Introduction to Particle Physics

Dr. F. Filthaut

April 18, 2018

These lecture notes are intended for use with the lecture course “Subatomaire Fysica”. The text in smaller font covers material that is not mandatory but that is hoped to provide useful extra background. Each chapter starts with a shaded area that lists the the most important pieces of information to take away from it.

Dr. F. Filthaut
Experimental High-Energy Physics
Institute for Mathematics, Astrophysics, and Particle Physics
Faculty of Science, Radboud University Nijmegen
e-mail: F.Filthaut@science.ru.nl
office: HG02.829, telephone: 024-3652308

Contents

Chapter 1

Introduction

This introductory chapter covers several topics that, a priori*, are not the main focus of the course, but that are nevertheless necessary for a proper understanding of the remaining material. Besides introducing the cross sections and decay widths (and their differential versions) which are are the main observables that are accessible experimentally, it recapitulates relevant elements of special relativity and relativistic kinematics (and in particular the Lorentz-covariant notation).*

1.1 Purpose of this course

Our present understanding of elementary particles and their interactions is best expressed in terms of the so-called *Standard Model* of particle physics. It deals with the particles mentioned below:

left-handed quarks:	$\begin{pmatrix} \mathrm{u} \\ \mathrm{d} \end{pmatrix}_{\mathrm{L}}$	$\begin{pmatrix} \mathrm{c} \\ \mathrm{s} \end{pmatrix}_{\mathrm{L}}$	$\begin{pmatrix} \mathrm{t} \\ \mathrm{b} \end{pmatrix}_{\mathrm{L}}$
right-handed quarks:	u_{R} d_{R}	c_{R} s_{R}	t_{R} b_{R}
left-handed leptons:	$\begin{pmatrix} \nu_{\mathrm{e}} \\ \mathrm{e}^- \end{pmatrix}_{\mathrm{L}}$	$\begin{pmatrix} \nu_{\mu} \\ \mu^- \end{pmatrix}_{\mathrm{L}}$	$\begin{pmatrix} \nu_{\tau} \\ \tau^- \end{pmatrix}_{\mathrm{L}}$
right-handed leptons:	$\nu_{\mathrm{e,R}}(?)$ $\mathrm{e}^-_{\mathrm{R}}$	$\nu_{\mu,\mathrm{R}}(?)$ μ^-_{R}	$\nu_{\tau,\mathrm{R}}(?)$ τ^-_{R}

Table 1.1: Elementary fermions in the Standard Model. Each quark flavour occurs in three "colours", not indicated here but generally denoted as *red*, *green*, and *blue*.

One of the main goals of this lecture course is to elucidate the structure encountered in Table 1.1 (*why* are the fermions ordered in doublets, and why the families? Are these three families all there are? What does the *handedness* mean?). Besides the structure, the *interactions* between

EM interaction:	γ
Weak interaction:	$W^{\pm}$, Z
Strong interaction:	$g_1 \ldots g_8$

Table 1.2: Interactions described by the Standard Model, and corresponding gauge bosons.

these particles will be discussed in detail. (Don't worry if at this point you don't know what all of this means: it should become clear during the course!)

The course will follow (more or less) the three interactions that are described by the Standard Model: the electromagnetic, the strong, and the weak interaction. The approach is a roughly historical one.

1.2 Observables

1.2.1 Particle Physics or High-Energy Physics?

The birth of the field of subatomic physics is most properly marked by Rutherford's famous experiments in 1908, involving α particles (produced by radioactive decays, already known at the time even if not understood in detail) scattered off a gold foil. These α particles (now known to be helium nuclei) had energies of $\mathcal{O}(1\ \mathrm{MeV})$. Correspondingly, their *de Broglie wavelength* $\lambda = h/p$ was sufficiently small to probe length scales much smaller than the 1 Å typical for atomic dimensions. This led to the discovery of the *nucleus*, with dimensions of the order of several femtometers (1 fm = 10^{-15} m).

As far as we can tell at present, the particles listed above are indeed point particles (and the Standard Model considers them as such). To verify this hypothesis to smaller and smaller distances, obviously, higher and higher momenta (and hence energies) have to be used. Thus the name high-energy physics or HEP.

Another reason for high energies is the equivalence of mass and energy, as embedded in Einstein's theory of Relativity. It is this equivalence that allows us to accelerate relatively low-mass particles like electrons and protons, and by making them collide with other particles create much heavier particles. The highest energies are therefore required to gain access to previously unexplored regions of particle physics.

1.2.2 Scattering Experiments

These high energies generally imply (ultra-)relativistic particles, and therefore very short interaction times. Except in very dense media (*e.g.* neutron stars, and to some extent in stars like our Sun), one cannot hope to confine more than two particles to a volume smaller than the size of a nucleus and within the corresponding timespan of $\sim 10^{-24}$ s. Therefore, the only scattering processes that are of general relevance to this field are *two-body scattering processes*.

The most important observable associated with such scattering processes is the (total or differential) *cross section*. This concept is well known from scattering theory in (non-relativistic) quantum mechanics, and in general, we will use the same approach here: we consider interactions that at large distances are sufficiently weak that long before and after the collision, the particles can effectively be regarded as *free*. This implies that a treatment in terms of plane-wave states is appropriate (obviously, the story is more complicated when bound states are involved).

Nevertheless, our treatment differs in several some respects. The first one follows from our venture into the relativistic domain: we are *not* anymore dealing with elastic scattering processes anymore, but instead need to consider *inelastic* processes as well. Such inelastic collisions involve a conversion between mass and (kinetic) energy, so that a proper treatment using relativistic kinematics is required!

The second is that we will not in general be interested in the behaviour of a particle in the presence of a macroscopic potential. Rather, we will be considering *particle collisions*, implying that the potential of interest in fact originates from another microscopic object (*i.e.*, the other particle). An important consequence is that it is no longer the (magnitude of the) momentum $|\vec{p}|$ and kinetic energy E_{kin} of a single particle that can be considered to be conserved. Instead, this is in fact another reason for the requirement of a proper treatment using relativistic kinematics.

Since this course deals with interactions of subatomic particles, it is perhaps not surprising that the cross sections pertaining to representative interactions are generally small. It is for this reason that they are typically expressed in so-called *barns*: $1\text{b} = 10^{-28}\text{m}^2$. This is well matched to *e.g.* the size of the atomic nucleus.

Differential cross section

The concept of a differential cross section is made most intuitive by considering a *classical* scattering process: that of point particles scattering elastically off hard spheres. This situation is displayed in Fig. 1.1. In this classical situation, it is "obvious" that the total scattering cross

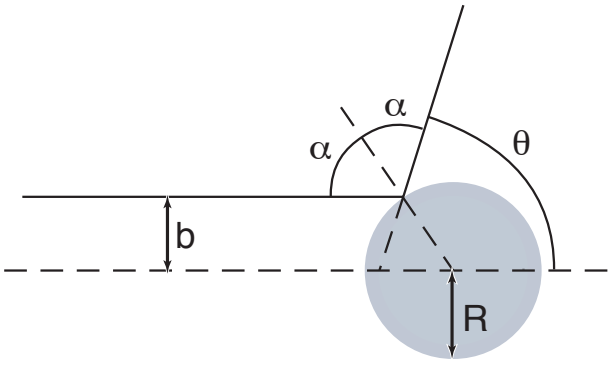

Figure 1.1: Elastic scattering of point a particle off a hard sphere.

section is πR^2: it is only for $b < R$ that any scattering occurs. But besides the total cross section, this situation is amenable to more detailed analysis. The elastic nature of the scattering process

relates the *scattering angle* θ to the angle of incidence on the sphere α by

$$\theta = \pi - 2\alpha.$$

In addition we find that the *impact parameter* of the collision is related to α by

$$b = R\sin\alpha;$$

Combining the above two equations yields

$$b = R\sin(\frac{1}{2}(\pi - \theta)) = R\cos(\theta/2)$$

The total cross section of the sphere, as seen by incident point particles, can now be seen as resulting from the contributions from different impact parameters (and different azimuthal angles: we are considering a 3D problem!),

$$\mathrm{d}\sigma = b\mathrm{d}b\mathrm{d}\phi.$$

We now rewrite this in terms of the contributions to the total scattering cross section from different scattering angles:

$$\begin{aligned} \mathrm{d}\sigma &= b\frac{\mathrm{d}b}{\mathrm{d}\theta}\mathrm{d}\theta\mathrm{d}\phi = \frac{1}{2}R^2\cos(\theta/2)\sin(\theta/2)\mathrm{d}\theta\mathrm{d}\phi \\ &= \frac{1}{4}R^2\sin\theta\mathrm{d}\theta\mathrm{d}\phi = \frac{1}{4}R^2\mathrm{d}\cos\theta\mathrm{d}\phi = \frac{1}{4}R^2\mathrm{d}\Omega. \end{aligned}$$

(Here, in the second step we have tacitly omitted the minus sign and simultaneously reversed the θ integration boundaries; the last step makes use of the fact that the solid angle $\mathrm{d}\Omega$ subtended by an infinitesimal cone at angles (θ,ϕ) is just $\mathrm{d}\Omega = \sin\theta\mathrm{d}\theta\mathrm{d}\phi$.) So in this case we find that

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = R^2/4 \tag{1.1}$$

(and it is easy to verify that integrated, this yields the total cross section we started out with). So the differential cross section, in this case, identifies the contributions to the total cross section from scattering processes at different angles. Also in a proper relativistic context, in the simplest case (final states featuring two particles only), the scattering angles θ and ϕ exhaust the kinematic degrees of freedom. In final states containing more particles, the situation becomes more complex as the magnitudes of particle momenta in general also vary (but we will not be doing detailed computations for such cases anyway).

(Note that cross sections for specific interactions may be *infinite*! A good example of this is given by elastic scattering processes of charged particles as described by Quantum Electrodynamics, which will be discussed in some detail in Chapter 2.)

Luminosity

The *measurent* of (total or differential) cross sections typically consists of aiming a beam of particles at a target (or two beams of particles at each other). The number and distribution of particles in such a beam will both affect the number of interactions that are observed; therefore this has to be dealt with.

The above case of point particles scattering off hard spheres is again a good example to explain this with. Let us assume first that a beam consists of some number of point particles N_1, randomly distributed over the beam's cross-sectional area A, and that furthermore $A \gg \sigma$. In that case, the number of interactions is simply given by the fraction of particles that happen to intersect the sphere:

$$N_{\text{int}} = N_1 \frac{\sigma}{A}.$$

The second step consists of assuming that distributed over the area A there is not just one sphere, but some number N_2. We then have

$$N_{\text{int}} = N_1 N_2 \frac{\sigma}{A} = \frac{N_1 N_2}{A} \sigma,$$

where the trivial reordering of the last step serves to make a more explicit separation in the variables describing the experimental details on the one hand, and the process-specific variable σ on the other hand. (Note that the above assumes implicitly that the number of spheres is still sufficiently small as not to lead to multiple scattering of individual point particles – this would distort the measurements!)

The final step follows from the notion that not all scatters (necessarily) take place all at once. Instead, they may occur over some time interval T, and in that case the total number of interactions is usually written as

$$N_{\text{int}} = \sigma \int_{t_0}^{t_0+T} \mathscr{L}(t) \mathrm{d}t, \tag{1.2}$$

where the quantity $\mathscr{L}(t)$ collects the experimental details at a given time t. It is called the (instantaneous) *luminosity*; the time integral in Eqn. 1.2 is usually denoted as the *integrated luminosity*. (Note that in the above, the total cross section σ may be replaced everywhere with the differential cross section for some process: the formalism is exactly the same.)

1.2.3 Particle Decays

As seen in the beginning of this chapter, our present understanding is that there are three generations of matter particles (spin-1/2 fermions). However, in practice we "observe" only the fermions of the first generation: the electron, the up and down quarks (the latter being bound in protons and neutrons, a topic we will get to later), and – with a lot of effort – the (electron?) neutrino. The reason for this is simple: they are the *lightest* existing fermions. All other fermions are sufficiently heavy that they are *unstable*: they decay. (Feynman's adagio that in physics, "everything that is not forbidden is mandatory" applies here!)

Total decay width

Consider a collection of (identical) unstable particles. The basic thesis is that the probability for one particle to decay in a short time dt does not depend either on the presence of other particles, or on the time the particle has lived so far. This means that in this short time dt, the number of particles $N(t)$ is expected to change as

$$\mathrm{d}N(t) = -\Gamma N(t)\mathrm{d}t,$$

where Γ is a constant of proportionality. This differential equation is easily solved to yield

$$N(t) = N(0)e^{-\Gamma t}. \tag{1.3}$$

This statement about the behaviour of an *ensemble* of particles can also be recast as the probability that any *individual* particle still exists after a time t. This exponential behaviour is entirely general.

It is easily verified that Eqn. 1.3 leads to an *average* lifetime $\langle t \rangle = 1/\Gamma$. In the following, this average lifetime will generally be denoted as τ. (It is much more relevant than the lifetime of individual particles, by the stochastic nature of the decay process.)

The quantity Γ is called the particle's *total decay width*. It has yet another important meaning: it turns out that the mass of individual unstable particles is not exactly the same (*i.e.*, a delta function) but instead describes a resonance curve of finite width. This width is exactly Γ.

Partial and differential decay width

But one can go further: by Feynman's adagio, it may well be possible for a particle to decay in more than one way. For instance, the Λ particle (which we will encounter later) has two decay modes:

$$\begin{aligned} \Lambda &\rightarrow \mathrm{p} + \pi^- \\ \Lambda &\rightarrow \mathrm{n} + \pi^0 \end{aligned}$$

The total decay width is then simply the *sum* of the corresponding partial decay widths:

$$\Gamma \equiv \Gamma_{\mathrm{tot}} = \Gamma(\Lambda \rightarrow \mathrm{p} + \pi^-) + \Gamma(\Lambda \rightarrow \mathrm{n} + \pi^0).$$

The decay time distribution is still given by Eqn. 1.3, with Γ still equal to the total decay width. The only refinement is that the Λ particle of this example, when it decays, will do so to the individual final states with *branching fractions* according to the partial decay widths: $B_i \equiv \Gamma_i/\Gamma_{\mathrm{tot}}$ (so the branching fractions necessarily sum up to unity). The relevance of this refinement is that knowing a particle's partial decay widths may teach us something about the physics of the decay.

The final refinement is that even the partial decay width itself is, in principle, obtained as an integral over decay variable distributions. In the case of a spin-0 particle decaying to two other particles, there is no preferential direction and the outgoing particles are monochromatic, so this isn't very interesting. However, particles of nonzero spin (like the Λ above, which has spin 1/2)

may be polarized, and then the decay angle with respect to the polarization axis may again teach us something about the physics involved in the decay. In that case, we have

$$\Gamma(\Lambda \to \mathrm{p} + \pi^-) = \int \mathrm{d}\Omega \frac{\mathrm{d}\Gamma(\Lambda \to \mathrm{p} + \pi^-)}{\mathrm{d}\Omega}.$$

1.2.4 Masses

In the case of the fundamental fermions of Table 1.1, their masses cannot be computed from first principles (at least not without a deeper understanding than the Standard Model), so knowing these may not *a priori* of great interest. But most cross sections and decay widths will in fact depend on the masses of the particles involved, and they may therefore be of indirect importance. In addition, quarks and/or antiquarks are usually bound in *hadrons* like the nucleons, and knowing the masses of these compound particles allows us to learn something about the interaction between their constituents.

1.3 Relativistic Kinematics

The practical use of relativistic kinematics will turn out to be important in all aspects of experimental particle physics. After a brief discussion of the covariant notation commonly used, Lorentz transformations, some common Lorentz frames, and some often-used relativistic invariants will be discussed in some detail. The same material is covered somewhat more extensively in Chapter 3 of the book by Griffiths.

1.3.1 Notation

We will be making extensive use of the so-called *covariant notation*. In this notation, we write *e.g.* the space-time coordinates as coordinates of a 4-dimensional vector called *Lorentz vector* or *four-vector*:

$$x^\mu = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

Note the μ being written as a superscript. This denotes a *contravariant* vector, as opposed to a *covariant* one (which we will construct shortly). The "length" squared of this Lorentz vector is now given by

$$x^2 \equiv \Sigma_{\mu,\nu} g_{\mu\nu} x^\mu x^\nu = (ct)^2 - x^2 - y^2 - z^2.$$

What we have implicitly defined here is the so-called *metric tensor* $g_{\mu\nu}$, which in special relativity always takes the form (if one insists on considering it as a matrix – some caveats will be

discussed below)

$$g_{\mu\nu} = \mathrm{diag}(1,-1,-1,-1) \equiv \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

(the blanks denote 0 elements). The indices here are denoted by subscripts, indicating that the metric tensor is a *covariant* tensor. A further notational issue is the *Einstein summation convention*, which states that the explicit summation signs over contra- and covariant indices (also called *contraction* of these indices) can be dropped. So the equation above could also be written as

$$x^2 = g_{\mu\nu} x^\mu x^\nu.$$

A contravariant vector can be transformed into a covariant one by contracting it with the metric tensor:

$$x_\mu = x^\nu g_{\mu\nu},$$

and similarly, a covariant vector can be transformed into a covariant vector:

$$x^\mu = x_\nu g^{\mu\nu}.$$

The contravariant tensor $g^{\mu\nu}$ has the same form as its covariant relative (it's left as an exercise to show this). Our vector's length squared can now also be written as

$$x^2 = x^\mu x_\mu.$$

On a practical note, it is hopefully obvious from the above discussion that covariant and contravariant indices must not be confused. Nor is it very helpful to think as a contravariant vector as being a column vector, and a covariant vector as being a row vector: the two forms of the metric tensor above are not consistent with that picture. Also, it is conceivable to think of tensors that have multiple contravariant and/or covariant indices (the metric tensor is one such case; see below for another example).

Finally, while the above discussion has dealt with a *position* four-vector only, it is important to realise that relativistic kinematics is by no means restricted to working with coordinates only. In fact, much more used in the context of scattering processes are *four-momenta*. These are often written as p^μ (or something similar) and are defined as

$$p^\mu = \begin{pmatrix} E/c \\ p_x \\ p_y \\ p_z \end{pmatrix}.$$

By construction, $p_\mu p^\mu$ is a Lorentz scalar, and it is easy to see that it is in fact equal to m^2c^2, just from the familiar equation

$$E^2 = \vec{p}^2 c^2 + m^2 c^4.$$

Concerning the mass occurring in this formula, it is perhaps useful to point out that we will always be referring to *rest mass* and never to *relativistic mass*: the latter concept is really useful only to yield formulas taking the same shape as in non-relativistic mechanics; and there are limitations to even that application.

1.3.2 Lorentz Transformations

The basis of special relativity is that quantities such as the above length are invariant under Lorentz transformations. To make this explicit, such quantities are called *Lorentz scalars*, and they do not carry any Lorentz indices. This in contrast to *e.g.* the above Lorentz vector, which transforms as

$$x^\mu \to x^{\mu\prime} = \Lambda^\mu{}_\nu x^\nu.$$

For definiteness, a "standard" choice for the transformation Λ is a boost $-\beta$ along the z axis (*i.e.* in the boosted system, a particle that moves with a velocity βc along the z axis in the lab frame is at rest):

$$\Lambda^\mu{}_\nu = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \quad \text{with} \quad \gamma = \frac{1}{\sqrt{1-\beta^2}}.$$

One benefit of the covariant notation is immediately obvious: just from the indices (which are denoted using Greek characters) it is obvious how a quantity transforms under Lorentz transformations. Similarly, for a tensor quantity we would have

$$A^{\mu\nu} \to A^{\mu\nu\prime} = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma A^{\rho\sigma}.$$

For a covariant vector, we would have

$$x_\mu \to x'_\mu = \Lambda_\mu{}^\nu x_\nu.$$

Note the placement of the contra- and covariant indices: this matrix is *different* from the $\Lambda^\mu{}_\nu$ encountered above. In fact, it is easy (and left as an exercise) to show that

$$\Lambda_\mu{}^\nu = g_{\mu\rho}\Lambda^\rho{}_\sigma g^{\sigma\nu}$$

(and this is also consistent with the conversion from covariant indices to contravariant ones, and vice versa).

It is noteworthy that the metric tensor (let's use its contravariant form $g^{\mu\nu}$ here) is itself a contravariant tensor: this implies that it should follow the appropriate transformation rules under Lorentz transformations. However, we also write the metric tensor as a constant! This paradox is resolved by turning it into an equation:

$$g^{\mu\nu} = \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma g^{\rho\sigma}.$$

This condition has to hold for any Lorentz transformation. In other words, it *defines* the possible Lorentz transformations. Or in group theoretical terms: it defines the *Lorentz group*. Note that,

although the above has dealt with boosts (*i.e.* transformations involving both spatial and time coordinates) exclusively, the Lorentz group also comprises the usual rotations (which involve only spatial coordinates).

1.3.3 Collisions

The scattering of point particles off hard spheres discussed in Sect. 1.2.2 is not quite representative for the situations considered in following chapters: the hard sphere is a *macroscopic* body, while in the following we will invariably be concerned with collisions between microscopic particles. This has one important consequence for the *kinematics* of scattering processes: rather than merely conserving energy, in particle collisions *the total four-momentum is conserved*.

When considering collision processes, two specific Lorentz frames are often used:

- the *laboratory frame*, representing the collisions as they would be observed "in the lab";
- given that the lab frame depends (by definition) on experimental circumstances, the *centre-of-mass* or CM frame is often used instead. In this frame, the total momentum (but not energy) is zero.

1.4 Natural Units

It will be convenient for many computations in this course to use so-called *natural units*, *i.e.*, setting $\hbar = c = 1$. (In more detail, we will be using so-called *Gaussian* units, following *e.g.* the book by Griffiths. This explains the absence of factors 4π *e.g.* in Sect. 2.2.2.)

Chapter 2

QED: Quantum Electrodynamics

The purpose of this chapter is twofold. First, it explains the conceptual issues in "marrying" special relativity and quantum mechanics, leading to the important conclusion that every particle has a corresponding anti-particle, and that an entirely appropriate treatment needs Quantum Field Theory (but the aim is not to cover this latter aspect in any depth; its most basic notion will suffice for our purposes).

Second, it introduces the topic of symmetries, and notably gauge symmetries*; this is a very important topic that will return also in later chapters (albeit in somewhat different forms). A hopefully fairly intuitive introduction of this concept is provided, taking electrodynamics as a starting point. These two purposes "meet" in the construction of Quantum Electrodynamics, which is discussed both for the case of scalar (spin-0) particles and for the more realistic case of spin-1/2 particles. Along the way, the concepts of Feynman diagrams and the associated Feynman rules are also introduced.*

2.1 Negative-Energy States: Antiparticles

2.1.1 Setting the Stage: Non-Relativistic Quantum Mechanics

In non-relativistic Quantum Mechanics, it was seen that (in the "standard" position representation) essentially everything can be derived by the substitution

$$E \rightarrow i\frac{\partial}{\partial t} \tag{2.1}$$

$$\vec{p} \rightarrow -i\vec{\nabla} \tag{2.2}$$

(remember that we have set $\hbar = 1$). This substitution directly converts the classical Hamiltonian

$$H = \frac{\vec{p}^2}{2m} + V(\vec{x})$$

into the Schrödinger equation

$$i\frac{\partial}{\partial t}\psi(\vec{x}) = \left(-\frac{\vec{\nabla}^2}{2m} + V(x)\right)\psi(\vec{x})$$

acting on the wave function $\psi(\vec{x})$.

Once we have found a wave function $\psi(\vec{x})$ satisfying the Schrödinger equation, we can also take the complex conjugate expression:

$$-i\frac{\partial}{\partial t}\psi^*(\vec{x}) = \left(-\frac{\vec{\nabla}^2}{2m} + V(x)\right)\psi^*(\vec{x})$$

We then multiply the original Schrödinger equation by $\psi^*(\vec{x})$, and its conjugate by $\psi(\vec{x})$. Subtracting the two yields

$$i\left(\psi^*(\vec{x})\frac{\partial}{\partial t}\psi(\vec{x}) + \psi(\vec{x})\frac{\partial}{\partial t}\psi^*\right) = -\frac{1}{2m}\left(\psi^*(\vec{x})\vec{\nabla}^2\psi(\vec{x}) - \psi(\vec{x})\vec{\nabla}^2\psi^*(\vec{x})\right).$$

It is easily seen that this can be written alternatively as

$$i\frac{\partial}{\partial t}|\psi(\vec{x})|^2 = -\frac{1}{2m}\vec{\nabla}\cdot\left(\psi^*(\vec{x})\vec{\nabla}\psi(\vec{x}) - \psi(\vec{x})\vec{\nabla}\psi^*(\vec{x})\right).$$

Thus, this leads us to the *continuity equation*

$$\frac{\partial}{\partial t}\rho(\vec{x}) + \vec{\nabla}\cdot\vec{j}(\vec{x}) = 0,$$

with

$$\begin{aligned}\rho(\vec{x}) &= |\psi(\vec{x})|^2 \quad \text{and} \\ \vec{j}(\vec{x}) &= \frac{-i}{2m}\left(\psi^*(\vec{x})\vec{\nabla}\psi(\vec{x}) - \psi(\vec{x})\vec{\nabla}\psi^*(\vec{x})\right)\end{aligned}$$

The quantity $\rho(\vec{x})$ occurring in this equation is *positive definite*, making the interpretation of $|\psi(\vec{x})|^2$ as the probability density of finding a particle at the position $\vec{x}$ a proper one.

2.1.2 Translation to the Relativistic Case

The approach in the case of *relativistic* Quantum Mechanics is exactly the same; however, this time it must be applied to the "Hamiltonian" of special relativity. Restricting ourselves to free particles, $V(x) = 0$, the basic classical equation is then

$$p_\mu p^\mu = m^2 \quad \text{or} \quad E^2 = \vec{p}^2 + m^2. \tag{2.3}$$

When we again make the substitutions of Eqn. 2.2, and make the resulting equation act on a wave function $\phi(x)$ (this notation combines the spatial and temporal dependence), the result is

$$\left(\frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 + m^2\right)\phi(x) = 0,$$

or, in explicitly covariant form:

$$(\partial_\mu \partial^\mu + m^2)\phi(x) \equiv (\Box + m^2)\phi(x) = 0. \tag{2.4}$$

This is the *Klein-Gordon equation*.

Unsurprisingly, for our case of free particles, this equation is easily solved to yield plane waves just like in the non-relativistic case:

$$\phi(x) = Ne^{-ip\cdot x} = Ne^{-i(Et - \vec{p}\cdot\vec{x})}, \tag{2.5}$$

with N an *a priori* arbitrary normalization constant, and the four-momentum components E and $\vec{p}$ satisfying our original classical Eqn. 2.3.

But here we are in trouble! For the solution to Eqn. 2.3 is

$$E = \pm\sqrt{\vec{p}^2 + m^2}.$$

While the solution with the $+$ sign gives us a "standard" picture, the solution with the $-$ sign cannot be ignored. As a consequence, the system has no ground state (it is unbounded from below), and hence no meaningful physical interpretation seems possible.

To make things worse, also the continuity equation becomes problematic. As in the non-relativistic case, it is obtained by taking also the complex conjugate of the Klein-Gordon equation and multiplying it with $\phi(x)$, and combining it with the original equation multiplied with $\phi^*(x)$. However, due to the fact that the Klein-Gordon equation involves a second order rather than a first order time derivative, this time we have to *subtract* the two. The result is

$$\frac{\partial}{\partial t}\left(i(\phi^*(x)\frac{\partial\phi(x)}{\partial t} - \phi(x)\frac{\partial\phi^*(x)}{\partial t})\right) + \vec{\nabla}\cdot\left(-i(\phi^*(x)\vec{\nabla}\phi(x) - \phi(x)\vec{\nabla}\phi^*(x))\right) = 0,$$

which can again be considered as a continuity equation, but with

$$\begin{aligned}\rho(x) &= i(\phi^*(x)\frac{\partial\phi(x)}{\partial t} - \phi(x)\frac{\partial\phi^*(x)}{\partial t}),\\ \vec{j}(x) &= -i(\phi^*(x)\vec{\nabla}\phi(x) - \phi(x)\vec{\nabla}\phi^*(x)).\end{aligned} \tag{2.6}$$

This can again be cast into explicitly Lorentz-covariant form:

$$\partial_\mu j^\mu(x) = 0, \quad \text{with} \quad j^\mu(x) = i(\phi^*(x)\partial^\mu\phi(x) - \phi(x)\partial^\mu\phi^*(x)). \tag{2.7}$$

When we now substitute the free-particle solution of Eqn. 2.5 in Eqn. 2.6, we find that

$$j^\mu(x) = 2p^\mu|N|^2. \tag{2.8}$$

In particular, we have $\rho(x) = 2E|N|^2$. So in the case of a negative-energy solution, we also find that $\rho(x)$ becomes negative, *i.e.*, it can no longer be interpreted as a probability density.

Finally, there is another problem with negative-energy solutions. Consider some localized spatial wavefunction at some time t. It is then straightforward to determine its Fourier spectrum, and in general it will be seen that this will contain both positive- and negative-energy components, which will have opposite time evolutions. Constructing the norm of the wavefunction would then contain oscillating terms; the corresponding "zitterbewegung" is not observed in reality. (This issue is discussed in more detail in Section 2-2-2 of Ref. [1].)

2.1.3 Field-theoretical Interpretation

A proper interpretation can only be given in the context of Quantum Field Theory. In that context, ϕ is a *field* rather than a wave function, and its plane-wave expansion leads to particle creation operators for the positive energies combined with antiparticle annihilation operators for the negative energies:

$$\phi(x) = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} \left(a(\vec{p}) e^{-ip\cdot x} + b^\dagger(\vec{p}) e^{ip\cdot x} \right)$$

The case of a real-valued classical field (we will discuss classical fields in a bit more detail in Section 2.1.4) then translates into a hermitian quantum field, *i.e.*, with $b(\vec{p}) = a(\vec{p})$. So in terms of plane-wave solutions, the action of the field is either to *create* a particle with four-momentum p^μ or to annihilate one with four-momentum p^μ (where it can be shown, although doing so is outside the scope of this course, that the positive-energy solution is associated with the *annihilation* of a particle, while the negative-energy solution is associated with the *creation* of a particle). Note that the four-momentum here is in both cases the *physical* (positive-energy) four-momentum.

However, here we have to watch the other desired properties of this field. Consider the case of a field representing a *charged* particle like the electron. The action of the field must be to change the charge by the *same* one unit, irrespective of whether creation or annihilation is concerned (motivating this requirement further would bring us too far in the realm of Quantum Field Theory; a construction can *e.g.* be found in the book by Peskin and Schroeder [2]). This means that in this case, the equality $b(\vec{p}) = a(\vec{p})$ cannot hold anymore (and hence also that ϕ cannot be a hermitian field). This can be achieved by making $b^\dagger(p)$ represent the creation of an *anti*-particle with four-momentum p^μ; we should therefore expect the existence of a particle much like the electron, but with opposite properties such of its charge. This particle is called the positron.

Of course, one would hope for experimental evidence of the existence of the positron. It was first observed in 1932, in a cloud chamber exposed to cosmic rays (see Fig. 2.1). Its discovery earned Anderson [3] the 1936 Nobel Prize. (The discovery followed the prediction of the positron by Dirac by only a year. Dirac used a different interpretation of negative-energy states, though, which is not appropriate for the description of bosons.)

So what about the continuity equation, and the fact that there doesn't appear to be a conserved quantity (*i.e.*, one occurring in a continuity equation) that can be associated with a probability

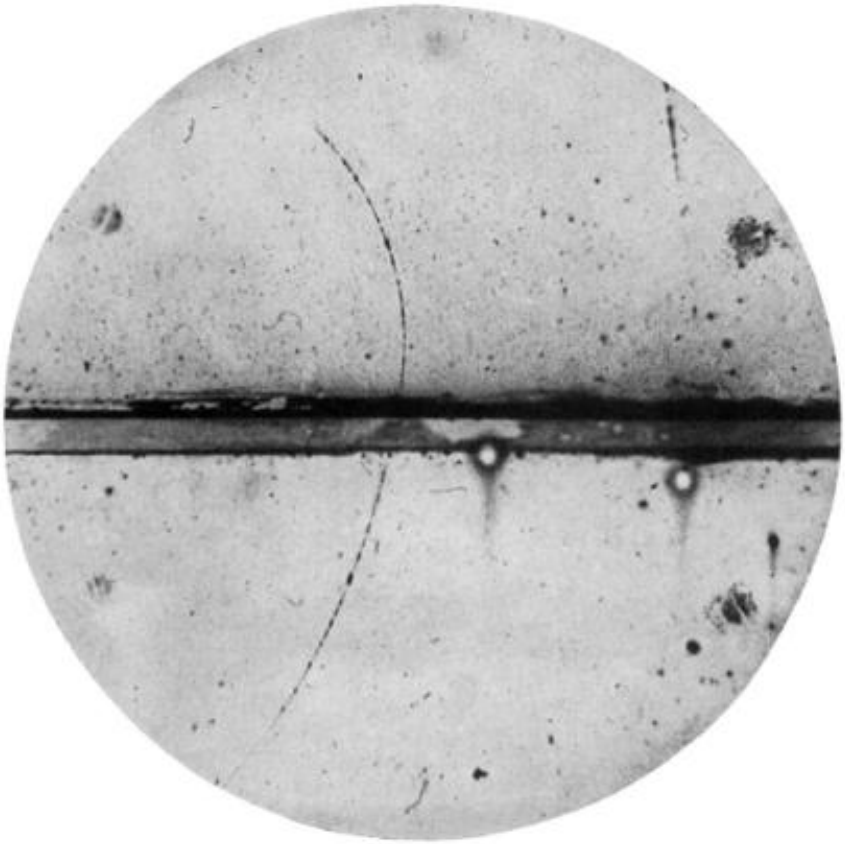

FIG. 1. A 63 million volt positron ($H\rho = 2.1 \times 10^5$ gauss-cm) passing through a 6 mm lead plate and emerging as a 23 million volt positron ($H\rho = 7.5 \times 10^5$ gauss-cm). The length of this latter path is at least ten times greater than the possible length of a proton path of this curvature.

Figure 2.1: Photograph made of a positron bent in a magnetic field and traversing (and losing energy in) a Pb plate. The positron hypothesis follows from (1) the *sign* of the curvature, indicating a positively charged particle; and (2) the track length after having traversed the plate and before being stopped, indicating a particle much lighter than a proton.

density? The fact of the matter is that the (conserved) probability density is a concept that is useful in non-relativistic quantum mechanics (non-conservation would correspond to the creation or disappearance of particles). However, in a relativistic context, *it is perfectly acceptable for (anti-)particles to be created or annihilated* (and the operator nature of quantum fields allows to describe such processes). So it doesn't make sense to ask for a conserved probability density.

