Isospin:
An Approximate Symmetry on the Quark Level

Molly S. Peeples

molly@mit.edu

Cambridge, MA 02139

April 27, 2004

Abstract

Isospin is an approximate symmetry which treats the up and down quarks as different eigenstates of the same particle. The mathematical structure for describing the isospin of a system is identical to that of angular momentum. We explore the implications of isospin. Specifically, we use isospin to predict the ratios of cross-sections in pion-nucleon scattering with incredible accuracy. We also derive the Gell-Mann–Okubo formula for baryons, $2(m_N + m_\Xi) = 3m_\Lambda + m_\Sigma$, which correctly predicts the m_Λ mass within 1%. This method of developing spurious symmetries is a powerful tool in quantum field theory.

1 Introduction and Motivation

Hadrons are particles comprised of quarks and governed primarily by the strong, or "hadronic" force. If we have a system of hadrons, then our overall Hamiltonian is $\mathcal{H} = \mathcal{H}^0_{\text{strong}} + \mathcal{H}'_{\text{other}}$, where $\mathcal{H}^0_{\text{strong}}$ is due to the strong interactions—the main governing force for quarks and nuclei. $\mathcal{H}'_{\text{other}}$ is a perturbative term that takes into account all of the other forces affecting our system. For instance, the Coulomb energy of the system is in this term because the strong force is independent of electromagnetic charge, and the electromagnetic force is about 10^5 times weaker than the strong force. $\mathcal{H}^0_{\text{strong}}$ itself can be broken up into symmetry-preserving and symmetry-breaking components, as we will see in §6. The symmetries we'll be studying are based on

treating different quarks as different eigenstates of the same particle, rather than as different particles themselves. These will be *approximate* symmetries; our goal is to get a feel for how good the approximation is. We know that perturbation theory is good when the perturbing factor is small compared to $\mathcal{H}^0$. In our case, the relevant energy scale is $\Lambda_{\text{QCD}} \approx 500\text{MeV}^*$, and the perturbative energies will be the differences in quark masses, as discussed in §§2.1.[3]

Once we've developed the basic particle physics background we'll need, we'll see how isospin—an approximate symmetry which holds with respect to $\mathcal{H}^0_{\text{strong}}$—can be used to accurately predict scattering cross-sections for pion-nucleon scattering. We'll then expand our model to include three quarks instead of only two. We will derive the Gell-Mann–Okubo formula, which provides an extremely good approximate relation between baryon masses.

2 Some Introductory Particle Physics

2.1 Quarks

As mentioned in §1, quarks are the elementary particles that make up hadrons. We'll only concern ourselves with the three lightest quarks in this paper, but for completeness, all six are summarized in Table 1. Quarks can combine to make two kinds of hadrons that we will consider in this paper. The first is the baryon, which is a bound system of three quarks, such as the proton. The second is the meson, which is a bound quark-antiquark pair, such as a kaon. Everyday matter is made up of baryons comprised of up and down quarks, that is, protons and neutrons. Particles composed of heavier quarks are much less stable because a much higher energy is required to keep them intact. Quarks are confined to hadrons, so unlike leptons—such as electrons and neutrinos—they are not observable as independent physical entities. Determining their masses is therefore much more a matter of applying theories than taking direct measurements. Methods similar to those used in §6 have been used to put limits on the masses of the lighter quarks. [8]

*Hadrons are typically of size $\sim \Lambda_{\text{QCD}}^{-1}$

Flavor	Symbol	≈Mass in MeV
up	$\|u\rangle$	4.2
down	$\|d\rangle$	7.5
strange	$\|s\rangle$	150
charm	$\|c\rangle$	1500
bottom	$\|b\rangle$	4700
top	$\|t\rangle$	176000

Table 1: Speculated quark masses.[8]

2.2 Internal Quantum Numbers

The typical 8.06 student is used to dealing with external quantum numbers, like n, ℓ, and m in Hydrogen. External quantum numbers can change—electrons in the same atom, for instance, cannot have the same set of numbers. Internal quantum numbers, on the other hand, correspond to properties inherent to the particle in question; they help us identify and label the particle. For instance, the total spin of an electron is *always* $\frac{1}{2}$, but the z-component can change. The total spin of the electron is an internal quantum number, but the z-component is external.

