module \( \mathcal{G}^o \) (because \( \left( [\mathbb{G}]^* \right)_{[\mathbb{G}]} \) is an isomorphism class of finite graphs is a basis of this \( \mathbb{k} \)-module). The map \( \psi \) thus sends the basis \( \left( [\mathbb{G}]^2 \right)_{[\mathbb{G}]} \) of the \( \mathbb{k} \)-module \( \mathcal{G}^o \) to the basis \( \left( \text{aut} (\mathbb{G}) \cdot [\mathbb{G}]^* \right)_{[\mathbb{G}]} \) of the \( \mathbb{k} \)-module \( \mathcal{G}^o \) (because it satisfies \( \psi \left( [\mathbb{G}]^2 \right) = \text{aut} (\mathbb{G}) \cdot [\mathbb{G}]^* \) for every finite graph \( G \)). Hence, the \( \mathbb{k} \)-linear map \( \psi \) sends a basis of its domain to a basis of its codomain. Therefore, the map \( \psi \) is a \( \mathbb{k} \)-module isomorphism. Combined with the fact that \( \psi \) is a Hopf algebra homomorphism, this yields that \( \psi \) is a Hopf algebra isomorphism. Proposition 7.32(c) is thus proven.

Now, all of Proposition 7.28, Proposition 7.30 and Proposition 7.32 are proven. Therefore, Exercise 7.33 is solved.

11.159. Solution to Exercise 7.44. Solution to Exercise 7.44. Proposition 7.28(b) yields that the elements \( [\mathbb{G}]^2 \), where \( [\mathbb{G}] \) ranges over all isomorphism classes of finite graphs, form a basis of the \( \mathbb{k} \)-module \( \mathcal{G}^o \). Hence, we can define a \( \mathbb{k} \)-linear map \( \Psi' : \mathcal{G} \rightarrow \mathbb{k} \) by requiring that

\[
\Psi' \left( [\mathbb{G}]^2 \right) = \sum_{\text{f:V→\{1,2,3,\ldots\}; } \exp f=E} x_f \quad \text{for every finite graph } G = (V, E)
\]

(because \( \sum_{\text{f:V→\{1,2,3,\ldots\}; } \exp f=E} x_f \) depends only on the isomorphism class \( [\mathbb{G}] \), but not on the graph \( G \) itself).

Consider this map \( \Psi' \). We shall show that \( \Psi' = \Psi \).
Indeed, let $G = (V, E)$ be any finite graph. Then, Proposition 7.28(a) yields $[G] = \sum_{H=(V,E')} \left[ H \right]^2$. Applying the map $\Psi'$ to both sides of this equality, we obtain

\[
\Psi'[G] = \sum_{H=(V,E') \mid E' \cap E = \emptyset} \left[ H \right]^2 = \sum_{H=(V,E') \mid E' \cap E = \emptyset} \Psi' \left( \sum_{\text{eqs } f \cap E = \emptyset \mid f} x_f \right)\]

(since the map $\Psi'$ is $k$-linear)

\[
= \sum_{H=(V,E') \mid E' \cap E = \emptyset} \sum_{f: V \rightarrow \{1,2,3,\ldots\} \mid \text{eqs } f = E'} x_f = \sum_{E' \text{ is a set of two-element subsets of } V \mid E' \cap E = \emptyset} \sum_{f: V \rightarrow \{1,2,3,\ldots\} \mid \text{eqs } f = E'} x_f
\]

(11.1227)

However, if $f : V \rightarrow \{1, 2, 3, \ldots \}$ is any map, then we have the following logical equivalence:

\[
(\text{eqs } f) \cap E = \emptyset \iff (f \text{ is a proper coloring of } G).
\]

(11.1228)
Hence, \((11.1227)\) becomes

\[
\Psi' [G] = \sum_{f: V \to \{1, 2, 3, \ldots\}; (eqs f) \cap E = \emptyset \atop f \text{ is a proper coloring of } G} \sum_{e} x_f = \sum_{f: V \to \{1, 2, 3, \ldots\}; f \text{ is a proper coloring of } G} \sum_{e} x_f
\]

(because of the equivalence \((11.1228)\))

\[
= \Psi [G] \quad \text{(since Proposition 7.36 yields } \Psi [G] = \sum_{f: V \to \{1, 2, 3, \ldots\}; f \text { proper colorings } f: V \to \{1, 2, 3, \ldots\})
\]