2.1.4 Principle of Least Action and Euler-Lagrange Equations; Noether Theorem

Accepting that we need a field-theoretical interpretation (per Sect. 2.1.3), we can now also use a different starting point for our computations than the Klein-Gordon equation. Going back to a classical single-particle system of a single degree of freedom $q(t)$, we can express the *action* S

as $S = \int_{t_0}^{t_1} \mathrm{d}t L(q,\dot{q})$, where L represents the Lagrangian. Demanding that S be stationary under arbitrary but small changes of $q(t)$ at each t results in the requirement

$$\delta S = \int_{t_0}^{t_1} \mathrm{d}t \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q} \right) = 0.$$

Interchanging the order of the time derivative and the δ operation and carrying out an integration by parts then results in the condition

$$\int_{t_0}^{t_1} \mathrm{d}t \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right) \right) \delta q = 0.$$

If this equality is to hold for arbitrary $\delta q(t)$, then we immediately arrive at the *Euler-Lagrange* equation

$$\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}}\right) = 0.$$

In a field-theoretical setting, things work in much the same way. The essential difference is that the Lagrangian L is obtained as the spatial integral of the *Lagrange density* $\mathscr{L}(\phi(x), \partial_\mu \phi(x))$, where $\partial_\mu \phi \equiv \frac{\partial}{\partial(\frac{\partial \phi}{\partial x^\mu})}$ refers to the time as well as spatial derivatives of ϕ. The action therefore becomes a four-dimensional integral – convenient since this allows us to express it in a covariant form. The arbitrary changes are then in the field $\phi(x)$, and the principle of least action becomes

$$\delta S = \int \mathrm{d}^4 x \left(\frac{\partial \mathscr{L}}{\partial \phi}\delta\phi(x) + \frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}\delta\partial_\mu \phi(x) \right) = 0. \tag{2.9}$$

The same manipulations as for the above single degree of freedom then lead to the Euler-Lagrange equation for the field:

$$\frac{\partial \mathscr{L}}{\partial \phi} - \partial_\mu \left(\frac{\partial \mathscr{L}}{\partial_\mu \phi}\right) = 0. \tag{2.10}$$

We will make use of this equation, as well as of properties of the Lagrange density, later in this and in other chapters. Note that in these lecture notes we will follow the common particle physicists' sloppiness and simply call $\mathscr{L}$ the Lagrangian. For now, suffice it to say that the Klein-Gordon equation can be recovered from the following choice of Lagrangian:

$$\mathscr{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{1}{2}m^2\phi^2. \tag{2.11}$$

The *Noether theorem* is related to so-called *internal symmetries*, which we will cover later in more detail, but which for now we can illustrate using the relativistic wavefunction ϕ of Section 2.1.2, which we subsequently concluded should really be treated as a quantum field. In the wavefunction picture, Quantum Mechanics dictates that the physics should not depend on any complex phase of ϕ. Now in the field theoretical context, it is quite well possible to posit a real scalar field; however as an alternative we can posit a *complex* scalar field ϕ, and still make the *assumption* that the physics described by the Langrangian indeed does not depend on the phase of ϕ. This is arguably the simplest example of an internal symmetry.

Under an infinitesimal phase change, which we will describe more generally as a *group transformation* (see Appendix B for more details), we can then write the transformation of the field ϕ as

$$\phi \to \phi' = \phi + i\alpha T\phi,$$

where α is the infinitesimal phase change, and T is the generator of the group transformation. In the case of phase changes, we know the transformation properties:

$$\phi \to \phi' = e^{i\alpha}\phi = \phi + i\alpha\phi, \tag{2.12}$$

so we simply have $T = 1$. We now require again that the action be invariant under this transformation, so we obtain the condition

$$\begin{aligned}
\delta S &= i\int \mathrm{d}^4x \left(\frac{\partial \mathscr{L}}{\partial \phi}\alpha T\phi + \frac{\partial \mathscr{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\alpha T\phi)\right) \\
&= i\int \mathrm{d}^4x \left(\partial_\mu\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu\phi)}\right)\alpha T\phi + \frac{\partial \mathscr{L}}{\partial(\partial_\mu\phi)}\partial_\mu(\alpha T\phi)\right) \\
&= i\int \mathrm{d}^4x \partial_\mu\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu\phi)}\alpha T\phi\right) = 0.
\end{aligned}$$

Requiring that this equality hold for any α and integration boundaries, we find that

$$\partial_\mu j^\mu = 0, \quad \text{with} \quad j^\mu = i\left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu\phi)}T\phi\right).$$

This is the essence of the Noether theorem: *every symmetry brings with it a conserved quantity*.

2.2 Perturbation Theory and Electromagnetic Interactions

2.2.1 Perturbation Theory

A theory describing only free particles is not terribly exciting... therefore, let us see how interactions can be incorporated. The aim here is not to be entirely rigourous, but rather to provide a heuristic introduction to the computation of scattering amplitudes that can be understood as a reasonably straightforward extension of (time-dependent) non-relativistic perturbation theory.

Suppose that the Hamiltonian of a system is described by

$$H = H_0 + V(\vec{x},t)$$

and that the system corresponding to the unperturbed Hamiltonian H_0 can be solved exactly,

$$H_0\phi_n = E_n\phi_n \quad \text{with} \quad \int \mathrm{d}^3x \phi_n^*(\vec{x})\phi_m(\vec{x}) = \delta_{nm}.$$

(Here we are assuming that the system leads to a set of *discrete* eigenstates. That limitation does not affect the following argument.) We now want to know the time evolution of a system that at

a time t is in the state $\psi(\vec{x})$. To this end, we decompose ψ in terms of the eigenfunctions of the unperturbed Hamiltonian:

$$\psi(\vec{x},t) = \sum_n a_n(t)\phi_n(\vec{x})e^{-iE_nt}.$$

Applying the Schrödinger equation then yields

$$\begin{aligned} i\frac{\partial\psi(\vec{x},t)}{\partial t} &= \sum_n \phi_n(\vec{x})e^{-iE_nt}\left(E_n a_n(t) + \frac{\mathrm{d}a_n(t)}{\mathrm{d}t}\right) \\ &= (H_0 + V(\vec{x},t))\psi = \sum_n (H_0 + V(\vec{x},t))a_n(t)\phi_n(\vec{x})e^{-iE_nt} \\ &= \sum_n (E_n + V(\vec{x},t))a_n(t)\phi_n(\vec{x})e^{-iE_nt} \\ \Rightarrow i\sum_n \frac{\mathrm{d}a_n(t)}{\mathrm{d}t}\phi_n(\vec{x})e^{-iE_nt} &= \sum_n V(\vec{x},t)a_n(t)\phi_n(\vec{x})e^{-iE_nt}. \end{aligned} \tag{2.13}$$

Now assume that the interaction $V(\vec{x},t)$ is switched off for large times $T \to \infty$, so that the decomposition into eigenstates of the unperturbed system is the "proper" thing to do for such large times. Multiplying Eqn. 2.13 by $\phi_f^*(\vec{x})e^{iE_ft}$ and integrating the result over all space then yields

$$\begin{aligned} \frac{\mathrm{d}a_f(t)}{\mathrm{d}t} &= -i\sum_n a_n(t)e^{-i(E_n - E_f)t} \cdot V_{fn}(t), \quad \text{with} \\ V_{fn}(t) &= \int \mathrm{d}^3x \phi_f^*(\vec{x})V(\vec{x},t)\phi_n(\vec{x}) \end{aligned}$$

This is just the well-known *Dyson series* from non-relativistic Quantum Mechanics.

Also the solution of this integro-differential equation proceeds in the same way as in non-relativstic Quantum Mechanics. In addition, suppose that before the interaction is switched on the system is in an eigenstate of the unperturbed Hamiltonian, *i.e.*, $a_n(-T) = \delta_{ni}$. Order by order, we have

$$\begin{aligned} a_f(t) &= \delta_{fi} \\ &+ (-i)\int_{-T}^{t} \mathrm{d}t' V_{fi}(t')e^{-i(E_i - E_f)t'} \\ &+ (-i)^2 \sum_n \int_{-T}^{t} \mathrm{d}t' V_{fn}(t')e^{-i(E_n - E_f)t'} \\ &\quad \cdot \int_{-T}^{t'} \mathrm{d}t'' V_{ni}(t'')e^{-i(E_i - E_n)t''} \\ &+ \ldots \end{aligned}$$

At this point, we formulate the above equation in a more covariant form by setting

$$\phi_n(x) \equiv \phi_n(\vec{x})e^{-iE_nt}.$$

Retaining only the lowest-order (nontrivial) transition, we then obtain

$$\begin{aligned} a_f(t) &= -i\int_{-T}^{t} \mathrm{d}t' \int \mathrm{d}^3x \left(\phi_f(\vec{x})e^{-iE_f t}\right)^* V(\vec{x},t') \left(\phi_i(\vec{x})e^{-iE_i t}\right) \\ &= -i\int_{-T}^{t} \mathrm{d}t' \int \mathrm{d}^3x \phi_f^*(x) V(x) \phi_i(x). \end{aligned}$$

Finally, considering this quantity far after the interaction, at $t = T$, and letting $T \to \infty$, this leads to the transition amplitude

$$T_{fi} = -i\int \mathrm{d}^4x \phi_f^*(x) V(x) \phi_i(x). \tag{2.14}$$

2.2.2 Covariant Formulation of Classical Electrodynamics

Before proceeding to the implementation in Eqn. 2.14, it is useful to pay some attention to the covariant formulation of classical electrodynamics. The starting point is the Maxwell equations:

$$\vec{\nabla}\cdot\vec{E} = \rho \qquad \text{(Gauss)}, \tag{2.15}$$

$$\vec{\nabla}\times\vec{B} - \frac{\partial\vec{E}}{\partial t} = \vec{j} \qquad \text{(Ampère)}, \tag{2.16}$$

$$\vec{\nabla}\cdot\vec{B} = 0 \qquad \text{(Gauss)}, \tag{2.17}$$

$$\vec{\nabla}\times\vec{E} + \frac{\partial\vec{B}}{\partial t} = 0 \qquad \text{(Faraday)}. \tag{2.18}$$

Eqn. 2.17 indicates that $\vec{B}$ can be written as

$$\vec{B} = \vec{\nabla}\times\vec{A},$$

where $\vec{A}$ is called the *vector potential*. Combining this with Eqn. 2.18, it follows that $\vec{E}$ can be written as

$$\vec{E} = -\vec{\nabla}\Phi - \frac{\partial\vec{A}}{\partial t},$$

with Φ the scalar potential. With this notation, it then follows that Eqn. 2.16 can be written as

$$\begin{aligned} \vec{\nabla}\times\vec{B} - \frac{\partial\vec{E}}{\partial t} &= \left(-\vec{\nabla}^2\vec{A} + \vec{\nabla}\cdot(\vec{\nabla}\cdot\vec{A})\right) + \vec{\nabla}\frac{\partial\Phi}{\partial t} + \frac{\partial^2\vec{A}}{\partial t^2} \\ &= \Box\vec{A} + \vec{\nabla}(\vec{\nabla}\cdot\vec{A} + \frac{\partial\Phi}{\partial t}) = \vec{j}. \end{aligned}$$

Finally, we have

$$\vec{\nabla}\cdot\vec{E} = -\vec{\nabla}^2\Phi - \frac{\partial}{\partial t}\left(\vec{\nabla}\cdot\vec{A}\right) = \rho.$$

When we add and subtract here a term $\frac{\partial^2\Phi}{\partial t^2}$, this last equation can be rewritten as

$$\Box\Phi - \frac{\partial}{\partial t}\left(\vec{\nabla}\cdot\vec{A} + \frac{\partial\Phi}{\partial t}\right) = \rho.$$

The two rewritten inhomogeneous equations now have a very similar form; defining

$$A^\mu = (\Phi, \vec{A}) \quad \text{and} \quad j^\mu = (\rho, \vec{j})$$

allows us to finally put the inhomogeneous equations into a manifestly covariant form:

$$\Box A^\mu - \partial^\mu(\partial_\nu A^\nu) = j^\mu,$$

which can also be written as

$$\partial_\mu F^{\mu\nu} = j^\nu, \quad \text{with} \quad F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{2.19}$$

The quantity $F_{\mu\nu}$ is called the *electromagnetic field tensor*, and it turns out that its elements are just $\vec{E}$ and $\vec{B}$.

(Of course, putting these equations into a nicely covariant-looking form does not guarantee the right -known- behaviour of $\vec{E}$ and $\vec{B}$ under Lorentz transformations. But that can be verified explicitly and turns out to be in good order.)

Even this nice formula can be simplified further. The field tensor $F_{\mu\nu}$ encodes the *physical* information. Therefore, a change in A^μ

$$A^\mu \to A'^\mu = A^\mu + \partial^\mu \chi, \tag{2.20}$$

with χ an arbitrary function, does not affect the physics. This is the *gauge* freedom of electromagnetism.

As a consequence, we can choose χ such that $\partial_\nu A^\nu = 0$: this is called the *Lorenz condition* (after Ludvig Lorenz). So finally

$$\Box A^\mu = j^\mu. \tag{2.21}$$

This choice for A^μ is also called the *Lorenz gauge*. It is to be emphasised again that the choice of gauge does *not* affect the physics of the system (and other choices are indeed used, such as the *Coulomb gauge*, in which $\vec{\nabla} \cdot \vec{A} = 0$).

A last ingredient that will be extremely useful in the following is the fact that the interaction of particles of charge q with an electromagnetic field can be described simply by the *minimal substitution*[1]:

$$p^\mu \to p^\mu - qA^\mu. \tag{2.22}$$

The usefulness of this substitution is that we can use it instead of a "proper" field theoretical treatment of gauge symmetries: the so-called *covariant derivative* corresponding to the $U(1)$ symmetry group relevant for this treatment of QED yields precisely the same result.

[1]The derivation of this property is lengthy and we will not venture into it here. More details can be found *e.g.* in Jackson [4], Chapter 12. Also one of the exercises offers a partial justification.

2.2.3 The covariant derivative, and implications of U(1) symmetry

As discussed in the exercises, the use of the minimal substitution allows for a derivation of the Lorentz force in *classical* electrodynamics. If we are to extend this validity to the realm of (non-relativistic) quantum mechanics, this results in a Schrödinger equation

$$\left(\frac{1}{2m}(-i\vec{\nabla}-q\vec{A})^2+qV\right)\psi(\vec{x},t)=i\frac{\partial\psi(\vec{x},t)}{\partial t}. \tag{2.23}$$

However, the requirement that the gauge transformation of Eqn. 2.20 should not affect the *physics* (*i.e.*, should leave the *form* of Eqn. 2.23 invariant) now has a nontrivial consequence. For it can be shown that this invariance is only achieved if simultaneously with Eqn. 2.20, also the *wavefunction* transforms:

$$\psi(\vec{x},t)\rightarrow\psi'(\vec{x},t)=e^{-iq\chi(\vec{x},t)}\psi(\vec{x},t). \tag{2.24}$$

An experimental justification of this (possibly surprising) result is given by the Aharonov-Bohm effect: this transformation is closely connected to the phase factor

$$q\int_C\vec{A}(\vec{x})\cdot\mathrm{d}\vec{x}$$

that enters in the computation of the interference pattern between the two paths of a charged-particle beam in magnetic field-free regions around a (long) solenoid (where C is the contour describing the two paths); it is easy to show that this phase factor is invariant under such transformations, however, as required for physically observable effects. The Aharonov-Bohm effect clearly demonstrates that A^μ is more than a mathematical construct but rather is at least as fundamental an object as the electric and magnetic field strengths embodied in $F_{\mu\nu}$: the effect exists even though the particles never encounter a nonzero field strength.

Although the above is done in the framework of non-relativistic quantum mechanics, exactly the same conclusion (Eqn. 2.24) holds in the relativistic case. In conclusion, we end up with a space- and time-dependent phase transformation of the wavefunction, which does not affect any physics. In group theoretical terms, the $U(1)$ symmetry group can be identified exactly with all possible phase transformations – hence the statement that QED implements a $U(1)$ symmetry.

But having drawn this conclusion, matters can in fact be turned around: let us suppose that we *require* that Eqn. 2.24 does not affect any physics. Then it can be shown that the quantum mechanical analogue of Eqn. 2.22,

$$i\partial_\mu\rightarrow iD_\mu\equiv i\partial_\mu-qA_\mu, \tag{2.25}$$

precisely achieves this. The quantity D_μ is called the *covariant derivative*.

Note that this phase change looks a lot like the one encountered in Eqn. 2.12. An essential difference is that rather than merely requiring invariance under *global* phase changes, we now impose this requirement even for *local* (*i.e.*, space and time dependent) phase changes. Another important difference, although we will not prove it here, is that we now require not merely the action to be invariant under the transformation, but also the Lagrangian itself!

Of course, all of the above hinges on the known properties of QED. However, it turns out that the *gauge principle* (starting here with the assumed phase transformation property of the wavefunction – or field – and constructing the appropriate covariant derivative, which then ultimately describes the interaction of charged particles with the electromagnetic field) is very powerful. The same principle will be used later to describe the strong and weak interactions.

Finally, note that while Eqn. 2.20 does not depend on the charge q of the fermion involved, the covariant derivative and the phase transformation do. This means that we can use the same principle (and with the same electromagnetic field!) for particles of different charge. In group theoretical terms, this means that different *representations* of the underlying phase symmetry are possible. This is a fact that will be exploited later on.

2.2.4 Transition Amplitudes

We now have all the required ingredients in hand to proceed further. In the Klein-Gordon equation, we make the minimal substitution of Eqn. 2.22; the resulting equation can be recast as

$$(\Box + m^2)\psi = -V\psi, \tag{2.26}$$

with the "potential" V given by

$$V\psi = -ie(\partial_\mu A^\mu + A^\mu \partial_\mu)\psi - e^2 A^2 \psi$$

(note the operator character of the derivative: it acts on ψ as well as on A). We will neglect the last term in this equation, on account of the fact that e is small. Retaining only the first two terms, we then have

$$\begin{aligned} T_{fi} &= -i\int \mathrm{d}^4x \phi_f^*(x) V(x) \phi_i(x) \\ &= i\int \mathrm{d}^4x \phi_f^*(x)(ie)(A^\mu \partial_\mu + \partial_\mu A^\mu)\phi_i(x). \end{aligned}$$

The last term is amenable to integration by parts, and neglecting the resulting surface integral the result becomes

$$T_{fi} = -i\int \mathrm{d}^4x j^\mu_{fi}(x) A_\mu(x) \quad \text{with} \quad j^\mu_{fi}(x) = -ie\left(\phi_f^*(x)\partial^\mu \phi_i(x) - (\partial^\mu \phi_f^*(x))\phi_i(x)\right). \tag{2.27}$$

Note that the quantity $j^\mu_{fi}(x)$ looks almost exactly like the quantity $j^\mu(x)$ in Eqn. 2.7. There is however a difference in that $j^\mu_{fi}(x)$ involves two *different* wavefunctions, those of both the initial and final states. The proper interpretation of $j^\mu_{fi}(x)$ is that of the current involved in the *interaction of a microscopic particle*. This is relevant in that the absorption or emission of a photon (we'll see later that this picture is appropriate) may affect the particle noticeably.

Eqn. 2.27 is appropriate for the description of the interaction of a particle with a general electromagnetic field. However, this is not the situation typically of interest in particle physics.

Rather, our interest is in scattering particles off each other, *i.e.*, in electromagnetic fields caused by *other particles*: the field satisfies

$$\Box A^\mu = j^{\mu(2)}_{fi} \tag{2.28}$$

relating it to the current of the other particle (which we will also assume to be an electron).

We will also restrict the further discussion to plane-wave initial and final states (as appropriate for our discussion of scattering experiments where long before and after the scattering process, the participating particles can be considered as free particles). In this case, the current $j^{\mu(2)}_{fi}$ takes on the simple form

$$j^{\mu(2)}_{fi}(x) = -e|N|^2 \left(p^{(2)}_i + p^{(2)}_f\right)^\mu e^{-i(p^{(2)}_i - p^{(2)}_f)\cdot x},$$

and it is not hard to see that Eqn. 2.28 is satisfied by

$$A^\mu(x) = -\frac{j^{\mu(2)}_{fi}(x)}{q^2} \quad \text{with} \quad q^\mu = (p^{(2)}_i - p^{(2)}_f)^\mu. \tag{2.29}$$

Therefore, the final transition amplitude is given by

$$\begin{aligned} T_{fi} &= -i\int \mathrm{d}^4x\, j^{\mu(1)}_{fi}(x)\frac{-g_{\mu\nu}}{q^2} j^{\nu(2)}_{fi}(x) \\ &= |N|^4 \int \mathrm{d}^4x e^{-i(p^{(1)}_i - p^{(1)}_f + p^{(2)}_i - p^{(2)}_f)\cdot x} \\ &\quad \cdot \left(ie(p^{(1)}_i + p^{(1)}_f)^\mu\right)\cdot\frac{-ig_{\mu\nu}}{q^2}\cdot\left(ie(p^{(2)}_i + p^{(2)}_f)^\nu\right) \\ &= |N|^4(2\pi)^4\delta^4(p^{(1)}_i + p^{(2)}_i - p^{(1)}_f - p^{(2)}_f)\cdot\left(ie(p^{(1)}_i + p^{(1)}_f)^\mu\right)\cdot\frac{-ig_{\mu\nu}}{q^2}\cdot\left(ie(p^{(2)}_i + p^{(2)}_f)^\nu\right). \end{aligned} \tag{2.30}$$

A few remarks are in order at this point:

1. For clarity, a label (1) has been attached to the current representing particle 1 (the particle that is scattered by the potential caused by particle 2). However, Eqn. 2.31 is clearly symmetric in the treatment of the two particles under consideration. This is in fact to be expected! For in our -now microscopic- setup, we are scattering two electrons off each other, and there really isn't any physics reason to treat them differently.

2. The factor $(2\pi)^4\delta^4(\ldots)$ arises from the integration over all of spacetime of the plane-wave exponents. Its effect is to impose conservation of four-momentum, as desirable for these scatterings. In fact, this is not at all particular to the process we are considering here, but is instead related to the assumption of asymptotically free states.

3. Implicit in Eqn. 2.31 is the assumption that the normalization N is independent of the momentum. This is in fact correct, but we will not bother with such normalization issues.

Therefore, in general we will be simplifying the discussion of the transition amplitude to that of the so-called *matrix element*, generically denoted by $\mathscr{M}$. Their relation is defined by

$$T_{fi} = -i(2\pi)^4 \delta^4(p_i^{(1)} + p_i^{(2)} - \sum_j p_j) N \mathscr{M}, \tag{2.31}$$

where the sum is over all particles in the final state, and N takes care of the above normalization. In this case, $\mathscr{M}$ is given by

$$-i\mathscr{M} = \left(ie(p_i^{(1)} + p_f^{(1)})^\mu\right) \cdot \frac{-ig_{\mu\nu}}{q^2} \cdot \left(ie(p_i^{(2)} + p_f^{(2)})^\nu\right). \tag{2.32}$$

Limitation

The thing that makes the above derivation heuristic is Eqn. 2.26, in which a "potential" term is added to the *equation of motion* for a free particle (and not to the free particle Hamiltonian). Clearly this is not a proper thing to do. Fortunately, it turns out that in a proper quantum field-theoretical context, we can use the actual Hamiltonian for a complex scalar field (which we lack the formalism to construct explicitly), and it can be shown that the expression for the transition amplitude is correct.

2.2.5 Feynman Diagrams and Feynman Rules

The transition amplitude of Eqn. 2.31 is our way to Feynman diagrams. Apart from the delta function and normalization factors, it contains three ingredients:

- two terms originating from the currents involving the two particles (and which are called the *couplings*);
- and one term that represents the electromagnetic field, as per Eqn. 2.29.

In addition, that same equation shows us that the four-momentum q^μ occurring in the term corresponding to the electromagnetic field corresponds precisely to the difference between the initial- and final-state particles, or in other words, their *momentum transfer*. This leads us to a very simple picture, especially given that we are aware of the particle nature of the photon: in this process, *a photon is exchanged* between the two electrons, absorbing four-momentum from one electron and transferring it to the other. The $-g_{\mu\nu}/q^2$ term is called the *photon propagator*.

This picture can in fact be translated easily to a graphical equivalent, as shown in Fig. 2.2, called the *Feynman diagram* corresponding to this amplitude. In it, the exchanged photon is clearly recognizable, as is its coupling to the electrons. The corresponding *Feynman rules* (given without proof – that is rather a topic for a course on Quantum Field Theory) then tell us how to go back from the diagram to the matrix element:

1. Each Feynman diagram consists of external and internal lines (in Fig. 2.2, the electron and photon lines, respectively) and of *vertices*, which are associated with the couplings of particles to each other.

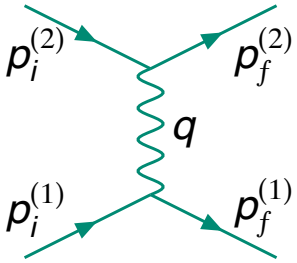

Figure 2.2: Graphical representation of the matrix element of Eqn. 2.32.

2. Each vertex involves a factor

$$(2\pi)^4\delta^4(\sum_i k_i)\cdot\left(ie(p_i+p_f)^\mu\right)$$

where the delta function expresses four-momentum conservation *at each vertex* (all the k_i are taken to be incoming; this is a generic feature of all Feynman diagrams) and in the coupling $e(p_i+p_f)^\mu$ the electron four-momenta "follow the arrows", as in Fig. 2.2.

3. Each internal photon (*i.e.*, each photon propagator) is represented by a "wavy" line and corresponds to a term

$$\int\frac{\mathrm{d}^4q}{(2\pi)^4}\frac{-ig_{\mu\nu}}{q^2},$$

where q^μ is the photon's four-momentum (meaning that each internal four-momentum is integrated over).

4. The result contains a factor $(2\pi)^4$ times a delta function expressing overall four-momentum conservation. This factor is discarded (but of course is to be kept in mind when doing actual computations); the result is equal to $-i\mathscr{M}$.

5. The complete matrix element for a given process (*i.e.*, for given -completely specified- initial and final states) in general corresponds to multiple Feynman diagrams, the individual matrix elements of which have to be summed. (In fact, to obtain the complete matrix element *all* possible Feynman diagrams need to be summed. This is a consequence of the Dyson series: we have restricted ourselves to the computation of the first term in perturbation theory, and ideally we would like to compute higher-order terms as well.)

It may be noted that the photon's four-momentum q^μ does *not* in general satisfy $q^2=0$. On the one hand this is good (as otherwise the transition amplitude would diverge), but on the other hand the question is how this relates to the masslessness of the photon!

The resolution of this issue rests on the fact that the interaction (*i.e.*, the exchange of the photon) takes place on very short timescales. On such timescales, the Heisenberg uncertainty principle dictates that a photon of (squared) "mass" q^2 may exist for an amount of time $\sim 1/\sqrt{|q^2|}$.

Such photons are called *virtual* (since they cannot propagate over macroscopic distances) or *off-shell*. In fact we will also encounter many examples of other off-shell particles being exchanged in interactions.

On a more practical note, while the process under consideration here is the elastic scattering of two *particles*, we could have equally well chosen to consider the scattering of a particle and an *anti-particle* instead (*e.g.*, electron-positron scattering). Now recall that in the Feynman-Stückelberg approach, anti-particles are (loosely speaking) considered as particles moving backward in time, and are associated with the negative-energy solutions. In Feynman diagrams, this difference between particles and anti-particles is expressed by *reversing the direction of the arrows*; so for anti-particles the direction of the arrows is always opposite the physical propagation in time. As a corollary, the conservation of (electrical) current implies that the arrows in a single "current line" (the external and internal lines featuring electrons and/or positrons) must always be in the same direction along the line.

Returning now to our computation of electron-electron scattering, it is not too hard to realize that the above Feynman rules give rise to *another* diagram, even at the lowest order in perturbation theory. Both of them are shown in Fig. 2.3. The second diagram arises because we are dealing with indistinguishable particles (this is why it is not immediately obvious that we did not find it straight from our original treatment of this process, in which we started out not making any assumptions as to the nature of the "other" particle). This process is called Møller scattering.

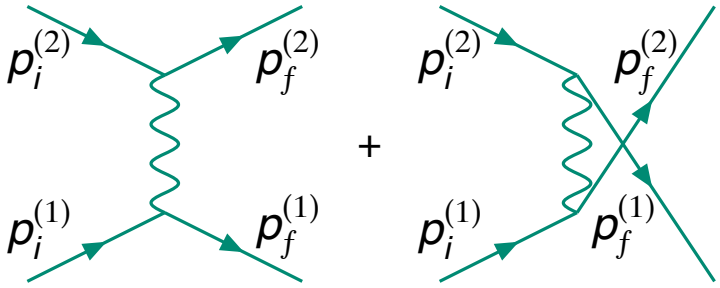

Figure 2.3: Diagrams contributing (in lowest order) to the Møller scattering process $e^- + e^- \to e^- + e^-$.

2.3 The Dirac Equation

2.3.1 Dirac's Attempt

As mentioned in Sect. 2.1, in a quantum mechanical setting there are two problems with the Klein-Gordon equation (perceived problems, as they are addressed by a proper field-theoretic treatment):

1. it involves a *second order* time derivative, giving rise to negative-energy states and a system that has no ground state;

2. and these same negative-energy states lead to a continuity equation that is not amenable to a probability interpretation.

Even if in the context of field theory there is no direct problem, Dirac's attempt to address the above "issues" by constructing an equivalent equation that only involves a *first order* time derivative has proven to be of great importance, as it leads us to a proper description of spin-1/2 particles (the discussion above has not mentioned spin at all, but of course we know that electrons are spin-1/2 particles).

The Dirac equation for free spin-1/2 particles (like the Schrödinger equation, in the position representation) is

$$(i\partial_\mu \gamma^\mu - m)\psi(x) = 0, \tag{2.33}$$

with the quantities γ^μ satisfying the anticommutation relation

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}. \tag{2.34}$$

Clearly this equation cannot be satisfied by ordinary numbers, and therefore a (four-dimensional) *matrix representation* is used. Multiple conventions are possible, but the one most often used (the *Björken and Drell* convention) is

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \gamma^j = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \tag{2.35}$$

where the σ^i represent the Pauli matrices (so also the right-hand side of Eqn. 2.34 formally needs to be multiplied by the 4×4 unit matrix $\mathbb{1}$). Also $\psi(x)$ cannot be a "simple" scalar-valued wavefunction anymore; instead it becomes a column vector of dimension four, called a *bi-spinor*.

(That the Dirac equation is sufficient can be seen by multiplying it from the left by $(i\partial_\nu\gamma^\nu + m)$. This simply yields the Klein-Gordon equation, so we have proven that it is a sufficient condition for the Dirac equation to be satisfied.)

Clearly, given that we are again considering free particles here, it is to be expected that the solutions to the Dirac equation are plane waves. Now in particular, let us consider those plane-wave solutions corresponding to a particle at rest. Given the 2×2 block form of the gamma matrices, write

$$\psi = \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix}.$$

In this case, the Dirac equation can be rewritten as

$$\begin{aligned} (i\frac{\partial}{\partial t} - m)\psi_A &= 0, \\ (-i\frac{\partial}{\partial t} - m)\psi_B &= 0. \end{aligned}$$

Clearly the solution $\psi_A \sim e^{-imt}$ corresponds to a "normal" positive-energy solution; however, $\psi_B \sim e^{+imt}$ again corresponds to a negative-energy solution. By now, however, aware of the antiparticle interpretation of $E<0$ states, we proceed undeterred.

Like in the case of the Klein-Gordon equation, we take the hermitian conjugate of the Dirac equation. The result is

$$-i\partial_\mu \psi^\dagger(x)\gamma^{\mu\dagger} - m\psi^\dagger(x) = 0.$$

We manipulate this by noting, from Eqn. 2.35, that $\gamma^{0\dagger} = \gamma^0$ and $\gamma^{i\dagger} = -\gamma^i$ (since the Pauli matrices are hermitian). Using Eqn. 2.34, this can be written concisely as $\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0$. So we have

$$-i\partial_\mu \psi^\dagger(x)\gamma^0\gamma^\mu\gamma^0 - m\psi^\dagger(x) = 0.$$

Next, we multiply the whole equation by -γ^0 from the right; the result is

$$i\partial_\mu \overline{\psi}(x)\gamma^\mu + m\overline{\psi}(x) = 0, \quad \text{with} \quad \overline{\psi}(x) \equiv \psi^\dagger(x)\gamma^0.$$

With this conjugate equation in hand, we proceed to construct again a continuity equation. This is easily done by multiplying Eqn. 2.33 by $\overline{\psi}(x)$ from the left, the conjugate equation by $\psi(x)$ from the right, and summing the result. This yields

$$i\partial_\mu(\overline{\psi}(x)\gamma^\mu\psi(x)) = 0.$$

Considering in particular the time component, we therefore find that we have

$$\overline{\psi}(x)\gamma^0\psi(x) = \psi^\dagger(x)\psi(x),$$

So we have found a solution where a probability interpretation makes sense! However, again because of the antiparticle interpretation we will not make further attempts in this direction, but instead consider this as a conserved *electric* current:

$$j_\mu = -e\overline{\psi}(x)\gamma_\mu\psi(x). \tag{2.36}$$

2.3.2 Spin-1/2 Particles

The virtue of the Dirac equation is that it allows for a description of spin-1/2 particles. This is perhaps to be expected already simply from the presence of the gamma matrices containing Pauli matrices (which also in non-relativistic Quantum Mechanics are associated with the spin operators for spin-1/2 particles). However, it can also be seen in more detail from considering the general Dirac equation, and again writing it in its 2×2 block form,

$$\psi(x) = \begin{pmatrix} u_A \\ u_B \end{pmatrix} e^{-ip\cdot x},$$

i.e., splitting off the plane-wave piece from the *spinors* u_A and u_B (at this stage we haven't yet specified whether the solution involves positive or negative energies). For nonzero momenta, we obtain *coupled* equations for the spinors:

$$\begin{aligned} (\vec{\sigma}\cdot\vec{p})u_B &= (E-m)u_A, \\ (\vec{\sigma}\cdot\vec{p})u_A &= (E+m)u_B. \end{aligned} \tag{2.37}$$

Restricting ourselves to the positive-energy solution, we can now choose two independent solutions for u_A:

$$u_A^{(1)} = \chi^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u_A^{(2)} = \chi^{(2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The second equation in Eqn. 2.37 then yields

$$u_B{}^{(1,2)} = \frac{(\vec{\sigma}\cdot\vec{p})}{E+m} u_A^{(1,2)}.$$

Similarly, in the case of negative-energy solutions, we choose two independent solutions for u_B, $u_B^{(1,2)} = \chi^{(1,2)}$, and find (from the first equation in Eqn. 2.37):

$$u_A{}^{(1,2)} = \frac{-(\vec{\sigma}\cdot\vec{p})}{-E+m} u_B^{(1,2)}.$$

The minus sign has been carried over to the four-momentum components here. The reason is that in this case, the *physical* four-momentum contains an extra minus sign relative to the four-momentum occurring in Eqn. 2.37. (Note that we might as well have started with the independent solutions for u_B in the case of positive-energy solutions, and vice versa. The important point is that in the non-relativistic limit, for positive-energy solutions, $u_A \gg u_B$, while for negative-energy solutions, $u_B \gg u_A$.)

The specific choice of quantization axis made in the above is not actually the most suitable one for general momenta. As discussed in one of the exercises, an alternative quantization axis is obtained by considering the eigenstates of the so-called *helicity operator* $\lambda \equiv \vec{S}\cdot\hat{p}$, where $\vec{S}$ is the spin operator as appropriate for spin-1/2 particles,

$$\vec{S} = \frac{1}{2}\vec{\Sigma} = \frac{1}{2}\begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}. \tag{2.38}$$

In terms of the *physical* four-momenta, the spinor solutions then become

$$\chi^{(1)}(p) = N' \begin{pmatrix} p_z + |\vec{p}| \\ p_x + ip_y \end{pmatrix}, \quad \chi^{(2)}(p) = N' \begin{pmatrix} -p_x + ip_y \\ p_z + |\vec{p}| \end{pmatrix}, \tag{2.39}$$

with $N' = 1/\sqrt{2|\vec{p}|(|\vec{p}| + p_z)}$ preserving the same normalisation as above.

