MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Physics Department

8.323: Relativistic Quantum Field Theory I

Prof. Alan Guth May 11, 2008

PROBLEM SET 8 SOLUTIONS

Problem 1: Commutation Relations for the Lorentz Group *(10 points)*†

(a)

$$
\begin{aligned}
[J^{\mu\nu}, J^{\rho\sigma}]^{\alpha}{}_{\beta} &= (J^{\mu\nu})^{\alpha}{}_{\eta}(J^{\rho\sigma})^{\eta}{}_{\beta} - (J^{\rho\sigma})^{\alpha}{}_{\eta}(J^{\mu\nu})^{\eta}{}_{\beta} \\
&= 2i\{g^{\mu\alpha}\delta^{\nu}{}_{\eta}\}_{\mu\nu} 2i\{g^{\rho\eta}\delta^{\sigma}{}_{\beta}\}_{\rho\sigma} - 2i\{g^{\rho\alpha}\delta^{\sigma}{}_{\eta}\}_{\rho\sigma} 2i\{g^{\mu\eta}\delta^{\nu}{}_{\beta}\}_{\mu\nu} \\
&= 4i^2\{g^{\mu\alpha}\,\delta^{\nu}{}_{\eta}\,g^{\rho\eta}\,\delta^{\sigma}{}_{\beta} - g^{\mu\eta}\,\delta^{\nu}{}_{\beta}\,g^{\rho\alpha}\,\delta^{\sigma}{}_{\eta}\}_{\substack{\mu\nu\\ \rho\sigma}} \\
&= 4i^2\{g^{\mu\alpha}\,g^{\rho\nu}\,\delta^{\sigma}{}_{\beta} - g^{\mu\sigma}\,\delta^{\nu}{}_{\beta}\,g^{\rho\alpha}\}_{\substack{\mu\nu\\ \rho\sigma}} \\
&= 4i^2\{g^{\mu\alpha}\,g^{\rho\nu}\,\delta^{\sigma}{}_{\beta} - g^{\nu\rho}\,\delta^{\mu}{}_{\beta}\,g^{\sigma\alpha}\}_{\substack{\mu\nu\\ \rho\sigma}} \\
&= 4i\{g^{\nu\rho}(J^{\mu\sigma})^{\alpha}{}_{\beta}\}_{\substack{\mu\nu\\ \rho\sigma}}
\end{aligned}
$$

(b) From the formula

$$
(J^{\mu\nu})^{\alpha}{}_{\beta} = i(g^{\mu\alpha}\,\delta^{\nu}{}_{\beta} - g^{\nu\alpha}\,\delta^{\mu}{}_{\beta}) \ ,
$$

we write J^{03} as the matrix

$$
(J^{03})^{\alpha}{}_{\beta} = \begin{array}{cc} & \beta \ \rightarrow \\ \begin{array}{c}\alpha\\ \downarrow\end{array} & \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \end{array}
$$

Then the infinitesimal transformation generated by J^{03} is

$$
1 - i\eta J^{03} = \begin{pmatrix} 1 & 0 & 0 & \eta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \eta & 0 & 0 & 1 \end{pmatrix} \ .
$$

We want to compare this to an infinitesimal Lorentz boost in the z-direction. A finite Lorentz boost in the z-direction has the form

$$
\Lambda = \begin{pmatrix} \cosh\eta & 0 & 0 & \sinh\eta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\eta & 0 & 0 & \cosh\eta \end{pmatrix} \ ,
$$

where η is the rapidity of the boost. The infinitesimal form of this boost is obviously the expression above for $1 - i\eta J^{03}$.