Let us now forget that we fixed \([G]\). Thus, we have shown that \(\Psi' [G] = \Psi [G]\) for any finite graph \(G\). In other words, the two maps \(\Psi'\) and \(\Psi\) are equal to each other on each element of the family \(([G])_G\) as an isomorphism class of finite graphs. Thus, the two maps \(\Psi'\) and \(\Psi\) must be identical (because these two maps are \(k\)-linear, and because the family \(([G])_G\) is an isomorphism class of finite graphs is a basis of the \(k\)-module \(\mathcal{G}\)). In other words, \(\Psi' = \Psi\). Now, every finite graph \(G = (V, E)\) satisfies

\[
\sum_{f: V \to \{1, 2, 3, \ldots\}} \Psi' ([G]_f^2) = \Psi' ([G]_f^2) = \sum_{eqs f = E} x_f.
\]

This solves Exercise 7.44.

11.160. Solution to Exercise 8.10. Solution to Exercise 8.10. We begin with some preparations.

Proof of \((11.1228)\): Let \(f : V \to \{1, 2, 3, \ldots\}\) be any map.

Let us first assume that \((eqs f) \cap E = \emptyset\). We shall then show that \(f\) is a proper coloring of \(G\).

Indeed, let us assume that there exists an edge \(e = \{v, v'\}\) in \(E\) such that \(f(v) = f(v')\). Consider this edge \(e = \{v, v'\}\). We are going to derive a contradiction.

The definition of \(eqs f\) yields \(eqs f = \{\{u, u'\} \mid u \in V, u' \in V, u \neq u' \text{ and } f(u) = f(u')\}\). But \(v \in V\) and \(v' \in V\) and \(v \neq v'\) (since \(\{v, v'\}\) is an edge of \(E\)) and \(f(v) = f(v')\). Hence, \(\{v, v'\} \in \{\{u, u'\} \mid u \in V, u' \in V, u \neq u' \text{ and } f(u) = f(u')\} = eqs f\).

Combined with \(\{v, v'\} \in E\), this yields \(\{v, v'\} \in (eqs f) \cap E = \emptyset\). But this is absurd, because the empty set \(\emptyset\) has no elements.

Hence, we have found a contradiction. Thus, our assumption (that there exists an edge \(e = \{v, v'\}\) in \(E\) such that \(f(v) = f(v')\)) was wrong. Hence, no edge \(e = \{v, v'\}\) in \(E\) has \(f(v) = f(v')\). In other words, \(f\) is a proper coloring of \(G\) (according to the definition of a “proper coloring”).

Let us now forget that we assumed that \((eqs f) \cap E = \emptyset\). We thus have proven the implication

\[
(eq f \cap E = \emptyset) \quad \Rightarrow \quad (f \text{ is a proper coloring of } G).
\]

Let us now assume that \(f\) is a proper coloring of \(G\). In other words, no edge \(e = \{v, v'\}\) in \(E\) has \(f(v) = f(v')\) (according to the definition of a “proper coloring”).

Now, let us prove that \((eqs f) \cap E = \emptyset\). Indeed, assume the contrary. Then, \((eqs f) \cap E \neq \emptyset\). Hence, there exists some \(g \in (eqs f) \cap E\). Consider this edge \(g\). We have

\[
g \in (eqs f) \cap E \subset eqs f = \{\{u, u'\} \mid u \in V, u' \in V, u \neq u' \text{ and } f(u) = f(u')\}.
\]

Hence, \(g = \{u, u'\}\) for some \(u \in V\) and \(u' \in V\) satisfying \(u \neq u'\) and \(f(u) = f(u')\). Consider these \(u\) and \(u'\). We have \(u \neq u'\) and \(\{u, u'\} = g \in (eqs f) \cap E \subset \mathcal{E}\). Hence, \(\{u, u'\} \in (eqs f) \cap E \subset \mathcal{E}\). Moreover, \(f(u) = f(u')\). Hence, the edge \(\{u, u'\}\) is an edge \(e = \{v, v'\}\) in \(E\) such that \(f(v) = f(v')\) (namely, for \(v = u\) and \(v' = u'\)). This contradicts the fact that no edge \(e = \{v, v'\}\) in \(E\) has \(f(v) = f(v')\). This contradiction proves that our assumption was wrong. Hence, we have shown that \((eqs f) \cap E = \emptyset\).

