4

The quark model: combinatorics and groups

The quark model was initially introduced to explain the emerging plethora of hadrons as bound states of quarks, and this is how we start with its study.

4.1 Bound states

Hadrons – mesons and baryons – are bound states of (anti-)quarks. Mesons are quark–antiquark bound states, held together by the strong nuclear force for which the mediating quanta (particles) are the *gluons*. Baryons consist of three quarks, bound by gluons. It is then clear that the three-particle bound states – baryons – are more complicated than mesons. Furthermore, even the description of mesons as bound states is hard in the case of the "light" quarks. For these one needs a relativistic description of bound states, which is considerably more complicated than the non-relativistic one, and needs more development. One can reliably discuss, using the methods of non-relativistic quantum theory, only the mesons consisting of the "heavier" quarks: the t-, b- and c-quarks and with less precision also the bound states with the s-quark.

The bound states may be analyzed in this way, as a non-relativistic system, if the mass of all constituents is sufficiently bigger than the binding energy. For example, the binding energy of the hydrogen atom (13.6 eV) is 2.66×10^{-5} times that of the electron rest energy, and 1.45×10^{-8} times the proton rest energy, so that the non-relativistic analysis of the hydrogen atom is very accurate. This analysis may then be adapted to many hadronic systems, and we first recall some of the important results about the hydrogen atom.

Unlike the hydrogen atom and the similar positronium, the bound states of quarks and antiquarks will additionally require combinatorial and group-theoretical results since, in addition to the electric charge and spin, quarks also have a "flavor" (u, d, c, s, b, t) and a "color" (red, yellow, blue). With this in mind, the Reader is referred to the group-theoretical results collected in Appendix A to begin with, and the literature [565, 258, 581, 256, 80, 260, 333, 447] for more complete explanations, proofs and detailed theory, which also offer a more complete and pedagogical organization.

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

4.1.1 The non-relativistic hydrogen atom without spin

Together with the linear harmonic oscillator, the hydrogen atom is the most frequently discussed system in all books on quantum mechanics [362, 363, 471, 328, 480, 134, 391, 407, 360, 472, 29, 339, 242, 3, 110, 324, for example] [☞ also Section 1.2.5]. It is well known that in this two-particle system one can separate the dynamics of the atom as a whole and the relative motion of the electron in the CM-system. Since the proton (nucleus) mass is 1,836.15 times larger than the electron mass, the so-called reduced mass, $\frac{m_e\,m_p}{m_e+m_p} \approx m_e$ differs from the electron mass only by a $\frac{1}{1{,}837.15}$ fraction, which may usually be neglected. Besides, the coordinate origin of the CM-system is very close to the proton location,[1] and one approximates the electron as moving within the electrostatic field of the stationary proton.

The Schrödinger equation is

$$i\hbar\,\frac{\partial}{\partial t}\,\Psi(\vec{r},t) = H\,\Psi(\vec{r},t), \qquad H = \Big[-\frac{\hbar^2}{2m_e}\vec{\nabla}^2 + V(r)\Big], \tag{4.1}$$

where H is the non-relativistic Hamiltonian and $V(r)$ is a central potential, which for the hydrogen atom is the Coulomb potential. This may be solved in spherical coordinates, looking for solutions in the form of so-called stationary states:

$$\Psi_{n,\ell,m}(\vec{r},t) = e^{-i\omega_{n,\ell,m}t}\,R_{n,\ell}(r)\,Y_\ell^m(\theta,\phi), \qquad \omega_{n,\ell,m} := \frac{E_{n,\ell,m}}{\hbar}, \tag{4.2}$$

since we know that

$$\vec{\nabla}^2(\cdots) = \frac{1}{r^2}\Big(\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\cdots\Big) - \frac{1}{r^2}L^2(\cdots) \;=\; \frac{1}{r}\Big(\frac{\partial^2}{\partial r^2}(r\,\cdots)\Big) - \frac{1}{r^2}L^2(\cdots), \tag{4.3}$$

$$L^2 Y_\ell^m(\theta,\phi) = \ell(\ell{+}1)\,Y_\ell^m(\theta,\phi). \tag{4.4}$$

The $Y_\ell^m(\theta,\phi)$ are the spherical harmonics, the eigenfunctions of the angular part of the Laplacian:

$$L^2(\cdots) := -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\Big(\sin\theta\frac{\partial}{\partial\theta}\cdots\Big) - \frac{1}{\sin^2\theta}\Big(\frac{\partial^2}{\partial\phi^2}\cdots\Big), \tag{4.5}$$

which is the square of the operator of so-called "dimensionless angular momentum," i.e., of the angular momentum divided by $\hbar^2$. The radial part of $\vec{\nabla}^2$ and the identity

$$\frac{1}{r^2}\Big(\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\,\cdots\Big) \;\equiv\; \frac{1}{r}\Big(\frac{\partial^2}{\partial r^2}r\,\cdots\Big) \tag{4.6}$$

suggest the substitution $R_{n,\ell}(r) = \frac{u_{n,\ell}(r)}{r}$, whereby the radial differential equation becomes

$$-\frac{\hbar^2}{2m_e}\frac{\mathrm{d}^2 u_{n,\ell}}{\mathrm{d}r^2} + \Big[V(r) + \frac{\hbar^2}{2m_e}\frac{\ell(\ell{+}1)}{r^2}\Big]u_{n,\ell} = E_{n,\ell}\,u_{n,\ell}, \tag{4.7}$$

which is effectively a one-dimensional problem, with $r \in [0,\infty)$ and where the effective potential is the sum of the "actual" potential and the "centrifugal barrier," $\frac{\hbar^2}{2m_e}\frac{\ell(\ell+1)}{r^2}$. For the hydrogen atom, we have the Coulomb potential,

$$V(r) = -\frac{1}{4\pi\epsilon_0}\frac{e^2}{r} = -\frac{\alpha_e\hbar c}{r}, \qquad \alpha_e := \frac{e^2}{4\pi\epsilon_0\hbar c}, \tag{4.8a}$$

[1] To be precise, the coordinate origin of the CM-system is at $1/1{,}837.15 = 5.443\,21{\times}10^{-4}$ of the distance between the proton center and the electron center, i.e., of the Bohr radius, i.e., $2.880\,42{\times}10^{-14}$ m from the proton center. Rutherford's experiment [☞ p. 45] showed that the atom nucleus must be smaller than about $2.7{\times}10^{-14}$ m – which we certainly expect for the simplest, hydrogen nucleus, with a single proton. Thus, the coordinate origin of the CM-system is just outside the proton that forms the atom nucleus, and more precise analyses must take into account the complementary motion of the proton in the CM-system.

for which the solutions are well known:

$$E_n = -\frac{1}{2}\alpha_e^2\, m_e c^2\, \frac{1}{n^2}, \qquad n = 1, 2, 3, \ldots \tag{4.8b}$$

$$\Psi_{n,\ell,m}(\vec{r}, t) = \sqrt{\Big(\frac{2}{n\, a_0}\Big)^3 \frac{(n-\ell-1)!}{2n[(n+1)!]^3}}\, e^{-r/(na_0)} \Big(\frac{2r}{na_0}\Big)^{\ell} L^{2\ell+1}_{n-\ell-1}\Big(\frac{2r}{na_0}\Big)\, Y^m_\ell(\theta, \phi), \tag{4.8c}$$

where

$$L^q_{k-q}(x) := (-1)^q \frac{\mathrm{d}^q}{\mathrm{d}x^q}\Big[e^x \frac{\mathrm{d}^k}{\mathrm{d}x^k}(e^{-x} x^k)\Big] \qquad \text{are the Laguerre polynomials,} \tag{4.8d}$$

$$a_0 := \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 0.529\times 10^{-10}\,\mathrm{m} \qquad \text{is the Bohr radius.} \tag{4.8e}$$

Recall that the complex phase – and so also the sign – of the wave-functions $\Psi_{n,\ell,m}(\vec{r}, t)$ is not measurable [☞ Chapter 5], so different Authors may use different sign conventions in these definitions (4.8c)–(4.8d) for convenience in some particular computations.

A discussion of this solution for the hydrogen atom may be found in every quantum mechanics textbook, and it is well known that Bohr's spectrum of the hydrogen atom (4.8b) is degenerate: Since the energy depends only on the principal quantum number n, states with different (permitted) values of the quantum numbers ℓ, m (and spin, s and m_s) have the same energy. Since [☞ Appendix A.3]

$$n = 1, 2, 3, \ldots, \qquad \ell = 0, 1, 2, \ldots (n-1), \qquad |m| \leqslant \ell,\ m \in \mathbb{Z}, \qquad s = \pm\tfrac{1}{2}, \tag{4.9}$$

it follows that the number of states with the same energy equals

$$\sum_{\ell=0}^{n-1} \sum_{m=-\ell}^{\ell} 2 = 2 \sum_{\ell=0}^{n-1} (2\ell+1) = 2\, n^2, \tag{4.10}$$

where the factor 2 stems from two possible values of spin. Since the potential is central, i.e., it depends only on the distance between the center of the Coulomb field and the electron that moves in that field, the system manifestly has rotational symmetry. In 3-dimensional space, rotation transformations form the *Spin*(3) group. This symmetry would explain the independence of the energy from the quantum number m (quantifying the direction of the angular momentum) and spin, but not the independence from ℓ, which quantifies the *intensity* of the angular momentum.[2]

Indeed, the hydrogen atom – and more generally, the Coulomb, i.e., the Kepler problem – has another symmetry, generated by the components of the so-called Laplace–Runge–Lenz vector:

$$\text{for}\ \ V(r) = -\frac{\varkappa}{r}, \qquad \vec{A} := \vec{p} \times \vec{L} - m_e \frac{\varkappa}{r}\, \vec{r}. \tag{4.11}$$

It may be shown that the Cartesian components of the Laplace–Runge–Lentz vector commute with the Hamiltonian H in equation (4.1). In turn, the dimensionless operators

$$L_i := \frac{1}{\hbar}(\vec{r}\times\vec{p})_i = -i\, \varepsilon_{ij}{}^k\, x^j \frac{\partial}{\partial x^k} \tag{4.12}$$

[2] The continuous group of rotations is generated by operators of the dimensionless angular momentum (4.12): Each rotation may be represented as the result of the action of the operators $R(\vec{\varphi}) := \exp\{\varphi^i L_i\}$, which change the direction of the atom. Since rotations are *symmetries*, it follows that the result of a rotation is not measurable, and the energy of the atom cannot depend on its direction. However, neither of these operators changes ℓ: $R(\vec{\varphi})Y^m_\ell = \sum_\mu c^m_\mu Y^\mu_\ell$. Therefore, the rotation symmetry does not explain the fact that states of the hydrogen atom with different ℓ nevertheless have the same energy (4.8b).

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

satisfy the relations

$$[L_j, L_k] = i\varepsilon_{jk}{}^l L_l, \qquad [L_j, A_k] = -i\varepsilon_{jk}{}^l A_l, \qquad [A_j, A_k] = \pm i\varepsilon_{jk}{}^l L_l, \text{ for } \{{}^{E<0,}_{E>0;} \tag{4.13}$$

where the operators

$$A_j = \frac{1}{\sqrt{2m_e H}}\left[\frac{\hbar}{2i}\varepsilon_j{}^{kl}\left(\frac{\partial}{\partial x^k}L_l + L_l\frac{\partial}{\partial x^k}\right) - \frac{m_e\,\varkappa}{\hbar}\,\hat{\mathrm{e}}_j\right] \tag{4.14}$$

are the components of the quantum dimensionless Laplace–Runge–Lenz vector, normalized by the energy of the stationary state upon which the operators A_j act. The structure of the symmetry group generated by the operators L_j and A_j depends on the choice of the sign in the third of the commutator relations (4.13). When acting on bound states (for which $E < 0$), the commutator relations (4.13) specify the continuous group $Spin(4)$; [☞ Section A.5]. When acting on $e^- + p^+ \to e^- + p^+$ scattering states (for which $E > 0$), L_j and A_j generate the group $Spin(1,3)$. The operators A_i change ℓ through its full range $\ell \in [0, n-1]$ while the operators L_j change m through its full range $m \in [-\ell, \ell]$, both in unit increments. This extended $Spin(4)$, i.e., $Spin(1,3)$, symmetry of the hydrogen atom implies that the energy, as obtained by Bohr's formula (4.8b), does not depend on ℓ, m and m_s.

Thus, the $Spin(4)$ symmetry fully explains the number and classification of hydrogen atom bound states. The lesson from this simple and very well known system is that symmetries may well be of great use in listing and classifying the possible states – very similar to the situation in Section 2.3.12.

Of course, as we know from the discussion in standard quantum mechanics textbooks, Bohr's formula (4.8b) is not the end of the story, and the value for energy acquires "corrections" owing to several different physical phenomena which we briefly review in the subsequent sections. It is well known that these corrections split the degeneracy, and so "break" the approximate $Spin(4)$ symmetry of the hydrogen atom to $Spin(3) \subset Spin(4)$. The quark model uses this correlation in reverse, and deduces some of the details of the quark dynamics from the hierarchy of approximate symmetries.

4.1.2 Relativistic corrections

The approach in Section 4.1.1 may easily be amended using stationary-state perturbation theory, and this "corrects" the energy values (4.8b). One of these corrections stems from the fact that the non-relativistic physics is of course only an approximation, and that the relativistic kinetic energy is

$$\begin{aligned} T_{\text{rel}} &= m_e c^2\big[\sqrt{1 + (\vec{p}/m_e c)^2} - 1\big] = m_e c^2 \sum_{k=1}^{\infty}\binom{\frac{1}{2}}{k}\left(\frac{\vec{p}^2}{m_e^2 c^2}\right)^k, \\ &\approx \frac{\vec{p}^2}{2m_e} - \frac{(\vec{p}^2)^2}{8m_e^3 c^2} + \frac{(\vec{p}^2)^3}{16 m_e^5 c^4} - \cdots . \end{aligned} \tag{4.15}$$

The first and second relativistic corrections to the Hamiltonian are then represented by the operators

$$H'_{\text{rel}} := -\frac{\hbar^4}{8m_e^3 c^2}(\vec{\nabla}^2)^2, \qquad H''_{\text{rel}} := +\frac{\hbar^6}{16 m_e^5 c^4}(\vec{\nabla}^2)^3, \tag{4.16}$$

and the first-order perturbative correction of the energy is

$$E_n^{(1,r_1)} = \langle n|H'_{\text{rel}}|n\rangle \;=\; \int \mathrm{d}^3\vec{r}\;\Psi^*_{n,\ell,m}(\vec{r})\,H'_{\text{rel}}\,\Psi_{n,\ell,m}(\vec{r}). \tag{4.17}$$

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

This is calculable by simple substitution of the wave-functions, the application of $H'_{\rm rel}$ on $\Psi_{n,\ell,m}(\vec{r})$, and computation of the ensuing integral. However, it is faster to use that

$$\frac{\hbar^2}{2m_e}\vec{\nabla}^2 = V(r) - H, \tag{4.18}$$

so that

$$H'_{\rm rel} = -\frac{1}{2m_e c^2}\big[V(r) - H\big]^2 = -\frac{1}{2m_e c^2}\big[V^2 - H\,V - V\,H + H^2\big], \tag{4.19}$$

$$\begin{aligned} H''_{\rm rel} &= +\frac{1}{2m_e^2 c^4}\big[V(r) - H\big]^3 \\ &= -\frac{1}{2m_e^2 c^4}\big[V^3 - V\,H\,V - V^2\,H + V\,H^2 - H\,V^2 + H^2\,V + H\,V\,H - H^3\big], \end{aligned} \tag{4.20}$$

the matrix elements of which are easier to compute, since $H^\dagger = H$ acts equally on $|n,\ell,m,m_s\rangle$ as well as on $\langle n,\ell,m,m_s|$, producing its eigenvalue, E_n, given by the relation (4.8b).

For second-order corrections, one ought to compute both the second-order perturbation correction stemming from $H'_{\rm rel}$ as well as the first-order perturbation correction stemming from $H''_{\rm rel}$. For the first of these two contributions, we must re-diagonalize the basis $|n,\ell,m,m_s\rangle$ to avoid the $\frac{1}{0}$-divergences in the formula (1.19), whereas the second contribution requires a little more attention for the term $\langle V\,H\,V\rangle$ owing to the fact that $(\vec{\nabla}^2\,\frac{1}{r}) = -4\pi\,\delta(r)$. However, this pointillist contribution is limited to the cases $\ell = 0 = m$, *[✎ why?]* which are not hard to compute separately. In the general case, we'll need the results [407, 471, 242, 472, 29, 328, 362, 363, 360, 3, for example]:

$$\langle r^2\rangle = n^4 a_0^2\Big[1 + \tfrac{3}{2}\big(1 - \tfrac{\ell(\ell+1)-\frac{1}{3}}{n^2}\big)\Big], \tag{4.21a}$$

$$\langle r\rangle = n^2 a_0\Big[1 + \tfrac{1}{2}\big(1 - \tfrac{\ell(\ell+1)}{n^2}\big)\Big], \tag{4.21b}$$

$$\langle r^{-1}\rangle = \frac{1}{n^2\,a_0}, \tag{4.21c}$$

$$\langle r^{-2}\rangle = \frac{1}{(\ell+\frac{1}{2})n^3\,a_0^2}, \tag{4.21d}$$

$$\langle r^{-3}\rangle = \frac{1}{\ell(\ell+\frac{1}{2})(\ell+1)n^3\,a_0^3}. \tag{4.21e}$$

The first-order perturbative energy correction stemming from $H'_{\rm rel}$ is

$$\begin{aligned} E_n^{(1,r_1)} &= \langle n,\ell,m,m_s|H'_{\rm rel}|n,\ell,m,m_s\rangle = -\frac{1}{2m_e c^2}\Big[\langle V^2\rangle - 2E_n^{(0)}\langle V\rangle + (E_n^{(0)})^2\Big] \\ &= -\frac{1}{2}\alpha_e^4\,m_e c^2\,\frac{1}{4n^4}\Big[\frac{4n}{(\ell+\frac{1}{2})} - 3\Big]. \end{aligned} \tag{4.22}$$

4.1.3 Magnetic corrections

Besides their electric charges, the electron and the proton both also have an intrinsic (dipole) magnetic field: $\vec{\mu}_e$ and $\vec{\mu}_p$, respectively. Since the electron and the proton move one with respect to the other, the motion of the electron produces a current that, by the Biot–Savart law, creates a magnetic field proportional to the angular momentum of the electron about the proton, $\vec{B} \propto \vec{L}$, and this magnetic field interacts with the intrinsic magnetic dipole of the proton. Of course, it would be nonsense saying that this same magnetic field, caused by the motion of the electron, also interacts

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

with the intrinsic magnetic dipole field of the electron: In its own coordinate system, the electron of course does not move, and so produces neither an electric current nor a magnetic field.

However, in the electron's rest-frame it is the proton that moves. This then produces a current and a corresponding magnetic field $\vec{B}'$, which interacts with the intrinsic magnetic dipole of the electron. To relate $\vec{B}'$ and $\vec{B}$, one must transform the vector of this "rotating" magnetic field from the electron's coordinate system into the proton's. Since the electron's coordinate system rotates about the proton, one must iterate this transformation from moment to infinitesimally adjacent moment, approximated by successive infinitesimal Lorentz boosts. The resulting effect is called Thomas precession and provides the relation $\vec{B}' = \frac{1}{2}\vec{B}$ [296].