"Good" internal quantum numbers are conserved in all interactions. The electromagnetic charge Q is a "good" example of this; an interaction in which the total charge of the system has changed has never been observed. Many quantum numbers simply count useful quantities. For instance, the baryon number B is always conserved. Baryons get a baryon number of $+1$, while antibaryons get a baryon number of -1; everything else has $B = 0$. Another way of thinking of this is to say that every quark has a baryon number of $+1/3$, while antiquarks have a baryon number of $-1/3$. S, or strangeness, is similar, but instead counts the number of strange quarks—in units[†] of $-1/3$. Strangeness is a "mostly good" quantum number; it is conserved in all interactions except for ones governed by the weak force.[‡] As we will see in §5, the hypercharge $Y \equiv B + S$ of a particle will be more useful than talking about just the particle's strangeness. As for antimatter, a particle has the same mass as its antiparticle, but opposite quantum numbers. For instance, $|\bar{u}\rangle$ has the same mass as $|u\rangle$, but a $Q = -2/3$, $B = -1/3$, and $S = 0$. (The bar in $|\bar{u}\rangle$ notes that this

[†]An unfortunate remnant of history

[‡]We observed this when we studied kaon decay in 8.05.

is the antiparticle.)[2, 4]

2.3 Invariance

The concept of symmetry is crucial to physics. When we say something is symmetric, we mean that it it doesn't change under a transformation. There is a famous theorem by Noether that states that there is an inherent connection between transformations (and thus symmetries) and conservation laws. Applied to physics, this is the basis for conservation of momentum, which arises from the fact that the origin for our coordinates and the directions of our axes are arbitrary—we can transform the system by translating or rotating it, but the physics is still the same. Similarly, an abitrary time $t = 0$ point leads to a conservation of energy. Charge conservation is due to gauge transformations; our zero-point of A or ϕ is arbitrary because we can alter both expressions with an aribitrary $f(\vec{x}, t)$ without changing the physics of the system. [1]

We can also talk about transformations in eigenspace. Suppose we have a two-state system, $|\uparrow\rangle$ and $|\downarrow\rangle$. We know that any 2×2 Hermitian operator can be written, in matrix form, as $\hat{\mathcal{A}} = a_0\mathbb{I} + a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 = a_0\mathbb{I} + \vec{a} \cdot \vec{\sigma}$. The generators for this matrix, the Pauli matrices σ_i are said to be generators for SU_2; with the identity $\mathbb{I}$, they form a basis for Hermitian 2×2 matrices. (SU_n, sometimes written as SU(n), means Special Unitary group in n dimensions—the n here is the same n in the $n \times n$ matrices/n-state system we're talking about. The "special" means that they have determinant 1. As far as the "unitary" is concerned, remember that for any Hermitian operator $\mathcal{A}$, $e^{i\mathcal{A}}$ is unitary.) It is very easy here to use the language of rotations because SU_2 also describes rotations in three-dimensional space. For instance, our system might be invariant under a rotation of $|\uparrow\rangle$ to $|\downarrow\rangle$, by which we mean that if we exchange *all* of our $|\uparrow\rangle$ with $|\downarrow\rangle$, then the physics our equations describe is the same as before the transformation—it is symmetric. If we increase n such that we have a 3 or 4 state system, the language is the same, but our math is a little messier because we are then dealing with the 3 or 4 dimensional analogues of the Pauli matrices. We will also find that a system with a greater number of states is much less likely to be invariant under transformations.

3 Isospin and SU_2

As you can tell by glancing at Table 1, the masses of the up and down quarks are both very small and very close to one another, relative to $\Lambda_{\mathrm{QCD}} \approx 500\mathrm{MeV}$. Because of this, we can then describe them as if they are different eigenstates of the same particle, but with different *isospin*. As the name suggests, this is analogous to spin.