Let us now forget that we assumed that \(f\) is a proper coloring of \(G\). We thus have proven the implication

\[
(f \text{ is a proper coloring of } G) \quad \Rightarrow \quad ((eqs f) \cap E = \emptyset).
\]

Combining this with \((11.1229)\), we obtain the equivalence

\[
((eqs f) \cap E = \emptyset) \quad \iff \quad (f \text{ is a proper coloring of } G).
\]

This proves \((11.1228)\).
We shall regard permutations as words, by identifying every \( \pi \in S_n \) with the word \((\pi(1), \pi(2), \ldots, \pi(n))\). For every \( n \in \mathbb{N} \), the lexicographic order on words thus defines a total order on \( S_n \); we will be using this order in the following when we make statements like “\( \sigma < \tau \)” for \( \sigma \) and \( \tau \) being two permutations in \( S_n \).

We denote the empty word by \( \emptyset \). This empty word is identified with the trivial permutation in \( S_0 \). We shall refer to \( \emptyset \) as the empty permutation; every other permutation will be called a nonempty permutation.

Let \( S \) denote the disjoint union \( \bigsqcup_{n \in \mathbb{N}} S_n \) of the posets \( S_n \). While each poset \( S_n \) is actually totally ordered (by the lexicographic order), the disjoint union \( S \) is not, since elements of different \( S_n \) are incomparable.\(^{928}\)

It is easy to see that

\[
(11.1230) \quad \alpha \cdot \gamma > \beta \cdot \gamma
\]

for any \( \alpha \in S \), \( \beta \in S \) and \( \gamma \in S \) satisfying \( \alpha > \beta \).\(^ {929}\)

For every \( k \in \mathbb{N} \) and \( \ell \in \mathbb{N} \) and any permutations \( u \in S_k \) and \( v \in S_\ell \), we define a permutation \( u \cdot v \) by \( u \cdot v = u \cdot v[k] \) (where “\( u \cdot v[k] \)” has to be read as “\( u \cdot (v[k]) \)” rather than as “\( (u \cdot v)[k] \)”\).\(^ {930}\) Thus, we have defined a binary operation \( \cdot \) on the set \( S \). It is easy to see that \( S \) becomes a monoid with respect to this binary operation \( \cdot \); the neutral element of this monoid is \( \emptyset \in S_0 \). Hence, products of the form \( w_1 \cdot w_2 \cdot \cdots \cdot w_k \) for \( k \)-tuples \((w_1, w_2, \ldots, w_k) \in S^k \) of permutations are well-defined (without specifying a bracketing). The monoid \((S, \cdot)\) is left-cancellative and right-cancellative.

Recall how we defined a connected permutation. Since we are regarding permutations as words, we can rewrite this definition as follows: A permutation \( p \in S_n \) is connected if and only if it is nonempty and no nonempty proper prefix\(^ {931}\) of \( p \) is itself a permutation. Hence, if a nonempty prefix of a connected permutation \( p \in S_n \) is itself a permutation, then this prefix must be \( p \).

It is easy to see that a permutation \( w \in S_n \) is connected if and only if \( n \) is a positive integer and there exist no nonempty permutations \( u \) and \( v \) satisfying \( w = u \cdot v \). It is furthermore easy to see that for every permutation \( w \in S \), there is a unique way to write \( w \) in the form \( w = w_1 \cdot w_2 \cdot \cdots \cdot w_k \) for some \( k \in \mathbb{N} \) and

\[^{928}\]We could define an ordering on all of \( S \) by restricting the total order on words defined in Definition 6.1; but we prefer not to.

\[^{929}\]Proof of (11.1230): Let \( \alpha \in S \), \( \beta \in S \) and \( \gamma \in S \) satisfy \( \alpha > \beta \). Notice that the \( n \in \mathbb{N} \) satisfying \( \alpha \in S_n \) and the \( n \in \mathbb{N} \) satisfying \( \beta \in S_n \) must be identical (because otherwise, \( \alpha \) and \( \beta \) would be incomparable in the poset \( S \)). In other words, the words \( \alpha \) and \( \beta \) have the same length.