Problem 2: Representations of the Lorentz Group *(15 points)*†

(a) We will make frequent use of the identity

$$\epsilon^{mij}\epsilon^{mk\ell} = \delta^{ik}\delta^{j\ell} - \delta^{i\ell}\delta^{jk} \ , \tag{2.1}$$

which we can use immediately to invert the definition of J^i:

$$J^i = \frac{1}{2}\epsilon^{ijk}J^{jk} \implies \epsilon^{ijm}J^m = \frac{1}{2}\epsilon^{ijm}\epsilon^{mk\ell}J^{k\ell} = \frac{1}{2}(\delta^{ik}\delta^{j\ell} - \delta^{i\ell}\delta^{jk})J^{k\ell} = J^{ij},$$

so

$$J^{ij} = \epsilon^{ijk}J^k. \tag{2.2}$$

Using curly brackets to denote antisymmetrization as in the previous problem,

$$\begin{aligned}
[J^i, J^j] &= \left[\frac{1}{2}\epsilon^{iab}J^{ab}, \frac{1}{2}\epsilon^{jcd}J^{cd}\right] = \frac{1}{4}\epsilon^{iab}\epsilon^{jcd}[J^{ab}, J^{cd}] \\
&= \frac{1}{4}\epsilon^{iab}\epsilon^{jcd}4i\left\{g^{bc}J^{ad}\right\}_{\substack{ab\\cd}} \ .
\end{aligned}$$

Now use the fact that b and c are spacelike, so $g^{bc} = -\delta^{bc}$. Furthermore, since the ϵ's are antisymmetric, the antisymmetrization of the quantity in curly brackets is superfluous. So

$$\begin{aligned}
[J^i, J^j] &= i\epsilon^{bia}\epsilon^{bjd}J^{ad} \\
&= i(\delta^{ij}\delta^{ad} - \delta^{id}\delta^{aj})J^{ad} \\
&= iJ^{ij} = i\epsilon^{ijk}J^k \ .
\end{aligned}$$

Furthermore

$$\begin{aligned}
[J^i, K^j] &= [\frac{1}{2}\epsilon^{iab}J^{ab}, J^{0j}] \\
&= \frac{i}{2}\epsilon^{iab}(g^{b0}J^{aj} - g^{a0}J^{bj} - g^{bj}J^{a0} + g^{aj}J^{b0}) \\
&= \frac{i}{2}(\epsilon^{iaj}J^{a0} - \epsilon^{ijb}J^{b0}) \\
&= -i\epsilon^{ija}J^{a0} \\
&= i\epsilon^{ijk}K^k \ ,
\end{aligned}$$

and

$$\begin{aligned}
[K^i, K^j] &= [J^{0i}, J^{0j}] \\
&= i(g^{i0}J^{0j} - g^{00}J^{ij} - g^{ij}J^{00} + g^{0j}J^{i0}) \\
&= -iJ^{ij} \\
&= -i\epsilon^{ijk}J^k \ .
\end{aligned}$$

In summary,

$$[J^i, J^j] = i\epsilon^{ijk} J^k \ , \tag{2.3}$$

$$[J^i, K^j] = i\epsilon^{ijk} K^k \ , \tag{2.4}$$

and

$$[K^i, K^j] = -i\epsilon^{ijk} J^k \ . \tag{2.5}$$

(b) Now we check that

$$\vec{J_+} = \frac{1}{2}(\vec{J} + i\vec{K}) \tag{2.6a}$$

and

$$\vec{J_-} = \frac{1}{2}(\vec{J} - i\vec{K}) \tag{2.6b}$$

generate two copies of the rotation group:

$$\begin{aligned}
[J_+^i, J_-^j] &= \frac{1}{4}([J^i, J^j] - i[J^i, K^j] + i[K^i, J^j] + [K^i, K^j]) \\
&= \frac{1}{4}(i\epsilon^{ijk} J^k - i^2 \epsilon^{ijk} K^k + i(-i)\epsilon^{jik} K^k + (-i)\epsilon^{ijk} J^k) \\
&= 0 \ ,
\end{aligned} \tag{2.7}$$

and

$$\begin{aligned}
[J_+^i, J_+^j] &= \frac{1}{4}([J^i, J^j] + i[J^i, K^j] + i[K^i, J^j] - [K^i, K^j]) \\
&= \frac{1}{4}(i\epsilon^{ijk} J^k + i^2 \epsilon^{ijk} K^k + i(-i)\epsilon^{jik} K^k - (-i)\epsilon^{ijk} J^k) \\
&= \frac{1}{2} i\epsilon^{ijk}(J^k + iK^k) \\
&= i\epsilon^{ijk} J_+^k \ ,
\end{aligned} \tag{2.8}$$

and

$$\begin{aligned}
[J_-^i, J_-^j] &= \frac{1}{4}([J^i, J^j] - i[J^i, K^j] - i[K^i, J^j] + [K^i, K^j]) \\
&= \frac{1}{4}(i\epsilon^{ijk} J^k - i^2 \epsilon^{ijk} K^k + i^2 \epsilon^{jik} K^k - (-i)\epsilon^{ijk} J^k) \\
&= \frac{1}{2} i\epsilon^{ijk}(J^k - iK^k) \\
&= i\epsilon^{ijk} J_-^k \ .
\end{aligned} \tag{2.9}$$