Let us say we have an isospinor $\vec{\mathrm{I}} \equiv (\mathrm{I}_1, \mathrm{I}_2, \mathrm{I}_3)$. Like with angular momentum, the algebra of this system is described by

$$[\hat{I}_k, \hat{I}_l] = i\varepsilon_{klm}\hat{I}_m. \tag{1}$$

$|u\rangle$ has an isospin of $+1/2$, that is, $I_3 |u\rangle = \frac{1}{2} |u\rangle$, while $I_3 |d\rangle = -\frac{1}{2} |d\rangle$. Another way of writing this is

$$\hat{I}_3 = \begin{pmatrix} +1/2 & \\ & -1/2 \end{pmatrix} \text{ in the } \begin{pmatrix} u \\ d \end{pmatrix} \text{ basis.} \tag{2}$$

In general, we'll write our states as $|I\ I_3\rangle$, where

$$\hat{I}^2 |I\ I_3\rangle = I(I+1) |I\ I_3\rangle \tag{3a}$$

$$\hat{I}_3 |I\ I_3\rangle = I_3 |I\ I_3\rangle \tag{3b}$$

As with spin, there are $2I + 1$ possible eigenvalues of I_3, ranging from $-I$ to I. For $|u\rangle$ and $|d\rangle$, Equations (3) work more as definitions than anything interesting. But there is a lot we can learn by combining the states, i.e. creating particles.

It is postulated that the strong interactions are invariant under isospin rotations. That is to say, if *all* up quarks were replaced with down quarks, the strong interactions would be unchanged. This is an SU_2 symmetry, as explained in §2.3. As a quantum number, isospin is somewhat-good; it is conserved only by the strong force, but not conserved in all other interactions. (For instance, $|u\rangle$ and $|d\rangle$ will interact differently under electromagnetism because their electromagnetic charges differ.)

The proton and neutron differ by one up or down quark: $|p\rangle = |uud\rangle$ and $|n\rangle = |udd\rangle$. If our approximation is valid, then the proton and the neutron would not only have about the same mass (i.e., energy), but would also act the same in strong interactions. The first prediction is indeed very true; note that the p and n masses differ by less than 0.5% of a proton mass [8]. But what about the strong interactions?

Let us consider pion-nucleon scattering, $\pi N \to \pi N$. First of all, the nucleon, N, is a two-state system. Its isospin-up eigenstate is the proton p, and its isospin-down

eigenstate is the neutron, n. (That is, we are considering the proton and the neutron to be two different eigenstates of the same particle.) As for pions, they are mesons, which are quark-antiquark pairs. Quarks have opposite quantum numbers as their antiquark pairs, so $|\bar{u}\rangle$ and $\left|\bar{d}\right\rangle$ have isospin $-\frac{1}{2}$ and $+\frac{1}{2}$ respectively. The π-meson triplet is a isospin-1 system. Pions and nucleons are summarized in Table 2. [2, 6]

State	Composition	$\|I\ I_3\rangle$
$\|\pi^+\rangle$	$\left\|u\bar{d}\right\rangle$	$\|1\ +1\rangle$
$\|\pi^-\rangle$	$\|d\bar{u}\rangle$	$\|1\ -1\rangle$
$\|\pi^0\rangle$	$\frac{1}{\sqrt{2}}(\|u\bar{u}\rangle - \left\|d\bar{d}\right\rangle)$	$\|1\ 0\rangle$
$\|p\rangle$	$\|uud\rangle$	$\left\|\frac{1}{2}\ +\frac{1}{2}\right\rangle$
$\|n\rangle$	$\|udd\rangle$	$\left\|\frac{1}{2}\ -\frac{1}{2}\right\rangle$

Table 2: Summary of pion and nucleon quark composition and isospin.[2, 6]

If we take into account just elastic scattering (that is, we're not going to end up with any mesons or anything else that isn't just a pion or a nucleon), then we need only pay attention to only isospin conservation. Isospin must be conserved because this is a strong interaction. The total isospin can therefore either be $\frac{1}{2}$ or $\frac{3}{2}$. We can predict the scattering amplitudes (i.e., the amounts of mixture) $\mathcal{P}_{1/2}$ and $\mathcal{P}_{3/2}$ using the Clebsch-Gordan coefficients [6]:

$$\begin{aligned}
|\pi^+p\rangle &= |11\rangle\left|\tfrac{1}{2}\ \tfrac{1}{2}\right\rangle &&= \left|\tfrac{3}{2}\ \tfrac{3}{2}\right\rangle \\
|\pi^0p\rangle &= |10\rangle\left|\tfrac{1}{2}\ \tfrac{1}{2}\right\rangle &&= \sqrt{2/3}\left|\tfrac{3}{2}\ \tfrac{1}{2}\right\rangle - \sqrt{1/3}\left|\tfrac{1}{2}\ \tfrac{1}{2}\right\rangle \\
|\pi^-p\rangle &= |1\ -1\rangle\left|\tfrac{1}{2}\ \tfrac{1}{2}\right\rangle &&= \sqrt{1/3}\left|\tfrac{3}{2}\ -\tfrac{1}{2}\right\rangle - \sqrt{2/3}\left|\tfrac{1}{2}\ -\tfrac{1}{2}\right\rangle \\
|\pi^+n\rangle &= |11\rangle\left|\tfrac{1}{2}\ -\tfrac{1}{2}\right\rangle &&= \sqrt{1/3}\left|\tfrac{3}{2}\ \tfrac{1}{2}\right\rangle + \sqrt{2/3}\left|\tfrac{1}{2}\ \tfrac{1}{2}\right\rangle \\
|\pi^0n\rangle &= |10\rangle\left|\tfrac{1}{2}\ -\tfrac{1}{2}\right\rangle &&= \sqrt{2/3}\left|\tfrac{3}{2}\ -\tfrac{1}{2}\right\rangle + \sqrt{1/3}\left|\tfrac{1}{2}\ -\tfrac{1}{2}\right\rangle \\
|\pi^-n\rangle &= |1\ -1\rangle\left|\tfrac{1}{2}\ -\tfrac{1}{2}\right\rangle &&= \left|\tfrac{3}{2}\ -\tfrac{3}{2}\right\rangle
\end{aligned} \tag{4}$$

The j in $\mathcal{P}_j$ corresponds to the I in the $|I\ I_3\rangle$ on the righthand side. For example, $\langle\pi^-p|\,\mathcal{P}\,|\pi^-p\rangle$ is given by

$$\left\langle\pi^-p\right|\mathcal{P}\left|\pi^-p\right\rangle = \sqrt{\frac{1}{3}}\sqrt{\frac{1}{3}}\,\mathcal{P}_{3/2} + \sqrt{\frac{2}{3}}\sqrt{\frac{2}{3}}\,\mathcal{P}_{1/2} = \frac{1}{3}\mathcal{P}_{1/2} + \frac{2}{3}\mathcal{P}_{3/2} \tag{5}$$

Similarly, $\langle\pi^-p|\,\mathcal{P}\,|\pi^0n\rangle = \frac{\sqrt{2}}{3}\mathcal{P}_{3/2} - \frac{\sqrt{2}}{3}\mathcal{P}_{1/2}$. The amplitude squared gives us the probability of having a certain outcome (e.g., $|\pi^0n\rangle$) given a certain input (e.g., $|\pi^-p\rangle$), so these brackets are then related to the cross-sections σ by

$$\sigma^+ : \sigma^0 : \sigma^- = 9|\mathcal{P}_{3/2}|^2 : 2|\mathcal{P}_{3/2} - \mathcal{P}_{1/2}|^2 : |\mathcal{P}_{3/2} + 2\mathcal{P}_{1/2}|^2. \tag{6}$$

Here, σ^+ denotes the cross-section associated with a π^+ being involved, etc. As it so happens, the pion and the nucleon form a "resonance" state—the Δ—which then quickly decays. The Δ is an isospin $\frac{3}{2}$ particle, so it might be reasonable to assume $\mathcal{P}_{3/2} \gg \mathcal{P}_{1/2}$. This assumption lets us simplify (6) to

$$\sigma^+ : \sigma^0 : \sigma^- = 9 : 2 : 1. \tag{7}$$

This is an incredibly good approximation; the experimental ratios are $(9.53 \pm 0.63) : (2.0 \pm 0.1) : (1.0 \pm 0.1)$ for pion kinetic energy from 120 MeV to 300 MeV. This means that the scattering is most likely to take place in the $I = 3/2$ regime.[2, 6]

Our main tool in making this calculation was isospin—the assumption that $|u\rangle$ and $|d\rangle$ are merely different eigenstates of the same particle. (We also assumed that scattering is in the $I = \frac{3}{2}$ channel, but this only helped us with a final simplification.) We know isospin to be an approximate symmetry because $|u\rangle$ and $|d\rangle$ have different masses and different charges, but it is still an extremely powerful appoximation, especially once we consider $|s\rangle$ as well. To do this, however, we will have need a slightly more developed mathematical background.