Now, the words \( \alpha \) and \( \beta \) have the same length and satisfy \( \alpha > \beta \) in the lexicographic order. Hence, every word \( \delta \) satisfies \( \alpha \cdot \delta > \beta \cdot \delta \) (where \( \cdot \) denotes concatenation). Applying this to \( \delta = \gamma[k] \), we obtain \( \alpha \cdot \gamma[k] > \beta \cdot \gamma[k] \). Now, the definition of \( \alpha \cdot \gamma \) yields \( \alpha \cdot \gamma = \alpha \cdot \gamma[k] > \beta \cdot \gamma[k] = \beta \cdot \gamma \) (by the definition of \( \beta \cdot \gamma \)), and thus (11.1230) is proven.

\[^{930}\]It is really obvious that this is indeed a permutation, because the word \( u \cdot v[k] \) has each integer from 1 to \( k + \ell \) appear exactly once in it.

\[^{931}\]A proper prefix of a word \( w \) is defined as a word \( u \) such that there exists a nonempty word \( v \) satisfying \( w = uv \).
some $k$-tuple $(w_1, w_2, \ldots, w_k) \in \mathcal{E}S^k$ of connected permutations. We refer to the tuple $(w_1, w_2, \ldots, w_k)$ as the connected decomposition of $w$.

For every $n \in \mathbb{N}$ and $w \in \mathcal{E}_n$, define an element $F_{\text{conn}}$ of FQSym by setting $F_{\text{conn}} = F_{w_1}F_{w_2} \cdots F_{w_k}$, where $(w_1, w_2, \ldots, w_k)$ is the connected decomposition of $w$. It is clear that this $F_{\text{conn}}$ is an element of FQSym. Hence, for every $n \in \mathbb{N}$, the family $(F_{\text{conn}})_{w\in \mathcal{E}_n}$ is a family of elements of FQSym. We shall prove that it is a basis of the $k$-module FQSym. Once this is proven, the claim of the exercise will quickly follow (as we will see later).

We need the following easy fact:

**Lemma 11.44.** Let $u \in \mathcal{S}$ and $v \in \mathcal{S}$. Then,

$$F_uF_v = F_{u \sqcup v} + (a \text{ sum of terms } F_t \text{ for } t \in \mathcal{S} \text{ satisfying } t > u \sqcup v).$$

(The symbol $\sqcup$ refers to the partial order on the poset $\mathcal{S}$.)

**Proof of Lemma 11.44.** Let $k$ and $\ell$ be the nonnegative integers satisfying $u \in \mathcal{S}_k$ and $v \in \mathcal{S}_\ell$. Write the words $u$ and $v\ [k]$ in the forms $u = (u_1, u_2, \ldots, u_k)$ and $v\ [k] = (p_1, p_2, \ldots, p_\ell)$, respectively. Let $(c_1, c_2, \ldots, c_{k+\ell})$ denote the concatenation $u \cdot v\ [k] = (u_1, u_2, \ldots, u_k, p_1, p_2, \ldots, p_\ell)$. By the definition of $u \sqcup v\ [k]$, we therefore have

$$(11.1231) \quad u \sqcup v\ [k] = \{(c_{w(1)}, c_{w(2)}, \ldots, c_{w(k+\ell)}) : w \in \text{Sh}_{k, \ell}\}_{\text{multiset}}.$$
From (8.1), we have

\[ F_u F_v = \sum_{w \in u \cup v[k]} F_w \]

\[ = \sum_{w \in Sh_{k,\ell}} F(c_w(1), c_w(2), \ldots, c_w(k+\ell)) \quad \text{(by (11.1231))} \]

\[ = F_u \square v + \sum_{w \in Sh_{k,\ell}; w \neq \text{id}} F(c_w(1), c_w(2), \ldots, c_w(k+\ell)) \text{ (here, we have split off the addend for } w = \text{id} \text{ from the sum, since } \text{id} \in Sh_{k,\ell}) \]

\[ (11.1232) \]

But it is easy to see that

\[ (c_w(1), c_w(2), \ldots, c_w(k+\ell)) > u \square v \quad \text{for every } w \in Sh_{k,\ell} \text{ satisfying } w \neq \text{id}. \]
Hence, (11.1232) becomes
\[ F_u F_w = F_u \Box w + \sum_{w \in \mathcal{Sh}_{k,t}; \ w \neq \text{id}} F(c_w(1), c_w(2), \ldots, c_w(k+\ell)) = (\text{a sum of terms } F_t \text{ for } t \in \mathcal{G} \text{ satisfying } t > u \Box v) \]
(by (11.1233))
\[ = F_u \Box v + (\text{a sum of terms } F_t \text{ for } t \in \mathcal{G} \text{ satisfying } t > u \Box v). \]
This proves Lemma 11.44. \qed

**Corollary 11.45.** Every \( n \in \mathbb{N} \) and \( w \in \mathcal{G}_n \) satisfy
\[ F_w^{\text{conn}} = F_w + (\text{a sum of terms } F_t \text{ for } t \in \mathcal{G}_n \text{ satisfying } t > w). \]

---

**Proof of (11.1233):** Let \( w \in \mathcal{Sh}_{k,t} \) be such that \( w \neq \text{id} \). The map \( w \) is a permutation, hence bijective and thus injective.