$\vec{J_+}$ and $\vec{J_-}$ do indeed separately satisfy the commutation relations of angular momentum.

(c) The rotation group is a subgroup of the Lorentz group, and under the rotation group the states $|m_1, m_2\rangle$ transforms exactly like states of two spin-$\frac{1}{2}$ particles. From nonrelativistic quantum mechanics we already know how to couple these states to form states of $J_{\rm tot} = 0$ and $J_{\rm tot} = 1$:

$$\begin{aligned} |J_{\rm tot}{=}1,\ M{=}1\rangle &= |{+}{+}\rangle \\ |J_{\rm tot}{=}1,\ M{=}0\rangle &= \frac{1}{\sqrt{2}}\big[|{+}{-}\rangle + |{-}{+}\rangle\big] \\ |J_{\rm tot}{=}1,\ M{=}{-1}\rangle &= |{-}{-}\rangle\,, \end{aligned} \tag{2.10}$$

and

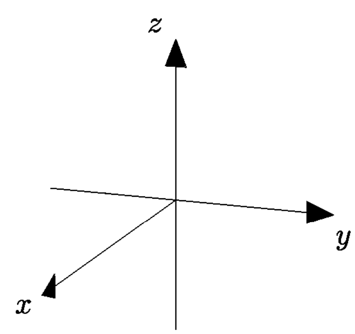
$$|J_{\rm tot}{=}0,\ M{=}0\rangle = \frac{1}{\sqrt{2}}\big[|{+}{-}\rangle - |{-}{+}\rangle\big]\ . \tag{2.11}$$

We also probably remember that the three $|J_{\rm tot}{=}1,\ M\rangle$ states form a vector representation, so that one can find linear combinations $|x\rangle$, $|y\rangle$, and $|z\rangle$ that rotate like the components of a cartesian vector. The precise definitions of $|x\rangle$, $|y\rangle$, and $|z\rangle$, are not likely to be remembered, so we will derive them. The state $|z\rangle$ must be invariant under rotations about the z-axis, so this identifies it as

$$|z\rangle = |J_{\rm tot}{=}1,\ M{=}0\rangle = \frac{1}{\sqrt{2}}\big[|{+}{-}\rangle + |{-}{+}\rangle\big]\ . \tag{2.12}$$

The phase of the above ket is not fixed by its rotational properties, but clearly we can arbitrarily define the overall phase of the four kets that we seek: $|0\rangle$, $|1\rangle \equiv |x\rangle$, $|2\rangle \equiv |y\rangle$, and $|3\rangle \equiv |z\rangle$. Thus we can adopt Eq. (2.12) as our phase convention.

As can be seen on the diagram at the right, the state $|x\rangle$ can be obtained by rotating $|z\rangle$ by 90° (or $\pi/2$) counterlclockwise about the y-axis, and the state $|y\rangle$ can obtained by rotating $|z\rangle$ by 90° clockwise about the x-axis. In the spin-$\frac{1}{2}$ representation,



$$\begin{aligned} D_y(\pi/2) &= e^{-i\frac{\pi}{2}J_y} = e^{-i\frac{\pi}{4}\sigma_y} \\ &= \cos\frac{\pi}{4} - i\sin\frac{\pi}{4}\sigma_y = \frac{\sqrt{2}}{2}\left[1 - i\sigma^2\right] \\ &= \frac{\sqrt{2}}{2}\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right] = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\ , \end{aligned} \tag{2.13}$$

and

$$\begin{aligned} D_x(-\pi/2) &= e^{i\frac{\pi}{2}J_x} = e^{i\frac{\pi}{4}\sigma_x} \\ &= \cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\sigma_x = \frac{\sqrt{2}}{2}\left[1 + i\sigma^1\right] \\ &= \frac{\sqrt{2}}{2}\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}\right] = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} . \end{aligned} \tag{2.14}$$