With two intrinsic magnetic dipoles $\vec{\mu}_e, \vec{\mu}_p$, and the "orbital" magnetic field $\vec{B}$, there then exist three additions to the hydrogen atom Hamiltonian:

$$
\begin{aligned}
H_{S_eO} &= -\vec{\mu}_e \cdot (\tfrac{1}{2}\vec{B}), \qquad H_{S_pO} = -\vec{\mu}_p \cdot \vec{B}, \\
H_{S_eS_p} &= -\frac{\mu_0}{4\pi}\Big[\Big(3(\vec{\mu}_e\cdot\hat{r})(\vec{\mu}_p\cdot\hat{r}) - \vec{\mu}_e\cdot\vec{\mu}_p\Big)\frac{1}{r^3} + \frac{8\pi}{3}\vec{\mu}_e\cdot\vec{\mu}_p\,\delta^3(\vec{r}\,)\Big],
\end{aligned}
\tag{4.23}
$$

taking the dipole–dipole interaction term from standard texts such as Ref. [296].

Digression 4.1 One of the original motivations for the Abraham–Lorentz model of the electron was also the attempt to explain – with classical physics – the origin of the electron's intrinsic dipole moment. In this model, the electron was supposed to be a teeny electrically charged sphere. If that sphere rotated, the charge distribution on the sphere would also rotate and so produce a circular current, which would in turn produce a magnetic field by the Biot–Savart law. This is the source of the idea that the electron rotates about its own axis, has *spin* (= intrinsic angular momentum), and that its intrinsic magnetic dipole moment is a consequence of this rotation and proportional to this spin. For a classical rotating electric charge q for which the charge and mass (m) distribution coincide, the magnetic dipole is proportional to the angular momentum:

$$
\vec{\mu} = \frac{q}{2m}\vec{L},
\tag{4.24a}
$$

and $\mu_e := e/2m_e$ is called the Bohr magneton (for the electron).

In fact, this identification is completely backwards: It is the electron's magnetic moment that may be measured and so has a *real* physical meaning; the rotation of the electron about its own axis – *spin* – is a fictitious quantity, *defined* through the relation (4.24a) in terms of the intrinsic magnetic moment. This backwards-engineered explanation stems from G. E. Uhlenbeck and S. A. Goudsmit, who measured the magnetic dipole moment of the electron in 1925, then concluded that this magnetic moment stems from a rotation of the electron about its own axis [528]; all along, they assumed the electron to be represented as an electrically charged sphere, following the Abraham–Lorentz model [☞ Digression 3.13 on p. 123].

Besides, the operators L_j that generate rotations close the $Spin(3) \cong SU(2)$ algebra, which has two classes of representations: tensors and spinors [☞ Digression A.2 on p. 465]. It is easy to show that 360°-rotations around any axis map tensor functions into themselves, but spinors into their negative multiple. Because of this property, physically observable quantities cannot be spinors. Since all real functions over the phase space are observables in classical physics, it follows that there is no room for spinors in classical physics. In quantum physics, however, wave-functions (and abstract state vectors in the

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

Hilbert space) are not directly observable, and so can be spinorial representations. In this sense, half-integral spin is an exclusively quantum-mechanical phenomenon.

It then also follows that the classical relation (4.24a), based on the fictive rotation of a fictive sphere in the Abraham–Lorentz model of the electron need not hold for the electron, which *is* a spin-$\frac{1}{2}$ particle.[3] Indeed, Dirac's relativistic theory of the electron provides for the electron's magnetic moment a result that is twice as large as the classical value, which is further corrected by quantum field theory effects:

$$\vec{\mu}_e = 2\Big[1 + \underbrace{\frac{\alpha_e}{2\pi} + \cdots}_{\text{quantum field theory}}\Big]\frac{(-e)}{2m_e}\,\vec{S}. \tag{4.24b}$$

Of course, if ever the electron turns out to show a structure, this relation will have to be revisited, just as the proton's magnetic moment is today determined using the fact that it is composed of three quarks, as well as a variable number of gluons (which hold those three quarks in the bound state) and virtual quark–antiquark pairs.

For our purposes, write [☞ relation (4.24b)]

$$\vec{\mu}_e = -g_e\mu_B\vec{S}_e, \qquad \mu_B := \frac{e}{2m_e}, \qquad g_e = 2.002\,319\,304\,361\,1(46) \approx 2, \tag{4.25}$$

$$\vec{\mu}_p = +g_p\mu_N\vec{S}_p, \qquad \mu_N := \frac{e}{2m_p}, \qquad g_p = 2.7928, \tag{4.26}$$

where μ_B and μ_N are, respectively, Bohr's (electron) and nucleon magnetic moments. Note that the electron "g-factor," g_e, is measured to a precision of 12 significant figures, and is in full agreement with the result of quantum electrodynamics [293]. The value $\frac{1}{2}(g_e-2)$ is also referred to as the "anomalous magnetic moment," in the sense that g_e deviates from the "bare" value of 2 in the Dirac theory of the electron; this should not be confused with the (quantum) anomalies mentioned elsewhere in this book [☞ Section 7.2.3].

Inserting the expressions for the magnetic dipole moments, the three additions (4.23) to the hydrogen atom energy become, after a little algebra (see Refs. [362, 363, 407, 471, 328, 242, 472, 29, 360, 3] for example),

$$H_{S_eO} = -\Big(\frac{g_e(-e)}{2m_e}\hbar\vec{S}_e\Big)\cdot\Big(\frac{1}{2}\frac{e}{4\pi\epsilon_0 m_e r^3}\hbar\vec{L}\Big) \approx \frac{e^2}{4\pi\epsilon_0}\frac{\hbar^2}{2m_e^2c^2}\frac{1}{r^3}\,\vec{L}\cdot\vec{S}_e, \tag{4.27a}$$

$$H_{S_pO} = \frac{g_p\,e^2}{4\pi\epsilon_0}\frac{\hbar^2}{m_e m_p c^2}\frac{1}{r^3}\,\vec{L}\cdot\vec{S}_p, \tag{4.27b}$$

$$H_{S_eS_p} \approx \frac{g_p\,e^2}{4\pi\epsilon_0}\frac{\hbar^2}{m_e m_p c^2}\Big[\Big(3(\vec{S}_e\cdot\hat{r})(\vec{S}_p\cdot\hat{r}) - \vec{S}_e\cdot\vec{S}_p\Big)\frac{1}{r^3} + \frac{8\pi}{3}\vec{S}_e\cdot\vec{S}_p\,\delta^3(\vec{r}\,)\Big], \tag{4.27c}$$

where "≈" indicates the use of the approximation (4.25). Comparing the constant pre-factors (which all have the same units), one expects the latter two contributions to be of the same order of magnitude, and about $m_p/2g_pm_e \approx 329$ times smaller than the first, so the first perturbation dominates.

[3] The adjective "spin-j" simply specifies that the particle (or wave, or field) has a characteristic orientation of sorts, so that its representative, its wave-function, transforms as one of the representations of the angular momentum algebra, $|j,m\rangle$ with $|m| \leqslant j$; the special case $j=0$ denotes rotation invariance [☞ Appendix A.3]. Conceptually, a "spin-j field" is simply a generalization of the electric and magnetic spin-1 (vector) fields, and "spin-j particles" are the quanta of spin-j fields.

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

Furthermore, one expects that $\langle\frac{1}{r^3}\rangle \sim (a_0)^{-3}$ and $\langle L\cdot\vec{S}_e\rangle = O(1)$, so that the first of these three contributions is of the order

$$\langle H_{S_eO}\rangle = \frac{e^2}{4\pi\epsilon_0}\frac{\hbar^2}{2m_e^2c^2}\left\langle\frac{1}{r^3}\,\vec{L}\cdot\vec{S}_e\right\rangle \sim \frac{\alpha_e\hbar^3}{2m_e^2c}\cdot\frac{1}{a_0^3} = \frac{\alpha_e^4 m_e c^2}{2}. \tag{4.28}$$

This result is of the same order as the relativistic correction (4.22), although its origins are completely different; cf. equation (4.16). Using the result (4.21e), we have

$$E_n^{(1,SO)} = \alpha_e^4 m_e c^2\frac{j(j{+}1)-\ell(\ell{+}1)-\frac{3}{4}}{4n^3\ell(\ell{+}\frac{1}{2})(\ell{+}1)} = \alpha_e^4 m_e c^2\frac{1}{4n^3}\begin{cases}\frac{1}{(\ell+\frac{1}{2})(\ell+1)},\\ -\frac{1}{\ell(\ell+\frac{1}{2})},\end{cases} \tag{4.29}$$

where we used the relation

$$\vec{J} := \vec{L}+\vec{S} \qquad\Rightarrow\qquad \vec{L}\cdot\vec{S} = \tfrac{1}{2}\big[J^2-L^2-S^2\big]. \tag{4.30}$$

The corrections (4.29) and (4.22) are indeed very similar, and add up:

$$E_n^{\rm fs} = E_n^{(1,r_1)} + E_n^{(1,SO)} = -\alpha_e^4\, m_e c^2\,\frac{1}{4n^4}\left[\frac{2n}{(j+\frac{1}{2})}-\frac{3}{2}\right], \qquad \begin{cases} j=\ell+\frac{1}{2},\\ j=\ell-\frac{1}{2};\end{cases} \tag{4.31}$$

providing for the so-called fine structure of the hydrogen atom spectrum.

— ❦ —

It remains to compare the contributions:[4]

$$E_n^{(1,r_2)} = \langle H''_{\rm rel}\rangle \sim \frac{1}{m_e^2c^4}\left\langle\Big(\frac{e^2}{4\pi\epsilon_0\, r}\Big)^3\right\rangle \sim \frac{1}{m_e^2c^4}\frac{(\alpha_e\hbar c)^3}{n^3\,a_0^3} \sim \frac{\alpha_e^6\, m_e c^2}{n^3}; \tag{4.32}$$

$$E_n^{(2,r_1)} = \sum_{n'\cdots\neq n\cdots}\frac{|\langle n',\cdots|H'_{\rm rel}|n,\cdots\rangle|^2}{E_n^{(0)}-E_m^{(0)}} \sim \frac{|E_n^{(1,r1)}|^2}{|E_n^{(0)}|} \sim \frac{\big(\alpha_e^4 m_e c^2/n^3\big)^2}{\alpha_e^2 m_e c^2/n^2} \sim \frac{\alpha_e^6 m_e c^2}{n^4}; \tag{4.33}$$

$$E_n^{(1,S_pO)} = \langle H_{S_pO}\rangle = \frac{g_p\, e^2}{4\pi\epsilon_0}\frac{\hbar^2}{m_e m_p c^2}\left\langle\frac{1}{r^3}\,\vec{L}\cdot\vec{S}_p\right\rangle \sim \frac{g_p\,\alpha_e\hbar^3}{m_e m_p c}\cdot\frac{1}{n^3a_0^3} \sim g_p\Big(\frac{m_e}{m_p}\Big)\frac{\alpha_e^4 m_e c^2}{n^3}; \tag{4.34}$$

$$E_n^{(1,S_eS_p)} = \langle H_{S_eS_p}\rangle \sim \frac{g_p\, e^2}{4\pi\epsilon_0}\frac{\hbar^2}{m_e m_p c^2}\left\langle\vec{S}_e\cdot\vec{S}_p\frac{1}{r^3}\right\rangle \sim g_p\Big(\frac{m_e}{m_p}\Big)\frac{\alpha_e^4 m_e c^2}{n^3}. \tag{4.35}$$

It is not hard to see that

$$E_n^{(1,r_2)} : E_n^{(2,r_1)} : E_n^{(1,SpO)} : E_n^{(1,SeSp)} \approx n\alpha_e^2 : \alpha_e^2 : g_p\Big(\frac{m_e}{m_p}\Big) : g_p\Big(\frac{m_e}{m_p}\Big). \tag{4.36}$$

Since $\alpha_e^2 \approx 5.33\times10^{-5}$ and $g_p\big(\frac{m_e}{m_p}\big) \approx 1.52\times10^{-3}$, the last two contributions are about 28 times larger than the first two. Therefore, we neglect the first two of these contributions in comparison with the latter two, the sum of which gives

$$E_n^{\rm hfs} = E_n^{(1,S_eS_p)} + E_n^{(1,S_pO)} = \Big(\frac{m_e}{m_p}\Big)\alpha_e^4\, m_e c^2\,\frac{g_p}{2n^3}\,\frac{\pm1}{(f+\frac{1}{2})(\ell+\frac{1}{2})}, \qquad \begin{cases} f=j+\frac{1}{2},\\ f=j-\frac{1}{2};\end{cases} \tag{4.37}$$

and similarly provide for the so-called *hyperfine* structure of the hydrogen atom spectrum. This last result introduces the so-called f-spin: $\vec{F} := \vec{J}+\vec{S}_p = \vec{L}+\vec{S}_e+\vec{S}_p$, as a vector sum of all three angular momenta.

[4] From Equation (4.27c), it is actually the last, δ-function part that contributes to the result (4.35) and produces the "21 cm hydrogen line," well-known in microwave radio astronomy [407].

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

When $\ell = 0$, the electron and proton spins are either parallel or antiparallel, giving the so-called triplet and singlet states: Denote $\vec{Z} := \vec{S}_e + \vec{S}_p$, so that the eigenvalue of Z^2 equals $z(z+1)$. When the electron and proton spins are parallel, $z = 1$ and $m_z = \pm 1, 0$; for antiparallel spins $z = 0$ and $m_z = 0$. In the result (4.37), $\ell = 0$, $j = s_e$ and $f = z$, so the numerical factor becomes

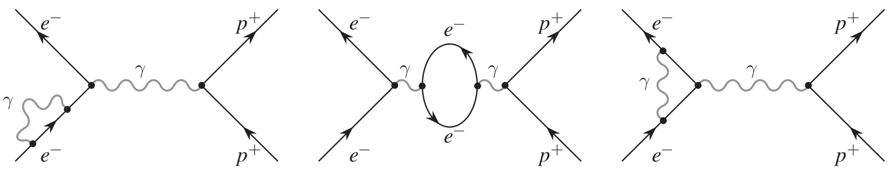
$$\frac{\pm 1}{(f+\frac{1}{2})(\ell+\frac{1}{2})} = \begin{cases} +\frac{4}{3}, \\ -4; \end{cases} \qquad \begin{cases} z = 1 \ \ \text{(triplet)}, \\ z = 0 \ \ \text{(singlet)}. \end{cases} \tag{4.38}$$

Owing to this split in the energies, a transition is possible (for the same n) between these states, emitting a photon of energy equal to this difference in energies, and a wavelength of 21.080 7 cm (for $n = 1$). This result (to first perturbative order) differs less than 1% from the precisely measured wavelength of 21.106 114 054 13 cm, very well known in microwave astronomy.

4.1.4 The Lamb shift

The corrections in the previous Sections 4.1.2–4.1.3 were computed with standard methods of non-relativistic quantum mechanics, in the approach that might be called *semi-quantum*, since the particles (electron and proton) receive a quantum treatment, while the binding (electromagnetic) field is treated classically.

There exist, however, measurable consequences of electromagnetic field quantization, for the computation of which field theory is needed. Consider here only qualitatively the following three Feynman diagrams:



(4.39)

The first of these Feynman diagrams describes the fact that, during “free” motion through “empty vacuum,” the electron interacts with a virtual photon, which changes its mass. The second diagram shows the reciprocal effect, whereby the photon mediating the interaction between the electron and the proton *en route* interacts with a virtual e^-–e^+ pair (is absorbed and then re-emitted by the pair), which effectively screens the electric charge of the nucleus and the electron in the orbit. The third diagram describes a correction to the nature of interaction of the orbiting electron and the mediating photon; this effectively changes the magnetic dipole moment of the electron and contributes to the gyromagnetic ratio (4.24b) by an amount proportional to α_e.

Suffice it here just to cite the resulting correction [407, 243, 150]:

$$E_n^{(QED)} = \begin{cases} \alpha^5 m_e c^2 \frac{1}{4n^3}\, k(n,0) & \ell = 0; \\ \alpha^5 m_e c^2 \frac{1}{4n^3} \Big[k(n,\ell) \pm \frac{1}{\pi(j+\frac{1}{2})(\ell+\frac{1}{2})} \Big], \quad j = \ell \pm \frac{1}{2}, & \ell \neq 0, \end{cases} \tag{4.40}$$

where $k(n,0)$ varies mildly from about 12.7 for $n = 1$ to about 13.2 for $n \to \infty$, while $k(n,\ell) \lesssim 0.05$ and also varies very mildly with n, ℓ. Note that, unlike the corrections considered in Sections 4.1.2–4.1.3 that are all proportional to an even power of the fine structure constant α_e, this quantum-electrodynamical contribution is proportional to α_e^5.

This contribution to the energies is called the Lamb shift. Comparing

$$E_n^{(1,r_2)} : E_n^{(2,r_1)} : E_n^{(1,S_pO)} : E_n^{(1,S_eS_p)} : E_n^{(QED)} \approx n\alpha_e^2 : \alpha_e^2 : g_p\Big(\frac{m_e}{m_p}\Big) : g_p\Big(\frac{m_e}{m_p}\Big) : \alpha_e, \tag{4.41}$$

$$\approx (5.33\times 10^{-5}\cdot n) : (5.33\times 10^{-5}) : (1.52\times 10^{-3}) : (1.52\times 10^{-3}) : (7.30\times 10^{-3}). \tag{4.42}$$

This simple dimensional analysis suggests the Lamb shift to be almost five times *larger* than the hyperfine splitting; the precise numerical results are however comparable. For example, the dipole–dipole interaction (4.35) produces the "21 cm hydrogen line" at about 1.42 GHz, while the Lamb shift permits the $2^2p_{1/2} \to 2^2s_{1/2}$ transition at about 1.06 GHz.

4.1.5 Positronium

The analysis of the hydrogen atom in Sections 4.1.1–4.1.4 is easy to adapt to many two-particle bound states, where the proton or the electron (or both) are replaced by other particles. Such systems are collectively called exotic atoms. Such systems include: muonic hydrogen $(p^+\mu^-)$, pionic hydrogen $(p^+\pi^-)$, muonium (μ^+e^-), etc. Amongst these, consider positronium, (e^+e^-). Together with the hydrogen atom, this gives a good foundation for understanding "quarkonium," i.e., mesons: positronium is an adequate template for mesons composed of a quark and an antiquark of roughly the same mass, while the hydrogen atom is an adequate template for mesons where the masses of the quark and the antiquark significantly differ.

Since $m_{e^+} = m_{e^-}$, the reduced mass is $\frac{m_{e^+}m_{e^-}}{m_{e^+}+m_{e^-}} = \frac{1}{2}m_e$. By the simple $m_e \mapsto \frac{1}{2}m_e$ substitution, we obtain the Bohr-like formula:

$$E_n(e^+e^-) = \tfrac{1}{2}E_n(H) = -\alpha_e^2 m_e c^2\,\frac{1}{4n^2}. \tag{4.43}$$

The wave-functions look identical to those for the hydrogen atom (4.8c), except that the Bohr radius is doubled:

$$a_0^{(pos)} = \frac{\frac{m_e m_p}{m_e+m_p}}{\frac{m_e m_e}{m_e+m_e}}\,a_0^{(H)} = \frac{2\,m_p}{m_e+m_p}\,a_0^{(H)} \approx 2a_0^{(H)}. \tag{4.44}$$

The first relativistic correction to the Hamiltonian is larger by a factor of 2, since both the electron and the positron contribute equally. However, $\langle(\vec{p}^2)^2\rangle \propto (m_e c)^4$, which is then diminished by a factor of $(\frac{1}{2})^4$ because of the smaller reduced mass. In total, the relativistic correction for positronium is an eighth of the corresponding correction for the hydrogen atom.