There exists some \( i \in \{1, 2, \ldots, k+\ell\} \) such that \( w(i) \neq i \) (since \( w \neq \text{id} \)). Let \( j \) be the smallest such \( i \). Thus, \( w(j) \neq j \), but (11.1234)
\[ \text{every } i \in \{1, 2, \ldots, k+\ell\} \text{ satisfying } i < j \text{ satisfies } w(i) = i. \]

We have \( w \in \mathcal{Sh}_{k,t} \), and therefore \( w^{-1}(1) < w^{-1}(2) < \cdots < w^{-1}(k) \) and \( w^{-1}(k+1) < w^{-1}(k+2) < \cdots < w^{-1}(k+\ell) \).

Thus, the restriction of the map \( w^{-1} \) to the set \( \{1, 2, \ldots, k\} \) and the restriction of the map \( w^{-1} \) to the set \( \{k+1, k+2, \ldots, k+\ell\} \) are strictly increasing.

Let us first show that \( j \leq k \). Indeed, assume the contrary. Then, \( j > k \). Hence, every \( i \in \{1, 2, \ldots, k\} \) satisfies \( i \leq k < j \) and therefore \( w(i) = i \) (by (11.1234)). Thus, \( w(\{1, 2, \ldots, k\}) = \{1, 2, \ldots, k\} \), so that \( w^{-1}(\{1, 2, \ldots, k\}) = \{1, 2, \ldots, k\} \) (since \( w \) is bijective). Notice also that \( j \in \{k+1, k+2, \ldots, k+\ell\} \) (since \( j > k \)). Now,
\[ w^{-1}\left(\{1, 2, \ldots, k+\ell\} \setminus \{1, 2, \ldots, k\}\right) \]
\[ = \{k+1, k+2, \ldots, k+\ell\} \setminus \{1, 2, \ldots, k\} \]
\[ = \{1, 2, \ldots, k\} \cup \{k+1, k+2, \ldots, k+\ell\} \]
\[ = \{1, 2, \ldots, k\} \cup \{k+1, k+2, \ldots, k+\ell\}. \]

Therefore, \( w^{-1} \) restricts to a map \( \{k+1, k+2, \ldots, k+\ell\} \to \{k+1, k+2, \ldots, k+\ell\} \). This latter map must be strictly increasing (since the restriction of the map \( w^{-1} \) to the set \( \{k+1, k+2, \ldots, k+\ell\} \) is strictly increasing), and therefore is the identity map (because the only strictly increasing map \( \{k+1, k+2, \ldots, k+\ell\} \to \{k+1, k+2, \ldots, k+\ell\} \) is the identity map).

In other words, \( w^{-1}(i) = \text{id}(i) = i \) for every \( i \in \{k+1, k+2, \ldots, k+\ell\} \). Applied to \( i = j \), this yields \( w^{-1}(j) = j \) (since \( j \in \{k+1, k+2, \ldots, k+\ell\} \)), whence \( w(j) = j \). But this contradicts \( w(j) \neq j \). This contradiction proves that our assumption was wrong. Thus, \( j \leq k \). Therefore,
\[ c_j = u_j \]
\[ = u(j) \leq k \]
\[ \in \mathcal{Sh}_{k,t}. \]

If we had \( w(j) < j \), then we would have \( w(w(j)) = w(j) \) (by (11.1234), applied to \( i = w(j) \)), which would lead to \( w(j) = j \) (since \( w \) is injective), which would contradict \( w(j) \neq j \). Hence, we cannot have \( w(j) < j \). Thus, we have \( w(j) \geq j \) (since \( w(j) \neq j \)). In other words, \( j < w(j) \).