Then, summing over repeated indices, we can write

$$\begin{aligned} |x\rangle &= e^{-i\frac{\pi}{2}J_y}\,|z\rangle \\ &= |m_1, m_2\rangle\,\langle m_1, m_2\,|e^{-i\frac{\pi}{2}J_y}|\,m_1', m_2'\rangle\,\langle m_1', m_2'\,|z\,\rangle \\ &= |m_1, m_2\rangle\,D_y(\pi/2)_{m_1 m_1'}\,D_y(\pi/2)_{m_2 m_2'}\,\langle m_1', m_2'\,|z\,\rangle \\ &= \frac{1}{\sqrt{2}}\,|m_1, m_2\rangle\,[D_y(\pi/2)_{m_1 +}\,D_y(\pi/2)_{m_2 -} + D_y(\pi/2)_{m_1 -}\,D_y(\pi/2)_{m_2 +}] \\ &= \frac{1}{2\sqrt{2}}\left[|{+}{+}\rangle\,(-1-1) + |{+}{-}\rangle\,(1-1) + |{-}{+}\rangle\,(-1+1) + |{-}{-}\rangle\,(1+1)\right] \\ &= \frac{1}{\sqrt{2}}\left[-\,|{+}{+}\rangle + |{-}{-}\rangle\right] , \end{aligned} \tag{2.15}$$

where we have used Eqs. (2.12) and (2.13). An analogous calculation gives

$$|y\rangle = e^{i\frac{\pi}{2}J_x}\;|z\rangle = \frac{i}{\sqrt{2}}\left[|{+}{+}\rangle + |{-}{-}\rangle\right] . \tag{2.16}$$

The 0-component of the four-vector must be rotationally invariant, and hence must be up to a phase equal to $|J_{\rm tot}{=}0,\ M{=}0\rangle$. So, using Eq. (2.11) we write

$$|0\rangle = \frac{\eta}{\sqrt{2}}\left[|{+}{-}\rangle - |{-}{+}\rangle\right] , \tag{2.17}$$

where η is an unknown phase. This phase is not arbitrary, however, because it must be consistent with the Lorentz group algebra. We can determine the phase by looking at any nonzero matrix element involving this state, so for example we can specialize the general relation

$$\langle\alpha\,|J^{\mu\nu}|\,\beta\rangle = (J^{\mu\nu})^{\alpha}{}_{\beta} = i\left(g^{\mu\alpha}\,\delta^{\nu}_{\beta} - g^{\nu\alpha}\,\delta^{\mu}_{\beta}\right) \tag{2.18}$$

to the case

$$\left\langle z\,\middle|J^{03}\middle|\,0\right\rangle , \tag{2.19}$$

where using Eqs. (2.6) we can write

$$J^{03} = K^3 = i\left(J_-^3 - J_+^3\right) \ . \tag{2.20}$$

Note that at this point we are choosing to let $|\beta\rangle$ denote a covariant vector (i.e., a lower index), while $\langle\alpha|$ denotes a contravariant vector, since β and α appear as lower and upper indices, respectively, in Eq. (2.18). One can nonetheless define $\langle\alpha|$ as the bra vector corresponding to $|\alpha\rangle$, in the usual way, since the inner product

$$\langle \alpha \,|\beta\,\rangle = \delta^\alpha_\beta \tag{2.21}$$

is consistent with both the Lorentz metric and the inner product that is standardly used with Dirac's bra-and-ket notation for linear vector spaces. The operators representing the Lorentz group, however, will not be unitary.