Significant differences from the contributions that provide the hyperfine structure to the spectrum of the hydrogen atom are: the ratio $\frac{m_{e^-}}{m_{e^+}} = 1$, the values $g_{e^+} = g_{e^-}$, and that the Thomas precession is now symmetric. The contributions analogous to (4.37) are now of the same order of magnitude as the fine structure contributions (4.31). The Lamb shift remains suppressed by a factor of α_e as compared with the contributions analogous to (4.31) and (4.37).

There exist, however, also two entirely novel effects, with no analogues in the analysis of the hydrogen atom:

Field latency In positronium, the center of the Coulomb field that acts on the electron moves with the positron, and *vice versa*. Since the changes in the Coulomb field propagate with the finite speed of light, this "tarrying" effect of field *latency* must be taken into account. This field latency may be computed in classical electrodynamics, and its contribution to the Hamiltonian is [59]

$$H_{\text{lat}} = -\frac{e^2}{4\pi\epsilon_0}\,\frac{1}{2m_e^2c^2}\,\frac{1}{r}\big(\boldsymbol{p}^2 + (\boldsymbol{p}\cdot\hat{\boldsymbol{r}})^2\big), \tag{4.45}$$

which gives the first-order perturbative contribution:

$$E_n^{(\text{lat})} = \langle H_{\text{lat}}\rangle = \alpha_e^4 m_e c^2\,\frac{1}{2n^3}\Big[\frac{11}{32n} - \frac{2+\epsilon}{\ell+\frac{1}{2}}\Big], \tag{4.46a}$$

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

where ϵ is a function of the electron and the positron spins:

$$\epsilon = \begin{cases} 0 & \text{for } j = \ell, \qquad\qquad s = 0, \\ -\frac{3\ell+4}{(\ell+1)(2\ell+3)} & \text{for } j = \ell+1, \\ \frac{1}{\ell(\ell+1)} & \text{for } j = \ell, \\ \frac{3\ell-1}{\ell(2\ell-1)} & \text{for } j = \ell-1, \end{cases} \Bigg\} \quad s = 1. \tag{4.46b}$$

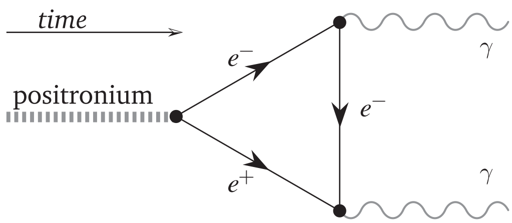
All spins contribute equally in positronium, so it seems reasonable to define $\vec{S} := \vec{S}_{e^-} + \vec{S}_{e^+}$, where $\vec{S}^2$ has eigenvalues $s(s+1)$ with $s = 0, 1$. Then we define $\vec{J} := \vec{L} + \vec{S}$, where $\vec{L}^2$ and $\vec{J}^2$ have eigenvalues $\ell(\ell+1)$ and $j(j+1)$.

Virtual annihilation In positronium, the electron and the positron may temporarily annihilate into a virtual photon which then, before the time alotted by Heisenberg indeterminacy relations, decays into an electron and a positron. Since the electron and the positron must be at the same location for this process, the contribution of the virtual annihilation must be proportional to $|\Psi(0)|^2$, and so can happen only when $\ell = 0$. *[✎ Why?]* Then, since the photon has spin 1, positronium also must have spin 1, i.e., it must be in the triplet state with $s = 1$ and parallel spins. The contribution to the energy of positronium is [243]

$$E_n^{(ann)} = \alpha_e^4 m_e c^2 \frac{1}{4n^3}, \qquad \ell = 0,\ s = 1. \tag{4.47}$$

Note that both new contributions (4.46) and (4.47) are of the same order of magnitude as the analogues of the fine and hyperfine structure contributions. The Lamb shift, as well as the analogues of the corrections (4.32)–(4.33) are then consistently negligible in comparison with the analogues of (4.31), (4.37), (4.40) and (4.46). The Lamb shift was shown to be $O(\alpha_e^5)$, and so contributes less than 1 % of the listed contributions, which are all $O(\alpha_e^4)$.

Real annihilation Positronium is an unstable bound state, as the comprising parts may also really annihilate and produce two or more photons. Just as in the previous discussion, since the electron and the positron must be at the same place to annihilate each other, the decay rate must be proportional to $|\Psi(\vec{0}, t)|^2$. For a two-photon decay we have the Feynman diagram



Strictly speaking, only one of the right-hand vertices is an annihilation; here we pick the lower one. The computation shows the result to be independent of this choice.

(4.48)

and it follows that $\mathfrak{M}$ must be proportional to $\frac{e^2}{4\pi\epsilon_0}$ and then also $\hbar^{-1}c^{-1}$, since $\mathfrak{M}$ is dimensionless. We will compute (5.179) in Section 5.3.2: $\mathfrak{M} = -\frac{4\,e^2}{\epsilon_0 \hbar c}$. Two-photon annihilation of positronium may also be interpreted as an $e^- + e^+ \to 2\gamma$ scattering, for which the effective cross-section, following the general result (3.127),

$$\sigma = \int \mathrm{d}^2\Omega \left(\frac{\hbar c}{8\pi}\right)^2 \frac{|\mathfrak{M}|^2}{(E_{e^-} + E_{e^+})^2} \left|\frac{\vec{p}_f}{\vec{p}_i}\right| = 4\pi \frac{(\alpha_e \hbar c)^2}{(m_e c^2)^2} \frac{c}{v} = 4\pi \frac{\alpha_e^2 \hbar^2}{m_e^2 c v}, \tag{4.49}$$

where v is the relative speed of the electron and the positron. By Conclusion 3.2 on p. 113, we have

$$\Gamma = \sigma\, v\, |\Psi(\vec{0}, t)|^2 = \left(4\pi \frac{\alpha_e^2 \hbar^2}{m_e^2 c v}\right)\left(\frac{1}{\pi a_0^{(pos)} n^3}\right) = \left(4 \frac{\alpha_e^2 \hbar^2}{m_e^2 c v}\right)\left(\frac{\alpha_e \left(\frac{1}{2} m_e\right) c}{\hbar\, n}\right)^3 = \frac{\alpha_e^5 m_e c^2}{2\, \hbar\, n^3}, \tag{4.50}$$

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

where the familiar expression for the $\ell = 0$ wave-function of the hydrogen atom is adapted for positronium by replacing the reduced mass, $m_e \to \frac{1}{2}m_e$. The positronium lifetime is then

$$\tau = \frac{1}{\Gamma} = \frac{2\,\hbar\,n^3}{\alpha_e^5 m_e c^2} = (1.24494\times 10^{-10}\,\mathrm{s}) \times n^3, \tag{4.51}$$

which is in excellent agreement with experiments [309].

4.1.6 Exercises for Section 4.1

✎ ***4.1.1*** *Compute the second-order perturbative energy correction E_n due to H'_{rel}.*

✎ ***4.1.2*** *Compute the second-order perturbative energy correction E_n due to H'_{S_eO}.*

✎ ***4.1.3*** *Compute the first-order perturbative energy correction E_n due to H''_{rel}.*

✎ ***4.1.4*** *Why can the electron and the positron making up positronium not annihilate into a single photon, i.e., why does the annihilation result in at least two photons?*

4.2 Finite symmetries

Symmetries with the structure of finite groups are widely used in solid state physics and crystallography. There are many such groups, and their structures may be very involved, and the applications very detailed and technically demanding.

In relativistic field theory, however, we are in general interested only in three rather simple finite symmetries:[5]

Parity P, which may be thought of simply as the mirror reflection of *one* of the Cartesian coordinates. In 3-dimensional space, this operation may always be followed by a 180° rotation in the mirror plane, which collectively flips the sign of all three coordinates. We then typically use this "more democratic" version of the parity operation: $P : \vec{r} \to -\vec{r}$ as well as $P : \vec{p} \to -\vec{p}$.

Time reversal T, which may be conceived classically as the simple operation $T : t \to -t$, and the physical meaning of which is simply that the process, under the action of T, runs backward in time.

Charge conjugation C, which may be thought of as the Hermitian conjugation of operators and (wave-)functions, and which physically swaps a particle for its antiparticle and *vice versa*. "Charge" here, foremost, means the electromagnetic charge (see Comment 5.2 on p. 169), but also the color in quantum chromodynamics, the *weak isospin* in weak nuclear interactions, and any and all charges related to symmetries other than spacetime coordinate transformations; see Chapters 5–7.

All three symmetries are of order 2, i.e., their successive applications (as defined here) result in the identity

$$P^2 = \mathbb{1}, \qquad T^2 = \mathbb{1}, \qquad C^2 = \mathbb{1}. \tag{4.52}$$

However, the application in quantum theory requires a little more care, as indicated in the subsequent three sections.

---

[5] As in Conclusion 2.6 on p. 78, the characteristics of the abstract mathematical model are also assigned to the concrete physical system that the model faithfully represents. This makes the symmetries of a state in the Hilbert space of the system also symmetries of the represented concrete physical system when in that state.

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

4.2.1 Parity

In 1956, Tsung-Dao Lee and Chen-Ning Yang studied the so-called "τ–θ" problem: two strange mesons that by this time had become known as τ and θ,[6] had all the same characteristics, except the difference in their decays:

$$\text{“}\theta^{+}\text{”} \to \pi^{+} + \pi^{0}, \qquad \text{“}\tau^{+}\text{”} \begin{cases} \nearrow \pi^{+} + \pi^{0} + \pi^{0} \\ \searrow \pi^{+} + \pi^{+} + \pi^{-} \end{cases} \tag{4.53}$$

The parity of a system of particles is the product of *intrinsic* parities of the individual particles times a factor $(-1)^{\ell}$, where ℓ is the total angular momentum of the system. It follows that the parity of the "θ"-particle is $+1$, and the parity of the "τ"-particle is -1, since the pion's parity is -1, and 2- and 3-pion states in the processes (4.53) have total angular momentum equal to the spin of the "θ^{+}"- and the "τ^{+}"-particles, i.e., $\ell = 0$. The existence of two particles that were identical in all except their parity characteristics was very unusual. Lee and Yang proposed that they are in fact one and the same particle, but that the P-symmetry is violated in these weak interactions. On a second glance, they realized that parity conservation has not been experimentally confirmed in weak processes, so they recommended several experimental tests.

The same year, Chien-Shiung Wu (known as "Madam Wu") successfully completed the first of such experiments, working with E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson. This experiment proved that there really do exist processes in Nature that are not invariant under the action of the parity operation. In the β-decay,

$$^{60}_{27}\mathrm{Co} \to {}^{60}_{28}\mathrm{Ni} + e^{-} + \bar{\nu}_{e}, \tag{4.54}$$

Madam Wu's group showed that most electrons are emitted in high correlation with the spin of the cobalt-60 nucleus. If $\vec{p}_e$ and $\vec{S}$, respectively, are the operators of the electron's linear momentum and the spin of the cobalt-60 nucleus, and $|\Psi\rangle$ the state of this nucleus before the decay, it follows that $\langle\Psi|\vec{p}_e\cdot\vec{S}|\Psi\rangle \neq 0$. Now, since parity P flips the sign of the linear momentum but not of spin, it follows that

$$\begin{aligned} \langle\Psi|\vec{p}_e\cdot\vec{S}|\Psi\rangle &= \langle\Psi|\mathbb{1}\,\vec{p}_e\cdot\mathbb{1}\,\vec{S}\,\mathbb{1}|\Psi\rangle = \langle\Psi|P^{-1}P\,\vec{p}_e\cdot P^{-1}P\,\vec{S}\,P^{-1}P|\Psi\rangle \\ &= \big(\langle\Psi|P^{-1}\big)\big(P\,\vec{p}_e P^{-1}\big)\cdot\big(P\vec{S}P^{-1}\big)\big(P|\Psi\rangle\big) = \langle\Psi'|\vec{p}_e^{\,\prime}\cdot\vec{S}'|\Psi'\rangle \\ &= \langle\Psi'|(-\vec{p}_e)\cdot(+\vec{S})|\Psi'\rangle = -\langle\Psi'|\vec{p}_e\cdot\vec{S}|\Psi'\rangle, \end{aligned} \tag{4.55}$$

where $|\Psi'\rangle = P|\Psi\rangle$ – whatever that action on the state $|\Psi\rangle$ may be[7] – and, of course

$$\vec{p}_e^{\,\prime} := P\,\vec{p}_e P^{-1} = -\vec{p}_e, \qquad \text{and} \qquad \vec{S}_e' := P\vec{S}P^{-1} = +\vec{S}. \tag{4.56}$$

If we assume that $[H, P] = 0$ (that parity is a symmetry of the system), it follows that:

1. either $|\Psi'\rangle = c|\Psi\rangle$ and so $\langle\Psi'| = c^{*}\langle\Psi|$, whereby the relation (4.55) would have to imply that $\langle\Psi|\vec{p}_e\cdot\vec{S}|\Psi\rangle = 0$, which was proven wrong by Madam Wu's experiment;
2. or $|\Psi\rangle \not\propto |\Psi'\rangle$ are degenerate states – which does not follow from otherwise successful nuclear models applicable to the cobalt-60 nucleus.

In fact, were we even to allow the possibility that the nuclear models err and $|\Psi'\rangle \neq c|\Psi\rangle$, so that $|\Psi'\rangle$ and $|\Psi\rangle$ are two distinct but degenerate states of the cobalt-60 nucleus, it may be shown [29]

[6] The early notation for the "τ"-particle must not be confused with the τ-lepton, which was experimentally detected only in 1975, almost two decades later.

[7] In any concrete representation, P acts on the wave-function by changing the argument of that wave-function, and also by changing both the integration measure and limits in expressions such as $\langle\Psi|\cdots|\Psi\rangle$. However, the result (4.55) is independent of its representation, and holds abstractly, as written here.

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

that the oscillation $|\Psi\rangle \leftrightarrow |\Psi'\rangle$ would be sufficiently fast to make the expectation value $\langle\Psi|\vec{p}_e\cdot\vec{S}|\Psi\rangle$ much smaller than was experimentally measured. Thus, $[H, P] \neq 0$ remains as the only possibility, i.e., that P is not a symmetry of Nature.

Once uncovered, P-violation was experimentally confirmed in more and more processes – and exclusively in processes mediated by the weak nuclear interaction. Moreover, it was discovered that *all* weak processes exhibit P-violation!

The most stunning consequence of P-violation is the fact that the "right-handed neutrinos" – if they even exist – behave significantly differently than the "left-handed" ones. That is, for every particle one may define its "helicity" [☞ Section 5.2.1 on p. 172] as the projection of the spin of the particle along the direction of its motion. For particles with non-vanishing mass, helicity cannot be Lorentz-invariant. A Lorentz boost can always transform into the particle's own coordinate system, wherein it does not move at all so the projection is undefined. It is, of course, also possible to "pass" the particle into a coordinate system in which the particle now moves in the opposite direction. All the while the spin remains unchanged, which then flips the sign of the helicity. However, a massless particle cannot be "passed," nor does there exist a coordinate system in which it is at rest, so that the helicity of a massless particle is Lorentz-invariant.

Particles with positive helicity are called "right-handed" (their spin – fictively – rotates in the direction of the fingers of the right hand when the thumb indicates the direction of motion), and the "left" particles have negative helicity. Experimental evidence to date shows that no more than about $1/10^{10}$ of all detected neutrinos are right-handed, and the mass of the observed (so almost entirely "left-handed") neutrinos is close to zero [☞ Ref. [293] and Section 7.3.2]. This extremely convincing asymmetry in Nature is of crucial importance to the structure of weak interactions [☞ Section 7.2].

For example, in the decay

$$\pi^- \to \mu^- + \bar{\nu}_\mu \tag{4.57}$$

analyzed in the pion's rest-frame, the muon and the antineutrino move in opposite directions. *[✎ Why?]* The relative angular momentum of the muon and the antineutrino must be orthogonal to the motion of the muon and the antineutrino, and so does not affect the definition of their helicities. The pion spin is zero, so the spin of the muon and the antineutrino must be antiparallel, which means that the antineutrino helicity is the same as that of the muon. Experiments confirm that all muons – and so also the antineutrinos – emerge with a right-handed helicity; the nearly complete absence (less than 1-in-10^{10}) of left-handed antineutrinos then provides for *maximal* parity violation.

By the way, in 1929, soon after the publication of Dirac's equation and the theory of the electron, Hermann Weyl gave a simpler equation suitable for spin-$\frac{1}{2}$ particles with no mass, and which uses the property that the helicity of such particles is Lorentz-invariant. Weyl's theory was neglected since the photon was the only known particle with no mass, and photons have spin 1. When Pauli, a year later, proposed the neutrino to preserve the energy conservation law, he ironically did not use Weyl's equation: Although he knew that the mass of the neutrino is small or even zero, this equation permits parity violation, which Pauli believed also to be a symmetry of Nature! Twenty-six years later, experiments proved him right about energy conservation but wrong about parity. The reason for such a convincing difference between left-handed and right-handed neutrinos remains one of the significant unexplained characteristics in elementary particle physics☝.

4.2.2 Charge conjugation and time reversal

Although the actions of the operations T and C clearly satisfy the relation (4.52), unlike P, these two operations are *anti-linear*, for example,

$$C\big(c\,\Psi(\vec{r},t)\big) = c^*\,C\big(\Psi(\vec{r},t)\big), \tag{4.58}$$

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

so that the proof of Conclusion A.1 on p. 461, does not apply. This makes the precise analysis of the C- and T-action nontrivial in quantum mechanics [29]; fortunately, we need not be concerned with these details. Note instead that in many cases the invariance of the dynamics with respect to the T- and/or C-operation simply implies a degeneracy, so pairs of states $|\Psi\rangle$ and $T|\Psi\rangle$, as well as $|\Psi\rangle$ and $C|\Psi\rangle$ have the same energy and the same lifetime. For example, if $|\Psi\rangle$ is used for the description of the electron, $C|\Psi\rangle$ must be assigned to the positron. The degeneracy of $|\Psi\rangle$ and $C|\Psi\rangle$ then means that the electron mass equals the positron mass, which is indeed true.

Of course, only chargeless particles may be eigenstates of the C-operation: it follows from the defining property (4.52) that, if $|\Psi\rangle$ is an eigenstate of the C-operation, then $|\overline{\Psi}\rangle = C|\Psi\rangle = \pm|\Psi\rangle$, so $|\Psi\rangle$ and $|\overline{\Psi}\rangle$ differ, at most, in the sign; one says that $|\Psi\rangle$ is its own anti-state, i.e., that the particle is its own antiparticle.

It may be shown [422] that the bound state of the spin-$\frac{1}{2}$ particle and its antiparticle is an eigenstate of the operator C with the eigenvalue $(-1)^{\ell+s}$, where ℓ is the angular momentum of the particle–antiparticle system, and s is their composite spin. For positronium, which at least virtually may annihilate into a single photon, it follows that $\ell + s = 1$ since the photon spin is 1. Since C is conserved in strong and electromagnetic processes and the electron–positron pair annihilation is evidently an electromagnetic process, it follows that the photon is a C-eigenstate, with the eigenvalue -1. Similarly, in the electromagnetic decay of the pion,

$$\pi^0 \to \gamma + \gamma, \tag{4.59}$$

there can be only an even number of photons, and the C-eigenvalue of the π^0-particle is $+1$. According to the quark model,

$$|\pi^0\rangle = \tfrac{1}{\sqrt{2}}\big(|u,\overline{u}\rangle + |d,\overline{d}\,\rangle\big) \tag{4.60}$$

is a linear combination of two particle–antiparticle bound states, so the formula $(-1)^{\ell+s}$ for the C-eigenvalue holds. Also, an n-photon system has the C-eigenvalue equal to $(-1)^n$.