Summarizing, ψ represents *four* independent degrees of freedom, two for $E > 0$ and two for $E < 0$. When doing practical calculations, the four solutions can be written (again in terms of the physical four-momenta) as

$$u^{(1,2)}(p) = N \begin{pmatrix} \chi^{(1,2)}(p) \\ \frac{(\vec{\sigma}\cdot\vec{p})}{E+m}\chi^{(1,2)}(p) \end{pmatrix}, \quad u^{(3,4)}(p) = N \begin{pmatrix} \frac{(\vec{\sigma}\cdot\vec{p})}{E+m}\chi^{(1,2)} \\ \chi^{(1,2)} \end{pmatrix}.$$

The final remark is that for antifermions, the spin operator as in Eqn. 2.38 acquires an additional minus sign. The reason for this is outside the scope of our mere quantum mechanical treatment, but a motivation, borrowed from Ref. [5], can be found in the fact that also the orbital angular momentum operator acquires this additional minus sign, and the spin operator must therefore follow suit (this follows from the exercise discussed above). As a result, the positive- and negative-helicity solutions for antifermions are respectively given by

$$v^{(1)}(p) \equiv u^{(4)}(p) \quad \text{and} \quad v^{(2)}(p) \equiv u^{(3)}(p).$$

2.3.3 Perturbation Theory

The step from free to *interacting* spin-1/2 particles is made in exactly the same fashion as in the case of spin-0 particles: by means of the minimal substitution (see Sect. 2.2.2). In that case, the Dirac equation is modified to

$$(i\partial_\mu\gamma^\mu - m)\psi(x) = (i\frac{\partial}{\partial t}\gamma^0 + i\vec{\nabla}\cdot\vec{\gamma} - m)\psi(x) = -eA_\mu\gamma^\mu\psi(x). \tag{2.40}$$

The reason for separating the time and spatial components is that we can use this equation to construct explicitly a Hamiltonian suited for spin-1/2 particles. To do so, multiply (from the left) by γ^0; we then have

$$i\frac{\partial}{\partial t}\psi(x) = (-i\vec{\nabla}\cdot\gamma^0\vec{\gamma} + \gamma^0 m)\psi(x) - eA_\mu\gamma^0\gamma^\mu\psi(x),$$

the right-hand side of which nicely has the form $H = H_0 + V$, so that we can identify the term $-eA_\mu\gamma^0\gamma^\mu$ with a perturbing potential V. Inserting this in Eqn. 2.14, we obtain

$$\begin{aligned} T_{fi} &= -i\int \mathrm{d}^4x\,\psi_f^\dagger(x)(-eA_\mu\gamma^0\gamma^\mu)\psi_i(x) \\ &= -i\int \mathrm{d}^4x\, j^\mu_{fi}(x)A_\mu(x), \quad \text{with} \quad j^\mu_{fi}(x) = -e\overline{\psi}_f(x)\gamma^\mu\psi_i(x). \end{aligned}$$

Also here, we restrict ourselves to plane-wave states, and assume that the electromagnetic field is generated by another particle. This implies that we can again insert Eqn. 2.29 – this time of course with a current that is appropriate for spin-1/2 particles. From here, it is not hard to see that also the rest of the computation of the transition amplitude proceeds as for scalar particles.

2.3.4 Feynman Rules for Spin-1/2 Particles

Without further ado, we quote here the Feynman rules appropriate for the computation of matrix elements in QED:

1. The basic "building blocks" of Feynman diagrams are again propagators and vertices.
2. Each photon propagator again corresponds to a factor
$$\int \frac{\mathrm{d}^4q}{(2\pi)^4}\frac{-ig_{\mu\nu}}{q^2}.$$
3. Each fermion propagator corresponds to a factor
$$\int \frac{\mathrm{d}^4q}{(2\pi)^4}\frac{i(\not{q}+m)}{q^2-m^2}.$$
Note that we have introduced here the notation $\not{a} \equiv a_\mu\gamma^\mu$ for *any* a_μ.

4. Each vertex corresponds to a factor

$$(2\pi)^4\delta^4(\sum_i k_i)\cdot ie\gamma^\mu,$$

where all four-momenta are again taken to be towards the vertex.

5. External lines now need to be dealt with more precisely, as the fermions can be labeled by their spins, and we also allow for external photon lines corresponding to specific spin states:

incoming fermion: u	outgoing fermion: $\bar{u}$
incoming antifermion: $\bar{v}$	outgoing antifermion: v
incoming photon: ε^μ	outgoing photon: $\varepsilon^{\mu *}$

6. All appropriate Feynman diagrams should again be summed. A small refinement compared to the case of "scalar QED", however, is that when combining matrix elements that differ only in the exhange of two identical fermions, a relative minus sign must be added. (This is because wavefunctions must be fully antisymmetric under exchange of any two identical fermions.)

7. The overall $(2\pi)^4\delta^4(\ldots)$ is again discarded, and the result is again $-i\mathscr{M}$.

They are shown here mostly for completeness, as we will not attempt to perform complete calculations of Feynman diagrams; nevertheless, it is important to be aware of the differences with the "scalar QED" case.

Polarization states of spin-1 bosons

The polarization vectors ε^μ mentioned in the above merit some further discussion. Let us first discuss the case of *massive* spin-1 bosons. In this case, one can transform to the particle's rest frame, so that the polarization vectors from a non-relativistic treatment are appropriate:

$$\vec{\varepsilon}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{\varepsilon}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{\varepsilon}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

for plane polarization states, and (taking the z axis as our quantisation axis)

$$\vec{\varepsilon}_{\lambda=1} = \frac{-1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \quad \vec{\varepsilon}_{\lambda=-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}, \quad \vec{\varepsilon}_{\lambda=0} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

for circularly polarized states. These polarization states are orthonormal:

$$\vec{\varepsilon}^*_\lambda \cdot \vec{\varepsilon}_{\lambda'} = \delta_{\lambda\lambda'}.$$

Next, we promote these polarization vectors to proper *four-vectors* and require that they remain orthonormal:

$$\begin{aligned} \varepsilon(p;\lambda)\cdot p &= 0, \\ \varepsilon^*(p;\lambda)\cdot\varepsilon(p;\lambda') &= -\delta_{\lambda\lambda'}. \end{aligned}$$

For a boost *e.g.* along the z axis, the *transverse* polarization states ($\lambda = \pm 1$) do not change under this transformation. However, the $\lambda = 0$ ("longitudinal") polarization state does change. From the orthonormality conditions it is not hard to see that for a momentum $p_\mu = (E,0,0,p)$, a vector

$$\varepsilon_\mu(p;\lambda=0) = \frac{1}{m}(p,0,0,E)$$

is required, where m is the particle mass. (Note: it is far from obvious to see how the polarization vectors transform under *general* Lorentz transformations! Suffice it to say that a proper covariant expression can be found, in the form of the so-called *Pauli-Lubanski* vector.) A final useful property of these polarization vectors is

$$\sum_\lambda \varepsilon^\mu(p;\lambda)\varepsilon^{\nu *}(p;\lambda) = -g^{\mu\nu} + p^\mu p^\nu/m^2.$$

(This can either be verified explicitly, or by realising that the result cannot depend anymore on any *specific* polarization vector, and hence only terms proportional to $g^{\mu\nu}$ and $p^\mu p^\nu$ remain. The orthonormality conditions can then be used to determine the corresponding coefficients.)

Let us now consider the case of *massless* spin-1 bosons. As discussed in Sect. 2.2.2, the QED gauge freedom allows for transformations $A_\mu \to A_\mu - \partial_\mu \chi$, with χ an arbitrary function. Specialising to plane waves

$$A^\mu \propto \varepsilon^\mu e^{-iq\cdot x},$$

these gauge transformations amount to changes of the polarization vectors

$$\varepsilon^\mu \to \varepsilon'^\mu = \varepsilon^\mu + \alpha q^\mu.$$

(Note that this does not violate the orthogonality condition $\varepsilon \cdot q = 0$: after all, for on-shell massless particles we have $q^2 = 0$.) This means that we can in fact *choose* χ such that $\varepsilon^0 = 0$. Given the Lorenz condition, this implies $\vec{\varepsilon}\cdot\vec{q} = 0$. So only the *transverse* polarization states survive (but of course this is well known from classical electrodynamics!).

2.4 The Electron's Magnetic Moment

As a final application of our manipulations involving spin-1/2 particles, consider the interaction of an electron with an external magnetic field. The non-relativistic quantum mechanical treatment of this phenomenon is to posit an interaction term $\vec{\mu}\cdot\vec{B}$ in the total Hamiltonian, leading to the *Zeeman splitting* in the presence of a (weak) static magnetic field. In this term, $\vec{\mu}$ is the

electron's magnetic moment. It is typically expressed in terms of the Bohr magneton $\mu_B \equiv e/2m$ as

$$\vec{\mu} = g\mu_B \vec{S},$$

where $\vec{S}$ is the electron's spin vector. For electrons in a quantum mechanical treatment, we have $\vec{S} = \frac{1}{2}\vec{\sigma}$, $\vec{\sigma}$ denoting the Pauli matrices as usual. In summary, we find a term in the Hamiltonian equal to

$$\frac{1}{2} g\mu_B \vec{\sigma} \cdot \vec{B}. \tag{2.41}$$

The issue is that in a "simple" quantum mechanical context, the *Landé factor* g cannot be computed from first principles. The following calculation shows that QED *does* provide a prediction for g – and a correct one at that!

We start again from the Dirac equation with the interaction with an electromagnetic field added through the minimal substitution, as in Eqn. 2.40. Writing in 2×2 block form, we have (cf. Eqn. 2.37)

$$\begin{aligned} \vec{\sigma} \cdot (\vec{p} + e\vec{A}) u_B &= (E + eA^0 - m) u_A, \\ \vec{\sigma} \cdot (\vec{p} + e\vec{A}) u_A &= (E + eA^0 + m) u_B. \end{aligned}$$

Combining these equations yields

$$\left(\vec{\sigma} \cdot (\vec{p} + e\vec{A})\right)^2 u_A = ((E + eA^0)^2 - m^2) u_A.$$

Next, we simplify the left-hand side of this equation, but *keeping in mind the operator character of* $\vec{p}$! Using the identity

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = \vec{a} \cdot \vec{b} + i(\vec{a} \times \vec{b}) \cdot \sigma,$$

this yields

$$\begin{aligned} \left(\vec{\sigma} \cdot (\vec{p} + e\vec{A})\right)^2 &= (\vec{p} + e\vec{A})^2 + i((\vec{p} + e\vec{A}) \times (\vec{p} + e\vec{A})) \cdot \sigma \\ &= (\vec{p} + e\vec{A})^2 + e(\vec{\nabla} \times \vec{A}) \cdot \sigma \\ &= (\vec{p} + e\vec{A})^2 + e\vec{B} \cdot \vec{\sigma}. \end{aligned}$$

Next, we consider the right-hand side of the equation for u_A, in the non-relativistic limit. This implies that the kinetic energy and A^0 are small compared to m, so

$$((E + eA^0)^2 - m^2) = ((m + (E + eA^0 - m))^2 - m^2) \approx 2m(E + eA^0 - m).$$

With that approximation and dividing by $2m$, we obtain

$$(E - m) u_A = \left(\frac{(\vec{p} + e\vec{A})^2}{2m} - eA^0 + \frac{e}{2m} \vec{\sigma} \cdot \vec{B} \right) u_A. \tag{2.42}$$

The last term clearly corresponds to the interaction of a magnetic moment with an external magnetic field, with $g = 2$ (by comparison with Eqn. 2.41).

So is the equation $g = 2$ exact? Not quite, in fact. The static external magnetic field is "merely" one form of an electromagnetic field, and as such the interaction that is of important at the diagrammatic level is that of an electron with the photon, *i.e.*, a diagram consisting essentially only of the $ee\gamma$ vertex (this is possible kinematically since the external magnetic field represents *virtual* rather than real photons). But higher-order perturbative corrections, exemplified in Fig. 2.4, need to be applied.

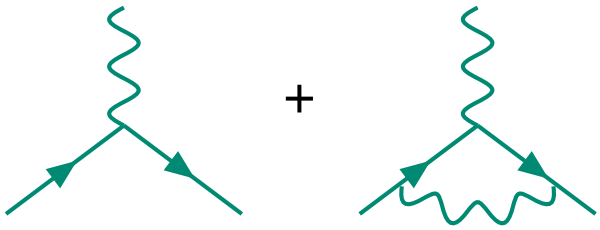

Figure 2.4: Fundamental vertex and "vertex correction" diagram describing the interaction of electrons with electromagnetic fields.

In fact, the electron's *anomalous magnetic moment* $a_e \equiv (g_e - 2)/2$ has been computed very accurately:

$$a_e = \frac{1}{2}\left(\frac{\alpha}{\pi}\right) - 0.328478965\left(\frac{\alpha}{\pi}\right)^2 + 1.1761\left(\frac{\alpha}{\pi}\right)^3 + \ldots$$

It is one of the great achievements of QED that the measured and predicted values of a_e agree with each other, within exceedingly small uncertainties of several parts in 10^9.

Chapter 3

QCD: The Strong Interaction

This chapter covers two manifestations of the strong interactions. First, it considers its static *aspects in the form of the properties (masses, production cross sections, and other properties) of hadrons – the particles susceptible to the strong interaction. In the course of this discussion, two symmetries emerge:* strong isospin*, and by extension (but as a symmetry that is more strongly violated) what is called "flavour* $SU(3)$*"; and "colour", which is at the basis of Quantum Chromodynamics and which is implemented as a* $SU(3)$ *gauge symmetry. In addition, also the* discrete *parity inversion (P) and charge conjugation (C) symmetries are introduced.*

The second part of the chapter deals with dynamic *aspects of the strong interaction, in the form of interactions that probe the internal structure of hadrons (and in particular the proton) more directly. They lead to the conclusion that hadrons are not elementary particles but rather consist of so-called* partons *(as quarks, antiquarks, and gluons are collectively denoted). Several elements are then needed to describe interactions of hadrons: the "hard" interactions of the partons (their strong interactions are described by QCD; the parton density functions (which, loosely speaking, describe how likely it is to pull these partons from the hadrons); and fragmentation (the process of turning (anti-)quarks and gluons to the hadrons that can be detected experimentally again). Along the way, the basic phenomenology of electron-positron, electron-proton, and hadron-hadron collisions will be covered.*

3.1 From Hadrons to the Quark Model

3.1.1 Isospin

One very striking feature of the nucleons (proton and neutron) is the similarity of their masses, 938.28 MeV for the proton and 939.57 MeVfor the neutron. Also there is no evidence (from studies of the nucleus) that they behave differently under the strong interaction. Of course they *do* interact differently with electromagnetic fields (after all, the proton is electrically charged while the neutron is neutral), and also behave differently under the weak interaction (*e.g.* in the form of β decay). However, on short distance scales (smaller than $\mathcal{O}(1\text{ fm})$), the strong

interaction is indeed so much stronger that we will ignore these differences for now.

What that leaves us with is two particles which – as far as the strong interaction is concerned – are really no different. Here, Heisenberg used the analogy with the non-relativistic quantum mechanical formulation of spin. In this formulation, one uses the fact that the physics describing the interaction of a particle's spin with other parts of a Hamiltonian system depends only on the *relative* orientation of the spin with respect to those other parts, and not on some absolute orientation (this is expressed much more elegantly in the form of the *Noether theorem*, which -applied to this case- states that invariance of interactions under rotation (whenever this applies) is directly related to conservation of angular momentum; see Section 2.1.4 for more details). This can be exploited by *choosing* a quantisation axis and calling this the z axis of the system, corresponding to definite eigenstates of the S_z operator. Now, in the case of spin-1/2 particles, $\vec{S}$ is represented by the Pauli matrices, and $S_z = \frac{1}{2}\sigma_3$. The analogy consists in associating the proton and neutron with the eigenstates of σ_3:

$$\mathrm{p} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathrm{n} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

in a so-called *isospin doublet*. So *isospin* (short for "isotopic spin") is the equivalent of spin, but it is an *internal symmetry*, involving different particles rather than different spatial orientations of the same particle.

Again in analogy to spin, rotations in "isospin space" are generated by the isospin operator $\vec{I}$: a rotation over an angle ω about an axis $\hat{e}$ can be written as

$$\mathscr{R}(\omega, \hat{e}) = e^{-i\omega\hat{e}\cdot\vec{I}}.$$

Also, isospin eigenstates can be (fully) classified according to their total isospin I and the eigenvalue of the I_3 operator, so: $|I, I_3\rangle$. In the above case, we are dealing with a *doublet* representation as we know of no charge -1 or +2 particles with otherwise the same properties (mass, spin) as the proton and neutron. Hence the proton and neutron are also denoted by $|\frac{1}{2}, \frac{1}{2}\rangle$ and $|\frac{1}{2}, -\frac{1}{2}\rangle$, respectively.

(We now know that isospin as discussed here is not really an appropriate formalism for the discussion of interactions at high energies. However, at low energies it is indeed a useful formalism, as we will see later in this section. In addition, the concept of internal symmetries turns out to be extremely useful in other areas.)

Comparison with QED

The concept of *internal symmetries* was in fact introduced already in Sect. 2.2.2; see in particular Eqn. 2.24. The difference is that the QED case concerns itself only with phase transformations, while isospin symmetry deals with the transformations between different particle types. This difference is artificial, however: what we referred to as a wavefunction in Sect. 2.2.2 should more properly be regarded as a (quantum) *field*. Again we will not dwell on all differences between the two; but at this point it is useful to consider the fact that fields do not need to be complex but might as well be real; the complex nature exhibited by Eqn. 2.24 is equivalent to dealing with rotations between the two real fields that are the real and imaginary parts of the original, complex field.

3.1.2 The Pion

From elastic scattering experiments it is known that the strong interaction between nuclei has a finite range, $R \approx 1.2$ fm. Given that we now know that interactions between particles can be thought of as being mediated by other particles, it becomes natural to ask what particles could mediate the strong interaction. To answer this question, consider *static* solutions to the Klein-Gordon equation, *i.e.*:

$$\vec{\nabla}^2\phi(\vec{x}) - m^2\phi(\vec{x}) = \delta^3(\vec{r}),$$

where the non-zero right-hand side represents the interaction's point source. It can be shown that

$$\phi(r) = \frac{1}{4\pi}\frac{1}{r}e^{-mr}$$

is a Green's function satisfying this equation. This indicates an interaction of finite range $R = 1/m$, and hence it becomes very natural to associate the 1.2 fm range of the strong interaction with a particle of mass ~ 160 MeV. This is the *Yukawa hypothesis*, after Yukawa who first proposed it in the 1930's.

Such a particle has in fact been found! In 1937, two groups (Anderson and Neddermeyer, and Street and Stevenson) found evidence for its exististence in studies of cosmic rays; an example (but from a later study by Powell *et al.*) is shown in Fig. 3.1. The charged pion, $\pi^{\pm}$, turns out to have a mass of 139.6 MeV, sufficiently close to the 160 MeV above to qualify as the mediator of the strong interaction between nucleons. (In fact the same photographs show the pion's decay – its lifetime is now well known, $\tau_{\pi^{\pm}} \approx 26$ ns – to yet another new particle. This is the muon $\mu^{\pm}$, which we will discuss in more detail in the context of the weak interaction. Also, at least one neutral particle must be involved in this decay, to conserve momentum. This, too, belongs to the topic of the weak interaction.)

But this is not the whole story: in 1950, an electrically neutral partner to the $\pi^{\pm}$ was discovered. The π^0 is most easily recognized through its decay to two photons, $\pi^0 \to \gamma\gamma$ and a much shorter lifetime, $\tau_{\pi^0} \approx 8 \cdot 10^{-17}$ s. Its mass (135.0 MeV) is only slightly lower than that of the charged pion. This implies that the three pions nicely fit into an *isospin triplet*, *i.e.*, states with $I =$ and $I_3 = -1, 0, 1$. Combining now the idea of the pion mediating the strong interaction between nucleons with that of isospin symmetry, we find that we can make use of all pions to mediate this interaction. An example is shown in Fig. 3.2.

Now that we are aware of the existence of pions, it is time to learn more about their *properties*. A variety of methods have been devised to do so, and here we mention a few.

Spin of the pion: detailed balance. This method uses the reaction

$$\pi^+ + \mathrm{d} \to \mathrm{p} + \mathrm{p} \tag{3.1}$$

where d denotes the deuteron (the p-n bound state), as well as its reverse reaction. As seen in one of the exercises, the differential cross section for such $2 \to 2$ body processes is given by

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{64\pi^2 s}\mathscr{S}\frac{|\vec{p}_f|}{|\vec{p}_i|}|\mathscr{M}|^2, \tag{3.2}$$

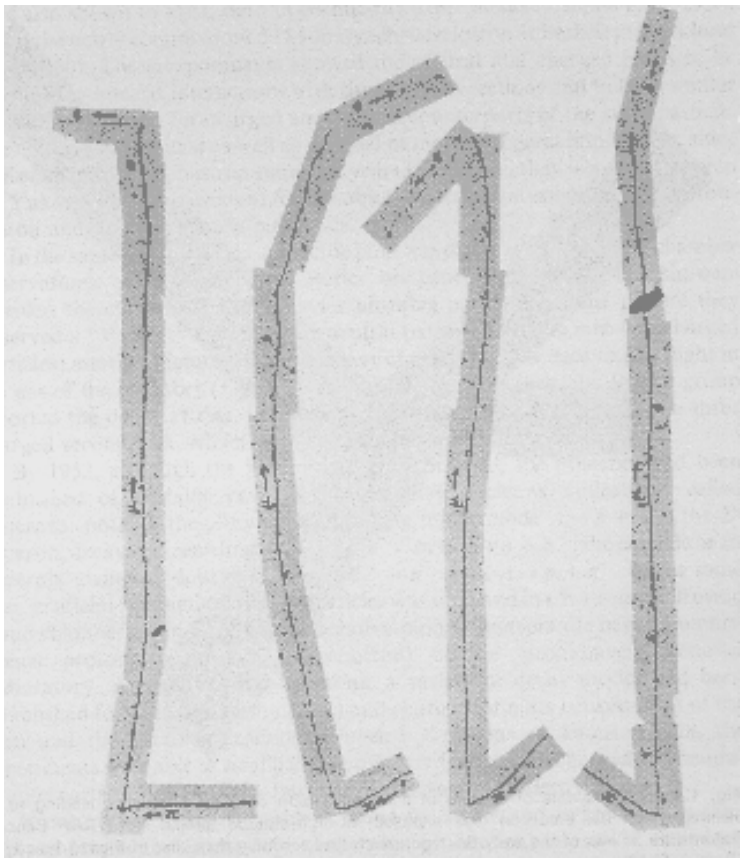

Figure 3.1: Trajectories of charged particles in emulsion plates exposed to cosmic rays. From the ionization density and their range (the distance they travel before being stopped) the masses of the particles involved can be inferred.

where $|\vec{p}_f|$ ($|\vec{p}_i|$) is the magnitude of the momentum of (either of) the particles in the final (initial) state, and $\mathscr{S}$ is a "spin factor".

At this point it is necessary to specify more precisely the meaning of this spin factor. It is due to the fact that the matrix element $\mathscr{M}$ depends on the spins of the particles in the initial and final states. A fairly common situation is to start from unpolarized particles in the initial state, and not to measure the final-state spins. In that case, the procedure is to

- average over initial-state spin states; and
- sum over final-state spin states.

As a result, for the above scattering process $\mathscr{S}$ takes the form

$$\mathscr{S} = (2s_{\mathrm{p}} + 1)^2$$

if we replace simultaneously $\mathscr{M}^2$ with its spin-averaged value; for the reverse process it takes the value $(2s_\pi + 1)(2s_{\mathrm{d}} + 1)$.

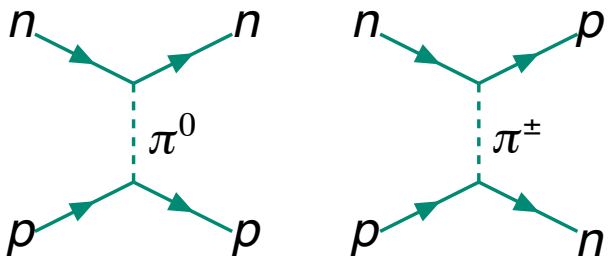

Figure 3.2: The strong interaction between a proton and a neutron, as mediated by pions.

The issue now is that when comparing the (differential) cross section for the process of Eqn. 3.1 with that for its reverse reaction, this spin-averaged squared matrix element value (usually denoted by $\langle|\mathscr{M}^2|\rangle$) remains the same, at least assuming *time reversal symmetry* (which we will not elaborate on further, except to say that it is closely related the *parity inversion* and *charge conjugation* symmetries that are discussed below). So the cross section ratio (of course it needs to be evaluated at the same centre-of-mass energy) becomes

$$\frac{\sigma(\pi^+ + \mathrm{d} \to \mathrm{p} + \mathrm{p})}{\sigma(\mathrm{p} + \mathrm{p} \to \pi^+ + \mathrm{d})} = \frac{(2s_\mathrm{p}+1)^2}{(2s_\pi+1)(2s_\mathrm{d}+1)} \frac{|\vec{p}_\mathrm{pp}|^2}{|\vec{p}_{\pi\mathrm{d}}|^2}.$$

With the spins of the proton ($s_\mathrm{p} = 1/2$) and the deuteron ($s_\mathrm{d} = 1$) known, the actual measurement of the cross section ratio implies that $s_\pi = 0$.

Parity of the pion. Parity is the discrete symmetry associated with spatial inversion, *i.e.*, the replacement of spatial coordinates with their inverse, $\vec{r} \to -\vec{r}$. In a quantum mechanical context, the parity operator (denoted $\mathscr{P}$) also commutes with the orbital angular momentum and spin operators $\vec{L}$ and $\vec{S}$, and hence with $\vec{J} = \vec{L} + \vec{S}$: this implies that states can be simultaneously parity and angular momentum eigenstates.

A case that is easily dealt with is the parity associated with bound states. The angular wave functions corresponding to states of definite L and m are described by spherical harmonic functions $Y_L^m(\theta, \varphi)$:

$$Y_L^m(\theta, \varphi) \propto e^{im\varphi} P_L^m(\cos\theta),$$

with

$$P_L^m(x) \equiv (1-x^2)^{|m|/2} \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^{|m|} P_L(x), \quad P_L(x) \equiv \frac{1}{2^L L!} \left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^L (1-x^2)^L.$$

It is readily seen that parity conjugation ($\theta \to \pi - \theta$, $\varphi \to \varphi + \pi$) changes the wave function by a factor $(-1)^L$.

However, in situations where particles can be created or annihilated, it turns out that also an *intrinsic* parity eigenvalue needs to be associated with particles.

- From the Dirac equation, it can be shown that the intrinsic parity of fermions and their antiparticles is opposite (but no definite assignment can be made; by convention, the *particles* are assigned $P = +1$).

- Also in the case of the *baryons* (hadrons of half-integer spin; the hadrons of integer spin are collectively known as *mesons*), a convention is necessary. The choice made is $P(\mathrm{p}) = P(\mathrm{n}) = +1$.

The above has been used to determine the parity of the charged pion from the reaction $\pi^- + \mathrm{d} \to \mathrm{n} + \mathrm{n}$, in a setup where the pion is captured at very low velocity. In this situation, the pion's orbital angular momentum is likely to be $L = 0$. This implies that the *total* angular momentum in the initial state is also $J = 1$. The same must therefore hold in the final state, implying that the *nn* system has either $L = 0$ and $S = 1$, or $L = 1$ and $S = 0, 1$:

- in the $L = 0$, $S = 1$ case, the neutrons are in a *symmetric* state, which is forbidden by the Pauli exclusion principle;
- the same holds for the case that $L = 1$ and $S = 0$. Therefore, only $L = S = 1$ is allowed.

But $L = 1$ in the final state implies that its parity is $P(\mathrm{nn}) = -1$. Conservation of parity then means that $P(\pi) \cdot P(\mathrm{d}) = P(\pi) = -1$ (as the deuteron has $P(\mathrm{d}) = +1$).

Charge Conjugation. There is another discrete symmetry operation that is best covered at this point, namely *charge conjugation*. This changes particles into their antiparticles, while leaving their other properties (momentum, position, and spin) unchanged. Now if a particle is *its own antiparticle*, they become eigenstates of the charge conjugation operator:

$$C|\psi\rangle = \pm|\psi\rangle,$$

as is easily realized from the fact that applying C twice is simply the identity operation. But its scope is broader than that:

- also particle-antiparticle bound states can be assigned a definite C eigenvalue. In the case of $\pi^+\pi^-$ bound states, for instance, one has

  $$C|\pi^+\pi^-; L\rangle = (-1)^L|\pi^+\pi^-; L\rangle,$$

  as can be seen from the fact that in this case, charge conjugation has the same effect as parity conjugation;
- within the context of field theory, it can be shown that fermion-antifermion bound states satisfy

  $$C|f\bar{f}; L, S\rangle = (-1)^{L+S}|f\bar{f}; L, S\rangle.$$

The above can be applied easily to determine the C eigenvalue of the π^0: it decays to $\gamma\gamma$ final states (this is an electromagnetic decay, as is to be expected from the presence of photons). Given that such decays conserve C (this can be shown explicitly for QED), the π^0 must have $C(\pi^0) = (C(\gamma))^2 = 1$. (Also the strong interaction conserves C, a fact that we will state without proof.)

Given the electromagnetic decay $\pi^0 \to \gamma\gamma$, it might be expected that the decay to three photons ($\pi^0 \to \gamma\gamma\gamma$) is suppressed relative to the two-photon decay mode by an order α. However, experimentally this decay has never been observed, and it is limited to $\Gamma(\pi^- \to \gamma\gamma\gamma)/\Gamma(\pi^- \to \gamma\gamma) < 3 \cdot 10^{-8}$. This can be explained if $C(\gamma) = -1$.

3.1.3 Resonances

With increasing beam energies attainable through more and more powerful accelerators, the 1950's brought a slew of new particles observable as *resonances* (with total decay widths of tens to hundreds of MeV) in the cross section for scattering processes. For sufficiently low centre-of-mass energies, $\sqrt{s} \lesssim 2$ GeV, isospin symmetry provides an adequate description of these processes. An important case in point is the so-called Δ *resonances*, with masses $m_\Delta \approx 1232$ MeV. The total decay width of these particles is approximately 118 MeV, corresponding to very short lifetimes (of order 10^{-23} s).

We consider the *total* pion-nucleon scattering cross sections, ignoring final states other than πN (this is a good approximation at low energy). With the pion and nucleon isospin assignments known, we can now use the *Clebsch-Gordan coefficients* (see, *e.g.*, Appendix A) to decompose the πN states into states of definite total isospin:

$$\begin{aligned}
\pi^+\mathrm{p}: \qquad |\tfrac{1}{2},\tfrac{1}{2}\rangle|1,1\rangle \quad &= |\frac{3}{2},\frac{3}{2}\rangle, \\
\pi^+\mathrm{n}: \qquad |\tfrac{1}{2},-\tfrac{1}{2}\rangle|1,1\rangle \quad &= \sqrt{\frac{1}{3}}|\frac{3}{2},\frac{1}{2}\rangle + \sqrt{\frac{2}{3}}|\frac{1}{2},\frac{1}{2}\rangle, \\
\pi^0\mathrm{p}: \qquad |\tfrac{1}{2},\tfrac{1}{2}\rangle|1,0\rangle \quad &= \sqrt{\frac{2}{3}}|\frac{3}{2},\frac{1}{2}\rangle - \sqrt{\frac{1}{3}}|\frac{1}{2},\frac{1}{2}\rangle, \\
\pi^0\mathrm{n}: \qquad |\tfrac{1}{2},-\tfrac{1}{2}\rangle|1,0\rangle \quad &= \sqrt{\frac{2}{3}}|\frac{3}{2},-\frac{1}{2}\rangle + \sqrt{\frac{1}{3}}|\frac{1}{2},-\frac{1}{2}\rangle, \\
\pi^-\mathrm{p}: \qquad |\tfrac{1}{2},\tfrac{1}{2}\rangle|1,-1\rangle \quad &= \sqrt{\frac{1}{3}}|\frac{3}{2},-\frac{1}{2}\rangle - \sqrt{\frac{2}{3}}|\frac{1}{2},-\frac{1}{2}\rangle, \\
\pi^-\mathrm{n}: \qquad |\tfrac{1}{2},-\tfrac{1}{2}\rangle|1,-1\rangle \quad &= |\frac{3}{2},-\frac{3}{2}\rangle.
\end{aligned}$$

Furthermore we exploit the fact that the conservation of isospin by the strong interaction implies that transition amplitudes depend *only* on the total isospin, and denote

$$\langle I, I_3|V|I', I_3'\rangle = \delta_{I,I'}\delta_{I_3,I_3'}A_I.$$

(where A_I may still depend on *kinematical* quantities – and it typically does). This can be used to parametrize the various possible transition amplitudes:

$$\begin{aligned}
\langle\pi^+\mathrm{p}|V|\pi^+\mathrm{p}\rangle &= A_{\frac{3}{2}}, \\
\langle\pi^0\mathrm{n}|V|\pi^-\mathrm{p}\rangle &= \frac{\sqrt{2}}{3}(A_{\frac{3}{2}} - A_{\frac{1}{2}}), \\
\langle\pi^-\mathrm{p}|V|\pi^-\mathrm{p}\rangle &= \frac{1}{3}A_{\frac{3}{2}} + \frac{2}{3}A_{\frac{1}{2}}.
\end{aligned}$$

Squaring these transition amplitudes, it is easy to verify that the following results are obtained

for the cross sections:

$$\begin{aligned}
\sigma_{\text{tot}}(\pi^+\text{p}) &\sim |A_{\frac{3}{2}}|^2, \\
\sigma_{\text{tot}}(\pi^-\text{p}) = \sigma_{\text{tot}}(\pi^-\text{p}\to\pi^-\text{p}) + \sigma_{\text{tot}}(\pi^-\text{p}\to\pi^0\text{n}) &\sim \frac{2}{9}|A_{\frac{3}{2}} - A_{\frac{1}{2}}|^2 + \frac{1}{9}|A_{\frac{3}{2}} + 2A_{\frac{1}{2}}|^2 \\
&= \frac{1}{3}|A_{\frac{3}{2}}|^2 + \frac{2}{3}|A_{\frac{1}{2}}|^2.
\end{aligned}$$

If a resonance occurs, it is often easily discernible in the total cross section for pion-nucleon scattering, as a function of centre-of-mass energy. Our central hypothesis is now that a resonance has either $I = \frac{1}{2}$ or $I = \frac{3}{2}$. Given the resonant behaviour, this implies that the $|A_{\frac{1}{2}}| \gg |A_{\frac{3}{2}}|$ in the first case, and $|A_{\frac{3}{2}}| \gg |A_{\frac{1}{2}}|$ in the second. Neglecting the smaller term, this then has direct consequences for the ratio of total πN scattering cross sections:‘

$$\frac{\sigma_{\text{tot}}(\pi^+\text{p})}{\sigma_{\text{tot}}(\pi^-\text{p})} = \begin{cases} 3 & \text{if } |A_{\frac{3}{2}}| \gg |A_{\frac{1}{2}}| \\ 0 & \text{if } |A_{\frac{1}{2}}| \gg |A_{\frac{3}{2}}| \end{cases}$$

This cross section ratio has been measured, and found to be very close to 3, implying that the $\Delta(1232)$ resonance has $I = 3/2$. If this indeed the case, then the resonances observed in $\pi^{\pm}$p scattering must be the $I_3 = 3/2$ (Δ^{++}) and $I_3 = -1/2$ (Δ^0) members of a quartet, and also Δ^+ and Δ^- particles are to be expected. These have indeed been found! The $\Delta(1232)$ resonances have been analyzed further, and found to have also $J = 3/2$. This will turn out to be of some importance later in this chapter.