Combining Eqs. (2.17–2.20), and also Eq. (2.12), one has

$$\begin{aligned}\left\langle z\,\middle|K^3\middle|\,0\right\rangle &= i\left(g^{03}\,\delta^3_0 - g^{33}\,\delta^0_0\right) = i\\ &= i\left\langle z\,\middle|J^3_- - J^3_+\middle|\,0\right\rangle\\ &= -\frac{i}{2}\left\langle z\,\middle|\sigma^3_+ - \sigma^3_-\middle|\,0\right\rangle\\ &= -\frac{i\eta}{\sqrt{2}}\,\langle z|\,\Big[|+-\rangle + |-+\rangle\Big] = -i\eta \ .\end{aligned} \tag{2.22}$$

By equating the first and last lines of the above equation, one finds $\eta = -1$ and therefore

$$|0\rangle = \frac{1}{\sqrt{2}}\left[|-+\rangle - |+-\rangle\right] \ . \tag{2.23}$$

Summarizing,

$$\boxed{\begin{aligned}|0\rangle &= \frac{1}{\sqrt{2}}\left[|-+\rangle - |+-\rangle\right]\\ |1\rangle &= \frac{1}{\sqrt{2}}\left[-|++\rangle + |--\rangle\right]\\ |2\rangle &= \frac{i}{\sqrt{2}}\left[|++\rangle + |--\rangle\right]\\ |3\rangle &= \frac{1}{\sqrt{2}}\left[|+-\rangle + |-+\rangle\right] \ .\end{aligned}} \tag{2.24}$$

Of course your answers might differ from these by an overall phase.

Problem 3: The Baker-Campbell-Hausdorff Theorem *(20 points)*¶

(a)

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}U_I(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left\{e^{iH_0t}e^{-i(H_0+\epsilon\Delta H)t}\right\} \\ &= e^{iH_0t}\left[iH_0 - i(H_0+\epsilon\Delta H)\right]e^{-i(H_0+\epsilon\Delta H)t} \\ &= -i\,\epsilon\, e^{iH_0t}\,\Delta H\, e^{-iH_0t}U_I(t)\ ,\end{aligned} \tag{3.1}$$

so

$$Q(t) = -i\ e^{iH_0t}\Delta H e^{-iH_0t}\ . \tag{3.2}$$

(b) $U_I(0) = e^0e^0 = 1$, while $U_I(t)|_{\epsilon=0} = e^{iH_0t}e^{-iH_0t} = 1$. Now consider $\frac{\mathrm{d}}{\mathrm{d}t}U_I(t) = \epsilon Q(t)U_I(t)$, and expand in ϵ by using $U_I(t) = 1 + \mathcal{O}(\epsilon)$, obtaining:

$$\frac{\mathrm{d}}{\mathrm{d}t}U_I(t) = \epsilon Q(t) + \mathcal{O}(\epsilon^2)\ . \tag{3.3}$$

Using $U_I(0) = 1$, this equation can be integrated to give:

$$U_I(t) = 1 + \int_0^t \mathrm{d}t'\,\frac{\mathrm{d}}{\mathrm{d}t'}U_I(t') = 1 + \epsilon\int_0^t \mathrm{d}t'\,Q(t') + \mathcal{O}(\epsilon^2)\ . \tag{3.4}$$

(c) From Eq. (3.4) we know that the first order contribution to $U_I(t)$ is given by

$$\Delta U_1(t) = \int_0^t dt'Q(t')\ . \tag{3.5}$$

Using Eq. (3.2) and the identity shown as Eq. (3.13) on the Problem Set, we can expand $Q(t)$ as

$$Q(t) = -\sum_{n=0}^{\infty} \frac{(i)^{n+1} t^n}{n!} \Omega_n(H_0, \Delta H) \ , \tag{3.6}$$

where Ω_n denotes the iterated commutator as defined in the Problem Set. We can then integrate Eq. (3.5) term by term, obtaining

$$\Delta U_1(t) = -\int_0^t dt' \sum_{n=0}^{\infty} \frac{(i)^{n+1} t'^n}{n!} \Omega_n(H_0, \Delta H) = -\sum_n \frac{(it)^{n+1}}{(n+1)!} \Omega_n(H_0, \Delta H) \tag{3.7}$$

which agrees with Eq. (3.16) of the Problem Set provided that

$$A_n(t) = -\frac{(it)^{n+1}}{(n+1)!} \ . \tag{3.8}$$

(d) Expanding the exponent of $e^{M(s+\epsilon\Delta s)}$ to first order in ϵ,

$$e^{M(s+\epsilon\Delta s)} = e^{M(s)+\epsilon\,\Delta s\, \frac{d}{ds} M(s)} + \mathcal{O}(\epsilon^2) \ . \tag{3.9}$$