On the other hand, the C conservation law is violated in weak interactions: The muons always emerge from the process (4.57) with right-handed helicity. Then

$$C\big(\pi^- \to \mu^-_R + \bar{\nu}_\mu\big) \;=\; \pi^+ \to \mu^+_R + \nu_\mu, \tag{4.61}$$

and the anti-muons would have to emerge also with 100% right-handed helicity from the π^+-meson decay, since the C-operation has no effect on the coordinate system, direction of particle motion and their spins. However, the same analysis as for the process (4.57) shows that the neutrino would have to have right-handed helicity – but such neutrinos may exist no more than 1-in-10^{10}! The neutrinos as well as the muons in the process (4.61) emerge with left-handed helicity. Thus, weak processes such as the $\pi^\pm$-meson decays (4.57) and (4.61) indicate that weak interactions maximally violate both P- and C-symmetry.

— ❦ —

Direct experimental verification of T-conservation or T-violation is much harder: No physical state can be a T-eigenstate, *[✎ why?]* so we cannot simply check the products of the eigenvalues on one side and the other of a process. The most direct verification would involve detailed measurement of parameters for a process $A + B + \cdots \to C + D + \cdots$, as well as the reverse process, $C + D + \cdots \to A + B + \cdots$, and then – taking into account the kinematic differences – comparing the effective cross-sections. That has indeed been done in a large number of electromagnetic and strong nuclear processes, and no trace of T-violation was found. The resulting equality, up to kinematic factors, between the "reversed" pairs of processes such as $A + B + \cdots \to C + D + \cdots$ and $C + D + \cdots \to A + B + \cdots$ is called the "principle of detailed balance."

On the other hand, the verification of the principle of detailed balance in weak nuclear processes is very hard to carry through: For example, the reversal of the weak decay $\Lambda^0 \to p^+ + \pi^-$

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

would require the fusion $p^+ + \pi^- \to \Lambda^0$. However, the collision $p^+ + \pi^-$ and its outcomes are (by far) so dominated by the strong nuclear interaction that the $p^+ + \pi^-$ fusion into Λ^0 simply cannot be experimentally detected, among the sea of all the different results obtained via the strong nuclear interaction. The only type of weak processes where neither the original nor the reversed process is swamped by strong and electromagnetic side-processes are the processes involving neutrinos. However, experiments with neutrinos are already very hard. Unlike other particles, neutrinos are very difficult to control in the lab, in part owing to their electromagnetic neutrality, and in part owing to the extremely small effective cross-section of their interaction with other matter.

4.2.3 The CPT-theorem and CP-violation

Note first that the operation P in 3-dimensional space is a mirror reflection of one of the Cartesian coordinates, e.g., $z \to -z$, so that the (x, y)-plane serves as the mirror. The Lorentz boost in the z-direction then mixes the z-coordinates and time, t. There exists an analytical continuation of this transformation that flips the sign of the z-coordinate and time t, and so turns the operation P into a T-operation. Finally, following the Feynman–Stückelberg interpretation of antiparticles as particles that travel backwards in time, T is equivalent to the C-operation (charge conjugation). The detailed treatment then shows that this connects the C-, P- and T-operations in any *local* field theory that (**1**) is Lorentz-invariant, (**2**) has a Lorentz-invariant ground state (vacuum), and (**3**) has a lower bound on the energy. Conversely, it follows that every non-invariance with respect to the combined CPT-transformation implies a violation of Lorentz-symmetry and/or of locality [350, 413, 300, 230, 102].

An alternative argument starts by noticing that

$$\begin{aligned} CPT\big(e^{\pm i(\vec{k}\cdot\vec{r}-\omega t)}\big) &= CP\big(e^{\pm i(\vec{k}\cdot\vec{r}-\omega(-t))}\big) = C\big(e^{\pm i(\vec{k}\cdot(-\vec{r})+\omega t)}\big) = \big(e^{\mp i(-\vec{k}\cdot\vec{r}+\omega t)}\big) \\ &= e^{\pm i(\vec{k}\cdot\vec{r}-\omega t)}, \end{aligned} \tag{4.62}$$

and then using the fact that plane waves form a complete set and are Lorentz-invariant, to argue that all spacetime-dependent Lorentz-invariant expressions (observables) are also CPT-invariant. The difficulty in this latter, seemingly much simpler argument lies in proving that there is no loss of generality and that the whole required linear combination of plane waves is also both CPT- and Lorentz-invariant.

Standard (Lagrangian) quantum field theory texts [425, 554] prove that the handful of typically used Lorentz-invariant Lagrangians are also CPT-invariant. Recently, however, a fully rigorous proof has been presented within the familiar framework of Lagrangian quantum field theories [220], applicable to all Lagrangians that depend polynomially on the fields and their derivatives.

— ❦ —

On the other hand, once it was proven in 1956 that P is not a symmetry of Nature, in 1957, Lev D. Landau suggested that the combined CP-transformation should be a symmetry of Nature. As we have seen, weak processes such as the $\pi^\pm$-meson decays (4.57) and (4.61) maximally violate both the P- and the C-symmetry, so it was reasonable to suppose that perhaps the *combined* CP-operation is a true symmetry of Nature. Thus, every naturally occurring process involving a collection of particles would have to have a "mirror image" involving antiparticles instead of particles.

However, James Cronin and Val Fitch surprised the physics community in 1964 by publishing their results showing unambiguously that, in neutral kaon decays, not even the combined CP-transformation is a symmetry of Nature.

Ironically, the possibility of CP-violation follows from a work by Murray Gell-Mann and Abraham Pais back in 1954, when they noticed that the K^0-meson cannot be its own antiparticle as its strangeness charge must equal $+1$; there then must exist a $\overline{K}^0$-meson with a strangeness charge of -1, which is evident from the following scattering events (the parenthetical indices denote strangeness):

$$\pi^-_{(0)} + p^+_{(0)} \to \Lambda^0_{(-1)} + K^0_{(+1)} \qquad \text{and} \qquad \pi^+_{(0)} + p^+_{(0)} \to p^+_{(0)} + \overline{K}^0_{(-1)} + K^+_{(+1)}. \tag{4.63}$$

These processes occur mediated by the strong interaction (so identified owing to the speed of the process), which one knows do preserve both the C- and the P-symmetry. Gell-Mann and Pais then noticed the possibility of the K^0–$\overline{K}^0$ transmutation with an explanation that is, today, easier to replace with the display of relevant Feynman diagrams shown in Figure 4.1.

We also know that the neutral kaons are pseudo-scalars, so:

$$CP|K^0\rangle = -C|K^0\rangle = -|\overline{K}^0\rangle \qquad CP|\overline{K}^0\rangle = -C|\overline{K}^0\rangle = -|K^0\rangle, \tag{4.64}$$

whereby the eigenstates of the CP-symmetry are

$$|K^0_+\rangle := \tfrac{1}{\sqrt{2}}\big(|K^0\rangle - |\overline{K}^0\rangle\big) \qquad \text{and} \qquad |K^0_-\rangle := \tfrac{1}{\sqrt{2}}\big(|K^0\rangle + |\overline{K}^0\rangle\big), \tag{4.65}$$

where

$$CP|K^0_\pm\rangle = (\pm 1)|K^0_\pm\rangle. \tag{4.66}$$

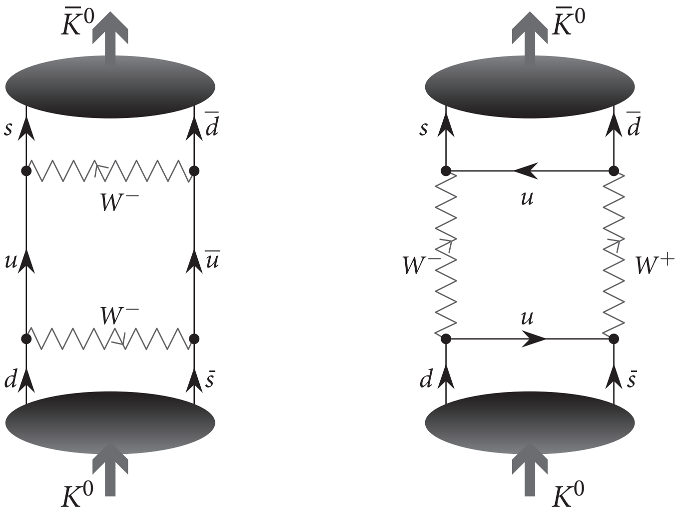


Figure 4.1 The $K^0 \to \overline{K}^0$ transmutation.

Now, neutral kaons decay (among other ways) into two or three pions, and we will neglect all other decay modes. Pions are pseudo-scalars so their intrinsic parity is -1; the parity of a two-pion system is then $+1$, and of a three-pion system, -1. Because of charge conservation, the total charge of both the two- and the three-pion systems in the neutral kaon decays must be zero, so their C-eigenvalue must be $+1$. It follows that $CP(2\pi) = +1$ but $CP(3\pi) = -1$, and it must be that

$$K^0_+ \to 2\pi, \qquad \text{and} \qquad K^0_- \to 3\pi, \tag{4.67}$$

if these (weak interaction) decays preserve the CP-symmetry. Since the two-pion decay has more energy,[8] one expects the K^0_+-state lifetime to be shorter than the K^0_--state lifetime. Indeed, the result (3.124) states that $\Gamma_{K^0_+} \propto \sqrt{1-(2m_{\pi^0}/m_{K^0})^2}$ in a decay into two particles of equal masses.

[8] From the data (2.31) it follows that about $(497.6 - 2\times 135.0) = 227.6\,\text{MeV}$ remains in the two-pion decay, and only about $(497.6 - 3\times 135.0) = 92.6\,\text{MeV}$ in the three-pion decay for the kinetic energy of the pions.

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

Analogously, $\Gamma_{K^0_-} \propto \sqrt{1-(3m_{\pi^0}/m_{K^0})^2}$ in a decay into three particles of equal masses. Thus, one expects $\Gamma_{K^0_+} > \Gamma_{K^0_-}$, and then also $\tau_{K^0_+} < \tau_{K^0_-}$. Although the ratio of these two lifetimes is not as simple, the experiments nevertheless indicate a significant difference and in the same direction as in this simplified estimate:

$$\tau^+ := \tau_{K^0_+} = 0.8958\times 10^{-10}\,\text{s} \qquad \text{and} \qquad \tau^- := \tau_{K^0_-} = 5.114\times 10^{-8}\,\text{s}. \tag{4.68}$$

Because of this difference, K^0_+ is called the "short" kaon ($K^0_S := K^0_+$), and K^0_- the "long" kaon ($K^0_L := K^0_-$).

Since K^0_- "lives" about 570 times longer than K^0_+, within a beam of neutral kaons (created by strong interactions, and so with a 50–50% K^0_+–K^0_- distribution) the "short" kaons quickly decay, leaving the beam as a "pure" K^0_--beam. Recall that the number of undecayed kaons diminishes exponentially, so that

$$\frac{N(K^0_+)}{N(K^0_-)} = \frac{e^{-t/\tau^+}}{e^{-t/\tau^-}} = \exp\Big\{-\frac{t}{\tau^+}+\frac{t}{\tau^-}\Big\} \approx \exp\Big\{-569.9\,\frac{t}{\tau^-}\Big\}, \tag{4.69}$$

which drops to 1.447×10^{-5} after just 1 ns.

Ten years after Gell-Mann and Pais's paper, Cronin and Fitch made use of this extraordinary property, and simply looked for two-pion decays in this decay-purified K^0_--beam. Although they only found about 1 two-pion decay to about 500 three-pion decays, this was a sufficient ($\approx$138-fold) discrepancy to prove that K^0_- nevertheless can also have a two-pion decay and so indicate the violation of *CP*-symmetry.

In addition, the K^0_--meson may also decay as

$$K^0_- \begin{cases} \pi^+ + e^- + \bar{\nu}_e, & (a) \\ \pi^- + e^+ + \nu_e, & (b) \end{cases} \tag{4.70}$$

where the a-type decay is the *CP*-image of the b-type decay. If the *CP*-transformation were a true symmetry of Nature, the probability of these decays would have to be precisely equal. Experiments, however, show a relative *difference* of about 3.3×10^{-3}, which also indicates a small but significant *CP*-violation.

Unlike the *P*- and the *C*-violation, the *CP*-violation is *small*: in the so-called Cabibbo–Kobayashi–Maskawa matrix (2.53)–(2.55) there exists precisely one parameter, δ_{13}, which parametrizes the *CP*-violation. The values of the parameter $\delta_{13} = (1.20 \pm 0.08)^\circ$ (as compared to $\max(\delta_{13}) = 180^\circ$) and its indirect appearance in computations result in the smallness of *CP*-violation, such as the $\sim \frac{1}{500}$ two-pion decays of what should be a $\leqslant \frac{1}{69\,000}$-pure $K^0_- = K^0_L$ beam. This smallness is hard to explain theoretically and remains one of the unsolved problems of elementary particle theory☝.

— ❦ —

There also exists the so-called "strong *CP*-problem": to wit, it is theoretically possible also for the strong nuclear interaction to violate *CP*-symmetry, but this is not the case. The theory of quantum chromodynamics has a parameter, ϑ,[9] which parametrizes possible strong *CP*-violation, whereupon it is a puzzle that $\vartheta \approx 0$, in all known experiments and to a high degree of precision [☞ Section 6.3.1].

Finally, the violation of *CP*-symmetry in the first seconds of the Big Bang is one of the three necessary requirements (as shown by Andrei Saharov) for an explanation of the fact that the

[9] This parameter indeed is an angle, but has no relation with the spherical coordinate of the same name.

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

universe that we observe today consists of matter, and not also of antimatter. This gives an unambiguous definition of the positive electric charge as that of the lepton emerging in the (somewhat but notably) more frequent "semi-leptonic" decay of the long-living neutral kaon (4.70). Thus, the existence of CP-violation is in fact a boon for us: If the CP-transformation were an exact symmetry of Nature, there could be no difference in the universe between matter and antimatter, the two would have annihilated in the first few seconds of the universe's existence, and we would not be here to notice this.

— ❦ —

It is worth noting that the C-, P-, T-, CP-, PT- and CT-symmetries are violated only in experiments that involve the weak interaction, and that these indeed are exact symmetries in all electromagnetic and strong nuclear processes.

4.2.4 Exercises for Section 4.2

✎ ***4.2.1*** *Suppose that the parity operation acts as $\mathsf{P} : |a\rangle \to |b\rangle$ and $\mathsf{P} : |b\rangle \to |a\rangle$ upon some two orthonormalized states, $|a\rangle, |b\rangle$. From these, try to construct the eigenfunctions of the P-operator and normalize them. Discuss the physical meaning of these eigenfunctions if they exist, or explain why an eigenstate of the P-operator cannot make sense physically.*

✎ ***4.2.2*** *Suppose that the time reversal operation acts as $\mathsf{T} : |\alpha\rangle \to |\beta\rangle$ and $\mathsf{T} : |\beta\rangle \to |\alpha\rangle$ upon some two orthonormalized states, $|\alpha\rangle, |\beta\rangle$. From these, try to construct the eigenfunctions of the T-operator and normalize them. Discuss the physical meaning of these eigenfunctions if they exist, or explain why an eigenstate of the T-operator cannot make sense physically.*

✎ ***4.2.3*** *Assuming that the CPT-transformation is an exact order-2 symmetry, prove that the eigenfunctions of the CP-operation are also T-eigenfunctions.*

4.3 Isospin

In 1932, Werner Heisenberg noticed that, for the purposes of describing atomic nuclei, it is possible to neglect the minute difference between the neutron mass and the proton mass:

$$\frac{m_n - m_p}{m_p} = \frac{939.566 - 938.272}{938.272} = 0.001\,379\,13. \tag{4.71}$$

It is even possible to ignore the fact that the proton is charged and the neutron is not: the strong nuclear interaction, which keeps the nucleus as a bound state, must be many times stronger than the electromagnetic repulsion of the protons in the nucleus. Thus, the proton and the neutron are regarded as two states of one particle, a *nucleon* (denoted N), just as the spin-$(\pm\frac{1}{2})$ electrons are both regarded as two polarizations of the *same* particle. In analogy with spin, Heisenberg then introduced a conserved quantity that Eugene Wigner named *isospin* in 1937 and for which he employed the corresponding mathematical formalism:

$$\vec{\mathsf{I}} : \quad \big[\, \mathsf{I}_j \,,\, \mathsf{I}_k \,\big] = i\varepsilon_{jk}{}^{m}\, \mathsf{I}_m, \tag{4.72}$$

just like the $\vec{J}$ in Appendix A.3. Following the digressions A.2 on p. 465 and A.3 on p. 467, we know that there exist eigenstates

$$|I, I_3\rangle : \quad \mathsf{I}^2|I, I_3\rangle = I(I{+}1)|I, I_3\rangle, \qquad \mathsf{I}_3|I, I_3\rangle = I_3|I, I_3\rangle, \qquad |I_3| \leqslant I, \tag{4.73a}$$

$$\mathsf{I}_\pm|I, I_3\rangle = \sqrt{I(I{+}1) - I_3(I_3{\pm}1)}\,|I, I_3{\pm}1\rangle, \qquad 2I,\ \triangle I \in \mathbb{Z}. \tag{4.73b}$$

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

4.3.1 Isospin, nucleons and pions

Heisenberg and Wigner introduced the isospin formalism for the purposes of nuclear physics, and here we identify

$$|p^+\rangle = |\tfrac{1}{2}, +\tfrac{1}{2}\rangle, \qquad |n^0\rangle = |\tfrac{1}{2}, -\tfrac{1}{2}\rangle. \tag{4.74}$$

Moreover, if the isospin "rotations" are a symmetry of strong interactions, then it follows that isospin is a conserved quantity in all strong nuclear processes, following Conclusion A.1 on p. 461. In 1932, the proposition of introducing such an *ad-hoc* and abstract symmetry as a further exact symmetry of strong interactions was an unusually bold move. However, such reliance on symmetries and the quantum version of Noether's theorem [☞ Conclusion A.1 on p. 461] has become one of the basic principles of fundamental physics in the twentieth century, and even grew into the gauge principle, which is the basis of contemporary understanding of interactions in general [☞ Chapters 5 and 6].

A few other (then known) hadrons are identified as

$$|\pi^+\rangle = |1, +1\rangle, \qquad |\pi^0\rangle = |1, 0\rangle, \qquad |\pi^-\rangle = |1, -1\rangle, \tag{4.75}$$

$$|\Delta^{++}\rangle = |\tfrac{3}{2}, +\tfrac{3}{2}\rangle, \quad |\Delta^+\rangle = |\tfrac{3}{2}, +\tfrac{1}{2}\rangle, \quad |\Delta^0\rangle = |\tfrac{3}{2}, -\tfrac{1}{2}\rangle, \quad |\Delta^-\rangle = |\tfrac{3}{2}, -\tfrac{3}{2}\rangle, \quad \text{etc.} \tag{4.76}$$

The relationship between the electric charge Q, the isospin "charge" I_3, the baryon number B and strangeness S for all hadrons was found before 1974 to be in agreement with the GNN formula (2.30). Thus, isospin symmetry, soon extended into the $SU(3)_f$ approximate symmetry, offers an excellent classification tool.