3.1.4 Strangeness

Around the time of the discovery of the charged pion, yet other particles were discovered in studies of cosmic-ray events in photographic emulsions and cloud chambers. Fig. 3.3 shows the earliest such photographs. From detailed studies of these events, it became obvious that the particles involved have masses well above m_π, some of them below m_p and some of them above. By present standards, they are long-lived, with lifetimes ranging from 200 ps to 80 ns. Table 3.1 provides a brief overview of these new particles.

The most peculiar feature of this class of particles was noticed in experiments using a 1.5 GeV pion beam colliding with hydrogen atoms in a cloud chamber: their production cross section is of similar magnitude as that for the productions of pions, protons, or resonances, even if their lifetimes are much longer than expected for decays mediated by the strong interaction (remember the large decay widths of Sect. 3.1.3)! In addition, they are produced *in pairs*.

These particles were called “strange”, and this nomenclature can be formalized by introducing *Strangeness* (usually denoted as S) as a new *quantum number*, and positing that S is conserved in the strong interaction responsible for their production, but *not* in their decay (which is now known to be due to the weak interaction – but more on that later).

Strangeness is known to be an *additive* quantum number, like isospin, and not a multiplicative one like parity. This follows from the fact that *e.g.* $\text{K}^+\Sigma^-$ and $\text{K}^0\Lambda$ pairs are produced in such

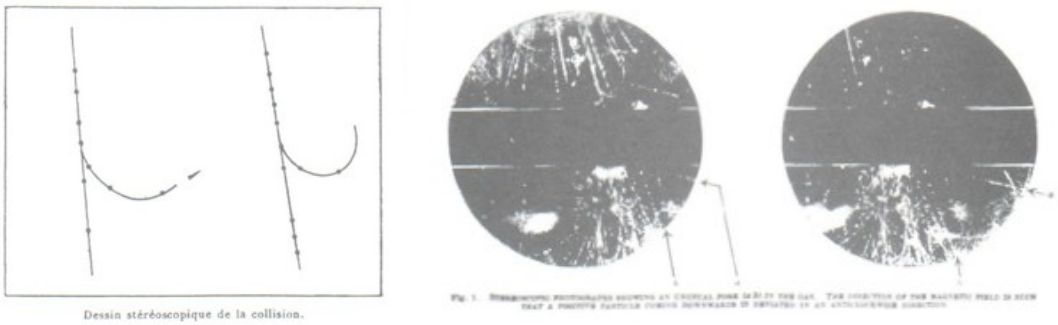

Figure 3.3: Discoveries of "strange" particles. Left: discovery of the charged kaon through its collision with an atomic electron. Right: observation of a long-lived neutral particle decaying to two charged particles (marked by arrows). This is now interpreted as the decay $K^0 \to \pi^+\pi^-$.

particle	mass (MeV)	lifetime (s)	(main) decay modes
spin 0			
K^-	494	$1.2 \cdot 10^{-8}$	$\mu^-\bar{\nu}_\mu$, $\pi^-\pi^0$
K^0_S	498	$9 \cdot 10^{-11}$	$\pi^+\pi^-$, $\pi^0\pi^0$
K^0_L	498	$5 \cdot 10^{-8}$	$\pi^\pm e^\mp \nu_e$, $\pi^\pm \mu^\mp \nu_\mu$, $\pi^+\pi^-\pi^0$, $\pi^0\pi^0\pi^0$
spin 1/2			
Λ	1116	$2.6 \cdot 10^{-10}$	$p\pi^-$, $n\pi^0$
Σ^+	1189	$8 \cdot 10^{-11}$	$p\pi^0$, $n\pi^+$
Σ^0	1193	$7 \cdot 10^{-20}$	$\Lambda\gamma$
Σ^-	1197	$1.5 \cdot 10^{-10}$	$n\pi^-$
Ξ^-	1322	$1.6 \cdot 10^{-10}$	$\Lambda\pi^-$
Ξ^0	1315	$3 \cdot 10^{-10}$	$\Lambda\pi^0$

Table 3.1: Properties of the lightest strange particles.

interactions, but *not* $K^-\Sigma^+$. (Keep in mind that the K^+ is the K^-'s antiparticle, and therefore must carry the opposite value of S compared to the K^-. This is not so for the Σ^+ compared to the Σ^-, as can easily be seen from the fact that they have unequal masses – see Table 3.1.)

3.1.5 From the Eightfold Way to the Static Quark Model

Besides these long-lived strange particles, also strange resonances turn out to exist, decaying (through the strong interaction) to the lightest strange particles. As a result, a great multitude of hadrons is now known. The question is how to make sense of this multitude?

The answer to this question came through careful studies of these particles to determine their quantum numbers (spin, isospin, parity). In particular, the assumption that isospin is part of a more extended internal symmetry turns out to be a fruitful one. It implies that all hadrons in the same *multiplet* (the equivalent of the singlets, doublets, and triplets we have seen above for isospin, but in a more involved structure due to the extended symmetry) should have the same

spin and parity! This leads to a lot more order. Gell-Mann and Ne'eman were the first to use this to construct the multiplets given in Figs. 3.4 and 3.5. The anti-baryons (the antiparticles corresponding to the baryons in Fig. 3.5) have also been observed (in the meson case, particles and antiparticles reside in the same multiplet).

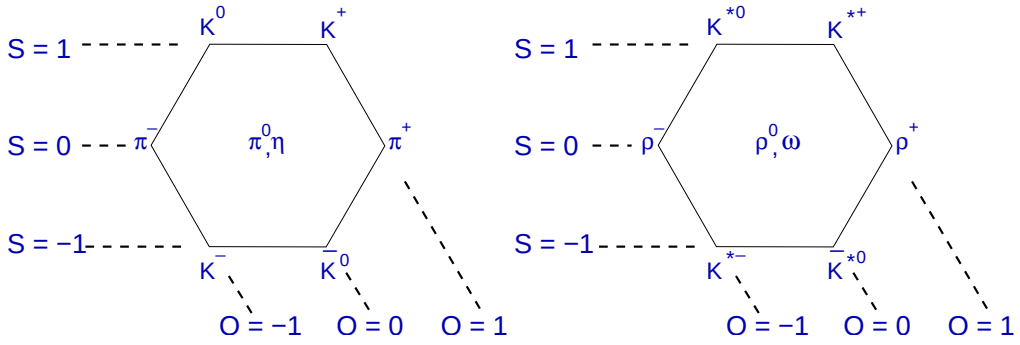

Figure 3.4: Pseudoscalar ($J^P = 0^-$, left) and vector ($J^P = 1^-$, right) meson octets.

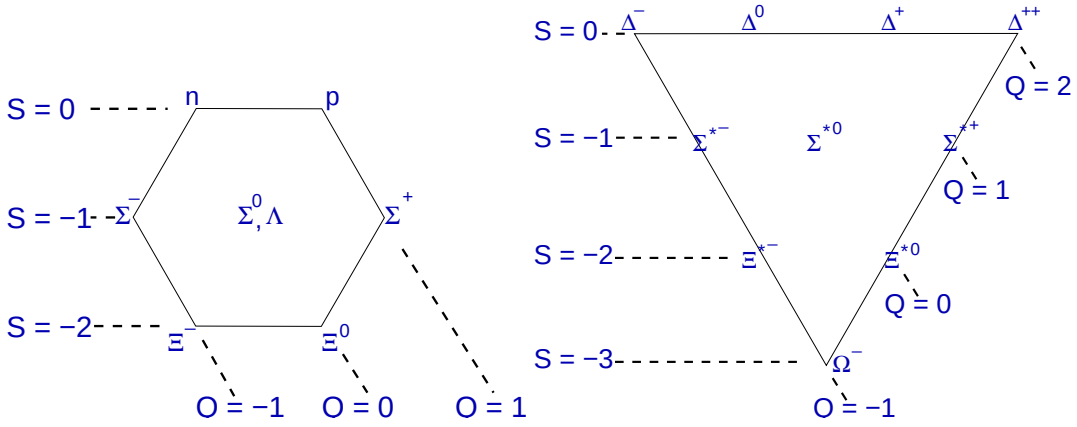

Figure 3.5: $J^P = \frac{1}{2}^+$ baryon octet (left) and $J^P = \frac{3}{2}^+$ baryon decuplet (right).

The symmetry represented by this "eightfold way" (the name refers in particular to the octets visible in these figures) is broken fairly badly by the large mass differences between particles with different strangeness (see *e.g.* Table 3.1) – significantly more so than isospin symmetry. Nevertheless, taking it seriously does allow us to make progress. For the multiplets observed are very indicative of an $SU(3)$ internal symmetry. On that basis, Gell-Mann in fact *predicted* the existence of the Ω^- baryon, as well as its mass. The subsequent discovery of this particle (with $\tau_{\Omega^-} = 8 \cdot 10^{-11}$ s and $m(\Omega^-) = 1672$ MeV) meant a great triumph for this ordering principle. The Ω^- decays to ΛK^- (68%), $\Xi^0\pi^-$ (24%), and $\Xi^-\pi^0$, consistent with the strangeness-changing decay of an $S = -3$ particle. Moreover, it was observed in the reaction $K^-p \to \Omega^- K^+ K^0$, establishing definitively that $S(\Omega^-) = -3$.

But if these multiplets are all proper representations of an $SU(3)$ symmetry (in group theoretical parlance), it begs the question whether there is a *fundamental* representation that can be

identified! Indeed this exists, and it is expressed by the *quark hypothesis* due to Gell-Mann and Zweig. According to this hypothesis, all mesons are quark-antiquark bound states, and baryons are bound states consisting of three quarks (and similarly, antibaryons consist of antiquarks). To account for the half-integer spin of baryons versus the integer spin of mesons, the quarks then need to have half-integer spins; we will assume that they are spin-1/2 particles. The meson $J^P = 0^-$ and 1^- octets are then obtained as $L = 0$ bound states with total spin $S = 0$ and $S = 1$, respectively (it is easily verified that this leads to the correct parity).

The fundamental representation is shown in Fig. 3.6. The isospin assignments of the quarks

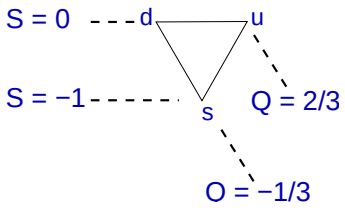

Figure 3.6: Fundamental representation of (flavour) $SU(3)$.

then become

$$\begin{aligned} \mathrm{u} &= |\tfrac{1}{2},\tfrac{1}{2}\rangle \\ \mathrm{d} &= |\tfrac{1}{2},-\tfrac{1}{2}\rangle \\ \bar{\mathrm{d}} &= -|\tfrac{1}{2},\tfrac{1}{2}\rangle \\ \bar{\mathrm{u}} &= |\tfrac{1}{2},-\tfrac{1}{2}\rangle \end{aligned}$$

(the minus sign for $\bar{\mathrm{d}}$ is a convention that does not affect the physics). The quark assignment of the particles in the pseudoscalar meson octet then becomes that of Table 3.2. There is an additional pseudoscalar meson, the η', which is an $SU(3)$ *singlet*: its quark content is $\frac{1}{\sqrt{3}}(\mathrm{u}\bar{\mathrm{u}}+\mathrm{d}\bar{\mathrm{d}}+\mathrm{s}\bar{\mathrm{s}})$.

π^+:	$\mathrm{u}\bar{\mathrm{d}}$	K^+:	$\mathrm{u}\bar{\mathrm{s}}$
π^0:	$\frac{1}{\sqrt{2}}(\mathrm{u}\bar{\mathrm{u}}-\mathrm{d}\bar{\mathrm{d}})$	K^0:	$\mathrm{d}\bar{\mathrm{s}}$
π^-:	$\mathrm{d}\bar{\mathrm{u}}$	$\overline{\mathrm{K}^0}$:	$-\mathrm{s}\bar{\mathrm{d}}$
η:	$\frac{1}{\sqrt{6}}(\mathrm{u}\bar{\mathrm{u}}+\mathrm{d}\bar{\mathrm{d}}-2\mathrm{s}\bar{\mathrm{s}})$	K^-:	$\mathrm{s}\bar{\mathrm{u}}$

Table 3.2: Quark content of pseudoscalar meson octet constituents

The case of the vector mesons is very similar, except that *mixing* takes place between the octet and the singlet. As a result, we have

$$\omega = (\mathrm{u}\bar{\mathrm{u}}+\mathrm{d}\bar{\mathrm{d}})/\sqrt{2} \quad \text{and} \quad \phi = \mathrm{s}\bar{\mathrm{s}}.$$

(This different quark content can be seen *e.g.* from the decay modes of the ϕ: it decays mostly to K^+K^-, $K^0\overline{K^0}$, and $\pi^+\pi^-\pi^0$. But we will not dwell on this mixing phenomenon further.)

Based on the above, it is reasonable to ask whether multiplets corresponding to higher values of L exist? Indeed they do: multiplets of up to $L = 2$ and $J = 2$ have meanwhile been identified, with particles of substantially higher mass (~ 1700 MeV) than for the pseudoscalar and vector mesons.

3.1.6 The Baryon Decuplet and Colour

Revisiting the baryon octet and decuplet of Fig. 3.5, we find that their interpretation in the context of the static quark model leads to a very interesting observation, in particular for the decuplet. We assume that the octet and decuplet correspond to $L = 0$, $S = 1/2$ and $L = 0$, $S = 3/2$ bound states, respectively. That means that these baryons are in a completely symmetric *spin* state. Choosing an arbitrary quantization axis, and denoting the spin component of the three quarks along this axis by arrows:

$$\begin{aligned}
|\tfrac{3}{2},\tfrac{3}{2}\rangle &= |\uparrow\uparrow\uparrow\rangle, \\
|\tfrac{3}{2},\tfrac{1}{2}\rangle &= \frac{1}{\sqrt{3}}\left(|\downarrow\uparrow\uparrow\rangle + |\uparrow\downarrow\uparrow\rangle + |\uparrow\uparrow\downarrow\rangle\right), \\
|\tfrac{3}{2},-\tfrac{1}{2}\rangle &= \frac{1}{\sqrt{3}}\left(|\uparrow\downarrow\downarrow\rangle + |\downarrow\uparrow\downarrow\rangle + |\downarrow\downarrow\uparrow\rangle\right), \\
|\tfrac{3}{2},-\tfrac{3}{2}\rangle &= |\downarrow\downarrow\downarrow\rangle.
\end{aligned}$$

But this is not all: also the *flavour* state is completely symmetric, *i.e.*

$$\begin{aligned}
\Delta^{++} &= |\mathrm{uuu}\rangle, \\
\Delta^{+} &= \frac{1}{\sqrt{3}}\left(|\mathrm{duu}\rangle + |\mathrm{udu}\rangle + |\mathrm{uud}\rangle\right),
\end{aligned}$$

etc. The result is that we have states that are *completely symmetric* under the exchange of any two particles. But quarks are fermions, so this violates the Pauli exclusion principle!

To mend this problem, a new degree of freedom has been invoked, called *colour*. This has little to do with the electromagnetic spectrum; rather, it posits yet another internal ($SU(3)$) symmetry with "orientations" in this internal space denoted by red, green, and blue. Each quark can then come in any of these three colours (or a linear superposition thereof). This can be used to construct a three-particle colour state that is *antisymmetric* under the exchange of any two quarks, namely

$$\frac{1}{\sqrt{6}}\left(|rgb\rangle - |rbg\rangle + |brg\rangle - |bgr\rangle + |gbr\rangle - |grb\rangle\right). \tag{3.3}$$

When combined with the symmetric flavour and spin states, this yields a state that is completely antisymmetric under the exchange of any two particles – exactly as required by Fermi-Dirac statistics. (This is of course a very *ad hoc* argument. However, we will see in the next section

that this colour $SU(3)$ interaction indeed provides a proper description of the strong interaction in many other respects.)

Eqn. 3.3 in fact describes a *colour singlet* state, meaning that it does not transform under a *colour* $SU(3)$ *transformation*, *i.e.*, a change of orientation in this new internal space. Also for mesons, such a colour singlet state can be constructed:

$$\frac{1}{\sqrt{3}}\left(|r\bar{r}\rangle + |g\bar{g}\rangle + |b\bar{b}\rangle\right). \tag{3.4}$$

Here, $\bar{r}$ *etc.* denote the colour of the antiquark (remember, *all* quantum numbers are reversed going from particles to antiparticles). It has been posited that *only colour-singlet states correspond to physical hadrons*. As it turns out, the question how (anti)quarks bind into hadrons cannot be answered using perturbation theory only. However, it can be argued that indeed, antisymmetric colour combinations for quarks and symmetric quark-antiquark colour combinations give rise to attractive interactions, whereas other combinations lead to repulsive interactions. (See Sect. 8.4 of the book by Griffiths for the details.)

3.1.7 Hadron Masses and Magnetic Dipole Moments

The quark model also allows us to say more about the masses of these hadrons. We start from two observations:

1. While the mass differences within isospin multiplets are generally small, large differences are seen between isospin multiplets with different strangeness S (which we now associate with the presence of different numbers of s (anti)quarks). A straightforward interpretation of this dependence is that the hadron mass is determined in part by the masses of its constituent quarks, with $m_\mathrm{s} \gg m_\mathrm{d} \approx m_\mathrm{u}$.

2. There is a substantial mass difference between vector and pseudoscalar mesons with the same quark content, and similarly between baryon decuplet and octet states with the same quark content. As we have seen above, the only difference between them is their spin state. A natural thing to do is therefore to introduce a spin-spin interaction term in the Hamiltonian describing these bound states, akin to the hyperfine splitting term in the hydrogen atom proportional to $\vec{\mu}_\mathrm{e} \cdot \vec{\mu}_\mathrm{p} \propto \vec{S}_\mathrm{e} \cdot \vec{S}_\mathrm{p}/m_\mathrm{e}m_\mathrm{p}$.

Baryons

For the baryons, this leads to a Hamiltonian

$$m_1 + m_2 + m_3 + A\left(\frac{\vec{S}_1 \cdot \vec{S}_2}{m_1 m_2} + \frac{\vec{S}_1 \cdot \vec{S}_3}{m_1 m_3} + \frac{\vec{S}_2 \cdot \vec{S}_3}{m_2 m_3}\right). \tag{3.5}$$

The consequences of this assumption for a few specific cases are readily calculated:

1. If all consituent quarks have the same mass m, Eqn. 3.5 reduces to
$$3m + A(\vec{S}_1 \cdot \vec{S}_2 + \vec{S}_1 \cdot \vec{S}_3 + \vec{S}_2 \cdot \vec{S}_3)/m^2.$$
Realizing also that in the absence of orbital momentum, the total spin J, or more precisely $\vec{J}^2$, can be written as
$$\vec{J}^2 = (\vec{S}_1 + \vec{S}_2 + \vec{S}_3)^2 = \vec{S}_1^2 + \vec{S}_2^2 + \vec{S}_3^2 + 2(\vec{S}_1 \cdot \vec{S}_2 + \vec{S}_1 \cdot \vec{S}_3 + \vec{S}_2 \cdot \vec{S}_3),$$
and with $\vec{S}^2_{1,2,3} = \frac{3}{4}$, we find that
$$(\vec{S}_1 \cdot \vec{S}_2 + \vec{S}_1 \cdot \vec{S}_3 + \vec{S}_2 \cdot \vec{S}_3) = \begin{cases} \frac{3}{4} & \text{for } J = \frac{3}{2} \\ -\frac{3}{4} & \text{for } J = \frac{1}{2} \end{cases}$$
From this (and with $m_\mathrm{u} \approx m_\mathrm{d}$) , it follows that
$$\begin{aligned} m_{\mathrm{n,p}} &= 3m_\mathrm{u} - \frac{3}{4}\frac{A}{m_\mathrm{u}^2}, \\ m_\Delta &= 3m_\mathrm{u} + \frac{3}{4}\frac{A}{m_\mathrm{u}^2}, \\ m_\Omega &= 3m_\mathrm{s} + \frac{3}{4}\frac{A}{m_\mathrm{s}^2}. \end{aligned}$$

2. In the case of the $J = 3/2$ decuplet, all spins are parallel; therefore we must have
$$(\vec{S}_i + \vec{S}_j)^2 = \frac{3}{2} + 2\vec{S}_i \cdot \vec{S}_j = 2 \Rightarrow \vec{S}_i \cdot \vec{S}_j = \frac{1}{4}.$$
From this, we find that
$$\begin{aligned} m_{\Sigma^*} &= 2m_\mathrm{u} + m_\mathrm{s} + \frac{A}{4}\left(\frac{1}{m_\mathrm{u}^2} + \frac{2}{m_\mathrm{u} m_\mathrm{s}}\right), \\ m_{\Xi^*} &= m_\mathrm{u} + 2m_\mathrm{s} + \frac{A}{4}\left(\frac{1}{m_\mathrm{s}^2} + \frac{2}{m_\mathrm{u} m_\mathrm{s}}\right). \end{aligned}$$

3. A harder case is that of the Λ and the Σ. But the Λ has isospin 0, meaning that its wavefunction is antisymmetric under the exchange of the u and d quarks (irrespective of their particle label). This means that also its spin wavefunction should be antisymmetric under the exchange of these two quarks, and hence $\vec{S}_\mathrm{u} \cdot \vec{S}_\mathrm{d} = -\frac{3}{4}$. For the Σ, with $I = 1$, the argument works just the other way around, so for the Σ^0 have $\vec{S}_\mathrm{u} \cdot \vec{S}_\mathrm{d} = \frac{1}{4}$. We therefore find a mass
$$\begin{aligned} & 2m_\mathrm{u} + m_\mathrm{s} + A\left(\frac{\vec{S}_\mathrm{u} \cdot \vec{S}_\mathrm{d}}{m_\mathrm{u}^2} + \frac{(\vec{S}_\mathrm{u} + \vec{S}_\mathrm{d}) \cdot \vec{S}_\mathrm{s}}{m_\mathrm{u} m_\mathrm{s}}\right) \\ = \; & 2m_\mathrm{u} + m_\mathrm{s} + A\left(\frac{\vec{S}_\mathrm{u} \cdot \vec{S}_\mathrm{d}}{m_\mathrm{u}^2} + \frac{(\vec{S}_\mathrm{u} + \vec{S}_\mathrm{d} + \vec{S}_\mathrm{s})^2 - (\vec{S}_\mathrm{u} + \vec{S}_\mathrm{d})^2 - \vec{S}_\mathrm{s}^2}{2m_\mathrm{u} m_\mathrm{s}}\right) \\ = \; & \begin{cases} 2m_\mathrm{u} + m_\mathrm{s} + \frac{A}{4}\left(-\frac{3}{m_\mathrm{u}^2}\right) & (\Lambda) \\ 2m_\mathrm{u} + m_\mathrm{s} + \frac{A}{4}\left(\frac{1}{m_\mathrm{u}^2} - \frac{4}{m_\mathrm{u} m_\mathrm{s}}\right) & (\Sigma) \end{cases} \end{aligned}$$

Clearly, with only three parameters (two quark masses and the interaction strength parameter A) and with a substantial number of baryons, there is an overconstrained system of equations, and the consistency of this model can be checked. With $m_{\mathrm{u}} = 308$ MeV, $m_{\mathrm{s}} = 482$ MeV, and $A = 0.0225$ GeV^3, we find the results in Table 3.3. The agreement is not perfect, but qualitatively the observed masses are reproduced well enough. (Remember that at this point, we do not know yet what interaction binds the quarks into baryons, so asking for perfect agreement is not very reasonable. But detailed – and complicated – computations of hadron masses have meanwhile been made in the framework of *Lattice Gauge Theory*, and they reproduce the observed spectrum of light hadrons very well.)

Particle	Predicted	Measured
n, p	0.89	0.939
Λ	1.08	1.116
Σ	1.15	1.193
Ξ	1.32	1.318
Δ	1.07	1.232
Σ^*	1.34	1.385
Ξ^*	1.50	1.533
Ω^-	1.68	1.673

Table 3.3: Baryon octet and decuplet predicted and measured masses (in GeV)

Mesons

The case of the mesons is easier. With the same basic assumptions we arrive at the following formula for the meson mass:

$$M(q\bar{q}) = m_q + m_{\bar{q}} + A' \frac{\vec{S}_q \cdot \vec{S}_{\bar{q}}}{m_q m_{\bar{q}}}. \tag{3.6}$$

With only two particles, the spin-spin interaction term is easily evaluated to be $-3/4$ for the $J^P = 0^-$ octet, and $1/4$ for the $J^P = 1^-$ octet. The only difficulty is with the fact that some of the mesons are superpositions of quark-antiquark states. For these cases, the operator nature of Eqn. 3.6 has to be exploited. For instance, for the η we have

$$\begin{aligned} m(\eta) &= \frac{1}{6}\langle \mathrm{u\bar{u}} + \mathrm{d\bar{d}} - 2\mathrm{s\bar{s}} | M(q\bar{q}) | \mathrm{u\bar{u}} + \mathrm{d\bar{d}} - 2\mathrm{s\bar{s}} \rangle \\ &= \frac{1}{6}\left(2(2m_{\mathrm{u}} - \frac{3}{4}\frac{A'}{m_{\mathrm{u}}^2}) + 4(2m_{\mathrm{s}} - \frac{3}{4}\frac{A'}{m_{\mathrm{s}}^2}) \right). \end{aligned}$$

With the same quark masses, but a different interaction strength parameter $A' \approx 0.06$ GeV^3, we find the results as presented in Table 3.4. The same conclusion holds: there is a good (qualitative) agreement.

Particle	Predicted	Measured
π	0.15	0.137
K	0.46	0.496
η	0.57	0.549
ρ	0.77	0.770
ω	0.77	0.782
K^*	0.87	0.892
ϕ	1.03	1.020

Table 3.4: Predicted and measured masses (in GeV) for the pseudoscalar and vector meson octets

Baryon magnetic dipole moments

With the static quark model in hand, we are now in a position to say more about the "odd" values for the proton and neutron magnetic dipole moments. These, as well as those of many *nuclei*, have been determined in a nuclear-physics context. With the definitions

$$\vec{\mu}_{\mathrm{n,p}} \equiv g_{\mathrm{n,p}} \mu_N \vec{S}, \quad \mu_N \equiv \frac{e}{2m_N},$$

i.e., defining the *nuclear magneton* in analogy with the Bohr magneton, we find that $g_{\mathrm{p}} \approx 5.6$ and $g_{\mathrm{n}} \approx -3.8$.

We will start by assuming that the quarks behave as electrons, *i.e.*, that they have a Landé factor $g = 2$, and that therefore the quark level magnetic moments are given by

$$\vec{\mu}_q = Q_q \frac{e}{m_q} \vec{S}_q = g Q_q \frac{m_N}{m_q} \mu_N \vec{S}_q \equiv g \mu_q \vec{S}_q.$$

The nucleon's magnetic moment can then be computed from

$$\vec{\mu}_{N,z} = g \sum_i \langle N, S_z = \frac{1}{2} | \mu_{i,z} | N, S_z = \frac{1}{2} \rangle.$$

Now it can be shown that the proton's spin$\otimes$flavour state can be described as

$$|\mathrm{p}, S_z = \frac{1}{2}\rangle = \frac{1}{3\sqrt{2}} \left(2|\mathrm{u}(\uparrow)\mathrm{u}(\uparrow)\mathrm{d}(\downarrow)\rangle - |\mathrm{u}(\uparrow)\mathrm{u}(\downarrow)\mathrm{d}(\uparrow)\rangle - |\mathrm{u}(\downarrow)\mathrm{u}(\uparrow)\mathrm{d}(\uparrow)\rangle\right)$$

(plus permutations with the down quark ending up having particle label 1 or 2; this leads effectively to a factor three in the computation). The result for the first term in the proton wavefunction therefore becomes

$$\begin{aligned} & 3 \cdot \sum_i \mu_i \frac{8}{18} \langle \mathrm{u}(\uparrow)\mathrm{u}(\uparrow)\mathrm{d}(\downarrow) | S_{i,z} | | \mathrm{u}(\uparrow)\mathrm{u}(\uparrow)\mathrm{d}(\downarrow)\rangle \\ = \quad & 3 \cdot \frac{4}{9} (2 \cdot \frac{1}{2} \mu_{\mathrm{u}} - \frac{1}{2} \mu_{\mathrm{d}}) = 3 \cdot \frac{2}{9} (2\mu_{\mathrm{u}} - \mu_{\mathrm{d}}), \end{aligned}$$

while for the second and third terms it becomes $3 \cdot \frac{1}{18}\mu_{\mathrm{d}}$. The final result is therefore $\frac{4}{3}\mu_{\mathrm{u}} - \frac{1}{3}\mu_{\mathrm{d}}$.

For the neutron, the computation proceeds in the same way, merely with the up and down quarks interchanged, *i.e.*, $\frac{4}{3}\mu_{\mathrm{d}} - \frac{1}{3}\mu_{\mathrm{u}}$. Now, our assumption above predicts $\mu_{\mathrm{d}} = -\frac{1}{2}\mu_{\mathrm{u}}$ (when we assume $m_{\mathrm{d}} = m_{\mathrm{u}}$, as earlier). Inserting this in the expression for the proton and neutron magnetic moments, we obtain a prediction for their ratio: $g_{\mathrm{n}}/g_{\mathrm{p}} = -2/3$, in excellent agreement with the observed ratio of -0.68.

3.1.8 Heavy-Quark Hadrons

It would take far too much time to say here all there is to say about hadrons. Suffice it to say that the static quark model, as arrived at above, is the starting point for more thorough investigations of (in particular) the proton and the neutron, which we will get to in the next section.

Before leaving the topic of quark-(anti)quark bound states, however, it is useful to point out that yet other quarks exist besides u, d, and s. The J/ψ meson, with a mass of ~ 3.1 GeV, was discovered in 1973 through its decay to $\mathrm{e}^+\mathrm{e}^-$ or $\mu^+\mu^-$ final states (or alternatively, the total cross section for $\mathrm{e}^+\mathrm{e}^-$ scattering exhibiting a strongly resonant behaviour at $\sqrt{s} \approx 3.1$ GeV). Despite its high mass it has a lifetime much too long to decay via the strong interaction (its total decay width is about 93 keV, to be compared with the tens to hundreds of MeV for lighter but strongly decaying hadrons), and it was rapidly interpreted as a $\mathrm{c\bar{c}}$ bound state, the c or *charm* quark being a new quark of substantially higher mass, $m_{\mathrm{c}} \approx 1.5$ GeV. Like for the lighter quarks, various combinations of c quarks with lighter quarks are possible (they are called D mesons), as well as baryons containing c quarks. The topic of *heavy quark spectroscopy* has taught us a lot about the strong interaction, and how it binds quarks in hadrons.

The same thing happened again in 1979, when the Υ particle was discovered, with a mass of ~ 9.5 GeV and interpreted as a $\mathrm{b\bar{b}}$ bound state, the b or *bottom* quark being about 4.75 GeVheavy. The same phenomena (B mesons containing a b quark and a lighter antiquark, baryons containing a b quark) have been observed. The B mesons in particular play an important role in present studies of the so-called *CP violation* in weak interactions, which we will touch upon in the next chapter.

The t or *top* quark was discovered in 1995. Its mass is sufficiently high (about 172 GeV) that it does not lead to bound states.

3.2 Dynamics of the Proton

All of the preceding material may be persuasive, but it can hardly be considered as actual *proof* for the existence of quarks. The task, therefore, is to devise an experiment that can demonstrate more directly their existence.

If quarks are indeed to be found as constituents of hadrons (and in the remainder of the section, we will mostly be limiting ourselves to a discussion of the *proton*), this would imply a situation reminiscent of that of nuclei (and upon further study, protons and neutrons) as constituents of atoms, as demonstrated by Rutherford's scattering experiments. This analogy is also

of direct use in that it makes a definite suggestion for the search for proton substructure, namely, the use of electron beams scattering off protons.

3.2.1 Elastic Electron-Proton Scattering

The first such studies were made in the 1950's (by McAllister and Hofstadter, using a 188 MeV electron beam [6]), and concentrated on *elastic* electron-proton scattering. Its Feynman diagram is shown in Fig. 3.7.

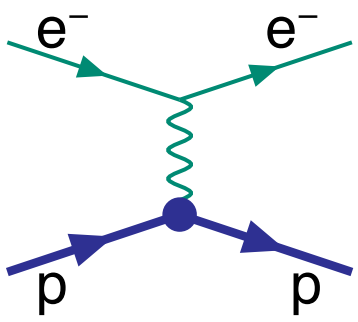

Figure 3.7: Feynman diagram for electron-proton elastic scattering

The formalism for the description of this scattering process (where the incoming electron, of energy E, is scattered through an angle θ off the proton which is at rest) is a reasonably straightforward extension of the Rutherford scattering formula,

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\text{Rutherford}} = \frac{\alpha^2}{16E^2\sin^4(\theta/2)}|F(q^2)|^2,$$

where the form factor $F(q^2)$ is the quantity that provides the actual information about the charge distribution of the object under study (more precisely, it is the Fourier transform of the charge distribution), and q^μ is the four-momentum of the photon exchanged in the process. Two improvements must be made on the Rutherford scattering formula:

1. the proton is not infinitely heavy (the mere fact that E/m_p is not very small tells us that it would be wrong to ignore this fact), but instead will receive a nonzero momentum. As a result, also the final-state electron energy is not equal to E anymore, but is rather given by

$$E' = \frac{E}{1+(2E/m_\mathrm{p})\sin^2(\theta/2)}. \tag{3.7}$$

   This leads to an extra factor E'/E in the cross section formula. Also, the expression for q^2 now necessarily also involves also the energy transfer between the electron and the proton, and (upon neglecting terms proportional to m_e/E) takes the form

$$q^2 = -4EE'\sin^2(\theta/2); \tag{3.8}$$

2. the proton is a spin-1/2 particle. As a result, the photon interacts both with the proton's *charge* and with its *magnetic moment*, and two form factors are now required to describe the interaction.

The result of this is the *Rosenbluth formula*:

$$\left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right)_{\text{Rosenbluth}} = \frac{\alpha^2}{16 m_\mathrm{p}^2 E^2 \sin^4(\theta/2)} \frac{E'}{E} \left(2K_1(q^2)\sin^2(\theta/2) + K_2(q^2)\cos^2(\theta/2)\right).$$

Here, the factors K_1 and K_2 are related to the form factors G_E and G_M for the specific couplings to the proton charge and magnetic moment, respectively. For completeness:

$$\begin{aligned} K_1(q^2) &= -q^2 G_M^2, \\ K_2(q^2) &= 4m_\mathrm{p}^2 \frac{G_E^2 - \left(q^2/4m_\mathrm{p}^2\right) G_M^2}{1 - \left(q^2/4m_\mathrm{p}^2\right)}. \end{aligned}$$

For a pointlike proton, $G_E = G_M = 1$. A measurement of this process indicated clearly the finite size of the proton, $r_\mathrm{p} \approx 0.7$ fm.