Let us next prove that \( w(j) > k \). Indeed, assume the contrary. Thus, \( w(j) \leq k \). Thus, \( w(j) \) and \( j \) are two elements of \( \{1, 2, \ldots, k\} \) (because \( w(j) \leq k \) and \( j \leq k \)) satisfying \( j < w(j) \). Hence, \( w^{-1}(j) < w^{-1}(w(j)) \) (since \( w^{-1}(1) < w^{-1}(2) < \cdots < w^{-1}(k) \)). \( w^{-1}(j) = w^{-1}(w(j)) = j \). Therefore, (11.1234) (applied to \( i = w^{-1}(j) \)) yields \( w(w^{-1}(j)) = w^{-1}(j) \), so that \( w^{-1}(j) = w^{-1}(w(j)) = j \). This contradicts \( w^{-1}(j) < j \). This contradiction proves that our assumption was wrong. Hence, \( w(j) > k \) is proven.

Now, write \( v \) in the form \((v_1, v_2, \ldots, v_k)\). Then, \( v[k] = (k + v_1, k + v_2, \ldots, k + v_k) \) (by the definition of \( v[k] \)). Since \( w(j) > k \), we have
\[ c_{w(j)} = p_{w(j) - k} \]
\[ = k + v_{w(j) - k} \]
\[ > k \geq c_j \]
(by (11.1235)).

So we have \( c_{w(j)} > c_j \), but for every \( i \in \{1, 2, \ldots, k+\ell\} \) satisfying \( i < j \), we have \( c_{w(i)} = c_i \) (because (11.1234) yields \( w(i) = i \)). Thus,
\[ (c_{w(1)}, c_{w(2)}, \ldots, c_{w(k+\ell)}) > (c_1, c_2, \ldots, c_{k+\ell}) \]
(by the definition of lexicographic order)
and thus (11.1233) is proven.
Proof of Corollary 11.45. We shall first prove that every \( k \in \mathbb{N} \) and every \((w_1, w_2, \ldots, w_k) \in G^k\) satisfy (11.1236)

\[ F_{w_1} F_{w_2} \cdots F_{w_k} = F_{w_1 \Box w_2 \Box \cdots \Box w_k} + \text{(a sum of terms} \ F_t \text{for} \ t \in G \text{satisfying} \ t > w_1 \Box w_2 \Box \cdots \Box w_k) \]  

Proof of (11.1236): We will prove (11.1236) by induction over \( k \):

- **Induction base:** For \( k = 0 \), the statement of (11.1236) is obvious. Hence, the induction base is complete.
- **Induction step:** Let \( K \in \mathbb{N} \). Assume that (11.1236) holds for \( k = K \). We need to prove that (11.1236) holds for \( k = K + 1 \).

Let \((w_1, w_2, \ldots, w_{K+1}) \in G^{K+1}\). By the induction hypothesis, we can apply (11.1236) to \( k = K \), and thus obtain

\[ F_{w_1} F_{w_2} \cdots F_{w_K} = F_{w_1 \Box w_2 \Box \cdots \Box w_K} + \text{(a sum of terms} \ F_t \text{for} \ t \in G \text{satisfying} \ t > w_1 \Box w_2 \Box \cdots \Box w_K) \].

In other words, there exists a finite family \((t_i)_{i \in I}\) of elements \( t \in G \) satisfying \( t > w_1 \Box w_2 \Box \cdots \Box w_k \) such that

\[ F_{w_1} F_{w_2} \cdots F_{w_K} = F_{w_1 \Box w_2 \Box \cdots \Box w_K} + \sum_{i \in I} F_{t_i} \]

Consider this family \((t_i)_{i \in I}\). We have

\[ t_i \Box w_{K+1} > w_1 \Box w_2 \Box \cdots \Box w_{K+1} \quad \text{for every} \ i \in I. \]
Now,
\[
F_{w_1} F_{w_2} \cdots F_{w_{K+1}} = \underbrace{(F_{w_1} F_{w_2} \cdots F_{w_K})}_{F_{w_1} \Box w_2 \Box \cdots \Box w_K} F_{w_{K+1}}
\]
which equals
\[
\underbrace{F_{w_1} \Box w_2 \Box \cdots \Box w_K}_{\text{a sum of terms } F_t \text{ for } t \in \mathcal{G} \text{ satisfying } t > (w_1 \Box w_2 \Box \cdots \Box w_K) \Box w_{K+1}} F_{w_{K+1}}
\]
by Lemma 11.44, applied to \( u = w_1 \Box w_2 \Box \cdots \Box w_K \) and \( v = w_{K+1} \).

\[
+ \sum_{t \in I} \underbrace{F_t F_{w_{K+1}}}_{\text{a sum of terms } F_t \text{ for } t \in \mathcal{G} \text{ satisfying } t > t_i \Box w_{K+1}}
\]
(formulas a sum of terms \( F_t \) for \( t \in \mathcal{G} \) satisfying \( t > t_i \Box w_{K+1} \))
by Lemma 11.44, applied to \( u = t_i \) and \( v = w_{K+1} \).