— ❦ —

However, isospin is also useful in dynamics: We know from quantum mechanics that addition of spins $\frac{1}{2}$ and $\frac{1}{2}$ produces the following possibilities, here applied to isospin:

$$\begin{cases} |1, +1\rangle_S = |\tfrac{1}{2}, +\tfrac{1}{2}\rangle|\tfrac{1}{2}, +\tfrac{1}{2}\rangle & = |p^+, p^+\rangle, & (4.77a) \\ |1, 0\rangle_S = \tfrac{1}{\sqrt{2}}\Big(|\tfrac{1}{2}, +\tfrac{1}{2}\rangle|\tfrac{1}{2}, -\tfrac{1}{2}\rangle + |\tfrac{1}{2}, -\tfrac{1}{2}\rangle|\tfrac{1}{2}, +\tfrac{1}{2}\rangle\Big) & = \tfrac{1}{\sqrt{2}}\big(|p^+, n^0\rangle + |n^0, p^+\rangle\big), & (4.77b) \\ |1, -1\rangle_S = |\tfrac{1}{2}, -\tfrac{1}{2}\rangle|\tfrac{1}{2}, -\tfrac{1}{2}\rangle & = |n^0, n^0\rangle; & (4.77c) \end{cases}$$

$$|0, 0\rangle_A = \tfrac{1}{\sqrt{2}}\Big(|\tfrac{1}{2}, +\tfrac{1}{2}\rangle|\tfrac{1}{2}, -\tfrac{1}{2}\rangle - |\tfrac{1}{2}, -\tfrac{1}{2}\rangle|\tfrac{1}{2}, +\tfrac{1}{2}\rangle\Big) = \tfrac{1}{\sqrt{2}}\big(|p^+, n^0\rangle - |n^0, p^+\rangle\big), \tag{4.77d}$$

where the subscript "S" denotes that the state is symmetric with respect to swapping the two nucleons, and "A" that it is antisymmetric. However, there exists only one two-nucleon bound state: the deuteron, the deuterium nucleus, which consists of a proton and a neutron. This implies that the isospin factor in the wave-function of the deuteron must be antisymmetric with respect to swapping the two nucleons. Were this factor symmetric, isospin "rotations" would guarantee the existence of all three symmetric states (4.77a)–(4.77c) – and it is well known that the bound state of neither two protons nor two neutrons exists in Nature.

This identifies the deuteron as the isospin $|0, 0\rangle$ state. Also, since $|0, 0\rangle$ is antisymmetric with respect to swapping the two nucleons and since the whole wave-function must be antisymmetric with respect to the swapping of any two (otherwise identical) fermions, it follows that the product of the remaining "spatial" and "spin" factors in the wave-function of the deuteron bound state must be symmetric. So, if the proton and neutron spins are parallel (evidently symmetric) or antiparallel and symmetrized, then the spatial factor in the wave-function also must be symmetric with respect to the exchange of two nucleons. If the spins are antiparallel and antisymmetrized, the spatial factor in the wave-function must also be antisymmetric.

Without the isospin formalism, which permits treating the proton and the neutron as two polarizations of the same particle, this indirect correlation of spins and spatial factors in the wavefunction could not have been derived.

— ❦ —

Finally, isospin also easily produces *relative* effective cross-sections of various processes mostly by way of the Wigner–Eckart theorem A.3 on p. 475. Consider, e.g., the three two-nucleon collisions:

$$\begin{aligned}
& & & \quad \boldsymbol{N_1} \quad\ \ \boldsymbol{N_2} \qquad\quad \boldsymbol{d} \qquad \boldsymbol{\pi} \\
a:&\ p^+ + p^+ \to d + \pi^+ &\leftrightarrow\ & |\tfrac{1}{2},+\tfrac{1}{2}\rangle|\tfrac{1}{2},+\tfrac{1}{2}\rangle \to |0,0\rangle|1,+1\rangle = |1,+1\rangle, && (4.78a)\\
b:&\ p^+ + n^0 \to d + \pi^0 &\leftrightarrow\ & |\tfrac{1}{2},+\tfrac{1}{2}\rangle|\tfrac{1}{2},-\tfrac{1}{2}\rangle \to |0,0\rangle|1,0\rangle \quad = |1,0\rangle, && (4.78b)\\
c:&\ n^0 + n^0 \to d + \pi^0 &\leftrightarrow\ & |\tfrac{1}{2},-\tfrac{1}{2}\rangle|\tfrac{1}{2},-\tfrac{1}{2}\rangle \to |0,0\rangle|1,-1\rangle = |1,-1\rangle, && (4.78c)
\end{aligned}$$

where d denotes the deuteron, the deuterium nucleus. On the other hand, combining (4.77a)–(4.77d), we have

$$\begin{aligned}
|p^+,p^+\rangle &= |\tfrac{1}{2},+\tfrac{1}{2}\rangle|\tfrac{1}{2},+\tfrac{1}{2}\rangle = |1,+1\rangle, && (4.79a)\\
|p^+,n^0\rangle &= |\tfrac{1}{2},+\tfrac{1}{2}\rangle|\tfrac{1}{2},-\tfrac{1}{2}\rangle = \tfrac{1}{\sqrt{2}}\big(|1,0\rangle + |0,0\rangle\big), && (4.79b)\\
|n^0,n^0\rangle &= |\tfrac{1}{2},-\tfrac{1}{2}\rangle|\tfrac{1}{2},-\tfrac{1}{2}\rangle = |1,-1\rangle. && (4.79c)
\end{aligned}$$

Then it follows that, up to factors independent of isospin and which are equal *[✎ why?]*,

$$\begin{aligned}
\mathfrak{M}_a &\propto \langle d,\pi^+|p^+,p^+\rangle = \langle 1,+1||\tfrac{1}{2},+\tfrac{1}{2}\rangle|\tfrac{1}{2},+\tfrac{1}{2}\rangle = 1, && (4.80a)\\
\mathfrak{M}_b &\propto \langle d,\pi^0|p^+,n^0\rangle = \langle 1,0||\tfrac{1}{2},+\tfrac{1}{2}\rangle|\tfrac{1}{2},-\tfrac{1}{2}\rangle = \langle 1,0|\Big(\tfrac{1}{\sqrt{2}}\big(|1,0\rangle + |0,0\rangle\big)\Big) = \tfrac{1}{\sqrt{2}}, && (4.80b)\\
\mathfrak{M}_c &\propto \langle d,\pi^-|n^0,n^0\rangle = \langle 1,-1||\tfrac{1}{2},-\tfrac{1}{2}\rangle|\tfrac{1}{2},-\tfrac{1}{2}\rangle = 1. && (4.80c)
\end{aligned}$$

Since $\sigma \propto |\mathfrak{M}|^2$, it follows that

$$\sigma_a : \sigma_b : \sigma_c = 2 : 1 : 2. \tag{4.81}$$

That is, it is twice as probable for a deuteron (and a pion) to emerge from the collision of two protons than from the collision of a proton and a neutron! (The collision of two neutrons is hard to arrange experimentally.)

Even more dramatic is the situation with pion–nucleon scattering. Listing all the possibilities consistent with charge conservation, we find six elastic pion–nucleon collisions:

$$\begin{aligned}
&(a)\ \pi^+{+}p^+ \to \pi^+{+}p^+, \quad (b)\ \pi^0{+}p^+ \to \pi^0{+}p^+, \quad (c)\ \pi^-{+}p^+ \to \pi^-{+}p^+, && (4.82a)\\
&(d)\ \pi^+{+}n^0 \to \pi^+{+}n^0, \quad (e)\ \pi^0{+}n^0 \to \pi^0{+}n^0, \quad (f)\ \pi^-{+}n^0 \to \pi^-{+}n^0, && (4.82b)
\end{aligned}$$

and four inelastic collisions resulting in a pion and a nucleon:

$$\begin{aligned}
&(g)\ \pi^+ + n^0 \to \pi^0 + p^+, \qquad (h)\ \pi^0 + p^+ \to \pi^+ + n^0, && (4.82c)\\
&(i)\ \pi^0 + n^0 \to \pi^- + p^+, \qquad (j)\ \pi^- + p^+ \to \pi^0 + n^0. && (4.82d)
\end{aligned}$$

Since $I(\pi) = 1$ and $I(N) = \frac{1}{2}$, the isospin of the incoming (initial) and of the outgoing (final) system may be either $\frac{3}{2}$ or $\frac{1}{2}$, and let $\mathfrak{M}_{3/2}$ and $\mathfrak{M}_{1/2}$ denote the corresponding so-called "reduced" amplitudes.[10] Using tables of Clebsch–Gordan coefficients we compute:

$$\pi^+ + p^+ : \qquad |1,1\rangle|\tfrac{1}{2},+\tfrac{1}{2}\rangle = |\tfrac{3}{2},+\tfrac{3}{2}\rangle, \tag{4.83a}$$

[10] By the Wigner–Eckart theorem, every amplitude may be factorized as a product of a reduced amplitude and a Clebsch–Gordan coefficient [☞ Section A.3.4 and Theorem A.3 on p. 475, as well as the textbooks [362, 363, 328, 471, 480, 134, 391, 407, 472, 360, 29, 242, 3, 110, for example] and the handbook [294]].

$$\pi^0 + p^+ : \quad |1,0\rangle|\tfrac{1}{2},+\tfrac{1}{2}\rangle = \sqrt{\tfrac{2}{3}}|\tfrac{3}{2},+\tfrac{1}{2}\rangle - \tfrac{1}{\sqrt{3}}|\tfrac{1}{2},+\tfrac{1}{2}\rangle, \tag{4.83b}$$

$$\pi^- + p^+ : \quad |1,-1\rangle|\tfrac{1}{2},+\tfrac{1}{2}\rangle = \tfrac{1}{\sqrt{3}}|\tfrac{3}{2},-\tfrac{1}{2}\rangle - \sqrt{\tfrac{2}{3}}|\tfrac{1}{2},-\tfrac{1}{2}\rangle, \tag{4.83c}$$

$$\pi^+ + n^0 : \quad |1,1\rangle|\tfrac{1}{2},-\tfrac{1}{2}\rangle = \tfrac{1}{\sqrt{3}}|\tfrac{3}{2},+\tfrac{1}{2}\rangle + \sqrt{\tfrac{2}{3}}|\tfrac{1}{2},+\tfrac{1}{2}\rangle, \tag{4.83d}$$

$$\pi^0 + n^0 : \quad |1,0\rangle|\tfrac{1}{2},-\tfrac{1}{2}\rangle = \sqrt{\tfrac{2}{3}}|\tfrac{3}{2},-\tfrac{1}{2}\rangle + \tfrac{1}{\sqrt{3}}|\tfrac{1}{2},-\tfrac{1}{2}\rangle, \tag{4.83e}$$

$$\pi^- + n^0 : \quad |1,-1\rangle|\tfrac{1}{2},-\tfrac{1}{2}\rangle = |\tfrac{3}{2},-\tfrac{3}{2}\rangle. \tag{4.83f}$$

For example, the processes (a) and (f) both have $I = \frac{3}{2}$ and the Clebsch–Gordan coefficients are 1:

$$(a)\ \pi^+ + p^+ \to \pi^+ + p^+ \quad \leftrightarrow \quad \mathfrak{M}_a = \langle\tfrac{3}{2},+\tfrac{3}{2}||\tfrac{3}{2},+\tfrac{3}{2}\rangle \times \mathfrak{M}_{3/2} = \mathfrak{M}_{3/2}, \tag{4.84}$$

$$(f)\ \pi^- + n^0 \to \pi^- + n^0 \quad \leftrightarrow \quad \mathfrak{M}_f = \langle\tfrac{3}{2},-\tfrac{3}{2}||\tfrac{3}{2},-\tfrac{3}{2}\rangle \times \mathfrak{M}_{3/2} = \mathfrak{M}_{3/2}, \tag{4.85}$$

and so we have that $\mathfrak{M}_a = \mathfrak{M}_f = \mathfrak{M}_{3/2}$. The remaining processes are a mixture of $\mathfrak{M}_{3/2}$ and $\mathfrak{M}_{1/2}$, such as

$$(c)\ \pi^- + p^+ \to \pi^- + p^+ \tag{4.86}$$

$$\begin{aligned}\mapsto \quad \mathfrak{M}_c &= \Big(\tfrac{1}{\sqrt{3}}\langle\tfrac{3}{2},-\tfrac{1}{2};\mathfrak{a}_3| - \sqrt{\tfrac{2}{3}}\langle\tfrac{1}{2},-\tfrac{1}{2};\mathfrak{a}_1|\Big)\Big(\tfrac{1}{\sqrt{3}}|\tfrac{3}{2},-\tfrac{1}{2};\mathfrak{a}_3\rangle - \sqrt{\tfrac{2}{3}}|\tfrac{1}{2},-\tfrac{1}{2};\mathfrak{a}_1\rangle\Big) \\ &= \tfrac{1}{3}\langle\tfrac{3}{2},-\tfrac{1}{2};\mathfrak{a}_3||\tfrac{3}{2},-\tfrac{1}{2};\mathfrak{a}_3\rangle + \tfrac{2}{3}\langle\tfrac{1}{2},-\tfrac{1}{2};\mathfrak{a}_1||\tfrac{1}{2},-\tfrac{1}{2};\mathfrak{a}_1\rangle = \tfrac{1}{3}\mathfrak{M}_{3/2} + \tfrac{2}{3}\mathfrak{M}_{1/2}\end{aligned} \tag{4.87}$$

$$(j)\ \pi^- + p^+ \to \pi^0 + n^0 \tag{4.88}$$

$$\begin{aligned}\mapsto \quad \mathfrak{M}_j &= \Big(\tfrac{1}{\sqrt{3}}\langle\tfrac{3}{2},-\tfrac{1}{2};\mathfrak{a}_3| - \sqrt{\tfrac{2}{3}}\langle\tfrac{1}{2},-\tfrac{1}{2};\mathfrak{a}_1|\Big)\Big(\sqrt{\tfrac{2}{3}}|\tfrac{3}{2},-\tfrac{1}{2};\mathfrak{a}_3\rangle + \tfrac{1}{\sqrt{3}}|\tfrac{1}{2},-\tfrac{1}{2};\mathfrak{a}_1\rangle\Big) \\ &= \tfrac{\sqrt{2}}{3}\langle\tfrac{3}{2},-\tfrac{1}{2};\mathfrak{a}_3||\tfrac{3}{2},-\tfrac{1}{2};\mathfrak{a}_3\rangle - \tfrac{\sqrt{2}}{3}\langle\tfrac{1}{2},-\tfrac{1}{2};\mathfrak{a}_1||\tfrac{1}{2},-\tfrac{1}{2};\mathfrak{a}_1\rangle \\ &= \tfrac{\sqrt{2}}{3}\mathfrak{M}_{3/2} - \tfrac{\sqrt{2}}{3}\mathfrak{M}_{1/2}.\end{aligned} \tag{4.89}$$

The labels "$\mathfrak{a}_3$" and "$\mathfrak{a}_1$" are arrays of all other quantifiers of the isospin-$\frac{3}{2}$ and isospin-$\frac{1}{2}$ states, respectively. The effective cross-sections of these processes are then related as

$$\sigma_a : \sigma_c : \sigma_f : \sigma_j \;=\; 9|\mathfrak{M}_{3/2}|^2 : |\mathfrak{M}_{3/2} + 2\mathfrak{M}_{1/2}|^2 : 9|\mathfrak{M}_{3/2}|^2 : 2|\mathfrak{M}_{3/2} - \mathfrak{M}_{1/2}|^2. \tag{4.90}$$

In a collision regime where either $\mathfrak{M}_{3/2} \gg \mathfrak{M}_{1/2}$ or $\mathfrak{M}_{3/2} \ll \mathfrak{M}_{1/2}$, this relationship simplifies:

$$\mathfrak{M}_{3/2} \gg \mathfrak{M}_{1/2} \Rightarrow \qquad \sigma_a : \sigma_c : \sigma_f : \sigma_j \;\approx\; 9:1:9:2, \tag{4.91}$$

$$\mathfrak{M}_{3/2} \ll \mathfrak{M}_{1/2} \Rightarrow \qquad \sigma_a : \sigma_c : \sigma_f : \sigma_j \;\approx\; 0:4:0:2. \tag{4.92}$$

4.3.2 Isospin in the quark model

In processes where it suffices to track only the u and d quarks, the application of the isospin formalism is very similar within the quark model. Writing only isospin factors,

$$|u\rangle = |\tfrac{1}{2},+\tfrac{1}{2}\rangle, \qquad |d\rangle = |\tfrac{1}{2},-\tfrac{1}{2}\rangle \tag{4.93}$$

and so we have the three-quark bound states:

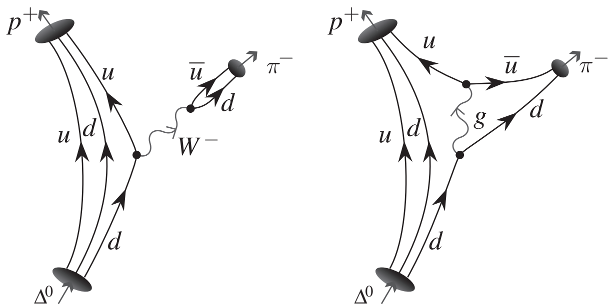
$$|\Delta^{++}\rangle = |u\rangle\otimes|u\rangle\otimes|u\rangle = |\tfrac{1}{2},+\tfrac{1}{2}\rangle|\tfrac{1}{2},+\tfrac{1}{2}\rangle|\tfrac{1}{2},+\tfrac{1}{2}\rangle = |\tfrac{3}{2},+\tfrac{3}{2}\rangle, \tag{4.94a}$$

$$|\Delta^{+}\rangle = |u\rangle\otimes|u\rangle\otimes|d\rangle = |\tfrac{1}{2},+\tfrac{1}{2}\rangle|\tfrac{1}{2},+\tfrac{1}{2}\rangle|\tfrac{1}{2},-\tfrac{1}{2}\rangle = |\tfrac{3}{2},+\tfrac{1}{2}\rangle, \quad |p^+\rangle = |\tfrac{1}{2},+\tfrac{1}{2}\rangle, \tag{4.94b}$$

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

$$|\Delta^0\rangle = |u\rangle \otimes |d\rangle \otimes |d\rangle \ = |\tfrac{1}{2}, +\tfrac{1}{2}\rangle |\tfrac{1}{2}, -\tfrac{1}{2}\rangle |\tfrac{1}{2}, -\tfrac{1}{2}\rangle = |\tfrac{3}{2}, -\tfrac{1}{2}\rangle, \quad |n^0\rangle \ = |\tfrac{1}{2}, -\tfrac{1}{2}\rangle, \tag{4.94c}$$
$$|\Delta^-\rangle = |d\rangle \otimes |d\rangle \otimes |d\rangle \ = |\tfrac{1}{2}, -\tfrac{1}{2}\rangle |\tfrac{1}{2}, -\tfrac{1}{2}\rangle |\tfrac{1}{2}, -\tfrac{1}{2}\rangle = |\tfrac{3}{2}, -\tfrac{3}{2}\rangle. \tag{4.94d}$$

The Δ^+ particle is not identical with the proton: The isospin factors differ in the value of I, but the full wave-functions also differ in the *spin* factors: Δ^+ has spin $\frac{3}{2}\hbar$ and the proton spin is $\frac{1}{2}\hbar$. Similarly, Δ^0 and n^0 have similar isospin factors – which is identified as the "bookkeeping" notation of the u-d content. So, for example, there exists the decay[11]



$\Delta^0 \to p^+ + \pi^-$

Although Δ^0 and n^0 contain the same quarks, n^0 does not have a sufficient mass for such a decay. (4.95)

From these Feynman diagrams we estimate that the right-hand contribution to the process amplitude is proportional to the square of the strong charge (because of the two gluon vertices) and the left-hand contribution is proportional to the square of the weak charge (because of the two W^--vertices). Owing to the immense difference in the strength of these interactions, the weak contribution is negligible and is not calculated. The right-hand diagram may be cut by any curve into the "initial" and "final" states, and the total isospin factor for the so-defined "initial" and "final" state vector computed. Across each such cut of the diagram, isospin is conserved.