3.2.2 Deep-Inelastic Scattering

Electrons of 188 MeV energy do not really allow to investigate the proton in other ways than by elastic scattering. But given that the proton is not a fundamental particle, one would hope that a more energetic probe would indeed allow to see the proton's constituents.

Progress in accelerator technology made such studies possible towards the end of the 1960's, when experiments with 20 GeV electrons could be carried out. The beauty of these experiments is that even without measuring what happens to the proton (and merely measuring the outgoing electron), it is possible to make quantitative statements about the proton. The relevant diagram is given in Fig. 3.8. It looks much like that in Fig. 3.7, aside from the fact that the proton is not necessarily left intact.

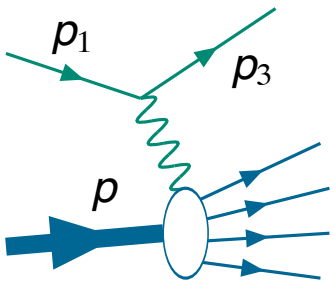

Figure 3.8: Feynman diagram for inelastic electron-proton scattering

That something qualitatively new happens, is easily seen in Fig. 3.9. The observed differential cross section is normalized to that for Mott scattering (as appropriate for the description

of the proton as a fundamental but infinitely heavy spin-1/2 particle). Here, the form factors as determined through the elastic cross section measurements of the preceding section have been used to predict the behaviour if *only* elastic scattering processes were taking place. This is clearly not the case!

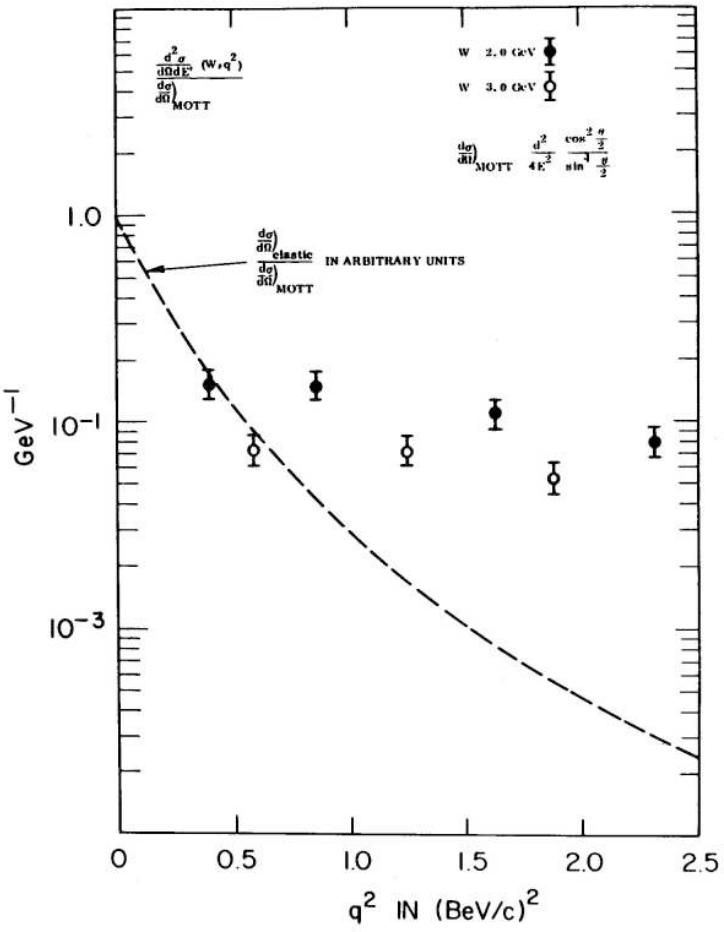

Figure 3.9: Differential cross section observed in inelastic electron-proton scattering, normalized to the cross section for Mott scattering. (BeV is an old name for GeV.)

Let us now try to understand this scattering process from a theoretical point of view, starting with two important observations:

1. if the proton is not left intact in this scattering process, Eqn. 3.7 does not hold anymore. Instead, E' and θ are to be regarded as *independent* kinematic variables;

2. from Fig. 3.8 it is obvious that part of this process, namely the coupling of the electron the photon, is well understood: this is simply QED.

This second observation is most conveniently exploited in the case of *unpolarized* electron beams. In this case, (spin-1/2) QED calculations tell us that the squared matrix element for

any hadronic final state will contain (for fixed p_3^μ; the four-momenta indicated in Fig. 3.8 are used here) a factor

$$L_{\mu\nu} \equiv 2\left(p_{1\mu}p_{3\nu} + p_{1\nu}p_{3\mu} + g_{\mu\nu}(m_\mathrm{e}^2 - p_1 \cdot p_3)\right).$$

This factor is called the *lepton tensor*.

To make up for the lepton tensor's Lorentz indices, the hadronic part must be of the structure $K^{\mu\nu}$ and may depend on various kinematic variables. But here we can exploit the fact that we don't need the specifics of the hadronic final state: this is tantamount to *integrating* over all hadronic degrees of freedom, and *summing* over all hadronic final states. The result then must still carry the same Lorentz indices, but it can only depend on the degrees of freedom not integrated out, *i.e.*, p^μ and q^μ! (Note that the proton "only" sees the photon, so the individual electron four-momenta are not relevant for the description of the hadronic part of the interaction.) In summary, the spin-averaged squared matrix element must look like

$$\langle|\mathscr{M}|^2\rangle = \frac{\alpha^2}{q^4} L_{\mu\nu} W^{\mu\nu}, \tag{3.9}$$

where the factor α^2 accounts for the photon coupling (even if the photon's interaction with the proton constituents is not known, we surmise that it has an interaction strength $\sim e$), and the factor $1/q^4 \equiv 1/(q^2)^2$ accounts for the photon propagator.

Considering the tensor $W^{\mu\nu}$ in more detail, a few more features are easily realized:

- given p^μ and q^μ as only remaining kinematic degrees of freedom, only they can be used to build independent terms (in addition to the metric tensor $g^{\mu\nu}$);
- the lepton tensor $L_{\mu\nu}$ is symmetric under the interchange of the indices μ and ν. That means that terms in $W^{\mu\nu}$ *anti*symmetric under this interchange will not give any contribution.

In addition, charge conjugation symmetry in QED requires that $q_\mu W^{\mu\nu} = q_\nu W^{\mu\nu} = 0$ (the proof of this statement is nontrivial, and we will not venture into it). As a consequence, $W^{\mu\nu}$ can be expressed as

$$W^{\mu\nu} = W_1 \cdot \left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\right) + \frac{W_2}{m_\mathrm{p}^2} \cdot \left(p^\mu - \left(\frac{p\cdot q}{q^2}\right) q^\mu\right)\left(p^\nu - \left(\frac{p\cdot q}{q^2}\right) q^\nu\right),$$

where the only kinematic dependence of $W_{1,2}$ can be on Lorentz *scalars*. But the only possible Lorentz scalars are the ones constructed from p^μ and q^μ, *i.e.*, p^2, q^2, and $p\cdot q$. Of these, $p^2 = m_\mathrm{p}^2$, so it isn't a variable. Hence only q^2 and $p\cdot q$ are left. In practice, the variables used are

$$Q^2 \equiv -q^2 \quad \text{and} \quad x \equiv \frac{Q^2}{2p\cdot q}. \tag{3.10}$$

It is evident that the mere requirement of Lorentz covariance imposes important constraints on $W^{\mu\nu}$! (Note that in the case of *elastic* scattering, the proton is left intact by definition. The

on-shell proton in the final state implies that $(p+q)^2 = p^2 + q^2 + 2p\cdot q = p^2 = m_\mathrm{p}^2$, and hence $q^2 + 2p\cdot q = 0$, implying $x = 1$.)

From here on, it is merely a matter of contracting these tensors and computing the results in the lab frame (in which the proton is at rest initially). It is a straightforward but tedious calculation to show that the (doubly differential) cross section becomes

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E'\mathrm{d}\Omega} = \frac{\alpha^2}{4E^2\sin^4(\theta/2)}\left(2W_1(x,Q^2)\sin^2(\theta/2) + W_2(x,Q^2)\cos^2(\theta/2)\right). \tag{3.11}$$

This expression implies that W_1 and W_2 can be measured separately by studying this differential cross section for different beam energies E. (This was not done in the first experiment; instead, it was assumed that W_1 was not vastly larger than W_2. Within that approximation, a measurement at small angles directly translates to a measurement of W_2, given the angular terms they multiply.)

The preceding does not do much more than set a theoretical *framework*, the "physics" of the photon-proton interaction being buried inside the $W_{1,2}$. However, it can be shown on theoretical grounds [7] that *if* the proton consists of point particles, then for asymptotically high values of Q^2 (this is the origin of the qualification "deep" in deep-inelastic scattering), $W_{1,2}$ should exhibit a simple "scaling" behaviour:

$$\begin{aligned} \lim_{Q^2\to\infty} m_\mathrm{p} W_1(x,Q^2) &= F_1(x), \\ \lim_{Q^2\to\infty} \frac{Q^2}{2m_\mathrm{p}x} W_2(x,Q^2) &= F_2(x), \end{aligned} \tag{3.12}$$

where $F_{1,2}(x)$ are called the proton's *structure functions*. This "Björken scaling" behaviour was indeed borne out by the first such experiment, as shown in Fig. 3.10!

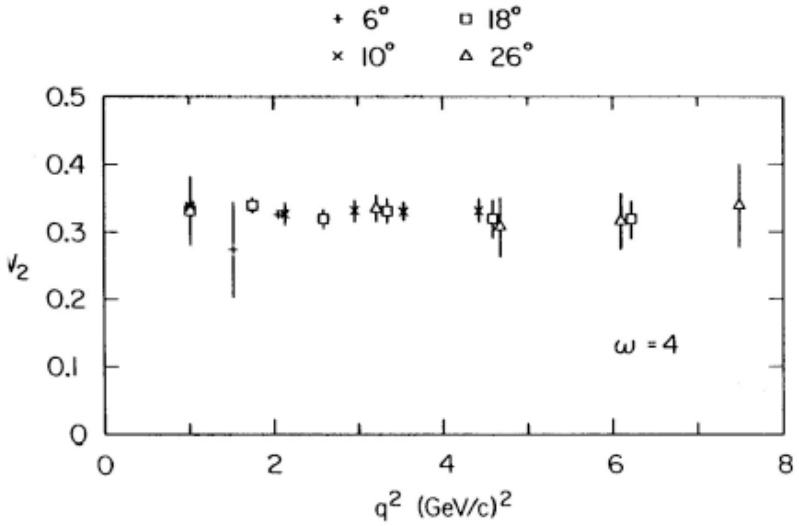

Figure 3.10: Behaviour of $\frac{Q^2}{2m_\mathrm{p}x}W_2(x,Q^2)$ as a function of Q^2, at a value of $x = 1/\omega = 0.25$.

But if the photon indeed interacts with point particles inside the proton, and according to the static quark model all hadrons consist of $q\bar{q}$ or qqq combinations, it becomes very reasonable to assume that the point particles observed in DIS are indeed these quarks! In that case, it becomes very reasonable to assume that their interaction with the photon is in fact completely specified, namely, just like that of the electron, but with different coupling strenghts due to their different (non-integer) electrical charge.

This, then, leads to an alternative description of the DIS process: a quark (or more generally, *parton*, after Feynman; we will get to the difference between them shortly. For now, it suffices to say that quarks are partons) carrying a *four-momentum fraction* x of that of the proton is scattered *elastically* by the electron. This "parton model" picture is shown Feynman diagrammatically in Fig. 3.11. Here, we have used the same variable x as earlier on in this section, *i.e.*, as one of the

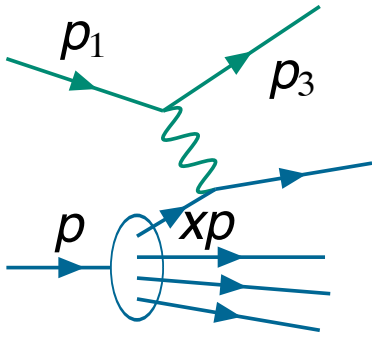

Figure 3.11: Deep-inelastic electron-proton scattering in the parton model.

two independent kinematic variables describing DIS. This is no coincidence, for the assumption of an *elastic* electron-parton scattering process presupposes that the parton is on-shell in both the initial and the final state. Therefore,

$$(xp+q)^2 = x^2p^2 + 2xp\cdot q + q^2 = (xp)^2 = x^2p^2 \Rightarrow 2xp\cdot q + q^2 = 0 \Rightarrow x = \frac{-q^2}{2p\cdot q}.$$

This means that we have found a physical interpretation for x (or alternatively, this tells us why it is convenient to work with x and Q^2 rather than some combination thereof). (We will not worry about the fact that this actually gives rise to a *variable* quark mass.)

The remaining ingredient required to describe the DIS process completely in the framework of the parton model is the *parton density function* or PDF $f_i(x)$: this tells us how many partons of type i are "available" at momentum fraction x. The squared matrix element of Eqn. 3.9 must therefore be replaced with

$$\frac{\alpha^2}{q^4}\sum_i Q_i^2 f_i(x) L_{\mu\nu}K^{\mu\nu},$$

where $K^{\mu\nu}$ is a copy of the lepton tensor, but with the *quark* initial- and final-state four-momenta rather than the electron ones.

Unsurprisingly, this complete specification implies a relation between $F_1(x)$ and $F_2(x)$. This relation turns out to be

$$F_2(x) = 2xF_1(x)$$

and is called the *Callan-Gross relation*. It is in good agreement with the experimental measurements.

The double differential cross section of Eqn. 3.11 becomes

$$\frac{\mathrm{d}\sigma}{\mathrm{d}E'\mathrm{d}\Omega} = \frac{\alpha^2}{E^2\sin^2(\theta/2)}\frac{F_1(x)}{2m_\mathrm{p}}\left(1+\frac{2EE'}{(E-E')^2}\cos^2(\theta/2)\right) \quad \text{with} \quad F_1(x) = \frac{1}{2}\sum_i Q_i^2 f_i(x). \tag{3.13}$$

So we see that a measurement of this cross section provides us with information on $F_1(x)$ (or equivalently, $F_2(x)$, which is more usually quoted), and hence of a combination of quark PDFs weighted by their (squared) charges.

Now *if* the static quark model is indeed "all there is" about the proton, it is to be expected that the proton's three quarks carry a fixed momentum fraction (maybe even 1/3, due to the fact that up and down quarks have very nearly the same mass?), perhaps smeared somewhat by the quarks' movement inside the proton. This is certainly *not* what is found! Instead, the whole continuum $0 < x < 1$ is covered.

Another testable consequence of this "naive" picture would be the expectation that

$$\int_0^1 x\mathrm{u}(x)\mathrm{d}x = 2\int_0^1 x\mathrm{d}(x)\mathrm{d}x$$

(here, $\mathrm{u}(x)$ is a shorthand for $f_\mathrm{u}(x)$, *etc.*), again because the up and down quarks have very nearly the same mass, and there are two up quarks in the proton, versus one down quark. As a consequence, it would follow that

$$\int_0^1 F_2(x)\mathrm{d}x = \int_0^1 x(\frac{4}{9}\mathrm{u}(x)+\frac{1}{9}\mathrm{d}(x))\mathrm{d}x = \int_0^1 x\mathrm{d}(x)\mathrm{d}x = \frac{1}{3}, \tag{3.14}$$

given the presupposed equal distribution of the momenta over the three quarks. Also this is not borne out by the results: experimentally, the integral is measured to be ~ 0.18, well below expectations. Clearly the static quark model is not sufficient to describe the dynamics of the proton.

3.3 Colour Revisited: Quantum Chromodynamics

At this point, it is time to take seriously the consequences of the colour hypothesis made in Sect. 3.1.6. In hindsight, this leads to a surprisingly "simple" interaction, but one with profound consequences.

3.3.1 QCD: the Theory of Quarks and Gluons

The breakthrough is to consider the strong interaction between quarks as an interaction between their *colour charges*. As said, the internal symmetry group associated with this interaction is $SU(3)$. All that needs to be done is to promote it to a *gauge* symmetry, just as QED was accomplished by applying the principle of minimal substitution to the electromagnetic interaction.

The result is *Quantum Chromodynamics* or QCD. The QCD gauge bosons are called *gluons*, and the $SU(3)$ symmetry dictates that there are eight of them. One important difference compared to QED is that the gluons themselves also carry a colour charge. In terms of the three-dimensional complex colour space, they can be represented (up to a common normalization factor $1/\sqrt{2}$) by the so-called *Gell-Mann matrices*, which form the $SU(3)$ equivalent of the Pauli matrices encountered for $SU(2)$:

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad \lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \qquad \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{3.15}$$
$$\lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \qquad \lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

The QCD interaction strength is expressed in terms of a new coupling constant, called g_s, or more usually in terms of $\alpha_s \equiv g_s^2/4\pi$ (analogously to the QED case). The Feynman rules for QCD involve the propagators for the (now coloured) quarks and for the gluons, as is to be expected. The *physics* of QCD is in the couplings between these quarks and gluons, and the QCD vertices are shown in Fig. 3.12. One thing that is immediately obvious from these vertices is that there

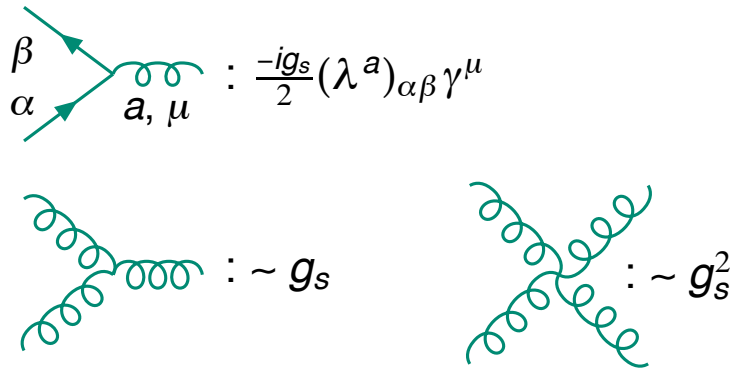

Figure 3.12: Vertices in QCD. In the case of the qqg vertex, all Lorentz and colour indices are given explicitly. Doing so for the other vertices would require delving deeper into group theory than is desirable here.

are other vertices besides just the coupling of a gluon to quarks. This difference is due to the fact that the gluons themselves carry colour charges, as already stated above, and it has profound consequences, as we will see next.

QCD's SU(3) symmetry

When promoting QCD to be a gauge symmetry, this implies that the "physics" of the colour interaction must be invariant under transformations

$$\psi(x) \to \psi'(x) = U(x)\psi(x), \tag{3.16}$$

where $\psi(x)$ represents a wavefunction or field with three components in colour space, and $U(x)$ is an arbitrary spacetime dependent $SU(3)$ transformation in this colour space.

Because this "physics" involves equations of motion, and hence derivatives of the fields, it becomes necessary again to introduce a covariant derivative, akin to the one introduced for QED in Eqn. 2.25. This time, however, it must look somewhat different to account for the fact the QCD implements an $SU(3)$ rather than a $U(1)$ symmetry: we have

$$D_\mu(x) = \partial_\mu + i\frac{g_s}{2}\lambda_a G_{\mu a}(x), \tag{3.17}$$

involving an implicit summation over the Gell-Mann matrices of Eqn. 3.15, which *generate* the $SU(3)$ symmetry group; the corresponding gluon fields are denoted $G_{\mu a}(x)$. (The factor 1/2 is an artifact related to the precise definition of the coupling constant g_s.)

Analogously to the case of QED (see Eqn. 2.20), also the gluon fields transform under the gauge transformation of Eqn. 3.16. However, the details differ. It turns out (we won't justify this) that under the gauge transformation we must have

$$D_\mu(x)\psi(x) \to D'_\mu(x)\psi'(x) = U(x)D_\mu(x)\psi(x) = \left(U(x)D_\mu(x)U^{-1}(x)\right)\psi'(x),$$

implying that

$$\begin{aligned} U(x)D_\mu(x)U^{-1}(x) &= U(x)\left(\partial_\mu + i\frac{g_s}{2}\lambda_a G_{\mu a}(x)\right)U^{-1}(x) \\ &= \partial_\mu + i\frac{g_s}{2}U(x)\lambda_a G_{\mu a}(x)U^{-1}(x) + U(x)\left(\partial_\mu U^{-1}(x)\right) \\ &= \partial_\mu + i\frac{g_s}{2}\lambda_a G'_{\mu a}(x). \end{aligned} \tag{3.18}$$

This is satisfied if

$$i\frac{g_s}{2}\lambda_a G'_{\mu a}(x) = i\frac{g_s}{2}U(x)\lambda_a G_{\mu a}(x)U^{-1}(x) + U(x)\left(\partial_\mu U^{-1}(x)\right).$$

By itself this expression may not look very instructive, so let us write it down specifically for *infinitesimal* transformations $U(x) \approx \mathbb{1} - i\alpha_a(x)\lambda_a(x)$. Writing the result to first order in the infinitesimal parameters $\alpha_a(x)$, we find that

$$\begin{aligned} \frac{g_s}{2}\lambda_a G'_{\mu a}(x) &\approx \frac{g_s}{2}\lambda_a G_{\mu a}(x) - i\alpha_a(x)[\lambda_a,\lambda_b]G_{\mu b}(x) + \partial_\mu\alpha_a(x)\lambda_a \\ &= \frac{g_s}{2}\lambda_a G_{\mu a}(x) + f_{abc}\lambda_c\alpha_a G_{\mu b}(x) + \partial_\mu\alpha_a(x)\lambda_a. \end{aligned}$$

In this result, we can recognise ingredients similar to those in Eqn. 2.20. But the non-abelian nature of $SU(3)$ leads to an additional term.

The non-abelian nature of $SU(3)$ is directly responsible for the three-gluon and four-gluon couplings shown in Fig. 3.12. To see this, we first need to construct the field tensor (in

analogy to Eqn. 2.19 for the QED case). Again, we lack the machinery for a full justification, but this can be constructed as

$$\begin{aligned} G_{\mu\nu a}\lambda_a &= \left(i\frac{g_s}{2}\right)^{-1}[D_\mu, D_\nu] \\ &= \frac{-2i}{g_s}[\partial_\mu + i\frac{g_s}{2}\lambda_b G_{\mu b}(x), \partial_\nu + i\frac{g_s}{2}\lambda_c G_{\nu c}(x)] \\ &= \lambda_a(\partial_\mu G_{\nu a}(x) - \partial_\nu G_{\mu a}(x)) + i\frac{g_s}{2}[\lambda_b, \lambda_c]G_{\mu b}(x)G_{\nu c}(x) \\ &= \lambda_a(\partial_\mu G_{\nu a}(x) - \partial_\nu G_{\mu a}(x) - \frac{g_s}{2}f_{abc}G_{\mu b}(x)G_{\nu c}(x)). \end{aligned} \tag{3.19}$$

Note that while D_μ is an operator, $G_{\mu\nu a}(x)$ is not. The same recipe can be used also to construct the field tensor for QED, leading to the expression 2.19 that in that case was derived from the Maxwell equations. Note also that $G_{\mu\nu a}(x)$ is not in fact gauge invariant, unlike QED's $A_{\mu\nu}(x)$! For using Eqn. 3.18 it is easy to see that under a gauge transformation,

$$[D_\mu, D_\nu] \to [D'_\mu, D'_\nu] = [U(x)D_\mu U^{-1}(x), U(x)D_\nu U^{-1}(x)] = U(x)[D_\mu, D_\nu]U^{-1}(x).$$

But this is to be expected, since the gluons themselves are coloured!

Finally, we quote the equation of motion for the gluon fields, which again looks more complicated than the QED case:

$$\partial^\mu G_{\mu\nu a}(x) - \frac{g_s}{2}f_{abc}G^\mu_b(x)G_{\mu\nu c}(x) = \frac{g_s}{2}\sum_f \bar{\psi}_f(x)\gamma_\nu\lambda_a\psi_f(x).$$

It doesn't take much effort to realise that this equation involves terms quadratic and cubic in the gluon fields. These terms correspond directly to the three- and four-gluon vertices of Fig. 3.12. The r.h.s. provides the QCD "current", also in analogy with Eqn. 2.19; the sum is over the quark flavours.

3.3.2 Running Coupling Constants

In Chapter 2, we discussed the concept of *radiative corrections*, which is simply related to the fact that to obtain the full transition amplitude for a given process, *all* Feynman diagrams featuring the same initial and final states must be accounted for. Returning to QED for a moment, and accounting for the first correction to the interaction between two electrons, the Feynman diagrams of Fig. 3.13 are obtained[1]. This process leads to a simple modification of the lowest-order amplitude which can be absorbed into an effective or *running* coupling constant. For $Q^2 \gg m_e^2$, its behaviour is given by

$$\alpha \to \alpha(Q^2) = \alpha(0)\left(1 + \frac{\alpha(0)}{3\pi}\ln(|Q^2|/m_e^2)\right).$$

[1] Actually, with our usual convention of time running from left to right, this diagram describes electron-positron elastic scattering. But the diagrams can be "crossed" (or alternatively, time assumed to be running from bottom to top) to obtain the said interaction.

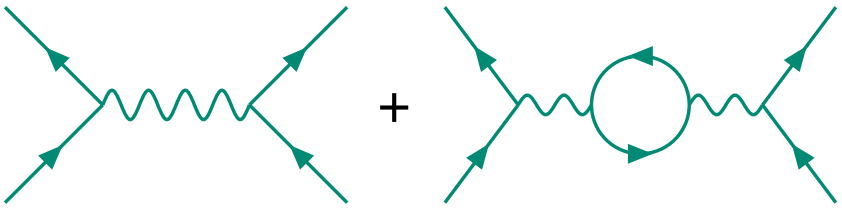

Figure 3.13: Lowest-order and vacuum polarization diagrams contributing to the process $e^+e^- \to e^+e^-$.

(Note that the procedure of "merely" adding these two diagrams is not sufficient: the result *diverges*, and only by a re-interpretation of parameters, a procedure called renormalization, do we end up at the stated result. We will not discuss this topic except very qualitatively.)

But if we need to include this diagram with one "vacuum bubble", so should we include that with two bubbles, *etc*. The result is an infinite series behaving as

$$1 + x + x^2 + x^3 + \ldots \to \frac{1}{1-x},$$

and hence we find a more correct expression for the running QED coupling:

$$\alpha \to \alpha(Q^2) = \frac{\alpha(0)}{1 - \frac{\alpha(0)}{3\pi}\ln(|Q^2|/m_e^2)}. \tag{3.20}$$

Quantitatively, we find that the value of the effective $\alpha(Q^2)$ rises slightly from its value at $Q^2 = 0$, $\alpha(0) \approx 1/137$, to $\alpha(Q^2 \approx 8 \cdot 10^3\ \text{GeV}^2) \approx 1/128$.

So why is all of this useful? The reason is that exactly the same procedure applies to the strong interaction. But the QCD equivalent of Fig. 3.13 features an additional diagram, as shown in Fig. 3.14 – made possible by the three-gluon vertex. Now it turns out that the contribution

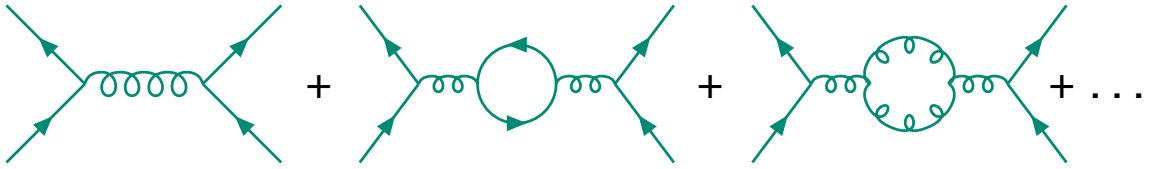

Figure 3.14: Lowest-order and vacuum polarization diagrams contributing to the process $q\bar{q} \to q\bar{q}$.

from the extra diagram (the one featuring the gluon loop) has the opposite sign to that of the quark loop, and in fact is larger. The result, again interpreted in terms of an effective coupling constant, is

$$\alpha_s \to \alpha_s(Q^2) = \alpha_0 \left(1 - (11N_c - 2N_f)\ln(|Q^2|/m^2)\right)$$

(where m becomes some average of the quarks that can circulate in the quark loop). Here, N_c is a formal notation for the number of colours, *i.e.*, $N_c = 3$, and N_f represents the number of

"active" quark flavours and is usually taken to be 4 or 5. So $\alpha_s(Q^2)$ *decreases* with increasing Q^2, contrary to the QED case! Now if we again attempt to resum the infinite series resulting from higher-order corrections, we encounter a problem: the resummed result diverges already for $Q^2 > 0$, namely at

$$1 + \frac{\alpha_s(0)}{12\pi}(11N_c - 2N_f)\ln(Q^2/m^2) = 0.$$

So resumming in this fashion is not meaningful. Fortunately, it turns out that it is indeed possible to express $\alpha_s(Q^2)$ relative to the (known?) value at a different scale, say, μ^2:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \frac{\alpha_s(\mu^2)}{12\pi}(11N_c - 2N_f)\ln(Q^2/\mu^2)}. \tag{3.21}$$

Comparing Eqns. 3.21 and 3.20, the significantly larger value of α_s (coupled with the larger coefficient in the denominator) implies that the Q^2 dependence of α_s is expected to be significantly more pronounced (if the divergence didn't already hint at that). This Q^2 dependence has been *measured*, and the QCD prediction of Eqn. 3.21 is found to be in good agreement with experiment. This is shown in Fig. 3.15.

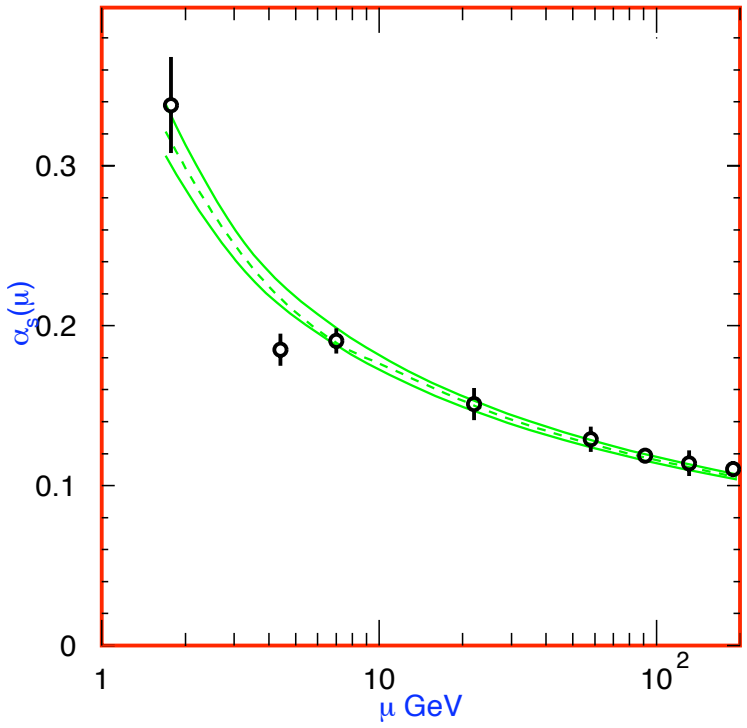

Figure 3.15: Measurements of α_s as a function of Q^2 compared with the best-fit QCD prediction.

This behaviour has two important consequences:

asymptotic freedom: this means that at asymptotically high Q^2, quarks (and gluons) can be considered as free particles (in the same sense that, *e.g.*, electrons are free);

confinement: on the other hand, for small $Q^2 \lesssim 1\ \mathrm{GeV}^2$ $\alpha_s(Q^2)$ becomes so large that even perturbation theory breaks down. This is why we cannot hope to find a perturbative approach suitable for the description of light hadrons as bound states (the situation is a bit easier for hadrons containing a b or c quark, which is why one can meaningfully study the spectroscopy of e.g. charmonium and bottomonium states).

3.3.3 Consequences for Deep-Inelastic Scattering

We are now in a position to understand why the structure functions $F_{1,2}(x)$ are not simply described by the static quark model. When accounting for higher-order corrections to the DIS process, we have to do more than just account for those Feynman diagrams featuring the same initial and final states, simply because our final state is not very well defined (remember that we sum and integrate over all the hadronic final states). As a consequence, the struck quark need not be one of the *valence* quarks (*i.e.*, one of the quarks that define the proton as a proton), but can also be a *sea* quark: these can be thought of as resulting from radiative corrections. The relevant parts of the associated Feynman diagrams of such corrections are given in Fig. 3.16.

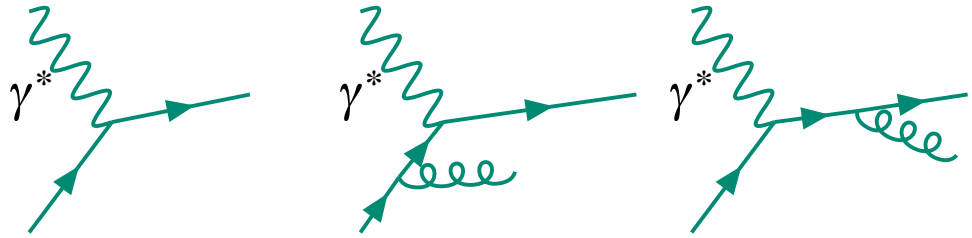

Figure 3.16: QCD radiative corrections in the DIS process. The asterisks emphasize the fact that the photon is *virtual*.

Of course, formally these are higher-order processes, and one may wonder whether it they are important to deal with here. However, in practice it turns out that they give a very sizeable contribution, in particular for *collinear* radiation (in which a single parton is split into two partons that both follow very nearly the direction of the original parton). Now it can be shown (for an explicit calculation see Chapter 10 of [8]) that these additional contributions are proportional to $\ln Q^2$. But this implies that the radiative corrections will also depend on Q^2, leading to *scaling violations* (*i.e.*, a violation of the "Björken scaling hypothesis" of Sect. 3.2.2)! That is, $F_{1,2}$ *do* depend on Q^2 (even if only logarithmically).

All of this is now well understood, and DIS measurements (combined with the measurements of other processes at hadron colliders sensitive to PDFs) are customarily used to extract information about the *individual* partons' PDFs. A summary of electron-proton DIS $F_2(x,Q^2)$ measurements is given in Fig. 3.17. In this plot, the scaling violation is fairly clearly visible. So why was "Björken scaling" thought to hold at first (see Fig. 3.10)? The answer is that (a) only a retricted Q^2 range was considered in these first experiments, and (b) the measurements were done at fairly large x. In particular, at the value $x = 0.25$ shown in Fig. 3.10 the Q^2 dependence is almost absent.

Another important consequence of QCD is that we are now in a position to explain why the expectation of Eqn. 3.14 is not borne out by the data: as a consequence of the higher-order corrections to the DIS process, the photon may well interact with another parton than one of the "initial" quarks. Therefore, the assumption underlying Eqn. 3.14 does not hold anymore.

As a closing word on PDFs, it should be stressed that their knowledge is *essential* for a precise prediction of the cross section for any process at hadron colliders (be it the Tevatron $p\bar{p}$ collider or the LHC proton-proton collider), as will be discussed in Sect. 3.5. So a very detailed knowledge of the proton is of great practical importance!

3.4 Jets

From the previous section, we have concluded that at high Q^2, the partons (quarks, antiquarks, and gluons) in the proton can effectively be regarded as *free* particles. The next question is how the corresponding final state in DIS (featuring a quark or antiquark at high momentum and large scattering angle) is compatible with the existence only of *colourless* hadrons as observable entities at any reasonable time after the interaction?