\[
= (w_1 \Box w_2 \Box \cdots \Box w_{K+1}) + (\text{a sum of terms } F_t \text{ for } t \in \mathcal{G} \text{ satisfying } t > (w_1 \Box w_2 \Box \cdots \Box w_K) \Box w_{K+1})
\]
+ \[
\sum_{t \in I} \underbrace{(F_t \Box w_{K+1})}_{\text{a sum of terms } F_t \text{ for } t \in \mathcal{G} \text{ satisfying } t > t_i \Box w_{K+1}}
\]
(since \( (w_1 \Box w_2 \Box \cdots \Box w_{K+1}) = w_1 \Box w_2 \Box \cdots \Box w_{K+1} \))

\[
= (w_1 \Box w_2 \Box \cdots \Box w_{K+1}) + (\text{a sum of terms } F_t \text{ for } t \in \mathcal{G} \text{ satisfying } t > w_1 \Box w_2 \Box \cdots \Box w_{K+1})
\]
+ \[
\sum_{t \in I} \underbrace{(F_t \Box w_{K+1})}_{\text{a sum of terms } F_t \text{ for } t \in \mathcal{G} \text{ satisfying } t > t_i \Box w_{K+1}}
\]
(since every \( t \in \mathcal{G} \) satisfying \( t > w_1 \Box w_2 \Box \cdots \Box w_{K+1} \) also satisfies \( t > w_1 \Box w_2 \Box \cdots \Box w_{K+1} \) \( \Box w_{K+1} \) (by (11.1237))

\[
= (w_1 \Box w_2 \Box \cdots \Box w_{K+1}) + (\text{a sum of terms } F_t \text{ for } t \in \mathcal{G} \text{ satisfying } t > w_1 \Box w_2 \Box \cdots \Box w_{K+1})
\]
+ \[
\sum_{t \in I} \underbrace{(F_t \Box w_{K+1})}_{\text{a sum of terms } F_t \text{ for } t \in \mathcal{G} \text{ satisfying } t > w_1 \Box w_2 \Box \cdots \Box w_{K+1}}
\]
(since \( w_1 \Box w_2 \Box \cdots \Box w_{K+1} \)).

In other words, (11.1236) holds for \( k = K + 1 \). This completes the induction step. Hence, (11.1236) is proven by induction.

Now, let \( n \in \mathbb{N} \) and \( w \in \mathcal{G}_n \). Let \( (w_1, w_2, \ldots, w_k) \) be the connected decomposition of \( w \). Then, \( F_w^{\text{conn}} = F_{w_1} F_{w_2} \cdots F_{w_k} \) (by the definition of \( F_w^{\text{conn}} \)) and \( w = w_1 \Box w_2 \Box \cdots \Box w_k \) (by the definition of a connected decomposition). Now,

\[
F_w^{\text{conn}} = F_{w_1} F_{w_2} \cdots F_{w_k}
\]

\[
= F_{w_1} \Box w_2 \Box \cdots \Box w_k + (\text{a sum of terms } F_t \text{ for } t \in \mathcal{G} \text{ satisfying } t > w_1 \Box w_2 \Box \cdots \Box w_k)
\]
(by (11.1236))

\[
= F_w + (\text{a sum of terms } F_t \text{ for } t \in \mathcal{G} \text{ satisfying } t > w)
\]
(since \( w_1 \Box w_2 \Box \cdots \Box w_k = w \)).

Since every \( t \in \mathcal{G} \) satisfying \( t > w \) must belong to \( \mathcal{G}_n \), this rewrites as

\[
F_w^{\text{conn}} = F_w + (\text{a sum of terms } F_t \text{ for } t \in \mathcal{G}_n \text{ satisfying } t > w).
\]
This proves Corollary 11.45.

Now, let us fix \( n \in \mathbb{N} \). The family \( (F_w)_{w \in \mathcal{S}_n} \) is a basis of the \( k \)-module \( \text{FQSym}_n \). Thus, the family \( (F_w^{\text{conn}})_{w \in \mathcal{S}_n} \) is a basis of the \( k \)-module \( \text{FQSym}_n \) as well (because Corollary 11.45 shows that the transition matrix from the basis \( (F_w^{\text{conn}})_{w \in \mathcal{S}_n} \) to the basis \( (F_w)_{w \in \mathcal{S}_n} \) is triangular (with respect to the relation \( > \)), with all diagonal entries being 1).