The ease of application of this combination of Feynman diagrams, isospin factors in state vectors and of quick estimates of relative strengths of the contributions to the amplitude of the process is the basic reason for the Feynman diagrams' popularity. Owing to the relative simplicity of the $SU(2)$ group, isospin factors here do not give much more information than the u-d content of the diagrams, but it is clear that they are nevertheless useful in estimates using the Wigner–Eckart theorem, just as in the results (4.78)–(4.90).

4.3.3 Exercises for Section 4.3

✎ ***4.3.1*** *Using equation (4.83) and following the derivation of equation (4.90), find the ratios between the probabilities for all ten pion–nucleon scattering processes (4.82).*

✎ ***4.3.2*** *Evaluate the result (4.90) in the limit when* $\mathfrak{M}_{3/2} = \mathfrak{M}_{1/2}$.

✎ ***4.3.3*** *Evaluate your solution to the problem 4.3.1 in the limit when* $\mathfrak{M}_{3/2} \gg \mathfrak{M}_{1/2}$.

✎ ***4.3.4*** *Evaluate your solution to the problem 4.3.1 in the limit when* $\mathfrak{M}_{3/2} = \mathfrak{M}_{1/2}$.

✎ ***4.3.5*** *Evaluate your solution to the problem 4.3.1 in the limit when* $\mathfrak{M}_{3/2} \ll \mathfrak{M}_{1/2}$.

[11] Section 4.4.2 will discuss the isospin details of π^- as a bound state of a d- and an anti-u-quark.

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

4.4 The eightfold way, the $SU(3)_f$ group and the u, d, s quarks

Some two decades after Heisenberg and Wigner introduced the isospin formalism and the $SU(2)$ group of symmetries, several elementary particle physics researchers realized that similar benefits might be derived from grouping the eight baryons in the plot (2.32). They all have spin $\frac{1}{2}$, and their masses are "layered" in isospin multiplets [293]:

S	I_3	Particles	δm
0	$-\frac{1}{2}, +\frac{1}{2}$	$n^0,\ p^+$	$\frac{m_n - m_p}{m_p} = 1.38 \times 10^{-3}$
-1	0	Λ^0	—
	$-1, 0, +1$	$\Sigma^-,\ \Sigma^0,\ \Sigma^+$	$\frac{\overline{\triangle m_\Sigma}}{m_{\Sigma^+}} = 4.49 \times 10^{-3}$
-2	$-\frac{1}{2}, +\frac{1}{2}$	$\Xi^-,\ \Xi^0$	$\frac{m_{\Xi^-} - m_{\Xi^0}}{m_{\Xi^0}} = 5.32 \times 10^{-3}$

(4.96)

where $\overline{\triangle m_\Sigma}$ denotes the average difference between the masses of the Σ^+-, Σ^0- and Σ^--baryons. The "layering" is similar for the spin-$\frac{3}{2}$ decuplet of baryons (2.35):

S	I_3	Particles	δm
0	$-\frac{3}{2}, -\frac{1}{2}, +\frac{1}{2}, +\frac{3}{2}$	$\Delta^-,\ \Delta^0,\ \Delta^+,\ \Delta^{++}$	$\frac{\max(\triangle m_\Delta)}{m_\Delta} = 8.117 \times 10^{-4}$
-1	$-1, 0, +1$	$\Sigma^{*-},\ \Sigma^{*0},\ \Sigma^{*+}$	$\frac{\max(\triangle m_{\Sigma^*})}{m_{\Sigma^*}} = 3.117 \times 10^{-3}$
-2	$-\frac{1}{2}, +\frac{1}{2}$	$\Xi^{*-},\ \Xi^{*0}$	$\frac{m_{\Xi^{*-}} - m_{\Xi^{*0}}}{m_{\Xi^*}} = 2.087 \times 10^{-3}$
-3	0	Ω^-	—

(4.97)

The relative difference between the average masses in any one layer is some 2 orders of magnitude bigger than the in-layer relative mass differences. For the octet (4.96):

$$\frac{\overline{m_{\Lambda,\Sigma}} - \overline{m_N}}{\overline{m_N}} = 0.2223, \qquad \frac{\overline{m_\Xi} - \overline{m_{\Lambda,\Sigma}}}{\overline{m_{\Lambda,\Sigma}}} = 0.1162; \tag{4.98}$$

and for the decuplet (4.97):

$$\frac{\overline{m_{\Sigma^*}} - \overline{m_\Delta}}{\overline{m_\Delta}} = 0.1238, \qquad \frac{\overline{m_{\Xi^*}} - \overline{m_{\Sigma^*}}}{\overline{m_{\Sigma^*}}} = 0.1075, \qquad \frac{m_\Omega - \overline{m_{\Xi^*}}}{\overline{m_{\Xi^*}}} = 0.0907. \tag{4.99}$$

Thus, the approximate isospin $SU(2)$ symmetry (which in the tabulations (4.96) and (4.97) mixes the baryons horizontally) is about a hundred times better than the $SU(3)_f$ symmetry that also includes strangeness (varying vertically in these tables). This agrees with the fact that the mass of the s-quark (as measured by deep inelastic scattering) is 2 orders of magnitude bigger than that of the u- and d-quarks; see Table 4.1 on p. 152, below.

However, not only was group theory practically unknown amongst physicists in the 1950s and 1960s, but there also existed an open animosity towards group theory as representative of "abstract mathematics." Wolfgang Pauli supposedly [577] even called group theory *Gruppenpest* (group pestilence, in German). Many results of angular momentum and isospin symmetry were obtained not using the abstract methods of group theory, but by direct computations.[12] Following

[12] Even today, the Clebsch–Gordan coefficients and the Wigner–Eckart theorem with concrete applications – the main tool in using $SU(2)$ symmetry in the past three-quarters of a century – are very rarely even mentioned in mathematical group theory textbooks. The computational methods developed mainly by physicists [565, 258, 581, 105] still have not penetrated the "mathematicians' circles." On the other hand, the "abstract mathematics" does slowly seep into fundamental physics, especially in superstring theory, and here often finds unexpected uses.

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

the same practice, Gell-Mann derived most results in the same, "pedestrian" way, and only later discovered the elegant arguments and derivations in the then known group theory. The "eight baryon problem," i.e., the problem of finding the right generalization of the isospin $SU(2)$ symmetry that would encompass these eight baryons, thus had a thorny path. Murray Gell-Mann's "eightfold way" is in fact a collection of results that were obtained by such pedestrian methods, mostly using isospin $SU(2)$-results, as well as several phenomenological relations between strangeness, the baryon number, charge, and other properties of particles that were observed in experiments.[13]

Today, of course, we know that the relevant group is $SU(3)_f$, which indeed has the isospin $SU(2)$ as a subgroup. In hindsight, the identification of the group was obstructed by the fact that there are no three baryons that would span the fundamental 3-dimensional representation of the $SU(3)_f$ group. The early proposal by Shoichi Sakata, whereby the Λ^0-baryon extends the isospin doublet p^+, n^0 into the $SU(3)$ triplet, could not replicate the success of the isospin $SU(2)$ classification, and was soon abandoned.

Digression 4.2 In some version of Sakata's proposal, the three baryons (p^+, n^0, Λ^0) – so-called "sakatons" – were supposed to "form" all other baryons and mesons: The mesons would be obtained resulting from sakaton–antisakaton combinations, and baryons resulting from a combination of three sakatons, in all possible combinations. Formally, that indeed does produce a reasonable list of hadronic states identified by their charges, isospin, strangeness, etc. However, it was not at all clear in what sense such "products" of baryons and antibaryons could represent much lighter mesons, or – even more puzzling – how the baryon states with the quantum numbers of the sakaton triplet, (p^+, n^0, Λ^0), could also be found within the list obtained by combining three sakatons. That would imply that any one of these three baryons could be represented as a system of three copies of these very same baryons – which clearly leads to an infinite regression and obstructs the identification of this scheme as a model in which hadrons are "really" bound states of "real" particles. Many of the supporters of the so-called "S-matrix approach" to strong interactions even openly accepted this infinitely regressive interpretation of Sakata's classification scheme.

Recall that in the 1960s – especially in the southwestern parts of the USA – variants of eastern philosophies became very popular and mixed with science [☞ e.g., Refs. [91, 487, 591]]. This additionally contributed to the prejudices against group theory and to the mystique of hadron classification. As one of the most prominent advocates of $SU(3)_f$ classification, Gell-Mann contributed to this confusion both by nomenclature ("eightfold way") and by avoiding to categorically decide for or against the infinitely regressive interpretation. In sharp contrast, Richard Feynman openly advocated the "real" particle-physicist approach, whereby hadrons are really bound states of more elementary particles, which he called *partons*, avoiding Gell-Mann's "quarks." During the 1960s, Gell-Mann gradually accepted Feynman's intuitive image – they were both at CalTech (California Institute of Technology, Pasadena) – which ultimately led to the final formulation and application of the quark model.

[13] Gell-Mann spent the 1959/60 academic year in Collège de France, looking for the right generalization of $SU(2)$, and never thought of asking the resident mathematicians – amongst whom was the world-famous Jean-Pierre Serre – for help. Only late in 1960, back at CalTech, did Gell-Mann get the help of a mathematician (Richard Block) in realizing that this generalization, the $SU(3)$ group, was already very well known amongst mathematicians [119, 577].

Gell-Mann's successful prediction of the Ω^--baryon's existence [☞ Section 4.4.3], complete with its quantum numbers and its mass approximately given by the relation (2.37), was essential for accepting his "eightfold way," i.e., the classifying application of the $SU(3)_f$ group of "flavors." The original idea for the eightfold way stemmed from "finding a home" for the isospin doublet p^+–n^0 not in the direct generalization – such as Sakata's triplet – but in the *octet* of spin-$\frac{1}{2}$ baryons (2.32). In comparison, it was clear that the nine spin-$\frac{3}{2}$ baryons (4Δ, $3\Sigma^*$ and $2\Xi^*$) had to form a bigger multiplet (2.35), so that the classification scheme also had to contain in a natural way multiplets bigger than the octet, but to *not contain* multiplets such as 4-plets, 5-plets, etc.

Besides, the classification of hadrons turned out much simpler upon accepting the quark model, where the u-, d- and s-quarks span the fundamental 3-dimensional representation [☞ Section A.1.4], and mesons and baryons are bound states of quarks. Using the $SU(3)_f$ group, it is fairly easy to show that the meson and baryon multiplets must have 8, 10, 27, 28, 35, ... particles, and not some other numbers – although the $SU(3)$ group also has representations of dimensions 3, 6, 15, 21, 24, ...

To wit, representations of the $SU(3)$ group also have the so-called "triality," which is additive modulo 3 [☞ Section A.4]. The elements of the fundamental, 3-dimensional representation – i.e., the u-, d- and s-quarks – have triality 1, antiquarks triality $-1 \cong 2$, and states with n quarks and $\overline{n}$ antiquarks then have triality $(n-\overline{n}) \pmod 3$. So, if both mesons and baryons must have triality of 0, this immediately rules out the $SU(3)$ representations of dimensions 3, 6, 15, 21, 24, ..., which were indeed never observed. The "triality-0" condition selects the representations of dimensions 1, 8, 10, 27, 28, 35, etc., of which, however, only the first three groupings have ever been observed.

In turn, using that mesons are quark–antiquark ($\mathbf{3}$-$\mathbf{3}^*$) bound states and since $\mathbf{3} \otimes \mathbf{3}^* = \mathbf{1} \oplus \mathbf{8}$, mesons may only form singlets and octets of the $SU(3)_f$ classification group. Similarly, using that baryons are three-quark ($\mathbf{3}$-$\mathbf{3}$-$\mathbf{3}$) bound states and since $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}$, baryons may only form singlets, octets and decuplets; see example A.6. Consequently, the triality-0 representations **27**, **28**, **35**, etc., may only appear as metastable multi-baryon and multi-meson states [☞ Appendix A.4.2].

Of course, just like the isospin $SU(2)$ symmetry, the $SU(3)_f$-transformations are only approximate symmetries, and with a bigger tolerance.[14] With the discovery of the J/ψ-particle and the c-quark, the $SU(3)_f$-symmetry was extended into the $SU(4)_f$-symmetry, which implies even bigger tolerance, etc. This progressively growing tolerance – i.e., measure of imprecision – of the classifying $SU(n)_f$-symmetry is reflected in the *effective masses* of quarks:[15] see Table 4.1 The basic

Table 4.1 Quark masses in MeV/c^2 [☞ Figure 2.1 on p. 76]

	Quark	Mass	Effective masses in Mesons	Effective masses in Baryons
Light	u	4.2	310	363
	d	7.5		
	s	150	483	538
Heavy	c	1,100	1,500	
	b	4,200	4,700	
	t	174,200	$\gtrsim$ 174,200	

[14] In this context, the tolerance of an approximate symmetry is the margin of permitted difference between the masses of the particles linked by the purported symmetry.

[15] By its definition, mass is the measure of the object's inertia. Since quarks cannot be isolated, neither can their inertia be measured as for free particles. Their effective mass is the measure of their inertia within the bound state (meson or baryon), and so is always affected by the interactions with the "rest" of that bound state, i.e., with the other quarks and gluons and depends on the specific bound state. For the "effective mass," one then always cites average values.

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

idea in applications of such a *phenomenologically* defined $SU(n)_f$-symmetry is simple: Let G be the group of approximate symmetries with a given tolerance, and $H \subset G$ a subgroup of approximate symmetries with a finer tolerance. The contributions to the Hamiltonian that are H-invariant but not G-invariant are treated as "corrections" to the initial Hamiltonian that is G-invariant. A larger tolerance level implies a larger group of approximate symmetries, and a smaller (finer) tolerance level reduces the group of operations that are accepted as approximate symmetries.

The best known example for this idea is the so-called Zeeman effect: non-relativistic treatment of the hydrogen atom with neglected spins (4.1)–(4.8e) is subjected to an *external* magnetic field $\vec{B}$, which adds to the Hamiltonian the "correction"

$$H_Z = -\vec{\mu} \cdot \vec{B} = \mu_B \vec{B} \cdot \big(g_\ell \mathbf{L} + g_s \mathbf{S}\big), \qquad g_\ell = 1, \;\; g_s = 2\big(1 + \tfrac{\alpha}{2\pi} + \cdots\big). \tag{4.100}$$

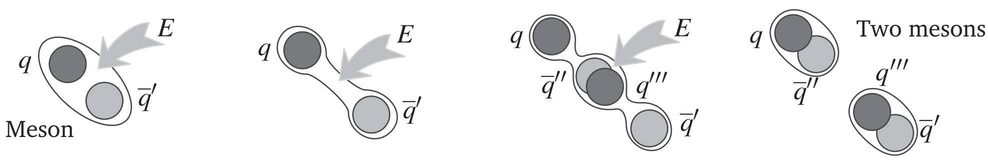
The basic Hamiltonian, without this correction, has $Spin(4)$ symmetry [☞ Section 4.1.1], whereas the Hamiltonian with the Zeeman addition only has $Spin(2) \subset Spin(4)$ symmetry. In particle-physicist parlance, one says that the Zeeman interaction with the external magnetic field – and so that external magnetic field itself – explicitly breaks the $Spin(4)$ symmetry of the hydrogen atom. By analogy, and because of the quark mass-hierarchy given in Table 4.1 on p. 152, the Hamiltonian terms, i.e., the mass contributions for mesons and baryons may be organized as:

1. the $SU(6)_f$-symmetric, original Hamiltonian;
2. the $SU(5)_f$-symmetric "corrections," where the t-quark is separated by contributions of order $m_t \approx 174.2\,\text{GeV}/c^2$;
3. the $SU(4)_f$-symmetric "corrections," where also the b-quark is separated by contributions of order $m_b \approx 4.7\,\text{GeV}/c^2$;
4. the $SU(3)_f$-symmetric "corrections," where also the c-quark is separated by contributions of order $m_c \approx 1.5\,\text{GeV}/c^2$;
5. the $SU(2)_I$-symmetric "corrections," where also the s-quark is separated by contributions of order $m_s \approx 0.5\,\text{GeV}/c^2$;
6. the final "corrections" also break the isospin $SU(2)_I$-symmetry, and finally separate the u- and d-quarks.

In practice, this approach is used in combination with other, more directly physics-inspired ideas. The next few sections will peek into some of those estimates.

4.4.1 Quarkonium

Mesons are bound states of a quark and an antiquark, so their analysis should follow the analysis of two-body bound states, akin to the hydrogen atom and positronium [☞ Section 4.1]. There is, however, a huge difference! In the hydrogen atom, the ratio of the binding energy and the rest energy of (either of) the bound particles is $13.6\,\text{eV}/510.999\,\text{keV} \approx 2.66 \times 10^{-5}$. In contradistinction, the binding energy of the quarks in mesons and baryons is in fact infinitely large, since the quarks cannot be extracted from these bound states. This involves the fact that in attempting to extract a quark one must invest amounts of energy that are at least comparable with the rest energy of the quarks themselves, whereby it is energetically more favorable to convert the invested energy into new–quark antiquark pairs



(4.101)

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

rather than further deforming the original bound state. Because of this possibility of creating quark–antiquark pairs, the process is essentially relativistic and definitely within the domain of field theory, where the number of particles is not conserved, as it is in standard quantum mechanics.

Besides, in the case of the hydrogen atom and exotic "atoms" such as muonium and positronium, the basic – Coulomb – potential is well known. In the case of strong interactions, however, there is no well-defined potential in the same sense: Recall that the Coulomb potential is a field that extends around the given electrically charged particle. In all points of space, it gives the information as to how the electrostatic force at that point would act upon a probing electric charge – *if and when* such a probing electric charge is placed at that point. In the case of electrodynamics, this mental construction has an excellent physical meaning, since it is physically possible to test the Coulomb field of a given particle with probing electric charges, which we really can move and place at will. Upon quarks, which are forever confined within mesons and baryons, we exert far less control.

In the case of mesons built of "heavy" quarks: c, b and t, it is possible to apply the analysis following the positronium template. Just before the discovery of the J/ψ-particle, Hugh David Politzer and Thomas Appelquist concluded that the c-quark – were it to exist following the logic of the so-called Glashow–Iliopoulos–Maiani (GIM) mechanism – would have to have non-relativistic $(c\,\bar{c})$ bound states akin to positronium, and which they called "charmonium." When the J/ψ-particle was experimentally discovered, it was immediately identified as the $1\,{}^3S_2$-state of charmonium,[16] and soon the other $n = 1, 2$ states (except $2\,{}^1P_1$, the detection of which poses exceptional experimental difficulties) were found.