The answer is given by the effective qq or $q\bar{q}$ potential, which takes the form

$$V(r) \sim -\frac{\alpha_s(r)}{r} \sim -\frac{\alpha_s}{r} + br,$$

where α_s has been re-parametrized as a function of r for the sake of the argument. This formula makes it very clear that the running of α_s is a very important effect! As a consequence, when a quark moves far from its production region, this potential makes it energetically favourable for an intermediate $q\bar{q}$ pair to be created, of which the $\bar{q}$ combines with the quark to form a colourless meson; there remains then a new quark, of lower momentum, which undergoes the same process. This is only stopped when there remain (anti)quarks of sufficiently low relative Q^2 to bind to a hadron. The result is a cascade of hadrons, in which the momentum of the original quark is distributed over a significant number of hadrons traveling more or less in the same direction of the original quark. This group of hadrons is (loosely) called a *jet*.

At this point, several remarks are in order:

- What is described above is called *fragmentation*. Its description is necessarily qualitative, as it is important for relatively low Q^2, a region where perturbation theory does not suffice. However, even if the kinematic distributions of hadrons produced in the fragmentation process cannot be described using perturbation theory, the effects on the jet distributions themselves *are* amenable to perturbation theory. Various *jet reconstruction algorithms*, with properties that allow for a meaningful comparison between theoretical predictions and measurements, have been exploited.

- The above describes well what happens at high-energy interactions. Consequently, jets were seen unambiguously in e^+e^- interactions at $\sqrt{s} = 35$ GeV at the PETRA collider, in hadronic final states. Examples of such events (or at least the charged particles in such events) are given in Fig. 3.18 (In fact, this figure displays events with *three* rather than

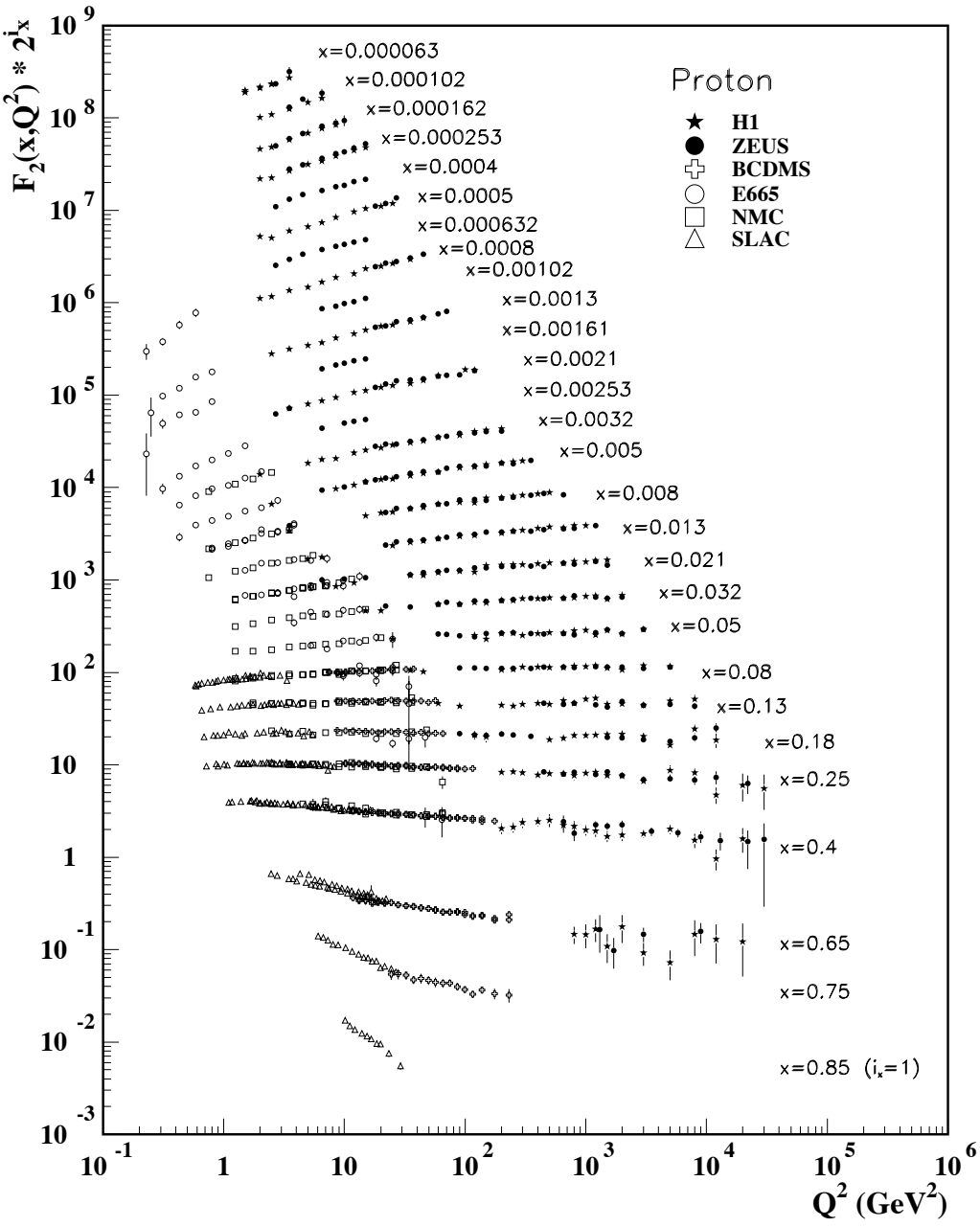

Figure 3.17: Measurements of the proton's $F_2(x, Q^2)$ in deep-inelastic scattering.

two jets. This has a straightforward interpretation in terms of gluon radiation, which may happen at large energy and angle compared to the quarks. Hence these events constituted the first *direct* evidence for the gluon). However, the first signs of jet-like features

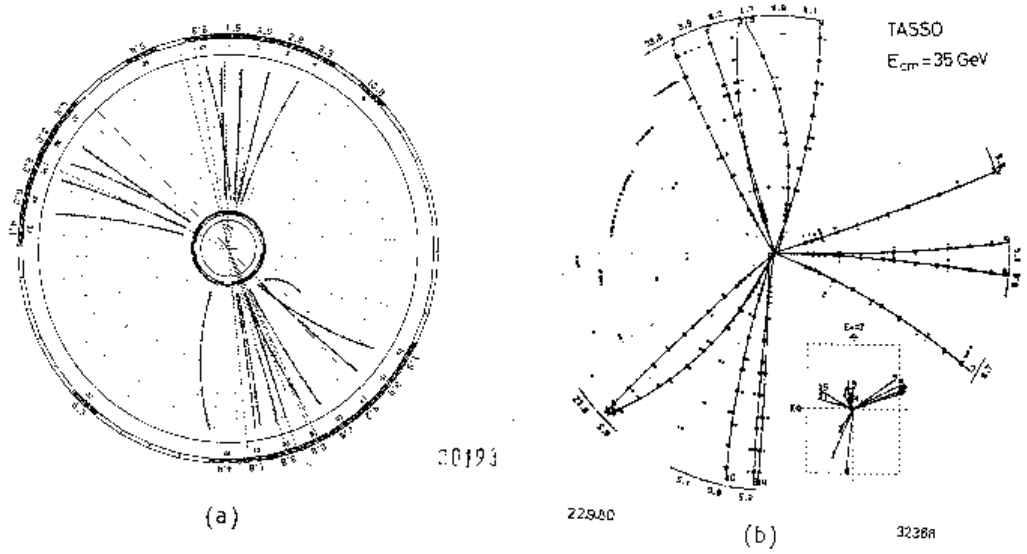

Fig. 3.24 Examples of three-jet events observed in the JADE and TASSO detectors as viewed along the beam axis. The insert in (b) shows a top view of the event.

Figure 3.18: Cross-sectional view (perpendicular to the beam direction) of three-jet events observed at the PETRA e^+e^- collider.

were visible already earlier: in 1975, the Mark-II collaboration at SLAC found that at $\sqrt{s} \lesssim 7.4$ GeV, the hadronic final states already exhibited jet-like features (disfavouring the hypothesis of *isotropic* emission of hadrons). Nevertheless, high energies certainly help, due to the fact that the momenta perpendicular to the quark directions are typically limited to $\mathcal{O}(300\,\text{MeV})$, while the quark momenta in e^+e^- experiments are typically of order $\sqrt{s}/2$.

Let us now look in a bit more detail at the process $e^+e^- \to$ hadrons. For sufficiently high $\sqrt{s}$, the cross section can actually be understood from first principles: the Feynman diagram responsible to lowest order for this process is given in Fig. 3.19 (left). This is simply a crossed version of the parton-level interaction in DIS! The only (posited) difference compared to electrons is the quarks' fractional electric charge.

This implies that we can compare the resulting cross section with that for the process $e^+e^- \to \mu^+\mu^-$: provided that the quark masses can be neglected compared to $\sqrt{s}/2$ (because otherwise a phase space suppression factor $\sqrt{1-4m_q^2/s}$ is incurred, but also because of resonant production

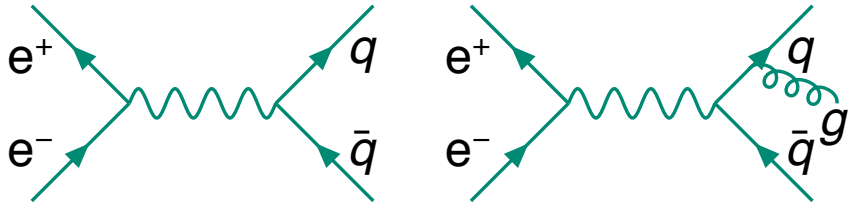

Figure 3.19: Feynman diagram describing hadron production in e^+e^- collisions in lowest order (left), and higher-order correction (right). For sufficiently high energy and angle compared to the quark and antiquark in the event, this final state may manifest itself as a three-jet event.

of bound states), the measurement of the ratio

$$R_{\text{had}} \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)}$$

"simply" counts quarks weighted with their charges. Measurements of R_{had} have been performed over a large $\sqrt{s}$ range, and the result is displayed in Fig. 3.20.

But there is more to be said about R_{had}:

- for normalization purposes, we use the $\mu^+\mu^-$ production cross section (muons, which we encountered in Sect 3.1.2, are "just like" electrons, apart from their larger mass, $m_\mu = 105.6$ MeV, and hence also couple to photons in the same way as electrons) and *not* the e^+e^- one (*i.e.*, the Bhabha scattering cross section). This is because of interference of the relevant s-channel diagram with a t-channel diagram in the Bhabha scattering case;

- if we go through the exercise of colour counting explicitly, *e.g.* for $\sqrt{s}$ well above 10 GeV, on the basis of the above we would expect that $R_{\text{had}} \approx \frac{11}{9}$. Clearly, this is far from the observed value of nearly 4. The reason for this is the fact that quarks carry *colour* charges, and this leads to different colour states in the final states that need to be summed over ($r\bar{r}$, $g\bar{g}$, $b\bar{b}$), resulting in an additional factor of three. So this measurement constitutes an additional validation of QCD!

- the above factor of three does not quite bring us into complete agreement with the measurement. The reason for this is that radiative corrections like in Fig. 3.19 (right) need to be applied (remember, we are making an *inclusive* measurement of hadron production), leading to a correction factor of the form

  $$1 + \frac{\alpha_s(s)}{\pi} + \ldots$$

  Rather than being merely a nuisance, this can be used to infer a value for $\alpha_s(s)$ instead! Indeed this is one of the methods used to measure $\alpha_s(Q^2)$. However, one can do more: one can isolate the events that contain three jets (as in Fig. 3.18). The production cross section for such events is evidently smaller, but to first order it is proportional to $\alpha_s(s)$, leading to more accurate measurements.

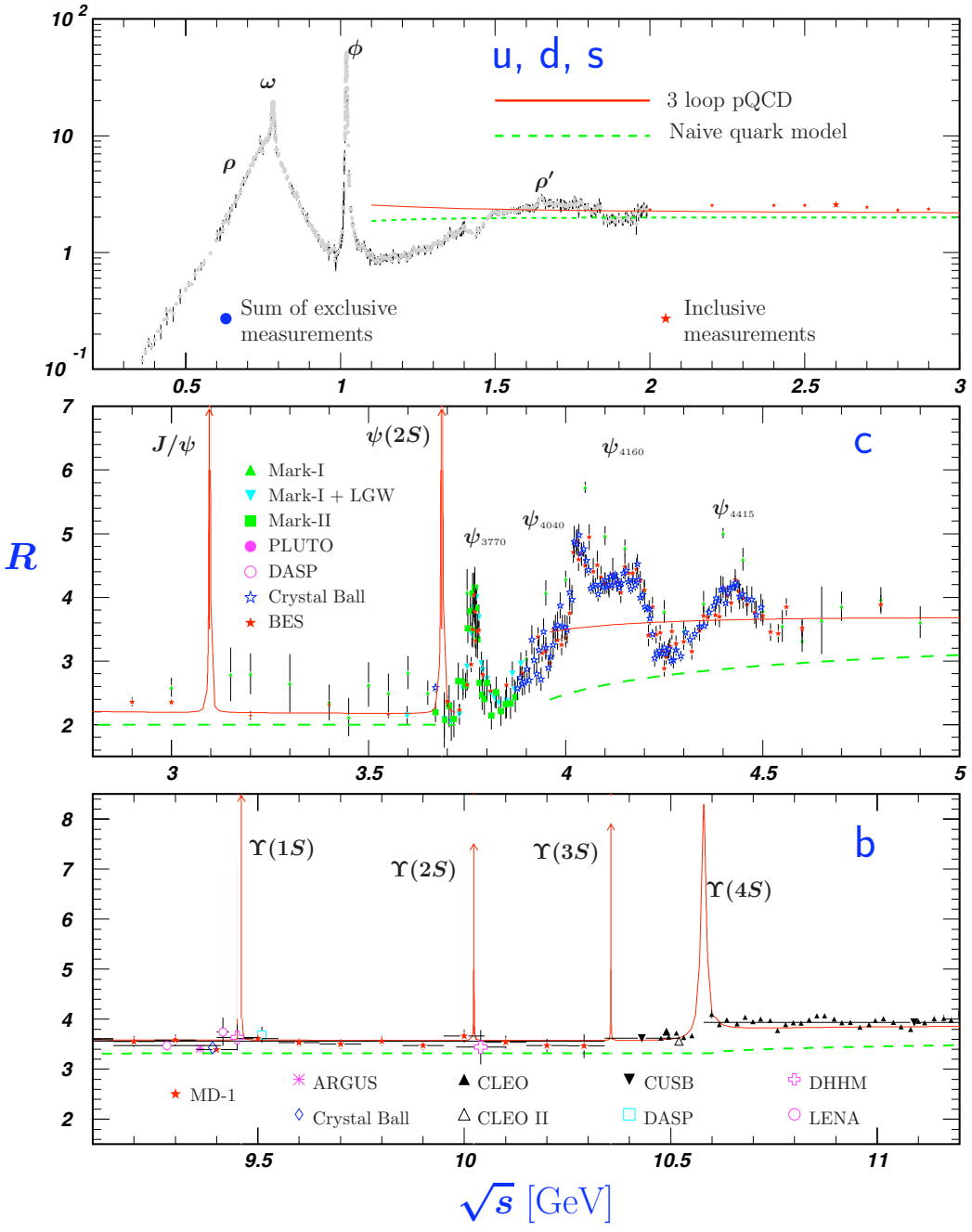

Figure 3.20: Measurements of R_{had} as a function of centre-of-mass energy

3.5 Interactions at Hadron Colliders

If protons are composite rather than elementary particles, what happens if they collide with each other rather than with electrons? The most prominent examples of such collisions at present are at the Tevatron at Fermilab (near Chicago) and at the Large Hadron Collider (LHC) at CERN: these provide $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV, and proton-proton collisions at centre-of-mass energies up to 14 TeV, respectively. Clearly the above question is one of great practical relevance.

But the answer is a simple one: just as in DIS, for high Q^2 it is the *partons* that can be considered to interact! This leads to quark-quark interactions, as shown in Fig. 3.21. However,

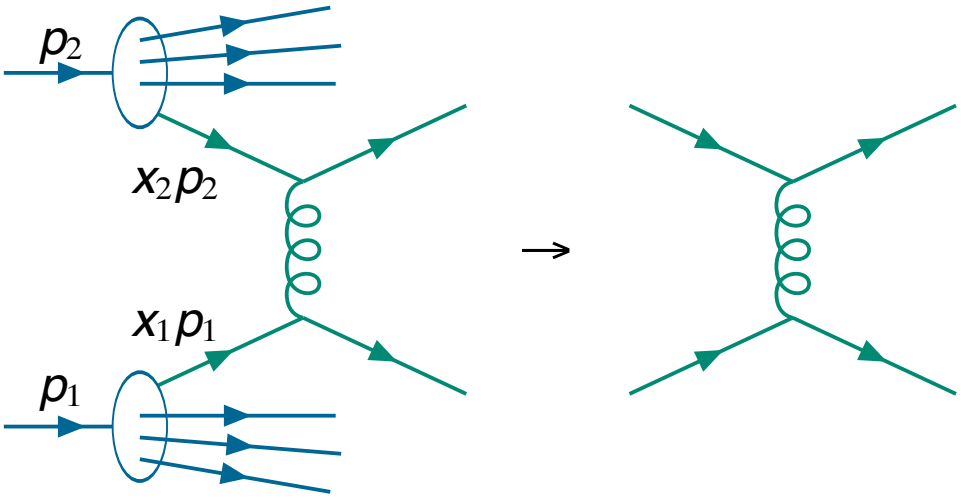

Figure 3.21: Feynman diagram describing a high-Q^2 $qq \to qq$ interaction at a hadron collider (left), and reduced diagram showing only the hard interaction itself.

this is not all. In DIS, only photon couplings with quarks and antiquarks need to be dealt with; however, this limitation arises merely from the fact that the photon only couples to electrically charged particles. Once the high-Q^2 or *hard* interaction involves the *strong* interaction, this limitation does not apply anymore. Instead, also other interactions are possible; some (lowest order) examples are shown in Fig. 3.22.

Note also that so far, we haven't specified whether we are considering proton-proton or $p\bar{p}$ collisions here. But in fact that does not matter! The only thing we need to account for is the appropriate PDFs. This procedure is called the *factorization hypothesis*, and it allows us to express the cross section for a given hard interaction as

$$\mathrm{d}\sigma = \sum_{i,j} \int \mathrm{d}x_1 \mathrm{d}x_2 f_i(x_1, \mu^2) f_j(x_2, \mu^2) \mathrm{d}\hat{\sigma}_{ij}(x_1 p_1, x_2 p_2, \alpha_s(\mu^2)), \tag{3.22}$$

where $\hat{\sigma}_{ij}$ expresses the hard scattering process at the *parton level*, which can (typically) be computed using perturbation theory, and the PDFs $f_{i,j}(x, \mu^2)$ account for the non-perturbative aspects (and as explained earlier in this chapter, they can be extracted from data). So a change of collider can be accounted for by "simply" changing PDFs and/or values of x.

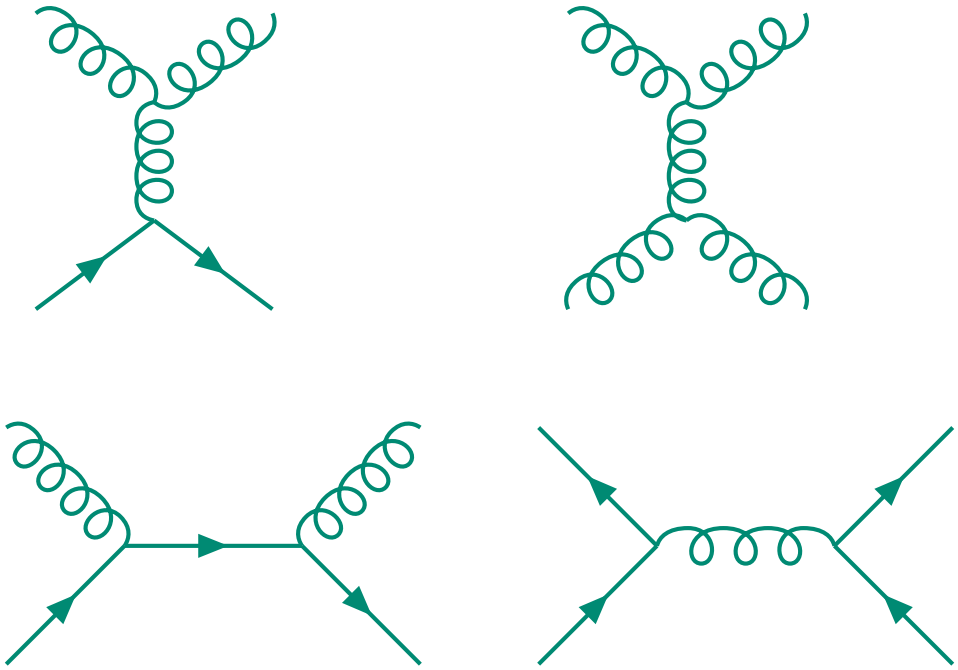

Figure 3.22: Other lowest-order QCD Feynman diagrams

Chapter 4

Weak Interactions

This chapter begins by discussing neutrinos, charged leptons, and their weak interactions. Another symmetry (called "weak isospin") appears to be relevant here. From the energy dependence of weak interaction cross sections, it will be argued that the particle(s) mediating the interaction must be heavy (as well as charged). The Brout-Englert-Higgs mechanism thought to be responsible for the "spontaneous" breaking of the weak isospin symmetry is then detailed, and some properties of its associated Higgs boson discussed.

The second part of the chapter describes the corresponding weak interactions of quarks; in this context, the Cabibbo angle is discussed, as well as its extension to three quark families, the Cabibbo-Kobayashi-Maskawa matrix. Finally, the violation of parity and even CP symmetry by the weak interaction is discussed.

In the final part of the chapter, neutrinos are considered once more. In particular, both the neutrino oscillation formalism and some of its most salient measurements are discussed in detail, with a slight emphasis on the different kinematic regimes, in which the oscillations can manifest themselves in quite different ways.

4.1 Leptons

4.1.1 Neutrino Properties

As discussed in the previous chapter, the discovery of the charged pion went hand in hand with that of the muon; see Fig. 3.1. However, looking at these photographs a bit closer, it appears that momentum is not conserved: the pion is stopped, and subsequently the muon is emitted upon the pion's decay. The *neutrino* was posited originally to explain the apparent lack of momentum conservation in the decay process. Based on what we know from the pion and muon (their mass and spin), it follows that the neutrino must be very light, and must have spin 1/2. (The same conclusion holds from measurements of neutron decay: if that were a two-body process, the final-state electron would be mono-energetic. Instead, a continuous electron energy spectrum is observed. An additional – neutral – particle is required to explain this. And the spin-1/2 nature again follows from the spins of proton, neutron, and electron.)

A truly weak interaction

The experimental observation of the neutrino did not take place until 1959. The reason for this is the weakness of the interactions involving neutrinos. In the "four-fermion" theory, first constructed by Fermi in the 1930's to describe weak interactions involving neutrinos, "crossing symmetry" could already be used to relate the matrix element for neutron decay to that for (anti)neutrino-proton scattering, as shown in Fig. 4.1. The long neutron lifetime ($\tau_n \approx 886$ s), together with known phase space factors, could be used to infer an interaction strength of $G_F = 1.166 \cdot 10^{-5}$ GeV^{-2}. This very small value of the so-called *Fermi constant* implied that the cross section for the scattering process must be very low. As a result, a strong source of

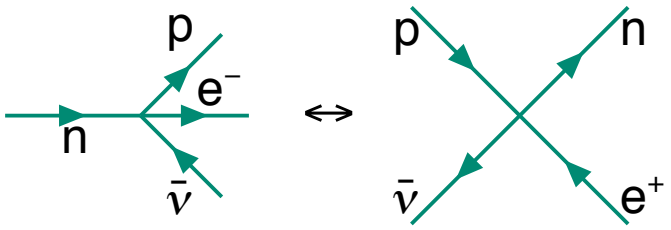

Figure 4.1: Crossing symmetry in the four-fermion theory

(anti)neutrinos was required. This was found by Cowan and Reines in the form of a nuclear reactor. They observed the reaction $\bar{\nu}p \to ne^+$, in a "delayed coincidence" setup, where the positron annihilates directly with atomic electrons to yield two photons of energy 511 keV, and the neutron is captured by Cd, giving rise to an excited state which decays through the emission of a third photon.

Antineutrinos versus neutrinos

In the above, when discussing the reactions in detail we referred to the *anti*neutrino rather than the neutrino. Given that the neutrino is an electrically neutral particle, could these not simply be identical? The answer is clearly "no", as established by Davis in 1959 in a so-called "radiochemical" experiment. He looked for the reaction

$$\bar{\nu} + {}^{37}\mathrm{Cl} \to \mathrm{e}^- + {}^{37}\mathrm{Ar},$$

where the antineutrino could be *defined* as the particle produced in nuclear reactors. The underlying reaction here would be $\bar{\nu} + \mathrm{n} \to \mathrm{p} + \mathrm{e}^-$. The absence of this reaction implies that ν and $\bar{\nu}$ are distinct particles. (This statement may be qualified to some extent later.)

Different neutrino types

As seen from the studies of cosmic rays, and also later from other studies of pions produced in scattering experiments, the charged pion decays almost exclusively as $\pi^- \to \mu^- \bar{\nu}$. On the other hand, the neutron decay process above *also* involves an antineutrino. This begs the question whether these two neutrinos are identical or different particles.

This was studied in the 1960's at Brookhaven, USA, by producing neutrino *beams*, by first producing pion beams of high energy (in interactions of a 15 GeV proton beam with a Be target) and subsequently letting these pions decay. The resulting (anti)neutrinos, of energy $\lesssim 3$ GeV, travel very nearly in the direction of the original pions.

If the resulting (anti)neutrinos are indeed the same as those involved in neutron decay, it is to be expected that in the interactions of high-energy neutrinos, equal amounts of electrons (or positrons) and muons will be produced. However, it is observed that only muons are produced in these interactions. Hence, they must be *different* neutrinos, and are denoted as ν_e and ν_μ.

Lepton doublets

The conclusion from the above is that the weak interactions as studied with leptons involve either e^- and ν_e, or μ^- and ν_μ. But we have already encountered this situation, where distinct particles somehow seem to "belong together" as far as their interactions are concerned: this was our reason for grouping the proton and the neutron into isospin doublets in Sect. 3.1.1. So we will do the same here, and create *weak isospin* doublets (the adjective serves to distinguish these doublets from those encountered in the context of the strong interaction, and which we will from now on denote as *strong isospin* doublets). If we also account for a third charged lepton (the τ lepton, with $m(\tau) = 1.777$ GeV and $\tau(\tau) = 291$ fs) and its accompanying neutrino ν_τ, we finally arrive at the three lepton doublets of the Standard Model:

$$L_e = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}, \quad L_\mu = \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}, \quad L_\tau = \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}. \tag{4.1}$$

Also the various decay processes described above can now be specified a little more precisely. For instance, muon and (leptonic) tau decays are given by

$$\begin{aligned} \mu^- &\to e^- \nu_\mu \bar{\nu}_e, \\ \tau^- &\to \mu^- \nu_\tau \bar{\nu}_\mu \quad \text{or} \\ &\quad\ e^- \nu_\tau \bar{\nu}_e \end{aligned}$$

(The τ lepton was discovered in 1975, in e^+e^- scattering experiments, and was produced in the interaction $e^+e^- \to \tau^+\tau^-$. The $\tau^\pm$ decay rapidly, even if they travel a measurable distance before decaying. Especially those decays where one decayed to a muon and the other to an electron, plus apparent missing energy caused by the neutrinos, were used to infer the existence of the τ.)

4.1.2 W and Z Bosons

Unitarity

The four-fermion theory does quite well in the description of *nuclear* processes involving the weak interaction. However, when considered in a high-energy context a theoretical problem appears. Let us consider again the $\bar{\nu}$-p scattering process shown in Fig. 4.1 (right panel), and assume for a moment that the proton and the neutron can be described as elementary spin-1/2

particles also at high energy (this is not what happens in reality, as is to be expected from the previous chapter; but it helps to make the point, and the argument works equally well when considering collisions with *quarks*). It turns out that in the four-fermion theory, the (spin-averaged, squared) matrix element for this process becomes

$$\langle|\mathscr{M}|^2\rangle \sim G_F^2 (p_1 \cdot p_3)(p_2 \cdot p_4),$$

where $p_{1,3}$ are the four-momenta of the initial-state p and $\bar{\nu}$, and $p_{2,4}$ are the four-momenta of the final-state particles. Clearly, when considering this process in the CM frame, all four-momenta are of order $\sqrt{s}$, and we find that $\langle|\mathscr{M}|^2\rangle \sim s^2$. Furthermore, for this two-body scattering process we know how to convert this to a (differential) cross section: see Eqn. 3.2. Integrating the result, we find that that *total* $\bar{\nu}$p scattering cross section behaves as $\sigma_{\text{tot}} \sim s$.

On the other hand, a formal and general consideration of scattering processes leads to the so-called *unitarity* or *Froissart bound*, indicating that the cross section for the above process cannot exceed π/s. The above linear dependence on s must therefore violate the Froissart bound at some value (in practice, this happens at $\sqrt{s} \approx 600$ GeV). Even if initial investigations of the weak interaction were far from attaining such energies, it was evident that Fermi's theory required modification.

Massive gauge bosons

Following the successful description of electromagnetic and strong interactions as being mediated by particles, it makes a lot of sense to suppose that the same paradigm is useful for the description of these $\bar{\nu}$p interactions as well. Given the lepton doublet structure of Eqn. 4.1, and now supposing that at high energy, it is rather a *quark* inside the proton that is struck, we arrive at the replacement of Fig. 4.2 (we will be discussing weak interactions of quark couplings in more detail in Sect. 4.3). The hypothesized intermediate particle is called the W boson.

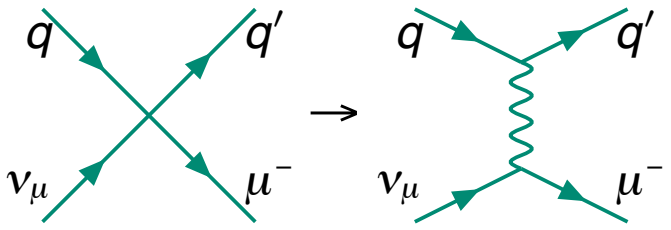

Figure 4.2: Modification of the neutrino-nucleon scattering amplitude by the insertion of a W boson.

But the W boson must have somewhat different properties than the photon and the gluons encountered before. First, it changes a neutral lepton (ν or $\bar{\nu}$) into a charged lepton or vice versa: this implies that it must be *charged*. Second, if it were a massless boson like the photon, the differential cross section for the process shown in Fig. 4.2 would receive a factor $1/t^2$ from the (squared) propagator. Integrating again over the scattering angle in the CM frame, this would lead to an additional $1/s^2$ dependence, and hence $\sigma_{\text{tot}} \sim 1/s$. But unlike Fermi's theory, this

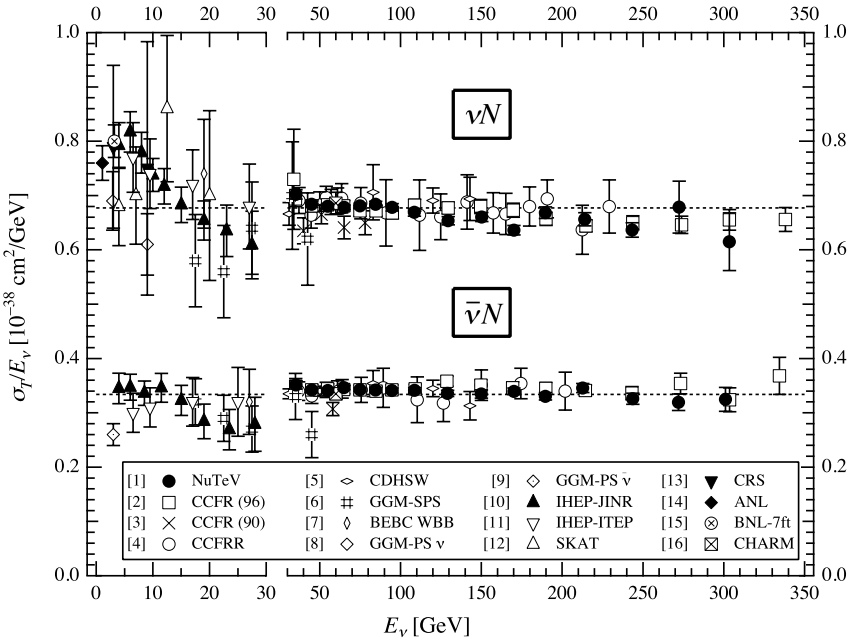

Figure 4.3: Measured ν_μN and $\bar{\nu}_\mu$N total charged current scattering cross sections versus neutrino beam energy ("charged current" means that a muon is observed in the final state).

result disagrees with experiment, as shown in Fig. 4.3! Addressing this requires a *massive* W boson, with a mass M_W sufficiently high that it does not lead to a discernible behaviour different from the one in Fermi's theory for the CM energies probed by these interactions. The propagator for this massive W boson is given by

$$\frac{g^{\mu\nu} - q^\mu q^\nu / M_\mathrm{W}^2}{q^2 - M_\mathrm{W}^2 + iM_\mathrm{W}\Gamma_\mathrm{W}}.$$

Focusing on the denominator (the numerator is not in fact very important for the high-energy behaviour of scattering processes), we see that for $|q^2| \ll M_\mathrm{W}^2$, this simply approaches a *constant* $\sim 1/M_\mathrm{W}^2$, while for $|q^2| \gg M_\mathrm{W}^2$, a behaviour consistent with the Froissart bound is obtained.

Also the decay of the muon can be recast in a process involving a W boson, as shown in Fig. 4.4. Neglecting the very small q^2 of the exchanged W boson, this relates the Fermi constant

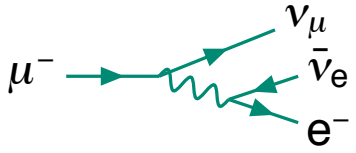

Figure 4.4: Decay of the muon.

to the "actual" coupling constant g:

$$G_F = \frac{1}{4\sqrt{2}} \frac{g^2}{M_{\mathrm{W}}^2}. \tag{4.2}$$

Unfortunately, while G_F is easily measured from this process, this relation does not *a priori* provide any further information regarding M_{W}.

At this point, we have not yet drawn any conclusion from the grouping of leptons into doublets. But again by analogy with the strong isospin doublets, which corresponded to an $SU(2)$ symmetry, we should expect to see signs of an $SU(2)$ symmetry at work in the weak interaction as well. But this also suggests that besides the charged $\mathrm{W}^{\pm}$ bosons, we should expect a third, neutral, partner – this is called the Z boson.