By using Eq. (3.7) with $H_0 = iM(s)$ and $\Delta H = i\,\Delta s\, \frac{d}{ds} M(s)$ we obtain

$$\begin{aligned} U_I(t{=}1) &= e^{-M(s)}\, e^{M(s)+\epsilon\,\Delta s\,\frac{\mathrm{d}}{\mathrm{d}s}M(s)} \\ &= 1 - \epsilon \sum_{n=0}^{\infty} \frac{i^{n+1}}{(n+1)!} \Omega_n \left(iM(s), i\,\Delta s\, \frac{\mathrm{d}}{\mathrm{d}s} M(s) \right) + \mathcal{O}(\epsilon^2) \\ &= 1 + \epsilon\,\Delta s \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!}\, \Omega_n \left(M(s), \frac{\mathrm{d}}{\mathrm{d}s} M(s) \right) + \mathcal{O}(\epsilon)^2 \ . \end{aligned} \tag{3.10}$$

Therefore

$$e^{M(s)+\epsilon\,\Delta s\,\frac{\mathrm{d}}{\mathrm{d}s}M(s)} = e^{M(s)} + \epsilon\,\Delta s\, e^{M(s)} \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!}\, \Omega_n \left(M(s), \frac{\mathrm{d}}{\mathrm{d}s} M(s) \right) + \mathcal{O}(\epsilon)^2 \ , \tag{3.11}$$

from which one sees that

$$\frac{\mathrm{d}}{\mathrm{d}s} e^{M(s)} = e^{M(s)} \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!}\, \Omega_n \left(M(s), \frac{\mathrm{d}}{\mathrm{d}s} M(s) \right) \ . \tag{3.12}$$

This matches Eq. (3.18) of the Problem Set, with

$$B_n = \frac{(-1)^n}{(n+1)!} \ . \tag{3.13}$$

(e) Defining

$$F(s) \equiv e^{-sB}\, e^{-sA} e^{s(A+B)+C(s)} \ , \tag{3.14}$$

we can calculate the derivative by using Eq. (3.12) to differentiate the final factor. Thus

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}F(s) &= \frac{\mathrm{d}}{\mathrm{d}s}\left[e^{-sB}e^{-sA}e^{s(A+B)+C(s)}\right] \\
&= -Be^{-sB}e^{-sA}e^{s(A+B)+C(s)} - e^{-sB}Ae^{-sA}e^{s(A+B)+C(s)} \\
&\quad + e^{-sB}e^{-sA}e^{s(A+B)+C(s)} \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!}\,\Omega_n\left(s(A+B)+C(s), A+B+\frac{\mathrm{d}}{\mathrm{d}s}C(s)\right) \\
&= -BF(s) - e^{-sB}\,A\,e^{sB}\,F(s) \\
&\quad + F(s)\sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!}\,\Omega_n\left(s(A+B)+C(s), A+B+\frac{\mathrm{d}}{\mathrm{d}s}C(s)\right) \ .
\end{aligned} \tag{3.15}$$

(f) Eq. (3.16) can be simplified by expanding $e^{-sB}\,A\,e^{sB}$ in a power series in s, as in Eq. (3.13) of the Problem Set, and then using the fact that $F(s) \equiv 1$, $\mathrm{d}F/\mathrm{d}s = 0$. The result is

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}F(s) = 0 = -B - \sum_{n=0}^{\infty}\frac{(-s)^n}{n!}\Omega_n(B,A) \\
+ \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!}\,\Omega_n\left(s(A+B)+C(s), A+B+\frac{\mathrm{d}C(s)}{\mathrm{d}s}\right) \ .
\end{aligned} \tag{3.16}$$

To continue, it is helpful to write the first term of each summation explicitly:

$$\begin{aligned}
0 = -B - A - \sum_{n=1}^{\infty}\frac{(-s)^n}{n!}\Omega_n(B,A) + A + B + \frac{\mathrm{d}C(s)}{\mathrm{d}s} \\
+ \sum_{n=1}^{\infty} \frac{(-1)^n}{(n+1)!}\,\Omega_n\left(s(A+B)+C(s), A+B+\frac{\mathrm{d}C(s)}{\mathrm{d}s}\right) \ .
\end{aligned} \tag{3.17}$$