The charmonium states are very well approximated following the positronium template, if the potential is modeled as

$$V_c = -\frac{4}{3}\frac{\alpha_s \hbar c}{r} + F_0\, r, \tag{4.102}$$

where F_0 is a coefficient of about 16 tons, and α_s is the strong interaction analogue of the fine structure constant; the coefficient $\frac{4}{3}$ will be computed in equation (6.68). This potential, of course, grows infinitely and so gives infinitely many bound states, the energy E_n of which asymptotically grows as $n^{3/2}$. *[✎ Why?]* However, the masses of the bound states with $n \geqslant 3$ are bigger than the "$D\overline{D}$-threshold" (the masses of the lightest D–$\overline{D}$ meson pair), so that such $(c\,\bar{c})$-states very quickly decay and are regarded as quasi-bound states. Table 4.2 lists a few lightest mesons that contain the c-quark.

— ❦ —

The story was repeated a few years later: in 1976, E. Eichten and K. Gottfried predicted that "bottomonium" would have to have even more true bound states than charmonium. When the first Υ-particle was detected in 1977, it was identified as the $1\,{}^3S_1$ state of the $(b\,\bar{b})$ system and during the next few years the existence of the $(b\,\bar{b})$-bound states with $n \leqslant 4$ was experimentally confirmed.

Finally, the $(t\,\bar{t})$ system was only recently detected owing to the much larger t-quark mass, and the "toponium" states are still relatively unexplored. Also, the "mixed" $(c\bar{b})$-, $(c\bar{t})$-, $(b\bar{t})$-states and their conjugates may be analyzed akin to the muonium $(\mu^+\,e^-)$. The first of these particles, $B_c^+ = (c\,\bar{b})$ and $B_c^- = (b\,\bar{c})$ are experimentally confirmed, with a mass of $6.276\,\text{GeV}/c^2$.

[16] The spectroscopic notation "$n\,{}^{2S+1}L_J$" gives the quantum numbers n, ℓ, s, j, where the letter gives the orbital angular momentum via the identification of $S, P, D, F, G, H, I, J, \ldots$ as $\ell = 0, 1, 2, 3, 4, 5, 6, 7, \ldots$

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

Table 4.2 Lightest mesons containing the c-quark; masses in MeV/c^2

Name	$n\,^{2S+1}L_J$	J^{PC}	Mass
η_c	$1\,^1S_0$	0^{-+}	2,980.3
J/ψ	$1\,^3S_1$	1^{--}	3,096.9
χ_{c0}	$1\,^3P_0$	0^{++}	3,414.8
χ_{c1}	$1\,^3P_1$	0^{++}	3,510.7
χ_{c2}	$1\,^3P_2$	0^{++}	3,556.2
η_c	$2\,^1S_0$	0^{-+}	3,637
J/ψ	$2\,^3S_1$	1^{--}	3,686.1

Charmonium: $(c\,\bar{c})$ states

Name	$(q\,\bar{q})$	J^P	Mass
D^+	$(c\,\bar{d})$	0^-	1,869.6
D^-	$(d\,\bar{c})$		
D^0	$(c\,\bar{u})$	0^-	1,864.8
$\bar{D}^0$	$(u\,\bar{c})$		
D^{*0}	$(c\,\bar{u})$	1^-	2,007.0
$\bar{D}^{*0}$	$(u\,\bar{c})$		
D^{*+}	$(c\,\bar{d})$	1^-	2,010.3
D^{*-}	$(d\,\bar{c})$		

4.4.2 Light mesons

Mesons that contain a light quark or antiquark automatically must be analyzed as relativistic bound states – for which there is no complete theoretical description✍.[17] Therefore, we remain content herein with classification.

The first fact worth noting is that although there are three quarks at our disposal, and so *nine* possible $(q\,\bar{q})$ bound states (fully neglecting spins, orbital angular momentum and dynamical details), mesons appear in groups of *eight* [☞ plot (2.31) as well as the result (A.76c)].

The reason for this is similar to the fact that with only two quarks, u and d, there exist not four but only three pions, $\pi^{\pm}, \pi^0$. The $SU(2)_I$ symmetry solves this puzzle by the method of isospin "addition." But, before that, the isospin of $\bar{u}$ and $\bar{d}$ must be established. In tensor notation, we have

$$\{t^1, t^2\} = \{u,\ d\} \quad \Rightarrow \quad \{t_1, t_2\} = \{\bar{u},\ \bar{d}\}. \tag{4.103}$$

However, since $\varepsilon_{\alpha\beta}$ is $SU(2)$-invariant *[✎ why?]*, we may identify $(t^\alpha)^\dagger = t_\alpha = \varepsilon_{\alpha\beta} t^\beta$, so that (regarding isospin properties only!)

$$t_1 = \varepsilon_{12} t^2 \ = t^2 \ \Rightarrow \ |\bar{u}\rangle = |d\rangle \quad = |\tfrac{1}{2}, -\tfrac{1}{2}\rangle, \tag{4.104a}$$

$$\text{and} \qquad t_2 = \varepsilon_{21} t^1 = -t^1 \ \Rightarrow \ |\bar{d}\rangle = -|u\rangle = -|\tfrac{1}{2}, +\tfrac{1}{2}\rangle. \tag{4.104b}$$

Thus,

$$\big\{|\bar{u}\rangle,\ |\bar{d}\rangle\big\} \otimes \big\{|u\rangle,\ |d\rangle\big\} = \big\{|\tfrac{1}{2}, -\tfrac{1}{2}\rangle, -|\tfrac{1}{2}, +\tfrac{1}{2}\rangle\big\} \otimes \big\{|\tfrac{1}{2}, +\tfrac{1}{2}\rangle, |\tfrac{1}{2}, -\tfrac{1}{2}\rangle\big\} \tag{4.105}$$

$$= \big\{|1, \pm 1\rangle,\ |1, 0\rangle\big\} \oplus \big\{|0, 0\rangle\big\} = \big\{|\pi^{\pm}\rangle,\ |\pi^0\rangle\big\} \oplus \big\{|\eta\rangle\big\}, \tag{4.106}$$

where

$$\begin{aligned}
|1, +1\rangle &= |\tfrac{1}{2}, +\tfrac{1}{2}\rangle|\tfrac{1}{2}, +\tfrac{1}{2}\rangle &&= -|\bar{d}\rangle|u\rangle, \\
|1, 0\rangle &= \tfrac{1}{\sqrt{2}}\big(|\tfrac{1}{2}, +\tfrac{1}{2}\rangle|\tfrac{1}{2}, -\tfrac{1}{2}\rangle + |\tfrac{1}{2}, -\tfrac{1}{2}\rangle|\tfrac{1}{2}, +\tfrac{1}{2}\rangle\big) &&= \tfrac{1}{\sqrt{2}}\big(-|\bar{d}\rangle|d\rangle + |\bar{u}\rangle|u\rangle\big), \\
|1, -1\rangle &= |\tfrac{1}{2}, -\tfrac{1}{2}\rangle|\tfrac{1}{2}, -\tfrac{1}{2}\rangle &&= |\bar{u}\rangle|d\rangle, \\
|0, 0\rangle &= \tfrac{1}{\sqrt{2}}\big(|\tfrac{1}{2}, +\tfrac{1}{2}\rangle|\tfrac{1}{2}, -\tfrac{1}{2}\rangle - |\tfrac{1}{2}, -\tfrac{1}{2}\rangle|\tfrac{1}{2}, +\tfrac{1}{2}\rangle\big) &&= \tfrac{1}{\sqrt{2}}\big(-|\bar{d}\rangle|d\rangle - |\bar{u}\rangle|u\rangle\big).
\end{aligned} \tag{4.107}$$

[17] N.B. One of the approaches is the so-called (MIT) "bag model": One approximates that quarks are free particles while within the meson, where they are confined by outside "pressure," which produces an impenetrable "bag." This "bag" is a 3-dimensional infinitely deep potential, the walls of which have a time-variable shape. Although the model is phenomenologically successful, it is clear that this is an *ad hoc* fiction where one needs to explain the dynamical origin of that "pressure." Another approach uses the so-called Gribov version of gauge theory of strong interactions and quark confinement [388].

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

Then[18]

$$\begin{cases} |\pi^+\rangle = -|\bar{d}\,u\rangle, \\ |\pi^0\rangle = \frac{1}{\sqrt{2}}\big(|\bar{u}\,u\rangle - |\bar{d}\,d\rangle\big), \\ |\pi^-\rangle = |\bar{u}\,d\rangle, \end{cases} \qquad \text{and} \qquad |\eta\rangle = -\frac{1}{\sqrt{2}}\big(|\bar{u}\,u\rangle + |\bar{d}\,d\rangle\big). \tag{4.108}$$

Note the signs that stem from the $SU(2)_I$-identification $|\bar{d}\rangle = -|\frac{1}{2},+\frac{1}{2}\rangle$, whereby $|\pi^0\rangle$ *looks* like an antisymmetric combination, but is not: a quark and an antiquark cannot be thought of as particles that are "identical up to some 'polarization' (or other selectable property)," so as to define the exchange (anti)symmetry. Instead, if we use u, d as the basis and $\bar{u}, \bar{d}$ as its conjugate basis, the three pion states (4.108) form a Hermitian matrix with no trace, whereas the η-state represents the trace of a Hermitian matrix.

So, define q^α so that $q^1 = u$ and $q^2 = d$, and

$$\pi^+ = (\bar{q}_\beta\,(\boldsymbol{\sigma}^+)_\alpha{}^\beta\,q^\alpha) \quad = (\bar{d}u), \tag{4.109a}$$

$$\pi^0 = \tfrac{1}{\sqrt{2}}(\bar{q}_\beta\,(\boldsymbol{\sigma}^3)_\alpha{}^\beta\,q^\alpha) = \tfrac{1}{\sqrt{2}}(\bar{u}u - \bar{d}d), \tag{4.109b}$$

$$\pi^- = (\bar{q}_\beta\,(\boldsymbol{\sigma}^-)_\alpha{}^\beta\,q^\alpha) \quad = (\bar{u}d), \tag{4.109c}$$

where $\boldsymbol{\sigma}^\pm := \frac{1}{2}[\boldsymbol{\sigma}^1 \pm i\boldsymbol{\sigma}^2]$ and $\boldsymbol{\sigma}^1, \boldsymbol{\sigma}^2, \boldsymbol{\sigma}^3$ are Pauli matrices:

$$\boldsymbol{\sigma}^1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \boldsymbol{\sigma}^2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad \boldsymbol{\sigma}^3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{4.109d}$$

the halves of which satisfy the defining relations (A.38a) of the $SU(2)$ algebra.

However, comparison with experiments does not single out an unambiguous candidate for $|\eta\rangle$: in fact, there exist two spin-0 (pseudo-scalar, $J^{PC} = 0^{-+}$) particles with isospin $|0,0\rangle$:

$$\eta: \ 547.853\,\text{MeV}/c^2, \qquad \eta': \ 957.66\,\text{MeV}/c^2, \tag{4.110}$$

as well as two spin-1 (vectorial, $J^{PC} = 1^{--}$) excitations:

$$\omega: \ 782.65\,\text{MeV}/c^2, \qquad \phi: \ 1019.455\,\text{MeV}/c^2. \tag{4.111}$$

Since the masses of the η- and ω-mesons are larger than the kaon masses (2.31), it is clear that the classification must also include the mesons that contain the s-quark, and also that η and η' (and similarly ω and ϕ) are linear combinations that also contain (s-$\bar{s}$)-contributions.

Generalizing the identification (4.109), by means of including the third quark, $q^3 = s$, and using Gell-Mann's matrices (A.71), we have

$$\pi^+ = (\bar{d}u), \quad \pi^- = (\bar{u}d), \quad \pi^0 = \tfrac{1}{\sqrt{2}}(\bar{u}u - \bar{d}d), \quad \eta = \tfrac{1}{\sqrt{6}}(\bar{u}u + \bar{d}d - 2\bar{s}s), \tag{4.112a}$$

$$K^+ = (\bar{s}u), \quad K^0 = (\bar{s}d), \quad \overline{K}^0 - = (\bar{d}s), \qquad K^- = (\bar{u}s), \tag{4.112b}$$

and

$$\eta' = \tfrac{1}{\sqrt{3}}(\bar{u}u + \bar{d}d + \bar{s}s). \tag{4.112c}$$

These $8+1$ states with antiparallel ($S = 0$) quark spins and orbital angular momentum $\ell = 0$ then have their total angular momentum $j = 0$.

[18] In fact, the *real* η^0 particle is a linear combination of not only $|u\rangle \otimes |\bar{u}\rangle$ and $|d\rangle \otimes |\bar{d}\rangle$, but also of $|s\rangle \otimes |\bar{s}\rangle$; see equations (4.112).

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

Combining the antiquark triplet $\{\bar{q}_\alpha,\ \alpha = 1,2,3\}$ and quark triplet $\{q^\alpha,\ \alpha = 1,2,3\}$ into 8+1 mesons (4.112) then precisely follows the $SU(3)$ decomposition (A.76c):

$$\mathbf{3}^* \otimes \mathbf{3} = \mathbf{8} \oplus \mathbf{1}. \tag{4.113}$$

Of course, this $SU(3)_f$-symmetry is approximate: in reality, kaons are heavier than the pions, as they contain the heavier s-quark instead of the lighter d-quark. Besides, the η-meson is an integral part of the $SU(3)_f$ octet and its mass is just barely larger than the kaon mass, reflecting the approximate nature of the $SU(3)_f$-symmetry. On the other hand, the η'-meson does not belong in the $SU(3)_f$ octet, but is a "singlet" – i.e., an $SU(3)_f$-invariant.

Excitation of these states where the sum of quark spins and orbital angular momentum equals 1 then produces the "vector" (spin-1) $\rho^\pm$-, ρ^0-, $K^{*\pm}$-, K^{*0}-, $\overline{K}^{*0}$- and ϕ-mesons. The total angular momentum of the bound state – with all contributions, orbital and spin – is the spin of the bound state as a particle. While the masses of the charged vector-mesons follow those of the charged pseudo-scalar mesons, ϕ- and ω-mesons mix "maximally." Experiments indicate that

$$\omega \neq \tfrac{1}{\sqrt{6}}(\bar{u}u + \bar{d}d - 2\bar{s}s), \qquad \text{but} \qquad \omega \approx \tfrac{1}{\sqrt{2}}(\bar{u}u + \bar{d}d); \tag{4.114a}$$

$$\phi \neq \tfrac{1}{\sqrt{3}}(\bar{u}u + \bar{d}d + \bar{s}s), \qquad \text{but} \qquad \phi \approx (\bar{s}s). \tag{4.114b}$$

The vector and pseudo-scalar mesons turn out to differ predominantly in the relative orientation of quark spins and both are well described as S-states, with no relative angular momentum. The difference in their masses should then stem from the spin–spin interaction, akin to the $\boldsymbol{S}_e$–$\boldsymbol{S}_p$ contribution (4.27c) to hyperfine structure in the spectrum of the hydrogen atom. Thus, the meson mass is parametrized as

$$M(\text{meson}) \approx m_q + m_{\bar{q}} + \frac{A}{m_q m_{\bar{q}}}\,\langle \boldsymbol{S}_q \cdot \boldsymbol{S}_{\bar{q}} \rangle\,, \tag{4.115}$$

where the coefficient A is some multiple of $|\Psi(\vec{0}, t)|^2$ that cannot be computed reliably for a relativistic system, and so is determined by comparing with experimental data. Using the well-known "trick":

$$\vec{\boldsymbol{S}} := \vec{\boldsymbol{S}}_q + \vec{\boldsymbol{S}}_{\bar{q}} \quad \Rightarrow \quad \vec{\boldsymbol{S}}_q \cdot \vec{\boldsymbol{S}}_{\bar{q}} = \tfrac{1}{2}\big(\boldsymbol{S}^2 - \boldsymbol{S}_q^2 - \boldsymbol{S}_{\bar{q}}^2\big), \tag{4.116}$$

the difference between the observed average masses of pseudo-scalar and vector mesons is rather well explained:

$$\vec{\boldsymbol{S}}_q \cdot \vec{\boldsymbol{S}}_{\bar{q}} = \begin{cases} \ \ \tfrac{1}{4}\hbar^2, & \text{for } S = 1 \text{ (vector mesons)}, \\ -\tfrac{3}{4}\hbar^2, & \text{for } S = 0 \text{ (pseudo-scalar mesons)}. \end{cases} \tag{4.117}$$

Using the effective (so-called "constituent") masses of quarks inside mesons from Table 4.1 on p. 152, and the best value of the parameter $A \approx \frac{4m_u^2}{\hbar^2}\,160\,\text{MeV}/c^2$, the masses of pseudo-scalar and vector mesons are obtained to within 1% from the experimental value [☞ Table 4.3] – except for the η'-meson, the mass of which poses an exceptional problem for the quark model [☞ commentary in Ref. [445]].

In this way the quark model with the $SU(3)_f$-symmetry predicts an infinitely growing ladder of meson $(\mathbf{8} + \mathbf{1})$ nonets, in good agreement with experiments up to the indicated discrepancies (4.114); Table 4.4 lists the first few nonets.

4.4.3 Baryons

The number of experimentally detected baryons composed of the u-, d- and s-quarks grows faster with mass than is the case with mesons. Foremost, this happens because of the fact that baryons are three-particle bound states, so that there exist combinatorially more different interactive contributions to the mass – such as (4.16) and (4.27a)–(4.27c), as well as (4.46a)

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

Table 4.3 Average masses of pseudo-scalar and vector mesons, in MeV/c^2. The η'-meson mass poses an exceptional problem for the quark model; see commentary in Ref. [445].

Meson	Computed	Measured	Meson	Computed	Measured
π	140	138	ρ	780	776
K	485	496	ω	780	783
η	559	549	K^*	896	892
η'	303	958	ϕ	1,032	1,020

Table 4.4 Lightest meson nonets in the $SU(3)_f$ quark model

			Nonet content			Mass*
ℓ	S	J^{PC}	$I=1$	$I=\frac{1}{2}$	$I=0$	(MeV/c^2)
0	0	0^{-+}	π	K	η,η'	500
	1	1^{--}	ρ	K^*	ω,ϕ	800
1	0	1^{+-}	B	Q_2	$H,?$	1,250
	1	0^{++}	δ	κ	ϵ,S^*	1,150
		1^{++}	A_1	Q_1	D,E	1,300
		2^{++}	A_2	K^*	f,f'	1,400

* Rough averages; see plot (2.31) and Ref. [293]

and (4.47) [☞ plots (2.32) and (2.35)]. These significantly complicate the computations and even just the estimates, hindering the experimental identification as to which baryon belongs to which multiplet.