4 SU_3

The Gell-Mann matrices λ_j form the standard basis for 3×3 hermitian matrices, the same way the Pauli matrices formed a basis in SU_2. (Remember that any 2×2 hermitian matrix can be written as $\hat{A} = a_0 \mathbb{I}_2 + \vec{a} \cdot \vec{\sigma}$. Now, we have that any hermitian 3×3 matrix can be written as $\hat{D} = b_0 \mathbb{I}_3 + \vec{d} \cdot \vec{\lambda}$.) Like the Pauli matrices, the λ_j are all traceless. Furthermore, $\exp(i\lambda_j)$ is unitary. If we let the identity matrix be λ_0, then there are nine matrices.[§] The first three matrices are easy; they're just the Pauli matrices again, such that $\lambda_j = \begin{pmatrix} \sigma_j & 0 \\ 0 & 0 \end{pmatrix}$:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{8a}$$

[§]Remember that when we included the identity matrix as σ_0 in U_2, we had $2^2 = 4$ basis matrices. We now have the same thing in U_3: $3^2 = 9$ matrices.

The next four Gell-Mann matrices are similarly constructed, but the extra 0's don't go in the third row and column. This gives us:

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \lambda_5 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \tag{8b}$$

The final Gell-Mann matrix,

$$\lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \tag{8c}$$

is special in that we can't construct it directly from the Pauli matrices, and because only it and λ_3 are diagonal. We'll come back to this point in §6.[4]

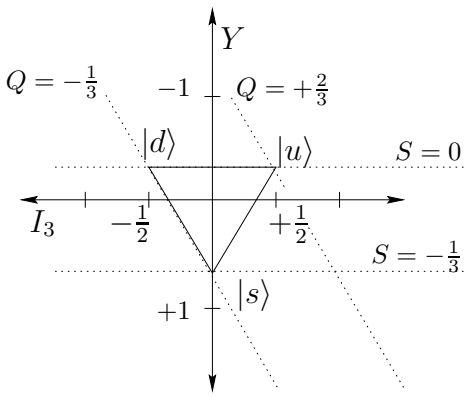
5 "The Eightfold Way"

Looking back at Table 1 in §2.1, we can also notice that the strange mass is small, at least compared to Λ_{QCD}. If we consider particles composed of just up, down, and strange quarks, there is much to be learned. Before we begin, however, we need a good way to visualize our particles. Gell-Mann introduced plotting Y versus I_3, as shown in the simple example of the $|u\rangle$, $|d\rangle$, and $|s\rangle$ triplet in Figure 1(a). Considering just these three quarks, the hypercharge, Y, and the electromagnetic charge, Q, are related by the Gell-Mann–Nishijima relation, $Q = I_3 + \frac{1}{2}Y$, which is in turn related to why hypercharge is called "*hyper*charge".[2]

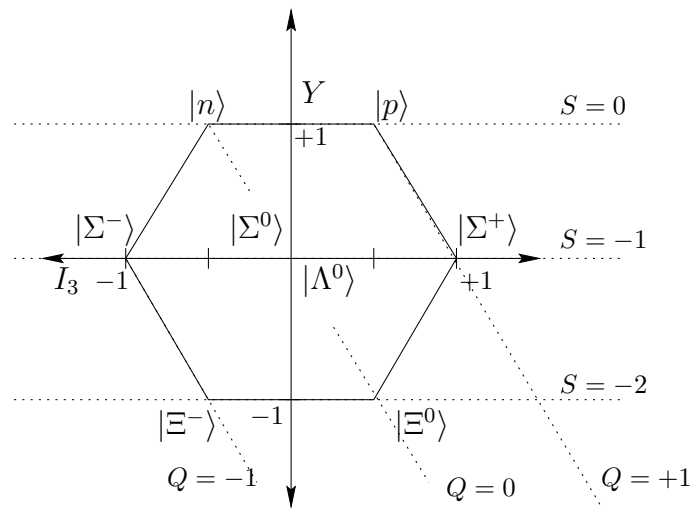
If we take into account baryons comprised of just the three lightest quarks, there are ten completely symmetric states (which form a decuplet), one completely antisymmetric state (the singlet), and eight states that are antisymmetric under one interchange (an octet).¶ Gell-Mann initially started studying this octet, as well as a similar octet of mesons, so this method of grouping hadrons has become known as "the eightfold way." We will be focussing on the octet of baryons, as shown in Figure 1(b).