Now, let us forget that we fixed \( n \). We thus have proven that the family \( (F_w^{\text{conn}})_{w \in \mathcal{S}_n} \) is a basis of the \( k \)-module \( \text{FQSym}_n \) for every \( n \in \mathbb{N} \). Hence, the family \( (F_w^{\text{conn}})_{w \in \mathcal{S}} \) (being the disjoint union of the families \( (F_w^{\text{conn}})_{w \in \mathcal{S}_n} \) for all \( n \in \mathbb{N} \)) is a basis of the \( k \)-module \( \bigoplus_{n \in \mathbb{N}} \text{FQSym}_n = \text{FQSym} \).

Now, it is easy to see that the family \( (F_w, F_{w_2} \cdots F_w)_{k \in \mathbb{N}; (w_1, w_2, \ldots, w_k) \in \mathcal{S}^k} \) is a reindexing of the family \( (F_w^{\text{conn}})_{w \in \mathcal{S}} \). Thus, the family \( (F_w^{\text{conn}})_{w \in \mathcal{S}} \) is a basis of the \( k \)-module \( \text{FQSym} \). In other words, \( \text{FQSym} \) is a free (noncommutative) \( k \)-algebra with generators \( (F_w)_{w \in \mathcal{S}} \). This solves Exercise 8.10.

Acknowledgements

The authors thank the following for helpful comments and/or teaching them about Hopf algebras: Marcelo Aguiar, Federico Ardila, Lou Billera, Richard Ehrenborg, Mark Haiman, Florent Hivert, Christophe Hohlweg, Jia Huang, Jang Soo Kim, Aaron Lauve, Dominique Manchon, John Palmieri, Alexander Postnikov, Margie Readdy, Nathan Reading, Hans-Jürgen Schneider, Richard Stanley, Muge Taskin, Jean-Yves Thibon.

References


Proof. For every \( k \in \mathbb{N} \) and \( (w_1, w_2, \ldots, w_k) \in \mathcal{S}^k \), the connected decomposition of the permutation \( w_1 \sqcup w_2 \sqcup \cdots \sqcup w_k \) is the \( k \)-tuple \((w_1, w_2, \ldots, w_k)\) (by the definition of a connected decomposition). Hence, for every \( k \in \mathbb{N} \) and \( (w_1, w_2, \ldots, w_k) \in \mathcal{S}^k \), we have

\[
F_{w_1 \sqcup w_2 \sqcup \cdots \sqcup w_k} = F_{w_1} F_{w_2} \cdots F_{w_k}.
\]

(by the definition of \( F_{w_1 \sqcup w_2 \sqcup \cdots \sqcup w_k} \)).

Now, recall that for every permutation \( w \in \mathcal{S} \), there is a unique way to write \( w \) in the form \( w = w_1 \sqcup w_2 \sqcup \cdots \sqcup w_k \) for some \( k \in \mathbb{N} \) and some \( k \)-tuple \((w_1, w_2, \ldots, w_k) \in \mathcal{S}^k \) of connected permutations. In other words, the map

\[
\bigcup_{k \in \mathbb{N}} \mathcal{S}^k \to \mathcal{S},
\]

\[
(w_1, w_2, \ldots, w_k) \mapsto w_1 \sqcup w_2 \sqcup \cdots \sqcup w_k
\]

is a bijection. Hence, the family \((F_{w_1 \sqcup w_2 \sqcup \cdots \sqcup w_k})_{k \in \mathbb{N}; (w_1, w_2, \ldots, w_k) \in \mathcal{S}^k}\) is a reindexing of the family \((F_w)_{w \in \mathcal{S}}\). Due to (11.1238), this rewrites as follows: The family \((F_{w_1} F_{w_2} \cdots F_{w_k})_{k \in \mathbb{N}; (w_1, w_2, \ldots, w_k) \in \mathcal{S}^k}\) is a reindexing of the family \((F_w^{\text{conn}})_{w \in \mathcal{S}}\). Qed.


146. W. Schmitt. Expository notes, specifically “A concrete introduction to category theory” and “Notes on modules and algebras”. http://home.gwu.edu/~wschmitt/


E-mail address: darijgrinberg@gmail.com

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139, USA

E-mail address: reiner@math.umn.edu

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MN 55455, USA