But how to observe such a neutral partner? The electromagnetic interaction is so much stronger than the weak interaction (at least effectively) that it will normally swamp any sign of Z bosons. The solution is to study again (anti)neutrino interactions, since electromagnetic interactions are absent in this case (the neutrino being neutral). Such interactions were indeed observed, first in 1973 by the Gargamelle experiment. An example is shown in Fig. 4.5. The

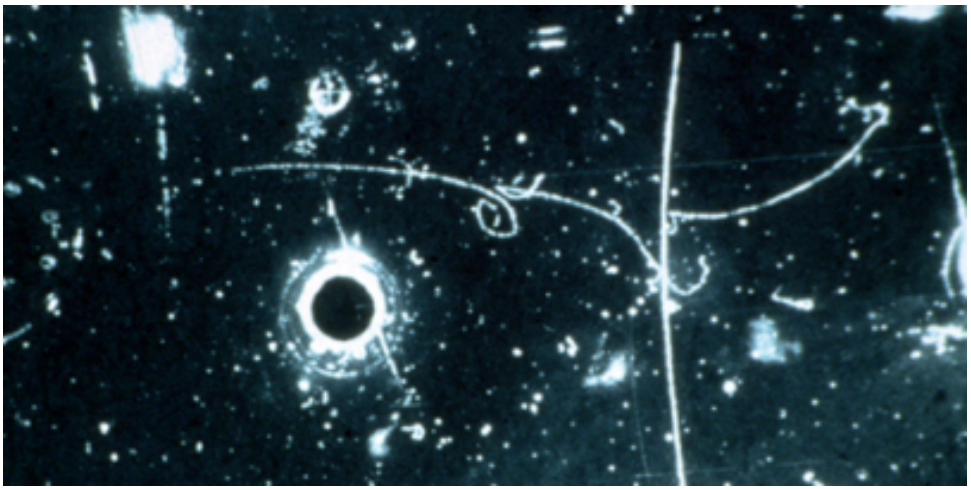

Figure 4.5: Bubble chamber photograph taken by the Gargamelle experiment. A neutrino beam arrives from the left; one neutrino collides with an atomic electron. The electron radiates two photons, both of which convert into $\mathrm{e}^+\mathrm{e}^-$ pairs, and is then stopped.

cross section for "neutral current" neutrino-nucleon interactions (involving the Z boson) is of the same order as that for "charged current" neutrino-nucleon interactions (involving the W boson). Therefore, their masses must be of the same magnitude.

4.2 Electroweak Unification and the Brout-Englert-Higgs Mechanism

As alluded to before, the photon as well as the gluons of QCD are strictly massless. On the other hand, the W and Z bosons are massive (and must in fact be heavy). It can be shown that it isn't possible to introduce direct mass terms in the Lagrangian to accommodate the W and Z boson masses: this would break the gauge invariance that is at the heart of the description of all fundamental interactions. So an alternative description is needed. It turns out that this indeed exists, in the form of *spontaneous symmetry breaking* coupled with the *Brout-Englert-Higgs mechanism*.

To set the scene, it turns out that this mechanism involves not only the $SU(2)$ symmetry but also electromagnetism – this is the origin of the "electroweak unification" in the title. This implies that we have four gauge bosons to account for ($W^{\pm}$, Z, γ), *i.e.*, one more than provided for by $SU(2)$. At first sight it would seem that the additional $U(1)$ symmetry needed would be "just" that of QED; but this turns out not to work. Instead, we need a *different* $U(1)$ symmetry, often labeled $U(1)_Y$ with the Y denoting "hypercharge". In this case, for a representation transforming under hypercharge and the "weak isospin" $SU(2)$ as

$$\begin{aligned} \psi &\rightarrow \psi' = e^{ig'\frac{Y}{2}b(x)}\psi \\ \psi &\rightarrow \psi' = e^{igT_i w_i(x)}\psi \end{aligned}$$

we find that the covariant derivative becomes

$$\partial_\mu \rightarrow D_\mu = \partial_\mu - igW_\mu^i T_i - ig'\frac{Y}{2}B_\mu. \tag{4.3}$$

Notes:

- The factor 1/2 in the hypercharge transformation is a convention only.
- The generators T_i depend on whether ψ represents a singlet, a doublet, or something else under the $SU(2)$ transformations. In practice, only $T_i = 0$ (for singlets, which do not transform at all) or $T_i = \sigma_i/2$ (for doublets) are relevant for our discussion.
- Similarly, the hypercharge Y quantum number can in principle be chosen independently for each doublet; we will see later that the same value of Y will be required for all lepton doublets, however. (It will turn out that for quark doublets, yet another value is required to make things "work"; and there are other details that will be covered later.)

With these ingredients we can now go back *e.g.* to the lepton doublet L_e. Since this is a doublet, its kinetic energy term in the Lagrangian must look like $\overline{L_e}(i\not{D})L_e$. Using the above expression for the covariant derivative as applied to doublets, and with hypercharge Y_L, we find

$$(\bar{\nu}_e, \bar{e}) \begin{pmatrix} i\not{\partial} - \frac{g}{2}\not{W}_3 - Y_L\frac{g'}{2}\not{B} & -\frac{g}{2}(\not{W}_1 - i\not{W}_2) \\ -\frac{g}{2}(\not{W}_1 + i\not{W}_2) & i\not{\partial} + \frac{g}{2}\not{W}_3 - Y_L\frac{g'}{2}\not{B} \end{pmatrix} \begin{pmatrix} \nu_e \\ e \end{pmatrix}. \tag{4.4}$$

In this formula, the relevant particle fields are simply represented by the particle names; so we could read $\mathrm{e} \equiv \psi_\mathrm{e}$ etc. Looking at the off-diagonal terms, one observes terms coupling the electron to the neutrino:

$$-\frac{g}{2}\left(\bar{\nu}_\mathrm{e}(W_1 - iW_2)\mathrm{e} + \bar{\mathrm{e}}(W_1 + iW_2)\nu_\mathrm{e}\right).$$

Since this coupling changes the charge of the lepton involved, we expect to find the (charged) W bosons here; so in detail, we must have

$$W^{\pm\mu} = (W_1^\mu \mp iW_2^\mu)/\sqrt{2}.$$

We will return to the diagonal terms later.

The Brout-Englert-Higgs mechanism consists in positing another *complex scalar doublet* field $\Phi(x)$, which transforms under weak isospin just like the lepton doublet, but which *a priori* (and also in practice!) has a hypercharge quantum number Y_h that differs from Y_L. This requires a kinetic energy term

$$(D_\mu\Phi)^\dagger(D^\mu\Phi), \quad \text{with} \quad D_\mu = \partial_\mu - i\frac{g}{2}\sigma_i W_{i\mu} - \frac{g'}{2}Y_h B_\mu.$$

In addition, however, we make a specific choice for the potential terms in the Lagrangian pertaining to Φ:

$$V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2. \tag{4.5}$$

In this expression, the quadratic term would be "just" another mass term if we had $\mu^2 > 0$. However, we instead choose $\mu^2 < 0$! This choice has rather dramatic consequences, due to the fact that the minimum of V will not anymore be at $\Phi = 0$ but rather at $|\Phi| = v/\sqrt{2}$, with $v = \sqrt{-\mu^2/\lambda}$. For then we are in a situation where although $|\Phi|$ is well defined, Φ itself is not! The same value of V is obtained for all values of Φ on a three-dimensional hypersphere in a four-dimensional space. But since the ground state or *vacuum* of the system must be a well-defined state, one specific orientation on this hypersphere must be chosen. (Note that in a *finite* system the conclusion would be different: the ground state would still have $\langle\Phi\rangle = 0$.) This situation (where the potential features the full -original- symmetry but where the choice of ground state breaks this symmetry) is called *spontaneous symmetry breaking*.

For the vacuum we choose the convention (for the interested: a different orientation in the above-mentioned four-dimensional space would result in different expressions for *e.g.* the lepton doublets)

$$\langle\Phi\rangle = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}.$$

Having done so, this allows us to choose a different parametrisation for Φ, by "expanding" around the ground state

$$\Phi = e^{ig\frac{\sigma_i}{2}\rho_i(x)} \begin{pmatrix} 0 \\ (v + h(x))/\sqrt{2} \end{pmatrix},$$

where $h(x)$, $\rho_i(x)$ describe the independent fields in this alternative formulation. By itself this formulation does not offer us anything new in terms of physics, but this changes upon realising

that we can use up almost all of the gauge degrees of freedom (that is, the freedom to apply an arbitrary $SU(2)$ and $U(1)_Y$ transformation *at each point in spacetime*) to "rotate away" the exponent in the above expression! So we are left with

$$\Phi = \begin{pmatrix} 0 \\ (v+h(x))/\sqrt{2} \end{pmatrix}. \tag{4.6}$$

(Note that three real parameters are involved in this rotation; so we still retain one degree of freedom.) Once we do this, the $D_\mu\Phi$ term can fairly easily be seen to reduce to

$$D_\mu\Phi = \begin{pmatrix} -i\frac{g}{2}(W_{1\mu} - iW_{2\mu}) \\ \partial_\mu + i\frac{g}{2}W_{3\mu} - i\frac{g'}{2}Y_h B_\mu \end{pmatrix} (v+h)/\sqrt{2},$$

so that the Brout-Englert-Higgs field's kinetic energy term becomes

$$\frac{1}{2}\partial_\mu h \partial^\mu h + \left(\frac{g}{2}\right)^2 W_\mu^- W^{\mu+}(v+h)^2 + (gW_{3\mu} - g'Y_h B_\mu)(gW_3^\mu - g'Y_h B^\mu)\frac{(v+h)^2}{8}. \tag{4.7}$$

Ignoring for a moment the terms involving the field $h(x)$, we see that we are left with one term proportional to $W_\mu^- W^{\mu+}$, which represents a mass term for the W bosons; and one other term which is proportional to the square of a linear superposition of the W_3^μ and B^μ fields. This last term implies that the latter two fields *mix* to yield the Z boson (which we know to be massive); the orthogonal combination then yields the photon, for which no mass term exists, *i.e.*, it remains massless. In detail, the superposition can be written as

$$\begin{aligned} A^\mu &= \sin\theta_{\rm w} W^{3\mu} + \cos\theta_{\rm w} B^\mu, \\ Z^\mu &= \cos\theta_{\rm w} W^{3\mu} - \sin\theta_{\rm w} B^\mu, \end{aligned}$$

with the *weak mixing angle* $\theta_{\rm w}$ defined as

$$\tan\theta_{\rm w} \equiv \frac{g'Y_h}{g}.$$

The masses of the W and Z bosons then become

$$M_W = \frac{1}{2}vg, \tag{4.8}$$

$$M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2 Y_h^2}, \tag{4.9}$$

so that we also have the relation $M_W = M_Z \cos\theta_{\rm w}$.

Now let us return to Eqn. 4.4, and in particular the diagonal term involving the neutrinos, $\bar{\nu}_{\rm e}(i\partial\!\!\!/ - \frac{g}{2}W\!\!\!/_3 - Y_L\frac{g'}{2}B\!\!\!/)\nu_{\rm e}$. A priori this term involves couplings of both the Z boson and the photon to the neutrino; the only way to ensure that there is no coupling to the photon is by requiring that the particular combination of W_3 and B yields again the Z boson. Comparing again with Eqn. 4.7, this implies that we must have $Y_L = -Y_h$. For the sake of definiteness, in the following we will

assume that $Y_h = 1$ (a different value could be accommodated by a different value of the coupling constant g', so this is not an actual constraint).

Turning now to the diagonal term involving the electrons, we then find that in this case we do have a coupling to both the Z boson and the photon. Focusing on the coupling to the photon, after a bit of rewriting this can be seen to be equal to

$$g \sin\theta_{\mathrm{w}}\, \bar{\mathrm{e}} \not{A} \mathrm{e}. \tag{4.10}$$

But we know that even if QED in this context only arises from mixing the $SU(2)$ and $U(1)$ symmetry transformations, we must still reproduce the couplings of the photon to charged particles in general and to electrons in particular. This implies that we must have

$$g \sin\theta_{\mathrm{w}} = e. \tag{4.11}$$

So we see that coupling of the weak interaction (g) is in fact *larger* than that of electromagnetism! The sole reason for the qualification "weak" lies in the high masses of the W and Z boson.

Lepton masses

It turns out that the Brout-Englert-Higgs mechanism is responsible not only for the masses of the W and Z bosons, but also for the leptons and quarks. To see this, let us first restrict ourselves to the leptons.

We need to make use of a property of the weak interactions that will be motivated by experimental results only a bit later: the $SU(2)$ and $U(1)$ combined electroweak interactions are sensitive to the *spins* of the leptons involved, and more particularly to their *chirality*. For massless particles, it can be shown that the action of the matrix $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$ is identical to that of the *helicity* operator λ encountered in one of the exercises. This allows us to view the left- and right-handed chirality projections

$$P_{R,L} = \frac{1}{2}(1 + \gamma^5)$$

as different states! For while it is always possible to carry out a Lorentz boost such that a left-handed *massive* particle in one Lorentz frame becomes a right-handed one in another frame (and *vice versa*, this is not possible for massless particles.

As will be shown later, the charged weak interaction deals with left-handed fermions and right-handed anti-fermions only. This implies that the lepton doublets $L_{\mathrm{e},\mu,\tau}$ encountered in Eqn. 4.1 refer to the left-handed components only; formally starting from the full four degrees of freedom for an "ordinary" lepton this can be written

$$L_{\mathrm{e}} = \begin{pmatrix} \nu_{\mathrm{e},L} \\ \mathrm{e}_L \end{pmatrix} = \begin{pmatrix} P_L \nu_{\mathrm{e}} \\ P_L \mathrm{e} \end{pmatrix}.$$

The absence of a charged weak interaction involving right-handed fermions then implies that these must be singlets under $SU(2)$: $\mathrm{e}_R = P_R \mathrm{e}$. It can be shown that the mass term $m_{\mathrm{e}}\bar{\mathrm{e}}\mathrm{e}$ ordinarily occurring in the Lagrangian can be written as

$$m_{\mathrm{e}}\bar{\mathrm{e}}\mathrm{e} = m_{\mathrm{e}}(\overline{\mathrm{e}_L}\mathrm{e}_R + \overline{\mathrm{e}_R}\mathrm{e}_L),$$

but such a term clearly would not be invariant under $SU(2)$ weak isospin transformations; this is a more mathematical argument why the left- and right-handed particles being separate entities is only possible if the particle involved is massless.

But this situation still does makes it possible to construct "Yukawa" interaction terms in the Lagrangian involving also the Higgs field. In particular, for the electron we have:

$$\mathscr{L}_{\mathrm{Yu,e}} = -y_{\mathrm{e}} \left(\overline{L_{\mathrm{e}}} \Phi \mathrm{e}_R + \overline{\mathrm{e}_R} \Phi^\dagger L_{\mathrm{e}} \right).$$

In this formula, we haven't yet specified how e_R transforms under $U(1)$ hypercharge transformations (*i.e.*, what hypercharge Y_R quantum number the right-handed electron carries). But in fact this follows simply from the fact that the Lagrangian must be invariant under hypercharge transformations; hence we must have $-Y_L + Y_h + Y_R = 0$. And since we already had $Y_h = -Y_L = 1$, this implies $Y_R = -2$.

Upon breaking again the electroweak symmetry and making the replacement of Eqn. 4.6, the above formula becomes

$$-y_{\mathrm{e}} \left(\overline{\mathrm{e}_L} \mathrm{e}_R + \overline{\mathrm{e}_R} \mathrm{e}_L \right) \frac{v+h}{\sqrt{2}} = -\frac{y_{\mathrm{e}}}{\sqrt{2}} \bar{\mathrm{e}} \mathrm{e} (v+h). \tag{4.12}$$

Neglecting the term proportional to h, this is easily seen to yield a mass term, with mass $m_{\mathrm{e}} = y_{\mathrm{e}} v/\sqrt{2}$. This same mechanism works also for the other leptons, where the higher masses of the muon and the tau lepton are generated by assuming higher values for the respective Yukawa coupling constants y_μ and y_τ.

W and Z boson masses

The above implies that *if* θ_{w} can somehow be measured, Eqn. 4.11 fixes g. From the known value of the Fermi constant and Eqn. 4.2, M_{W} can then be determined. Finally, Eqn. 4.9 can then be used to estimate M_{Z}.

A method to estimate θ_{w} in fact *does* exist, in the form of a measurement of (the ratio of) the total cross sections for neutral current (NC) and charged current (CC) neutrino-nucleon interactions. Again we lack the theoretical machinery to derive the result, and therefore quote it without justification:

$$\begin{aligned} R_\nu &\equiv \left(\frac{\sigma_{\mathrm{tot}}(\mathrm{NC})}{\sigma_{\mathrm{tot}}(\mathrm{CC})} \right)_\nu &&= \frac{1}{2} - \sin^2\theta_{\mathrm{w}} + \frac{20}{27} \sin^4\theta_{\mathrm{w}}, \\ R_{\bar\nu} &\equiv \left(\frac{\sigma_{\mathrm{tot}}(\mathrm{NC})}{\sigma_{\mathrm{tot}}(\mathrm{CC})} \right)_{\bar\nu} &&= \frac{1}{2} - \sin^2\theta_{\mathrm{w}} + \frac{20}{9} \sin^4\theta_{\mathrm{w}}. \end{aligned}$$

These cross sections have indeed been measured (see also Fig. 4.3), resulting in an estimate $\sin^2\theta_{\mathrm{w}} \approx 0.23$, from which it follows that $M_{\mathrm{W}} \approx 80$ GeV and $M_{\mathrm{Z}} \approx 92$ GeV.

Experimental proof of the W and Z bosons was obtained in 1983 and 1984 at the Sp$\bar{\mathrm{p}}$S collider, which collided protons and antiprotons at $\sqrt{s} = 630$ GeV. At these high energies, it is again the *partons* in the (anti)protons that can be considered to collide. The Feynman diagrams of Fig. 4.6 show how this allows the W and Z bosons to manifest themselves. The W and Z

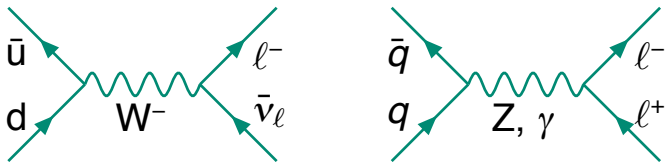

Figure 4.6: Production and (leptonic) decay of W and Z bosons in hadron colliders.

bosons have been studied in much more detail at the Large Electron-Positron or LEP collider. In its first phase, from 1990 to 1994, it operated at $\sqrt{s} \approx M_Z$, probing the Z resonance. Among many other measurements, a scan of the $\sqrt{s}$ dependence of fermion pair production cross sections (see Fig. 4.7, left diagram) resulted in a very precise measurement of the Z boson's mass: $M_Z = 92.188(2)$ GeV. The same measurements also established that there are no more than the three (light, $m_\nu < M_Z/2$) neutrino species already known.

In its second phase, LEP operated at higher energies, $\sqrt{s} \lesssim 208$ GeV. An important part of the physics programme was the study of W-boson pair production: as shown in the right diagram of Fig. 4.7, this probes the WWZ and WWγ couplings.

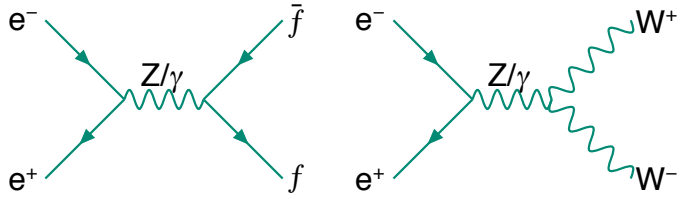

Figure 4.7: Processes studied at the LEP collider: fermion-pair production in the vicinity of the Z resonance (left), and W-boson pair production (right).

4.3 Quarks

4.3.1 The CKM Matrix

Remembering the quark content of the lightest mesons (see Table 3.2) and the lightest baryons, it is in fact easy to guess at the weak interaction processes responsible for the decay of the pion and of the neutron. They are shown in Fig. 4.8. (It should be stressed that this picture, referred to as the "spectator model", is *incomplete*, as it ignores the – non-perturbative – interactions between the (anti)quarks in the decaying hadron. But it reproduces the observed features of such decays quite well on a qualitative level.)

Consequently, it is very reasonable to suppose that also quarks form *weak isospin doublets*, just like the leptons. However, there are other decays that then are less easily understood. For instance, the K^- may decay just like the charged pion, *i.e.*, to $\mu^-\bar{\nu}_\mu$. But looking at the quark content of the K^-, we see that this must be $s\bar{u}$. How can the u quark be in two doublets at once?

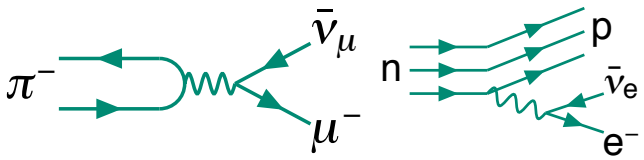

Figure 4.8: Feynman diagrams describing the decay of the pion (left) and of the neutron (right), in the spectator model.

The solution to this puzzle is to assume that the down-type quark occurring in the doublet is not simply d or s, but a linear combination of them:

$$\begin{pmatrix} \mathrm{u} \\ \mathrm{d}' \end{pmatrix} = \begin{pmatrix} \mathrm{u} \\ \mathrm{d}\cos\theta_c + \mathrm{s}\sin\theta_c \end{pmatrix}, \tag{4.13}$$

with θ_c the so-called *Cabibbo angle*. When accounting (in addition to u, d, and s) only for the charm quark, a similar linear combination $\mathrm{s}' = -\mathrm{d}\sin\theta_c + \mathrm{s}\cos\theta_c$ can be defined. The result is a *rotation* in the "plane" of the down-type quarks:

$$\begin{pmatrix} \mathrm{d}' \\ \mathrm{s}' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \end{pmatrix}. \tag{4.14}$$

The above can be used to obtain an estimate of θ_c from a comparison of the decay widths of the charged kaon and the charged pion. Assuming that the non-perturbative aspects are the same for both decays, we find that the ratio $\Gamma(\mathrm{K} \to \mu\nu)/\Gamma(\pi \to \mu\nu)$ yields an estimate of $\tan^2\theta_c$ (apart from phase space factors, which can be computed). The result is that $\theta_c \approx 13°$.

This rotation, compared to the introduction of only the single quark doublet of Eqn. 4.13, also solves another problem, namely that of the (*a priori*) unexpectedly small decay width for the decay $\mathrm{K}^0 \to \mu^+\mu^-$. If only the first quark doublet is taken into account, this is described by the diagram in Fig. 4.9. This is a higher-order, so-called "box" diagram, and its computation is

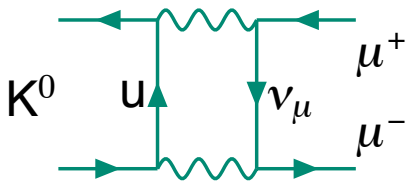

Figure 4.9: Feynman diagram describing the decay $\mathrm{K}^0 \to \mu^+\mu^-$, in the spectator model.

involved. However, it is straightforward to see that the amplitude for this diagram must involve a factor $\sin\theta_c\cos\theta_c$. When adding also the second doublet, we find that a second diagram becomes possible, with the up quark replaced with a charm quark; this yields a factor $-\sin\theta_c\cos\theta_c$. So *if* the up and charm quark masses were identical, these decay amplitudes would cancel completely! Turning the argument around, from the observed decay width a prediction for $m_c \approx 1.5$ GeV was made *before* the discovery of the J/ψ, in good agreement with the experimental value.

But as we now know, there are *three* rather than two generations of quarks. So the rotation of Eqn. 4.14 needs to be replaced with one appropriate for three generations of quarks:

$$\begin{pmatrix} \mathrm{d}' \\ \mathrm{s}' \\ \mathrm{b}' \end{pmatrix} = V_{\mathrm{CKM}} \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \\ \mathrm{b} \end{pmatrix}, \quad \text{with} \quad V_{\mathrm{CKM}} = \begin{pmatrix} V_{\mathrm{ud}} & V_{\mathrm{us}} & V_{\mathrm{ub}} \\ V_{\mathrm{cd}} & V_{\mathrm{cs}} & V_{\mathrm{cb}} \\ V_{\mathrm{td}} & V_{\mathrm{ts}} & V_{\mathrm{tb}} \end{pmatrix} \tag{4.15}$$

The matrix V_{CKM} is called the *Cabibbo-Kobayashi-Maskawa* matrix, after Cabibbo, and Kobayashi and Maskawa who extended the formalism. This extension is slightly less trivial than it seems at first sight: the quark fields are intrinsically *complex*, and V_{CKM} is therefore inherently *unitary*, leading to nine independent parameters. But five quark phases can be "rotated away", leaving four parameters. This leaves room for three angles and one irreducible complex phase. This complex phase is believed to be responsible for the CP violation phenomena discussed in Sects. 4.5 and 4.6.

A few more remarks about the CKM matrix are in order:

- Its effect is to cause a different between *weak interaction eigenstates* (defined as those states occurring in doublets) and *mass eigenstates* (those states that can be assigned a definite mass). The convention is to speak only about mass eigenstates; therefore the weak interaction vertex describing the coupling of quarks q and q' must involve a factor $V_{qq'}$ (or $V^*_{q'q}$, whichever is the up- or down-type quark).

- Given this situation, one may wonder whether this difference also affects the coupling of quarks to the Z boson in such a way that vertices result involving *different* down-type quark flavours (these would be called *flavour-changing neutral current* or FCNC interactions). For instance, we would find a coupling of down quarks to the Z boson which now becomes one of the corresponding weak interaction states (d'). At the quantum field level, this would correspond to a coupling

  $$\mathrm{d}'\bar{\mathrm{d}}'Z = (V_{\mathrm{ud}}\mathrm{d} + V_{\mathrm{us}}\mathrm{s} + V_{\mathrm{ub}}\mathrm{b})(V^*_{\mathrm{ud}}\bar{\mathrm{d}} + V^*_{\mathrm{us}}\bar{\mathrm{s}} + V^*_{\mathrm{ub}}\bar{\mathrm{b}})Z$$

  implying that FCNC interactions might indeed occur. However, we should in fact be dealing with a term $(\mathrm{d}'\bar{\mathrm{d}}' + \mathrm{s}'\bar{\mathrm{s}}' + \mathrm{b}'\bar{\mathrm{b}}')Z$; if this is written in terms of the couplings to mass eigenstates, it becomes

  $$\begin{aligned} \sum_i \mathrm{d}'_i\bar{\mathrm{d}}'_i Z &= \sum_{ijk} V_{ij}V^*_{ik}\mathrm{d}_j\bar{\mathrm{d}}_k Z = \sum_{ijk}(V^\dagger)_{ki}V_{ij}\mathrm{d}_j\bar{\mathrm{d}}_k Z \\ &= \sum_{jk}\delta_{jk}\mathrm{d}_j\bar{\mathrm{d}}_k Z = \sum_j \mathrm{d}_j\bar{\mathrm{d}}_j Z, \end{aligned}$$

  where we have used the unitarity of the CKM matrix. So no FCNC interactions result!

- A large number of measurements have been performed to estimate the values of individual CKM matrix elements (akin to the estimation of θ_c above). As a result, they are now

fairly well known. In particular, the *absolute values* of the matrix elements show a clear hierarchical pattern:

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \approx \begin{pmatrix} 0.97 & 0.23 & 0.004 \\ 0.23 & 0.97 & 0.04 \\ 0.008 & 0.04 & 1 \end{pmatrix} . \tag{4.16}$$

It should be noticed that the off-diagonal elements of the CKM matrix have a significant practical impact. Without them, it would still be possible to produce *e.g.* top quark pairs, and the top quark would still decay via the process $t \to Wb$; but there would be no way for the b quark to decay (and a similar situation would hold for the s quark)! So these off-diagonal elements are ultimately responsible for the fact that all "stable" matter involves quarks and (charged) leptons from the first generation only.

Finally, let us return to the topic of top quark decays. In Sect. 3.1.8, we commented on the fact that it is too massive to lead to bound states. In more detail, the important ingredient is the fact that its mass ($m_t \approx 172$ GeV) is substantially larger than that of the W boson ($m_W \approx 80.4$ GeV). As a consequence, the top quark can decay to an *on-shell* W boson, and the propagator suppression $\sim 1/M_W^2$ does not hold anymore. In fact, its total decay width can be computed to be of order 1.4 GeV, corresponding to a lifetime of $\mathcal{O}(10^{-24}\text{s})$. It is this lifetime which is too short to lead to the formation of bound states. The top quark was discovered in 1995 at the Tevatron $p\bar{p}$ collider, in interactions leading to $t\bar{t}$ final states. The subsequent decays of both top quarks, especially if at least one of them leads to a *leptonic* decay, provide a sufficiently distinctive feature that such final states can be recognized among the huge backgrounds. The Feynman diagrams responsible for the top quark production and decay at hadron colliders are shown in Fig. 4.10. (The V_{ts} and V_{td} matrix elements are sufficiently small that usually, only the decay to b quarks is accounted for.)

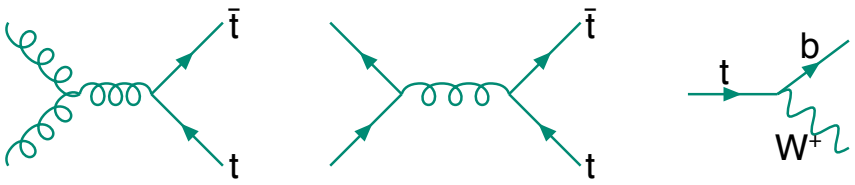

Figure 4.10: Feynman diagrams describing $t\bar{t}$ pair production at hadron colliders (left, middle), and top quark decay (right).

4.4 Parity violation

In Sect. 3.1.2, the discrete symmetry operations *parity conjugation* and *charge conjugation* were first introduced, along with the notion that particles may be eigenstates of the quantum mechanical operators corresponding to these operations (in the case of charge conjugation, this is only the case for states that remain unchanged under this operation). The implicit assumption is that

all *physics* is invariant under these operations; and this holds for the strong and electromagnetic interactions.

4.4.1 Evidence for parity violation

However, if we consider the decays of the charged kaon, focusing on some of the most important ones, we find the following:

$$\begin{aligned} \mathrm{K}^+ \to \quad & \mu^+\nu_\mu \qquad && (B = 0.635), \\ & \pi^+\pi^0 \qquad && (B = 0.207), \\ & \pi^+\pi^+\pi^- \qquad && (B = 0.055), \\ & \pi^+\pi^0\pi^0 \qquad && (B = 0.018). \end{aligned}$$

We encountered the first decay mode in Sect. 4.3.1, where its decay width was used to estimate θ_c. The three-pion decay modes have been studied extensively using partial-wave analysis, and found to involve no orbital angular momentum. However, focusing on the two-pion decay mode, we find that it violates parity! For given the spin-0 nature of the pions, the spin of the kaon is entirely determined by the orbital angular momentum in the decay. Since we know that the kaon has $J = 0$, we know that the parity in the final state is $(-1)^{L+2} = 1$. However, the kaon has $P = -1$, so parity must be violated in this weak decay!

(For the historically inclined: due to this parity violation, it was not immediately recognized that the particles decaying to two and three pions were the same. This is known as the "$\theta - \tau$ puzzle": two particles were known with the same lifetime, mass, and spin, but with apparently different parity.)

The first indisputable evidence for parity violation in weak interactions is due to a *nuclear* physics experiment, namely the study of the decay of a polarized cobalt isotope, ^{60}Co, which decays as $^{60}\mathrm{Co} \to \mathrm{e}^- + \bar{\nu}_\mathrm{e} + {}^{60}\mathrm{Ni}$. The crucial observation was that the electrons are emitted preferentially *opposite* the direction of the ^{60}Co spin (its nucleus has $J^P = 5^+$).

Why does this imply parity violation? Remember that parity conjugation means spatial inversion. This means that positions and momenta are reversed: $\vec{r} \to -\vec{r}$, $\vec{p} \to -\vec{p}$, but angular momenta are *not* (in the case of orbital angular momentum, this is easy to see: $\vec{r} \times \vec{p} \to \vec{r} \times \vec{p}$. More generally, spatial inversions and rotations commute, so it is possible to find states that are *simultaneous* spin and parity eigenstates). So we can consider the projection of spin onto the electron momentum vector, which again *does* change sign under parity conjugation. The (differential) decay rate evidently depends on such a term, which is a manifestation of parity violation.

It turns out that in fact, parity violation is inherent to the weak interaction! Once it was discovered, many of its manifestations were studied. In particular, one experiment studied the decay through electron capture $^{152}\mathrm{Eu} + \mathrm{e}^- \to {}^{152}\mathrm{Sm}^* + \nu_\mathrm{e}$. The excited ^{152}Sm, with $J = 1$, subsequently decays through the emission of a photon, resulting in a ^{152}Sm nucleus in the ground state, with nuclear spin $J = 0$. Now the electron capture takes place from the innermost shell (the K shell), implying $L = 0$ in the initial state. Also the initial ^{152}Sm has $J = 0$. So the *only* relevant spins in this process are those of the electron, the neutrino, and the photon. In particular, we

can consider the spin projection along the momentum of the neutrino (this choice of quantization axis is called *helicity*). Choosing this as our z axis, we find that

$$S_z(\mathrm{e}^-) = S_z(\gamma) + S_z(\nu_\mathrm{e}),$$

and to conserve spin, we find that the photon spin must be *opposite* to that of the neutrino. Now the captured electron is unpolarized, so that if parity is conserved, we expect to find unpolarized photons. However, it is found that always $S_z(\gamma) = +1$, and hence $S_z(\nu) = -\frac{1}{2}$. In other words, the neutrino is always *left-handed*! And this conclusion is not specific to this particular process, but is general (as is perhaps to be expected, given the "universal" way in which the neutrino couples to the W boson).

4.4.2 Separating left- and right-hand states

This left-handed nature can be generalised further: *all* couplings of fermions (antifermions) to the W boson are left-handed (right-handed) – in the limit that the fermions can be considered to be massless. (The actual coupling involves a factor $\gamma^\mu(1-\gamma^5)$, with $\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$, to be compared with the factor γ^μ in the Feynman rules of Sect. 2.3.4: this projection is onto *chiralilty* rather than helicity states. As a result, for massive fermions the situation is more complicated, but for massless fermions they are the same.)

It is for this reason that we find such counter-intuitive results as that for the decay of the charged pion π^-, which is almost exclusively to $\mu^-\bar{\nu}_\mu$, and not to $\mathrm{e}^-\bar{\nu}_\mathrm{e}$: merely on the basis of the available phase space, the branching fraction for the latter decay would be expected to be higher. What happens is that weak decay forces the $\bar{\nu}$ to be right-handed. Spin conservation (the pion has $J = 0$) then requires that the charged lepton is also right-handed. But if the charged lepton were massless, the weak decay would require the charged lepton to be left-handed, and the decay would not occur at all! This effect is larger than the phase space effect, resulting in the observed branching fractions.

A more fundamental consequence of this left-handed nature is that left- and right-handed particles are (in principle) to be regarded as *independent* degrees of freedom. For massive particles this does not seem to make much sense, since it is always possible to convert a left-handed particle into a right-handed one simply by applying a Lorentz boost and thus reversing its momentum direction (while leaving its spin untouched). However, it is to be kept in mind that the Standard Model does not allow for *ab initio* fermion masses; rather, they are generated by the Brout-Englert-Higgs mechanism, as already mentioned in Sect. 4.2. So this is not an issue. Summarizing, it turns out that our classification of leptons and quarks needs to be modified. This is how we finally end up with Table 1.1.

Parity violation in weak interactions has meanwhile been measured in great detail. For instance, the parity violation in the weak decays of charged leptons is now customarily used as a tool to analyze the polarization of those leptons. This was done *e.g.* at LEP when it operated at the Z resonance to study the polarization of τ leptons produced through $\mathrm{e}^+\mathrm{e}^- \to \tau^+\tau^-$; an example is shown in Fig. 4.11.