The naïve way of doing math with this system is to create an 8-dimensional vector to store all of eight of the baryons in. We could then let all of our operators be 8×8

¶For those of you with a bit of an algebra background, these correspond to irreducible representations of SU_3.



(a) The three lightest quarks.



(b) The baryon octet. $|\Sigma^0\rangle$ and $|\Lambda^0\rangle$ are at the origin.

Figure 1: Weight diagrams. Note that (a) is an equilateral triangle and (b) is a regular hexagon.[2, 4, 6]

matrices. But because this is big and clumsy, and because we can, we will condense it into a 3-dimensional representation, the B (baryon) matrix, where

$$B = \begin{pmatrix} \frac{\Sigma^0}{\sqrt{2}} + \frac{\Lambda}{\sqrt{6}} & \Sigma^+ & p \\ \Sigma^- & \frac{-\Sigma^0}{\sqrt{2}} + \frac{\Lambda}{\sqrt{6}} & n \\ \Xi^- & \Xi^0 & \frac{-2\Lambda}{\sqrt{6}} \end{pmatrix}. \tag{9}$$

B will transform like SU_3, that is, all of the math that has already been developed for SU_3 is also applicable to B. In particular, our old 8×8 operators can now be written as 3×3 unitary matrices. We can now describe what we had before as $\hat{A}_{8\times 8}\vec{b}$ as $U^\dagger_{3\times 3} B U_{3\times 3}$, where $\vec{b}$ is an eight-dimensional vector, B is as defined in (9), and $U_{3\times 3}$ is in fact unitary. This works because $3 \otimes \bar{3} = 1 \oplus 8$, where the bar denotes that we need to use the hermitian conjugate of one of the operators. (This tensor multiplication/addition is similar to what we did with angular momentum, where, for instance, we found that $\frac{1}{2} \otimes \frac{1}{2} = 0 \oplus 1$.)

If this symmetry were exact—that is, if $|u\rangle$, $|d\rangle$, and $|s\rangle$ were in fact just different flavors of the same particle—then all of these 8 baryons would have the same mass.

While their masses are close—they range from 940MeV for the nucleon to 1320MeV for Ξ—they are definitely not close enough to be "equal". However, we can still learn a lot by applying the SU_3 symmetry (all three quarks have the same mass), and then breaking it in such a way that isospin invariance still holds ($m_u = m_d \neq m_s$). [2, 4]

6 The Gell-Mann–Okubo Formula

We want to now try to learn something about the masses of the baryons shown in Figure 1(b). Following [4], we can write the Hamiltonian describing the strong interactons as the sum of two parts of different strength,

$$\mathcal{H}_S = \mathcal{H}_{VS} + \mathcal{H}_{MS}. \tag{10}$$

Think of Equation (10) as a perturbative equation. The first term is the contribution from the "very strong" force; it preserves the SU_3 symmetry that treats $m_u = m_d = m_s$. The second part is our perturbative, symmetry-breaking term; it is the contribution from a "medium-strong" force that treats $m_u = m_d \neq m_s$. Note that both $\mathcal{H}$ on the right hand side preserve isospin invariance. We don't know what exactly $\mathcal{H}_{MS}$ is, but

$$\mathcal{H}_{MS} = \lambda_8 O \tag{11}$$

is a good guess. Looking back to (8), we note that λ_8 is a good choice here if we want $|u\rangle$ and $|d\rangle$ to be treated equivalently. λ_8 is also diagonal, which will make the math easier. O is an operator describing the actual physics of the system; it has some combination of particle creation and annihilation operators, etc. in it that we don't know exactly. But we *do* know that its matrix representation must have a certain shape in order to preserve the symmetries we want it to.