A simple rearrangement then gives the desired form:

$$\frac{\mathrm{d}C(s)}{\mathrm{d}s} = \sum_{n=1}^{\infty}\frac{(-s)^n}{n!}\Omega_n(B,A) - \sum_{n=1}^{\infty} \frac{(-1)^n}{(n+1)!}\,\Omega_n\left(s(A+B)+C(s), A+B+\frac{\mathrm{d}C(s)}{\mathrm{d}s}\right) \ . \tag{3.18}$$

(g) To proceed we expand $C(s)$ in a power series, noting that $F(s) = 1$ and Eq. (3.14) imply that $C(0) = 0$. So

$$C(s) = C_1\, s + C_2\, s^2 + C_3\, s^3 + \dots\ . \tag{3.19}$$

By rewriting Eq. (3.18) at order s^0 we get:

$$C_1 = 0\ , \tag{3.20}$$

while at order s we find

$$2C_2 s = -s[B, A] = 0 \quad \Longrightarrow \quad C_2 = \frac{1}{2}[A, B]\ , \tag{3.21}$$

and at order s^2 we find

$$\begin{aligned} 3C_3 s^2 &= \frac{s^2}{2}[B,[B,A]] + \frac{1}{2}[s(A+B) + C_2 s^2, A + B + 2C_2 s] \\ &= \frac{s^2}{2}[B,[B,A]] - \frac{1}{2}\left(\frac{s^2}{2}[[A,B], A+B]\right)\ , \end{aligned} \tag{3.22}$$

from which we conclude

$$C_3 = \frac{1}{12}\Big([A,[A,B]] + [B,[B,A]]\Big)\ .$$

(h) We have seen that Eq. (3.18) provides a recursion relation which allows us to calculate C_k at least up to $k = 3$. To continue indefinitely, one would like to calculate C_k by examining the equation at order s^{k-1}, so the left-hand side will be $kC_k s^{k-1}$. If we can show that the right-hand side can be determined by knowing only $C_1 \dots C_{k-1}$, then the recursion can be carried out indefinitely. That is, we must show that C_k and all higher terms in the series contribute to the right-hand side only at order s^k and higher. Since $C(s)$ occurs in only two places on the right-hand side of Eq. (3.18), the issue is easy to study. The first argument of

$$\Omega_n\left(s(A+B) + C(s), A + B + \frac{\mathrm{d}C(s)}{\mathrm{d}s}\right)$$

will include the term $C_k s^k$ and higher terms, but these are already of order s^k or higher, and there is certainly no way that negative powers of s can arise in computing the iterated commutator. The second argument contains $\mathrm{d}C(s)/\mathrm{d}s$, which includes $s^k C_k s^{k-1}$, which looks like it might prevent the recursion from working. However, the contribution to the right-hand side is found by calculating the nth order commutator of this term with the first argument, where n is summed and has a minimum value of 1. Since the first argument is of order s, the leading possible contribution would be of order s^k, and is therefore harmless. Thus the recursion relation (3.18) determines the series C_k for all k.

To see that each C_k is a linear combination of iterated commutators of A and B, we proceed by induction on k. It is true for $k = 1, 2$, and 3, as we computed explicitly, so suppose that it is true up to $k - 1$.

We have just shown that Eq. (3.18) allows us to express C_k in terms of the commutators which appear on the right-hand side, which involve only the operators A, B, and $C_1 \ldots C_{k-1}$. By the induction hypothesis $C_1 \ldots C_{k-1}$ are all expressed as iterated commutators, so any commutator of these objects is itself an iterated commutator. QED.

I should mention that if one is looking for the most efficient way to compute the terms in the Baker-Campbell-Hausdorff expansion, Eq. (3.18) is certainly not it. Our goal in deriving this equation was to show that every term in the series can be written as an iterated commutator. If one did not care about showing this fact, then there are easier ways to expand both sides of

$$e^A e^B = e^{A+B+C}$$

so that C can be determined.

---

†Solution written by Joydip Kundu, Tom Faulkner, and Alan Guth.

¶Solution written by Guido Festuccia and Alan Guth.