Classification

As three-particle systems, baryons have *two* orbital angular momenta: the orbital angular momentum of any two of the three quarks about their center of mass, and then the orbital angular momentum of that two-quark system and the third quark about their joint center of mass. We will consider only states with $n=1$ and $\ell=0=\ell'$, the masses of which are easily shown to be the lowest.[19] In this case, the baryon spin stems exclusively from the sum of the quark spins, for which the addition of three spins of magnitude $\frac{1}{2}$ we have

$$\big\{|\tfrac{3}{2},\pm\tfrac{3}{2}\rangle,\ |\tfrac{3}{2},\pm\tfrac{1}{2}\rangle\big\}_S,\qquad \big\{|\tfrac{1}{2},\pm\tfrac{1}{2}\rangle_{[12]}\big\},\qquad \big\{|\tfrac{1}{2},\pm\tfrac{1}{2}\rangle_{[23]}\big\}. \tag{4.118}$$

The subscript S denotes total symmetry and, following Ref. [243], we use the basis

$$|\tfrac{1}{2},+\tfrac{1}{2}\rangle_{[12]}=\frac{1}{\sqrt{2}}\big(|\uparrow\downarrow\uparrow\rangle-|\downarrow\uparrow\uparrow\rangle\big),\qquad |\tfrac{1}{2},-\tfrac{1}{2}\rangle_{[12]}=\frac{1}{\sqrt{2}}\big(|\uparrow\downarrow\downarrow\rangle-|\downarrow\uparrow\downarrow\rangle\big); \tag{4.119}$$

$$|\tfrac{1}{2},+\tfrac{1}{2}\rangle_{[23]}=\frac{1}{\sqrt{2}}\big(|\uparrow\uparrow\downarrow\rangle-|\uparrow\downarrow\uparrow\rangle\big),\qquad |\tfrac{1}{2},-\tfrac{1}{2}\rangle_{[23]}=\frac{1}{\sqrt{2}}\big(|\downarrow\uparrow\downarrow\rangle-|\downarrow\downarrow\uparrow\rangle\big), \tag{4.120}$$

which are antisymmetric with respect to the exchange of the particles indicated in the subscript. It is not hard to show that

$$|\tfrac{1}{2},+\tfrac{1}{2}\rangle_{[13]}=\frac{1}{\sqrt{2}}\big(|\uparrow\uparrow\downarrow\rangle-|\downarrow\uparrow\uparrow\rangle\big)=|\tfrac{1}{2},+\tfrac{1}{2}\rangle_{[12]}+|\tfrac{1}{2},+\tfrac{1}{2}\rangle_{[23]}, \tag{4.121}$$

[19] The mass of a baryon as a bound state of three quarks equals the sum of the masses of the constituent quarks, minus the mass equivalent of the binding energy. Then, the strongest-bound baryons are also the lightest amongst the possible bound states of the given quarks.

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

and similarly for $|\frac{1}{2}, -\frac{1}{2}\rangle_{[13]}$. We introduced the abbreviations:

$$|\uparrow\rangle := |\tfrac{1}{2}, +\tfrac{1}{2}\rangle, \quad |\downarrow\rangle := |\tfrac{1}{2}, -\tfrac{1}{2}\rangle, \quad |\uparrow\downarrow\uparrow\rangle := |\tfrac{1}{2}, +\tfrac{1}{2}\rangle|\tfrac{1}{2}, -\tfrac{1}{2}\rangle|\tfrac{1}{2}, +\tfrac{1}{2}\rangle, \quad \text{etc.} \tag{4.122}$$

Using the approximate $SU(3)_f$-symmetry, u-, d- and s-quarks are treated as if they were different polarizations of the same fermion, so that Pauli's exclusion principle must be applied. That is, the *entire* wave-function of the bound state of three quarks must be antisymmetric with respect to the exchange of any two of the three quarks. The wave-function for the baryon then factorizes:

$$\Psi(\text{baryon}) = \Psi(\vec{r}, t)\, \chi(\text{spin})\, \chi(\text{flavor})\, \chi(\text{color}). \tag{4.123}$$

For states with $\ell = 0 = \ell'$, $\Psi(\vec{r}, t)$ must be a totally symmetric function since it cannot depend on angles, and so neither on the quarks' relative positions. On the other hand, the color factor depends on the additional degree of freedom: each quark has a linear combination of the three primary colors [☞ Section 2.3.13]. That is, every quark is in fact a triple of quarks that span the 3-dimensional representation of the $SU(3)_c$-symmetry,[20] and a bound state of three quarks *must* be $SU(3)_c$-invariant. Group theory applies to the $SU(3)_c$-symmetry as well as for $SU(3)_f$, and the decomposition (A.76f) provides for the fact that the $SU(3)_c$-invariant factor $\chi(\text{color})$ is totally antisymmetric.

Since the entire product (4.123) must be totally antisymmetric by Pauli's exclusion principle, and $\Psi(\vec{r}, t)$ is totally symmetric while $\chi(\text{color})$ is totally antisymmetric, it follows that the product $\chi(\text{spin})\chi(\text{flavor})$ must be totally symmetric.

Since the $\chi(\text{flavor})$ factor for the decuplet of the $SU(3)_f$-symmetry is totally symmetric [☞ decomposition (A.76f)], it follows that the $\chi(\text{spin})$ factor must also be totally symmetric. Writing out the first two of the kets (4.118):

$$|\tfrac{3}{2}, +\tfrac{3}{2}\rangle = |\uparrow\uparrow\uparrow\rangle, \qquad |\tfrac{3}{2}, +\tfrac{1}{2}\rangle = \tfrac{1}{\sqrt{3}}\big(|\uparrow\uparrow\downarrow\rangle + |\uparrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\rangle\big), \tag{4.124a}$$

$$|\tfrac{3}{2}, -\tfrac{3}{2}\rangle = |\downarrow\downarrow\downarrow\rangle, \qquad |\tfrac{3}{2}, -\tfrac{1}{2}\rangle = \tfrac{1}{\sqrt{3}}\big(|\uparrow\downarrow\downarrow\rangle + |\downarrow\uparrow\downarrow\rangle + |\downarrow\downarrow\uparrow\rangle\big), \tag{4.124b}$$

we know that these four spin states $|\frac{3}{2}, m_s\rangle$ are totally symmetric, so that these ten baryons *must* have spin-$\frac{3}{2}$. That is Gell-Mann's decuplet $(4\Delta, 3\Sigma^*, 2\Xi^*, \Omega^-)$, where the fast experimental confirmation of the predicted Ω^- baryon induced most researchers to finally accept the quark model.

The construction of the octet is a little more complicated, as we must find a totally symmetric (linear combination) of products of $\chi(\text{spin})$ and $\chi(\text{flavor})$ that, separately, have a mixed symmetry. Notice first that the product $\chi_{[12]}(\text{spin})\chi_{[12]}(\text{flavor})$ is *symmetric* with respect to the exchange of the first two particles, since each of the two factors is antisymmetric. Then, it follows that the linear combination

$$\chi_{[12]}(\text{spin})\, \chi_{[12]}(\text{flavor}) + \chi_{[13]}(\text{spin})\, \chi_{[13]}(\text{flavor}) + \chi_{[23]}(\text{spin})\, \chi_{[23]}(\text{flavor}) \tag{4.125}$$

is totally symmetric and provides the spin-flavor factor in the wave-function (4.123) for the octet of spin-$\frac{1}{2}$ baryons. In spite of the relationship (4.121) – whereby $|\frac{1}{2}, m_s\rangle_{[13]}$ is linearly dependent on $|\frac{1}{2}, m_s\rangle_{[12]}$ and $|\frac{1}{2}, m_s\rangle_{[23]}$, the bilinear terms in the expression (4.125) are linearly independent, and the full expression does not simplify.

— ❦ —

Without the additional color degree of freedom for quarks, i.e., without the totally antisymmetric $\chi(\text{spin})$ factor in the product (4.123), the product $\Psi(\vec{r}, t)\, \chi(\text{spin})\, \chi(\text{flavor})$ would have to be

[20] Unlike the approximate $SU(3)_f$-symmetry, the $SU(3)_c$-symmetry is exact.

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

totally antisymmetric. For the state with smallest mass where $n = 1$ and $\ell = 0 = \ell'$, the product $\chi(\text{spin})\,\chi(\text{flavor})$ would have to be totally antisymmetric. For spin-$\frac{1}{2}$ octets, one could construct such a state, but the spin-$\frac{3}{2}$ baryons would have to have a totally antisymmetric $\chi(\text{flavor})$ factor, which would have to be the $SU(3)_f$-invariant – and so a single spin-$\frac{3}{2}$ baryon, instead of the ten experimentally detected ones (2.35).

This is the core of the problem noticed by Oscar W. Greenberg in 1964. As a resolution, he proposed [229] that the quark annihilation and creation operators should satisfy para-fermionic rules (2.41c)–(2.41d). This turns out to effectively introduce an additional degree of freedom – the same as the one called "color" in the 1965 independent proposal by Han and Nambu, where quarks had integral electric charges, with values that depend on the color [☞ Digression 5.14 on p. 214]. Their model also predicted particles (that would soon be called gluons) that mediate transformations of the color charge in quarks, where these transformations have the structure of the exact $SU(3)_c$ symmetry group and are the source of the strong interaction [☞ Section 6.1]. The current version of this model with fractionally (and color-independently) charged quarks was proposed by Harald Fritzsch and Murray Gell-Mann in 1971, and was finalized in collaboration with William A. Bardeen in 1973 [32].

Masses

By the reasoning used so far that led to the approximation (4.115), for baryons we have

$$M(\text{baryon}) \approx m_1 + m_2 + m_3 + A' \sum_{i \neq j} \frac{1}{m_i m_j} \left\langle S_i \cdot S_j \right\rangle , \tag{4.126}$$

where the spin–spin contributions (leading to the hyperfine structure in the hydrogen atom spectrum) must be computed separately for baryons in isospin groups. Indeed, in the general case, the three masses are different and every spin–spin pair – which stems from the dipole–dipole magnetic interaction – must be considered separately.

In the decuplet case, the situation is simpler, as the $\chi(\text{flavor})$ factor and therefore also the $\chi(\text{spin})$ factor are both totally symmetric. Thus, the spins of any two quarks are parallel, and the well-known "trick"

$$\vec{S}_1 \cdot \vec{S}_2 = \tfrac{1}{2}\Big((\vec{S}_1 + \vec{S}_2)^2 - S_1^2 - S_2^2\Big) \tag{4.127}$$

provides for each pair of quarks in the baryon decuplet:

$$\left\langle \vec{S}_i \cdot \vec{S}_j \right\rangle = \tfrac{1}{2}\big(2 - \tfrac{1}{2}(\tfrac{1}{2}{+}1) - \tfrac{1}{2}(\tfrac{1}{2}{+}1)\big)\hbar^2 = \tfrac{1}{4}\hbar^2, \qquad \text{for the decuplet.} \tag{4.128}$$

The cases

$$M(\Delta) \approx 3m_u + \frac{3A'\hbar^2}{4m_u^2} \qquad \text{and} \qquad M(\Omega^-) \approx 3m_s + \frac{3A'\hbar^2}{4m_s^2} \tag{4.129}$$

are particularly simple, where the first estimate applies to $\Delta^{++}, \Delta^{+}, \Delta^{0}$ and Δ^{-} since $m_u \approx m_d$ [☞ Table 4.1 on p. 152]. The results

$$M(\Sigma^*) \approx 2m_u + m_s + \frac{A'\hbar^2}{4}\Big(\frac{1}{m_u^2} + \frac{2}{m_u m_s}\Big), \tag{4.130a}$$

$$M(\Xi^*) \approx m_u + 2m_s + \frac{A'\hbar^2}{4}\Big(\frac{2}{m_u m_s} + \frac{1}{m_s^2}\Big) \tag{4.130b}$$

are just a little more involved.

For the baryon octet, we first must look at the isospin sub-multiplets. For example, we know the Λ^0-baryon, a (u,d,s) bound state, has the isospin $|0,0\rangle$-factor antisymmetric with respect to

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

the $u \leftrightarrow d$ exchange. For the $\chi(\text{spin})\,\chi(\text{flavor})$ product to be symmetric, it must be that the spin factor is also antisymmetric, and it then follows that the u- and d-quark spins in the Λ^0-baryon are antiparallel. Similarly, we know that the $\Sigma^\pm$- and Σ^0-baryons, also (u, d, s) bound states, form an isospin triplet, $\{|1, \pm 1\rangle, |1, 0\rangle\}$, so that the u- and d-quark spins in the Σ-baryons must be parallel. Thus,

$$\left\langle \vec{S}_u \cdot \vec{S}_d \right\rangle = \begin{cases} \frac{1}{4}\hbar^2 & \text{in the } \Lambda^0\text{-baryon}, \\ -\frac{3}{4}\hbar^2 & \text{in the } \Sigma\text{-baryons}. \end{cases} \tag{4.131}$$

Also the generalization of the relation (4.127) gives

$$\vec{S}_1 \cdot \vec{S}_2 + \vec{S}_1 \cdot \vec{S}_3 + \vec{S}_2 \cdot \vec{S}_3 = \tfrac{1}{2}\Big((\vec{S}_1 + \vec{S}_2 + \vec{S}_3)^2 - S_1^2 - S_2^2 - S_3^2\Big), \tag{4.132}$$

and

$$\left\langle \vec{S}_1 \cdot \vec{S}_2 + \vec{S}_1 \cdot \vec{S}_3 + \vec{S}_2 \cdot \vec{S}_3 \right\rangle = \begin{cases} \frac{3}{4}\hbar^2 & \text{for the spin-}\frac{3}{2}\text{ decuplet}, \\ -\frac{3}{4}\hbar^2 & \text{for the spin-}\frac{1}{2}\text{ octet}. \end{cases} \tag{4.133}$$

Adding the corresponding terms and using that $m_d \approx m_u$, we have

$$M(p^+, n^0) \approx 3m_u - \frac{3A'\hbar^2}{4m_u^2}, \tag{4.134}$$

$$M(\Lambda) \approx 2m_u + m_s - \frac{3A'\hbar^2}{4m_u^2}, \tag{4.135}$$

$$M(\Sigma) \approx 2m_u + m_s + \frac{A'\hbar^2}{4}\Big(\frac{1}{m_u^2} - \frac{4}{m_u m_s}\Big), \tag{4.136}$$

$$M(\Xi) \approx 2m_u + m_s + \frac{A'\hbar^2}{4}\Big(\frac{1}{m_s^2} - \frac{4}{m_u m_s}\Big). \tag{4.137}$$

With the effective quark masses taken from Table 4.1 on p. 152, and choosing the constant $A' = (2m_u/\hbar)^2 \cdot 50\,\text{MeV}/c^2$, one obtains excellent approximations for the measured masses [☞ Table 4.5].

Table 4.5 The lightest baryon masses in MeV/c^2

Baryon	Computed	Measured	Baryon	Computed	Measured
p^+, n^0	939	939	Δ	1,239	1,232
Λ	1,116	1,114	Σ^*	1,381	1,384
Σ	1,179	1,193	Ξ^*	1,529	1,533
Ξ	1,327	1,318	Ω	1,682	1,672

Magnetic moments

In the absence of orbital angular momenta, $\ell = 0 = \ell'$, the baryon magnetic moment is – up to corrections of the type (4.24b) – simply the sum of the constituent quarks' magnetic moments:

$$\vec{\mu}(\text{baryon}) = \vec{\mu}^{(1)} + \vec{\mu}^{(2)} + \vec{\mu}^{(3)}. \tag{4.138}$$

For a spin-$\frac{1}{2}$ particle, with charge q and mass m, we have

$$\langle \mu_3 \rangle = \left\langle \frac{q}{m_e c} S_3 \right\rangle = \pm \frac{q\hbar}{2m_e c}, \tag{4.139}$$

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

and so

$$\mu_u := \langle \mu_3^{(u)} \rangle = \pm \frac{e\hbar}{3m_u c}, \qquad \mu_d := \langle \mu_3^{(d)} \rangle = \mp \frac{e\hbar}{6m_d c}, \qquad \mu_s := \langle \mu_3^{(s)} \rangle = \mp \frac{e\hbar}{6m_s c}, \tag{4.140}$$

and

$$\langle \mu_3(\text{baryon}) \rangle = \frac{2}{\hbar} \sum_{i=1}^{3} \Big\langle \text{baryon} | \mu_i \, S_3^{(i)} | \text{baryon} \Big\rangle . \tag{4.141}$$

So, to compute the baryon magnetic moment, one must compute the right-hand side contribution to the relation (4.141) for each baryon separately, by

1. writing out the baryon state explicitly as a linear combination (4.125), using the results (4.124) and (4.118),
2. substituting this in the right-hand side sum (4.141), term by term and for both the ket and the bra,
3. evaluating each term in the so-expanded sum,
4. and finally adding the partial results.

There are clearly many contributions, but the so-obtained values are in very good agreement with the experimental measurements, as shown in Table 4.6 [243].

Table 4.6 The baryon magnetic dipole magnitudes in the first octet, expressed in units of nuclear magneton, $\frac{e\hbar}{2m_p c} = 3.152 \times 10^{-13}\,\text{MeV}/c^2/\text{T}$

Baryon	$\langle \mu_3 \rangle$	Computed	Measured
p^+	$\frac{1}{3}(4\mu_u - \mu_d)$	2.79	2.793
n^0	$\frac{1}{3}(4\mu_d - \mu_u)$	−1.86	−1.913
Ξ^0	$\frac{1}{3}(4\mu_s - \mu_u)$	−1.40	−1.253
Ξ^-	$\frac{1}{3}(4\mu_u - \mu_s)$	−0.47	−0.69
Λ^0	μ_s	−0.58	−0.61
Σ^+	$\frac{1}{3}(4\mu_u - \mu_s)$	2.68	2.33
Σ^0	$\frac{1}{3}(2\mu_u + \mu_d - \mu_s)$	0.82	—
Σ^-	$\frac{1}{3}(4\mu_d - \mu_s)$	−1.05	−1.41

As a final note on hadron spectroscopy, the Reader should recall that most hadrons decay within $\sim 10^{-23}$ s of their creation within clusters of hundreds and thousands of simultaneous collision processes. The fact that such measurements on individual particles are in fact being performed is an impressive feat of resourcefulness, ingenuity and hard work.

4.4.4 Exercises for Section 4.4

✎ ***4.4.1*** *Estimate the relative magnitudes of the contributions analogous to (4.8b), (4.22), (4.28), (4.32), (4.33), (4.34), (4.35), (4.40), (4.46a) and (4.47), as functions of α_s, the strong interaction constant, for $(c\bar{c})$, $(b\bar{b})$ and $(t\bar{t})$ systems.*

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press

✎ ***4.4.2*** *Estimate the relative magnitudes of the contributions analogous to (4.8b), (4.22), (4.28), (4.32), (4.33), (4.34), (4.35), (4.40), (4.46a) and (4.47), as functions of α_s, the strong interaction constant, for $(c\bar{b})$, $(c\bar{t})$ and $(b\bar{t})$ systems.*[21]

✎ ***4.4.3*** *From the fact that the lifetime of charmonium states above the D–$\overline{D}$ threshold is $\sim 10^{-23}$ s and by comparing with the positronium lifetime (4.51), estimate α_s and show that $\alpha_s \sim O(\frac{1}{10}) - O(1)$.*

✎ ***4.4.4*** *Write out all collisions of the $\pi + \pi \to \pi + \pi$ type.*

✎ ***4.4.5*** *Find the relation between the probabilities of the four collisions:*

$$\pi^+ + \pi^+ \to \pi^+ + \pi^+, \qquad \pi^+ + \pi^0 \to \pi^+ + \pi^0, \tag{4.142a}$$

$$\pi^+ + \pi^- \to \pi^+ + \pi^- \quad \text{and} \quad \pi^+ + \pi^- \to \pi^0 + \pi^0. \tag{4.142b}$$

✎ ***4.4.6*** *Use the Wenzel–Brillouin–Kramers (WKB) approximation to prove that, for the potential (4.102), the bound-state energies $E_n \approx E_1 n^{3/2}$ for large enough n.*

✎ ***4.4.7*** *Determine the degeneracy of the states predicted by the non-relativistic treatment of the potential (4.102). (Hint: try verifying the maximal symmetry of this non-relativistic Hamiltonian using the relations (4.13) and explicit computation.)*

✎ ***4.4.8*** *Upon fully expanding the expression (4.125) and by explicitly tracing the action of swapping quarks, show that the complete expression is symmetric with respect to the exchange of any two of the three quarks.*

✎ ***4.4.9*** *Derive the relations (4.130).*

[21] N.B. The details of these systems are subject to contemporary research.

https://doi.org/10.1017/9781009291507.007 Published online by Cambridge University Press