To find the first order correction to the energy (or, equivalently, the masses of our baryons), we must calculate the matrix element $\langle B| \mathcal{H}_{MS} |B\rangle$. $\langle B| O |B\rangle$ will be invariant if we transform both B and O; that is, $\langle B| UU^\dagger OUU^\dagger |B\rangle = \langle B| O |B\rangle$. This is like rotating both our axes and a vector we are considering. On the other hand, if we want something that is invariant under a transformation of just B or O, then there are only two such quantities. That is, we find that there are only two independent factors in the matrix for $\mathcal{H}_{MS}$. As matrix elements, these are

$$\begin{aligned} B^\dagger_{ij}[\lambda_8]_{jk}B_{kl} &= \mathrm{Tr}(BB^\dagger\lambda_8) \\ \text{and} \quad B^\dagger_{ij}[\lambda_8]_{jk}B_{kl} &= \mathrm{Tr}(B^\dagger B\lambda_8). \end{aligned} \tag{12}$$

For instance, if we take $B \to U^\dagger BU$, then $\mathrm{Tr}(BB^\dagger\lambda_8) \to \mathrm{Tr}(U^\dagger BUU^\dagger B^\dagger U\lambda_8)$, which is *just* $\mathrm{Tr}(U^\dagger UBB^\dagger\lambda_8) = \mathrm{Tr}(BB^\dagger\lambda_8)$ because trace is cyclic.[3] (The fact that there are only two such independent quantities also follows from the Wigner-Eckhart theorem.) Combining Equations (12), we have

$$\begin{aligned}\langle B|\,\mathcal{H}_{MS}\,|B\rangle &= X\mathrm{Tr}(BB^\dagger\lambda_8) + Y\mathrm{Tr}(B^\dagger B\lambda_8)\\ &= X([B^\dagger B]_{11} + [B^\dagger B]_{22} - 2[B^\dagger B]_{33})/\sqrt{12}\\ &\quad Y([BB^\dagger]_{11} + [BB^\dagger]_{22} - 2[BB^\dagger]_{33})/\sqrt{12}.\end{aligned} \tag{13}$$

(Remember that we are summing over repeated indices.) Looking back at our definition of B in (9), we notice that this gives us a relationship between all eight particles. $[B^\dagger B]_{ii}$ gives us the sum of the squares of the elements in the ith row of B, while $[BB^\dagger]_{jj}$ gives us the sum of the squares of the elements in the jth column. Happily, when it is all combined, many terms drop out such that we get

$$\begin{aligned}\langle B|\,\mathcal{H}_{MS}\,|B\rangle &= X(|\Sigma|^2 + |\Xi|^2 - |\Lambda|^2 - 2|n|^2)/\sqrt{12}\\ &\quad + Y(|\Sigma|^2 + |n|^2 - |\Lambda|^2 - 2|\Xi|^2)/\sqrt{12}.\end{aligned} \tag{14}$$

Now we are ready to write a relationship between the masses of these four particles. Note that the masses and energies are equivalent here, that is, $m_\Sigma = \langle\Sigma|\,\mathcal{H}_S\,|\Sigma\rangle = \langle\Sigma|\,\mathcal{H}_{VS}\,|\Sigma\rangle + \langle\Sigma|\,\mathcal{H}_{MS}\,|\Sigma\rangle$, etc. If we add back in the mass m_0 common to all four from the $\mathcal{H}_{VS}$ contribution, we get a system of equations:

$$\begin{aligned} m_n &= m_0 - 2X/\sqrt{12} + Y/\sqrt{12}\\ m_\Sigma &= m_0 + X/\sqrt{12} + Y/\sqrt{12}\\ m_\Lambda &= m_0 - X/\sqrt{12} - Y/\sqrt{12}\\ m_\Xi &= m_0 + X/\sqrt{12} - 2Y/\sqrt{12}\end{aligned} \tag{15}$$

There are four equations and three unknowns (m_0, X, and Y); after a little massaging, we find that

$$2(m_N + m_\Xi) = 3m_\Lambda + m_\Sigma. \tag{16}$$

This is known as the Gell-Mann–Okubo formula.[5] We take $m_N = 940$, $m_\Sigma = 1190$, and $m_\Xi = 1320$, all in MeV, where m_Σ is the average mass of Σ^-, Σ^0, and Σ^+, and so forth. Plugging these values in to (16), we predict that the Λ mass will be 1110 MeV, which is an astonishingly good approximation. The experimental result is 1115 MeV; our approximation is good to within 1%.[6]

7 Conclusion

As we have seen, the perturbative method of creating symmetries and then breaking them works as an extremely good modelling technique. We were able to predict ratios of pion-nucleon scattering with incredible precision using just isospin symmetry, Clebsch-Gordon coefficients, and simple scattering theory. We also found the Λ mass within one percent of the experimental result using only isospin symmetry, the Wigner-Eckhart theorem, and initial mass values of three baryons that were good to only about 0.5%. While isospin is not physical—the up and down quarks are *not* different eigenstates of the same particle!—this approximation works because their energies are well below the characteristic energy of the strong force. This method furthermore lets us use an old, familiar language to discuss a new system, which helps us gain a better intuition and feel for what is really happening.

8 Historical Aside

Isospin was studied long before quarks were postulated. In 1920, when Rutherford first postulated existence of a neutral subatomic particle, he speculated that its mass would be about that of the proton. It wasn't until 1932, shortly after the neutron's discovery, that Heisenberg started playing with rotations of p and n using the Pauli matrices and all of the SU_2 algebra in the non-physical space we now call isospin space. In fact, it wasn't until 1937 that the term "isospin" began to develop. Wigner initially called it "isotopic spin"—which is actually a misnomer. Isotopes are not related at all when it comes to isospin because two isotopes have differing numbers of neutrons. *Isobars*, such as ^{57}Fe and ^{57}Co, etc., on the other hand, do form isospin multiplets because they do have the same number of nucleons.

When Heisenberg first introduced the concept of the nucleon, he was trying to figure out what exactly the neutron was. At first the idea was that it was a tightly-bound proton-electron state, but that did not take long to debunk. Once it was decided that the neutron was its own particle, it became even more important to figure out what kind of forces held together the atomic nucleus. The big revolution came when it was realized that this force had to be *charge independent*—that is, a proton-proton interaction (or neutron-neutron) had to be just as strong as a proton-neutron interaction. This was a big start in answering many questions, from why nuclei are more likely to have an even number of particles to why the proton and

neutron have masses that are so close. [4, 7]

In the early 1960s, Gell-Mann and others were trying to rationalize the observed mass spectra of hadrons, specifically why particles seemed to come in "families", like the one shown in Figure1(b). Gell-Mann postulated that these multiplets could be explained by an "orderly-broken symmetry"—isopsin.[5] He developed most of the math needed to explain the structures of these multiplets, only to later discover that this was a field well-known to mathematicians. It was also Gell-Mann who postulated that there were certain multiplets, but not others, because the hadrons were made of still smaller and more fundamental particles, which he called quarks. He then predicted the existence of the Ω^- particle, which was eventually discovered. The quark model has since held up, leading to the current view of isospin.[1]

9 Acknowledgements

I would like to thank Austin Roach, Joel Corbo, and Krishna Rajagopal for their help with this paper, with special thanks to Michael Forbes.

References

[1] Beiser, Arthur. Modern Physics. McGraw Hill, 2003.

[2] Fayyazuddin & Riazuddin. A modern introduction to particle physics. World Scientific, 2000.

[3] Forbes, Michael. Personal communication, April 2004.

[4] Georgi, Howard. Lie Algebras in Particle Physics: From Isospin to Unified Theories. Perseus Books, 1999.

[5] Gell-Mann, Murray and Yuval Ne'eman. The Eightfold Way: A Review. W. A. Benjamin, Inc., 1964.

[6] Griffiths, David. Introduction to Particle Physics. John Wiley & Sons, Inc., 1987.

[7] Wilkinson, D. H., editor. Isospin in Nuclear Physics. North-Holland Publishing Company, 1969.

[8] Particle Data Group. http://pdg.lbl.gov/