Finally, a comment concerning the Z boson is in order. While the $SU(2)$ weak interaction deals exclusively with left-handed fermions (and right-handed antifermions), the Z boson and the

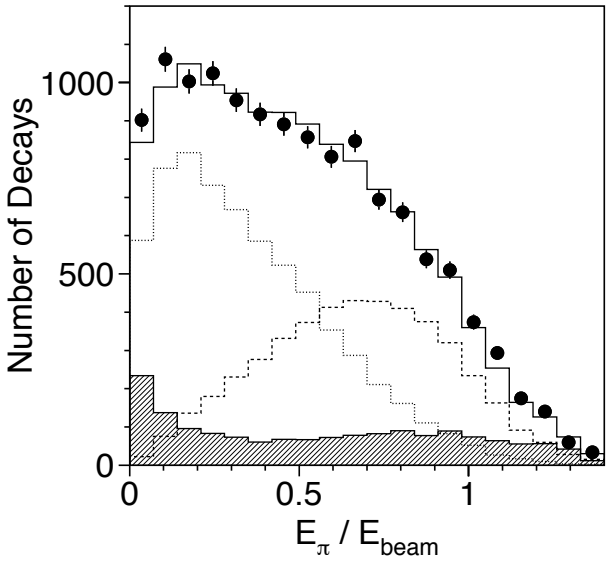

Figure 4.11: Pion energy distribution in $\tau^- \to \pi^- \nu_\tau$ decay candidates, as measured in the lab frame by the L3 experiment at LEP. There is a one-to-one correspondence between this energy and the pion emission angle in the τ^- rest frame. The expected background is denoted by the hatched area; the contributions from the two τ^- polarization states are given by the two open histograms.

photon result from a mixing of the $SU(2)$ and $U(1)_Y$ gauge bosons. Requiring that the couplings of fermions to the resulting photon remain those of QED then fixes their couplings to the Z boson; it is found that these couplings become considerably more complex, and we will not go into more detail concerning this issue.

4.5 Mixing and CP Violation in the neutral kaon system

In the preceding section, we focused primarily on parity conjugation, and did not pay much attention to charge conjugation. But this was in fact implicit in what we *did* discuss. Recalling that we said that only left-handed neutrinos and right-handed antineutrinos exist, it is not hard to relate those states to the separate or combined C and P conjugations, as shown in Fig. 4.12. So while the physics is not invariant under the separate C and P conjugations, it may be expected that it *is* invariant under the combined CP conjugation. We will see later in this section that this assertion is not confirmed by the experimental data.

4.5.1 Mixing

To see how this comes about, we return to the neutral kaons. These have appeared in the previous chapter in two contexts: once as particles of definite (strong) isospin and strangeness (K^0 and

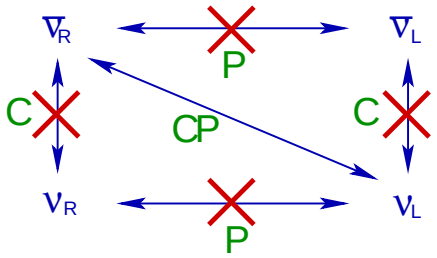

Figure 4.12: Cartoon describing the effect of applying charge and parity conjugations.

$\overline{\mathrm{K}^0}$) in the pseudoscalar meson octet in Fig. 3.4, and once as particles with definite lifetimes ($\mathrm{K_S}$ and $\mathrm{K_L}$) in Table 3.1. The K^0 and $\overline{\mathrm{K}^0}$ are related by CP conjugation:

$$CP|\mathrm{K}^0\rangle = |\overline{\mathrm{K}^0}\rangle \quad \text{and} \quad CP|\overline{\mathrm{K}^0}\rangle = |\mathrm{K}^0\rangle.$$

The inclusion of parity conjugation in these equations is necessary to deal with the kaon decays. (Given that the K^0 and $\overline{\mathrm{K}^0}$ are not their own antiparticles, an additional phase factor may be involved; however, that is merely a convention.)

How do these differences arise? The crucial point is that the (charged) weak interaction does not conserve flavour (and this affects both isospin and strangeness). This implies that as long as spin and four-momentum is conserved, and unbroken internal symmetries (colour and electrical charge) are respected, the weak interaction can mediate transitions between these two states. That this is indeed possible is shown by the Feynman diagrams of Fig. 4.13.

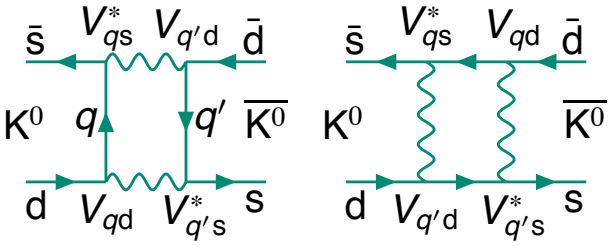

Figure 4.13: Feynman diagrams responsible for $\mathrm{K}^0 \leftrightarrow \overline{\mathrm{K}^0}$ oscillations. q and q' both represent up-type (u, c, or t) quarks.

This results in *mixing* of these states. A priori it seems reasonable to assume (we will formulate the physics more carefully in Section 4.6) that the physical particles, due to this mixing, are indeed

$$|\mathrm{K}_{1,2}\rangle = \frac{1}{\sqrt{2}}\left(|\mathrm{K}^0\rangle \pm |\overline{\mathrm{K}^0}\rangle\right).$$

We then find that K_1 and K_2 are the CP-even and CP-odd state, respectively:

$$\begin{aligned} CP|\mathrm{K}_1\rangle &= |\mathrm{K}_1\rangle, \\ CP|\mathrm{K}_2\rangle &= -|\mathrm{K}_2\rangle. \end{aligned} \tag{4.17}$$

Let us now consider the decays of these neutral kaons to two-pion ($\pi^+\pi^-$, $\pi^0\pi^0$) and three-pion ($\pi^+\pi^-\pi^0$, $\pi^0\pi^0\pi^0$) final states.

- Clearly, the two-pion decays must have $L = 0$ (since both the kaon and the pion are pseudoscalar mesons). Therefore, this final state must be CP-even (it is C-even, as detailed in Sect. 3.1.2, as well as P-even).

- The situation for the three-pion decays is somewhat more complicated. But it can nevertheless be addressed to a fair extent:

  – the orbital angular momentum can be written as $\vec{L} = \vec{L}_{12} + \vec{L}_3$, where $\vec{L}_{12}$ represents the angular momentum of the system of the first two pions, and $\vec{L}_3$ is the angular momentum of the third pion with respect to this system. But since the kaon has $J = 0$, we must have $L_3 = L_{12}$. Therefore the overall parity is $P = (-1)^{L_{12}+L_3+3} = -1$;

  – the $\pi^0\pi^0\pi^0$ final state has $C = C^3(\pi^0) = 1$ (see again Sect. 3.1.2). For the $\pi^+\pi^-\pi^0$ system we find $C = C(\pi^0) \cdot (-1)^{L_{12}}$. It is found *experimentally* that $L_{12} = 0$.

  As a result, it follows that the three-pion final state is CP-odd.

So if CP is respected, we expect two-pion and three-pion decays for the K_1 and the K_2, respectively. This has important consequences! For besides possible differences between the dynamics of the two decay processes, there is simply much less phase space available for the three-body decay ($m_\mathrm{K} - 3m(\pi) \approx 83$ MeV) than for the two-body decay ($m_\mathrm{K} - 2m(\pi) \approx 218$ MeV). There are other decays (not to CP eigenstates), but it turns out that the two-pion decay of the K_1 overwhelms all other decay modes, while a similar situation does *not* hold for the three-pion decay of the K_2. This leads to a large difference in lifetimes: as already indicated in Table 3.1, $\tau(\mathrm{K_S}) \approx 90$ ps, while $\tau(\mathrm{K_L}) \approx 50$ ns – a difference of almost three orders of magnitude! And also the branching fractions differ substantially:

$$\begin{aligned} \mathrm{K_S}: &\quad B(\pi^0\pi^0) \approx 0.31, \quad B(\pi^+\pi^-) \approx 0.69 \\ \mathrm{K_L}: &\quad B(\pi^0\pi^0\pi^0) \approx 0.21, \quad B(\pi^+\pi^-\pi^0) \approx 0.13, \quad B(\pi^\pm\ell^\mp\bar{\nu}_\ell(\nu_\ell)) \approx 0.66 \end{aligned}$$

Therefore it becomes straightforward to associate the K_1 with the $\mathrm{K_S}$, and the K_2 with the $\mathrm{K_L}$.

4.5.2 Strangeness oscillations

Let us focus for a moment on the decays to non-CP eigenstates, like the decay $\mathrm{K_L} \to \pi^\pm\ell^\mp\bar{\nu}_\ell(\nu_\ell)$. It is easily realized that we must have $\mathrm{K}^0 \to \pi^-\ell^+\nu_\ell$, so this final state *tags* the decaying meson as a K^0. This can be used to demonstrate the existence of *strangeness oscillations*.

To understand these, it is useful to keep in mind that the particle that is *produced* (typically in a strong or electromagnetic process) is a *flavour eigenstate* (K^0 or $\overline{\mathrm{K}^0}$), and *not* a mass eigenstate

(K_S or K_L). Letting $t = 0$ at the time of production, and supposing that the meson that is produced is a K^0, we find a superposition of the two mass eigenstates:

$$|K^0(t=0)\rangle = \frac{1}{\sqrt{2}}\left(|K_1(t=0)\rangle + |K_2(t=0)\rangle\right).$$

However, the time evolution of these mass eigenstates is different. Taking into account also the decays, we have:

$$|K_{1,2}(t)\rangle = e^{-im_{1,2}t}e^{-\Gamma_{1,2}t/2}|K_{1,2}(t=0)\rangle,$$

and therefore the time evolution of our original K^0 becomes

$$\begin{aligned}
|K^0(t)\rangle &= \frac{1}{\sqrt{2}}\left(e^{-im_1 t}e^{-\Gamma_1 t/2}|K_1(t=0)\rangle + e^{-im_2 t}e^{-\Gamma_2 t/2}|K_2(t=0)\rangle\right) \\
&= \frac{1}{2}\left(e^{-im_1 t}e^{-\Gamma_1 t/2} + e^{-im_2 t}e^{-\Gamma_2 t/2}\right)|K^0\rangle \\
&+ \frac{1}{2}\left(e^{-im_1 t}e^{-\Gamma_1 t/2} - e^{-im_2 t}e^{-\Gamma_2 t/2}\right)|K^0\rangle.
\end{aligned}$$

It is now straightforward to compute *rate* of particles that are produced as a K^0 to decay either as a K^0 or a $\overline{K^0}$:

$$\begin{aligned}
|\langle K^0|K^0(t)\rangle|^2 &= \frac{1}{4}\left(e^{-\Gamma_1 t} + e^{-\Gamma_2 t} + 2e^{-\frac{1}{2}(\Gamma_1+\Gamma_2)t}\cos(\Delta m t)\right), \\
|\langle \overline{K^0}|K^0(t)\rangle|^2 &= \frac{1}{4}\left(e^{-\Gamma_1 t} + e^{-\Gamma_2 t} - 2e^{-\frac{1}{2}(\Gamma_1+\Gamma_2)t}\cos(\Delta m t)\right),
\end{aligned}$$

with $\Delta m \equiv m_1 - m_2 (= 2\delta m)$. This oscillatory behaviour is clearly observed (see also Fig. 4.15). The mass difference is tiny, $5 \cdot 10^9 s^{-1} \approx 3 \cdot 10^{-12}$ MeV.

4.5.3 CP violation

Returning to the decays to CP eigenstates, we can now study these in more detail. We concluded that the K_1 meson always decays to two pions, while the K_2 meson always decays to three pions.

It is not straightforward to tell from an individual decay whether the decaying meson was indeed a K_1 or a K_2. But we can still look at the *distribution* of decay times for a given decay mode (this can be done by producing neutral kaon *beams*: longer decay times correspond to longer distances traveled) and verify that this distribution is the one expected for either the short-lived or long-lived kaon; or, given the large difference in average decay times, we can simply wait until all the short-lived kaons have decayed: it is then simply the decay modes of the long-lived kaon that are studied.

The latter approach was taken in an experiment carried out in 1964. Its conclusion was that some of the long-lived kaon decays are to two-pion final states ($B(K_L \to \pi^+\pi^-) \approx 2 \cdot 10^{-3}$ and

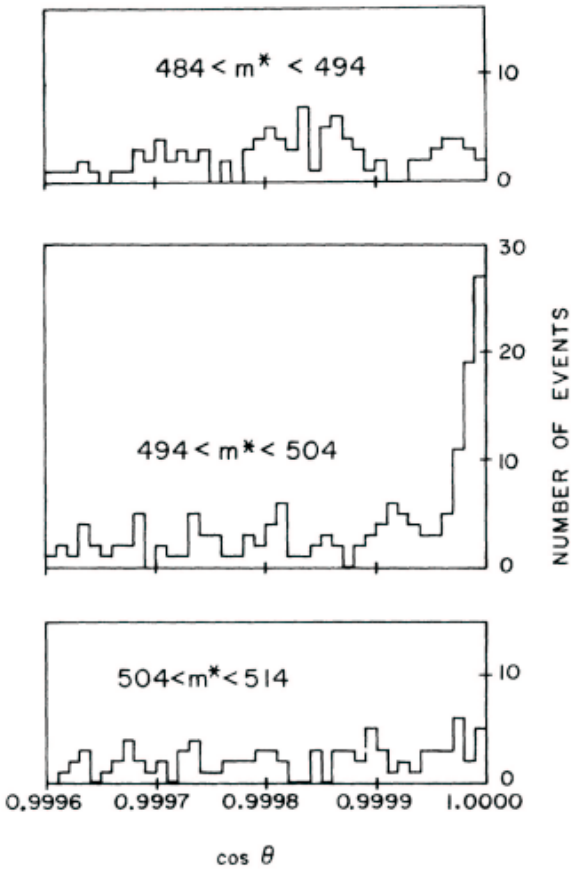

Figure 4.14: Demonstration of the decay $\mathrm{K_L} \to \pi^+\pi^-$.

$B(\mathrm{K_L} \to \pi^0\pi^0) \approx 9 \cdot 10^{-4}$). This implies that CP invariance is violated! The original discovery, in the $\pi^+\pi^-$ channel only, is shown in Fig. 4.14.

The observed amount of CP violation is small, and can be explained by a small admixture of $\mathrm{K_2}$ in the $\mathrm{K_S}$, and conversely, a small admixture of $\mathrm{K_1}$ in the $\mathrm{K_L}$:

$$\begin{aligned} |\mathrm{K_S}\rangle &= \frac{1}{\sqrt{1+|\varepsilon|^2}}\left(|\mathrm{K_1}\rangle + \varepsilon|\mathrm{K_2}\rangle\right), \\ |\mathrm{K_L}\rangle &= \frac{1}{\sqrt{1+|\varepsilon|^2}}\left(|\mathrm{K_2}\rangle + \varepsilon|\mathrm{K_1}\rangle\right), \end{aligned}$$

with $|\varepsilon| \approx 2.2 \cdot 10^{-3}$.

This mixing-induced CP violation also affects the decays to the non-CP eigenstates, because it implies that the magnitudes of the $\mathrm{K^0}$ and $\overline{\mathrm{K^0}}$ components in the $\mathrm{K_L}$ are no longer equal. This results in the *decay asymmetry*

$$\frac{\Gamma(\mathrm{K}_L \to \pi^-\ell^+\nu_\ell) - \Gamma(\mathrm{K}_L \to \pi^+\ell^-\bar{\nu}_\ell)}{\Gamma(\mathrm{K}_L \to \pi^-\ell^+\nu_\ell) + \Gamma(\mathrm{K}_L \to \pi^+\ell^-\bar{\nu}_\ell)} \approx 2\Re(\varepsilon).$$

Experimentally, the asymmetry is evaluated to be $2\Re(\varepsilon) \approx 3.3 \cdot 10^{-3}$. So a combination with the measurement of CP violation in the decays to CP eigenstates also provides information about the *phase* of ε. Fig. 4.15 shows the observed asymmetry – with its complete time dependence, which also shows the effect of the K_S.

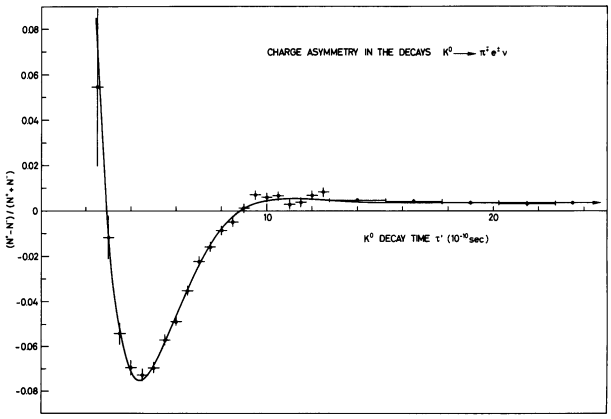

Figure 4.15: Time dependence of the charge asymmetry observed in the decay $K_L \to \pi^{\pm} e^{\mp} \bar{\nu}_e(\nu_e)$.

One may wonder whether CP violation is *only* induced through mixing, or whether other effects are also present. This has been studied by a precise comparison of the CP-violating effect in $\pi^+\pi^-$ and $\pi^0\pi^0$ final states (if CP violation is only induced through mixing, the measured CP violation should be identical in the two decay modes). It turns out that there is also a certain amount of *direct* CP violation, but this effect is again three orders of magnitude smaller than that induced by mixing. But we have not developed the theoretical machinery to interpret this phenomenon properly, and will not venture into it further.

4.5.4 K_S regeneration

From the above, it is evident that if we make a neutral kaon beam and let the kaons propagate for a sufficiently long time, we will obtain an essentially pure K_L beam. Yet it is possible to *regenerate* K_S from such a beam, by sending it through matter.

Kaons being hadrons, their usual interaction with matter will be a *strong* one. Therefore, the interaction should be considered in terms of the K^0 and $\overline{K^0}$ rather than the K_L. Part of these interactions are absorptive, *i.e.*, reduce the amount of neutral kaons:

$$\begin{aligned} K^0 + p &\to K^+ + n, \\ \overline{K^0} + n &\to K^- + p, \\ \overline{K^0} + p &\to \pi^+ + \Lambda. \end{aligned}$$

It is obvious that since the detector material is CP-asymmetric (it contains only nucleons and no anti-nucleons), the K^0 and $\overline{K^0}$ will be affected differently (see *e.g.* the third reaction above, which

has no analogue for the K^0). Representing the reduction in amplitude by factors f and $\bar{f}$ (with $\bar{f} \neq f$, from the above), we find that (neglecting the small CP violating effect in the mixing) the original

$$|K_L\rangle = \frac{1}{\sqrt{2}}\left(|K^0\rangle - |\overline{K^0}\rangle\right)$$

is transformed to

$$\frac{1}{\sqrt{2}}\left(f|K^0\rangle - \bar{f}|\overline{K^0}\rangle\right) = \frac{1}{2}(f+\bar{f})|K_L\rangle + \frac{1}{2}(\bar{f}-f)|K_S\rangle.$$

Indeed, unless $\bar{f} = f$ a K_S component is regenerated! This effect has been used in the NA48 experiment at CERN for a precise study of CP violation in the neutral kaon system.

Note that absorption of the K^0 and $\overline{K^0}$ is not even a necessary prerequisite for regeneration: also *forward elastic scattering* processes lead to different phase factors for the K^0 and $\overline{K^0}$ components that remain in the beam. (This is analogous to the description of refractive indices as resulting from the scattering of light in matter.)

4.6 Mixing and CP violation in neutral B-meson systems

4.6.1 Mixing

As is to be expected, mixing phenomena occur not only in the neutral kaon system, but also in other systems in which particle-antiparticle transitions can be mediated by the weak interaction: the B_d ($a\ \bar{b}d$ bound state), the B_s ($\bar{b}s$), and the D^0 ($c\bar{u}$) mesons, along with their antiparticles, all fall in this category. The diagrams responsible for the $B_d - \overline{B_d}$ transition are shown in Fig. 4.16. In the $B_d - \overline{B_d}$ system, the mixing phenomenon is again clearly present. It turns out that the

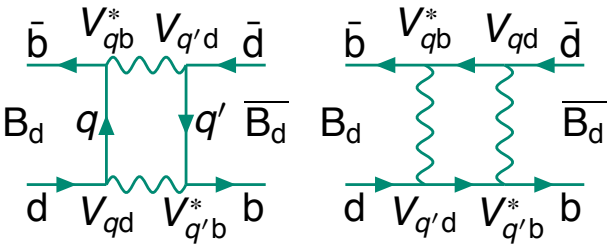

Figure 4.16: Feynman diagrams responsible for $B_d \leftrightarrow \overline{B_d}$ oscillations. q and q' both represent up-type (u, c, or t) quarks.

Feynman diagrams for this process, just as in the case of kaon oscillations, are dominated by the contributions from the (heavy) top quarks, leading to a dependence $\sim m_t^{3/2}$. A similar conclusion holds for the $B_s - \overline{B_s}$ system, where the oscillation frequency is again substantially higher due to the fact that $|V_{ts}| \gg |V_{td}|$ (see Eqn. 4.16). In contrast, in the $D^0 - \overline{D^0}$ system mixing has indeed been discovered, but only a few years ago: it is a very small effect, since the amplitude for the

transition is very small (which can be traced back to the fact that the virtual quarks occurring in the relevant diagrams are *down-type* rather than up-type quarks, combined with the fact that the b quark is much lighter than the top quark).

Let us now describe the quantum mechanics of these phenomena a bit more carefully, using the B_d–$\overline{B_d}$ system as an example. The most straightforward way to do so in the form of a two-state system (which necessarily neglects the final states that the B mesons decay to). The free-particle system can then be described by the Schrödinger equation

$$i\frac{\partial}{\partial t}\psi(t) = H\psi(t) = (M - i\Gamma/2)\psi(t), \psi(t) \equiv \begin{pmatrix} B(t) \\ \overline{B}(t) \end{pmatrix}, \tag{4.18}$$

and where M and Γ are hermitian matrices (note that this means that H is *not* hermitian! This is a consequence of our restriction to two states). On general grounds, we will furthermore assume that $M_{11} = M_{22} \equiv m$ and $\Gamma_{11} = \Gamma_{22} \equiv \overline{\Gamma}$. This can be seen as an eigenvalue problem. Solving it leads to eigenvalues $m - i\overline{\Gamma}/2 \pm \frac{1}{2}(\Delta m - i\Delta\Gamma/2)$, where Δm and $\Delta\Gamma$ must satisfy

$$\begin{aligned} \frac{1}{4}(\Delta m^2 - \frac{1}{4}\Delta\Gamma^2) &= |M_{12}|^2 - \frac{1}{4}|\Gamma_{12}|^2, \\ \frac{1}{4}\Delta m\Delta\Gamma &= \Re(M_{12}\Gamma_{12}^*). \end{aligned}$$

Now, a difference with the kaon system that is of great practical importance is the fact that the B meson is much heavier, and that consequently the phase space for many decay modes is large. Therefore, while it possible to construct CP eigenstates from the original B and $\overline{B}$ mesons, the decay modes to CP eigenstates only represent a small fraction of the total decay width. It turns out that in the B_d–$\overline{B_d}$ system, the *lifetime difference* is a small, rather than a huge effect as in the case of the neutral kaon system; we will ignore it and set $\Gamma_{12} = \Delta\Gamma = 0$. (Note that this argument does not hold in the case of the B_s–$\overline{B_s}$ system; but we will not discuss this system in detail.) In this case, the corresponding eigenvectors can be seen to be

$$|B_{H,L}\rangle = p|B\rangle \pm q|\overline{B}\rangle, \quad \text{with} \quad q/p = \sqrt{M_{12}^*/M_{12}}, \tag{4.19}$$

and where the H and L subscripts denote the heavier and lighter states (given that $\Delta\Gamma = 0$ it does not make sense to speak of a long-lived or a short-lived state).

With this solution, the time evolution of a meson produced (at time $t = 0$) as a B or as a $\overline{B}$ meson can be shown to be

$$\begin{aligned} |B(t)\rangle &= e^{-\overline{\Gamma}t/2}e^{-imt}\left(\cos(\Delta mt/2)|B\rangle + i\frac{q}{p}\sin(\Delta mt/2)|\overline{B}\rangle\right), \\ |\overline{B}(t)\rangle &= e^{-\overline{\Gamma}t/2}e^{-imt}\left(\cos(\Delta mt/2)|\overline{B}\rangle + i\frac{p}{q}\sin(\Delta mt/2)|B\rangle\right), \end{aligned} \tag{4.20}$$

leading to the oscillatory behaviour e.g. for a particle produced as a B meson:

$$\begin{aligned} |\langle B|B(t)\rangle|^2 &= e^{-\overline{\Gamma}t}\cos^2(\Delta mt), \\ |\langle \overline{B}|B(t)\rangle|^2 &= e^{-\overline{\Gamma}t}\sin^2(\Delta mt). \end{aligned}$$

4.6.2 CP violation

CP violation occurs also in the B-meson system. It manifests itself in the same way as in the neutral kaon system, as CP violation in the mixing of the neutral mesons (although this turns out to be a very small effect in the case of the B_d–$\overline{B_d}$ system, and in fact outside the scope of the above formalism as it requires $|q/p| \neq 1$) and as direct CP violation in specific decays. Rather than discussing these, however, in the following a different form of CP violation is discussed, namely, in the *interference between mixing and decay*.

This phenomenon can occur for decays of the B or $\overline{\mathrm{B}}$ meson to CP eigenstates, denoted by f in the following. The quantity of interest is the decay rate to f as a function of time, for particles produced as either a B or $\overline{\mathrm{B}}$ meson. Let us define

$$\begin{aligned} A_f &\equiv \langle f|\mathrm{B}\rangle, \\ \overline{A}_f &\equiv \langle f|\overline{\mathrm{B}}\rangle. \end{aligned}$$

From Eqn. 4.20, it is then relatively easy to arrive at

$$\begin{aligned} \langle f|\mathrm{B}(t)\rangle &= e^{-\overline{\Gamma}t/2}e^{-imt}A_f\left(\cos(\Delta mt/2)+i\lambda_f\sin(\Delta mt/2)\right), \\ \langle f|\overline{\mathrm{B}}(t)\rangle &= e^{-\overline{\Gamma}t/2}e^{-imt}\overline{A}_f\left(\cos(\Delta mt/2)+i\frac{1}{\lambda_f}\sin(\Delta mt/2)\right), \end{aligned} \tag{4.21}$$

where we have defined

$$\lambda_f \equiv \frac{q}{p}\frac{\overline{A}_f}{A_f}.$$

This is about as far as the general formalism gets us. We can say more, however, in cases where only a single diagram is responsible for the decay to the final state f. In this case, we have $|\overline{A}_f| = |A_f|$, λ_f becomes a pure phase, and the expressions for the decay rates (apart from an overall constant) become particularly simple:

$$\begin{aligned} |\langle f|\mathrm{B}(t)\rangle|^2 &= e^{-\overline{\Gamma}t}|A_f|^2\left(1+\Im(\lambda_f)\sin(\Delta mt)\right), \\ |\langle f|\overline{\mathrm{B}}(t)\rangle|^2 &= e^{-\overline{\Gamma}t}|A_f|^2\left(1-\Im(\lambda_f)\sin(\Delta mt)\right), \end{aligned} \tag{4.22}$$

so that the decay rate *asymmetry* becomes

$$\frac{|\langle f|\mathrm{B}(t)\rangle|^2-|\langle f|\overline{\mathrm{B}}(t)\rangle|^2}{|\langle f|\mathrm{B}(t)\rangle|^2+|\langle f|\overline{\mathrm{B}}(t)\rangle|^2} = 2\Im(\lambda_f)\sin(\Delta mt). \tag{4.23}$$

The above seems quite a restriction, but such final states in fact do exist! A noteworthy example is shown in Fig. 4.17. Here it is to be noticed that the kaon as shown in the final state is a K^0, which is not a CP eigenstate; but it will manifest itself as a K_S or a K_L (which implies an additional phase factor in the decay amplitude, which can however be absorbed in A_f). These decay processes have been investigated in great detail by the so-called *asymmetric B-factory* experiments BaBar (at SLAC, Stanford) and Belle (at KEK, Japan). Figure 4.18 shows the observed decay asymmetries, which are in good agreement with Eqn. 4.23. A few features are to be noticed:

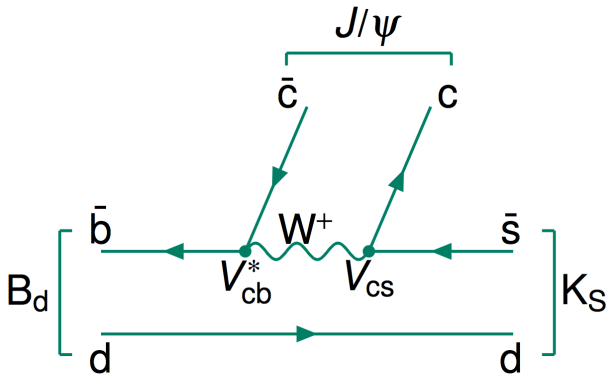

Figure 4.17: Feynman diagram representing the decay $B_d \to J/\psi K_{S,L}$.

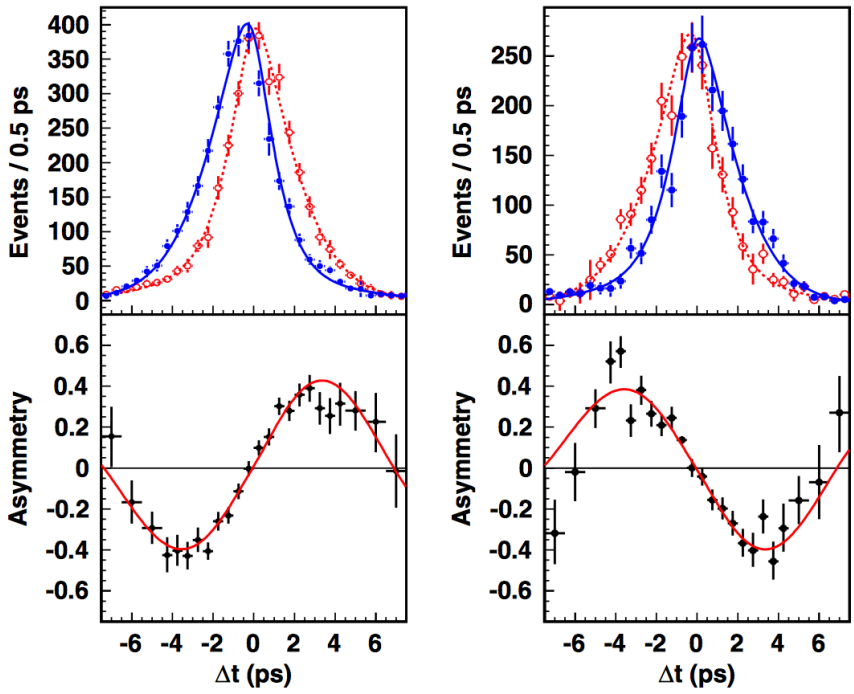

Figure 4.18: Observed decay rate asymmetries for the $B_d \to J/\psi K_S$ (left) and $B_d \to J/\psi K_L$ decays, as measured by the Belle experiment.

- The amplitudes for the two processes differ by a minus sign. It turns out that this is related to the CP eigenvalues of the $K_{S,L}$ (notice that here we ignore the fact that these aren't exactly the CP eigenstates).

- The plots show results not only for positive but also for negative times. This is related to the specific operation mode of these experiments: they exploit the decay $\Upsilon(4S) \to B_d \overline{B_d}$, in which one can show that until the first meson decay, the B mesons are in an *entangled* state in which there is always one B and one $\overline{B}$ meson; time $t = 0$ then represents the time when the "other" meson decays to a flavour eigenstate (labeling it either as a B or as a $\overline{B}$ meson). As a result, one does not look at absolute decay times but rather at their *differences*. The negative times are for cases where the "other" meson decays *after* the decay to the CP eigenstate.

Now it is all nice to observe CP violation in this fashion; however, it would be much nicer yet if it were possible to interpret the measurements (of $\Im(\lambda_f)$) in terms of fundamental parameters. Intuitively, it will be clear that the CKM matrix elements should somehow appear here, as they affect both the mixing parameters p and q and the decay amplitudes A_f and $\overline{A}_f$. We lack the theoretical machinery to derive the result rigourously, but it turns out that the result is

$$\begin{aligned} B \to J/\psi K_S: &\quad \lambda = e^{2i\beta} \\ B \to J/\psi K_L: &\quad \lambda = e^{-2i\beta} \end{aligned}$$

where the angle β is defined as

$$\beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right).$$

This angle is one of the angles of the so-called *unitarity trangle* that is the graphical representation of the equation

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0,$$

which is one of the 9 equations obtained when imposing unitarity on the CKM matrix elements.

There is much more to be said on the topic of CP violation measurements and related topics; suffice it to say that the interpretation in terms of CKM matrix elements thus far shows a consistent picture. This is shown graphically in Fig. 4.19, which shows a compilation of many measurements relevant to the above unitarity triangle.

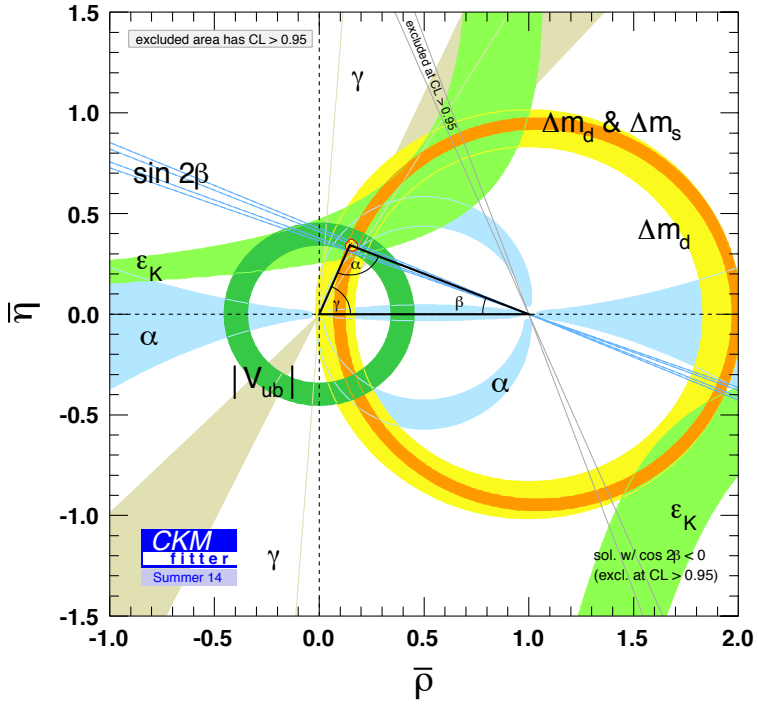

Figure 4.19: Unitarity angle constraints from CP violation and other measurements.

CP violation and the universe

On a grander scale, it is worth pointing out that the topic of CP violation is relevant also outside particle physics itself, namely, in the context of the matter-antimatter "puzzle": assuming that the Big Bang started from a symmetric state (equal amounts of matter and antimatter), an explanation is needed for why at present only matter is present; this phenomenon is called *baryogenesis*. (In fact, the situation can be quantified better: asuming particle-antiparticle annihilation processes to lead essentially to photon pairs, the *baryon-to-photon ratio* $n_B/n_\gamma \mathcal{O}(10^{-9})$ reflects the "final" asymmetry – which evidently is tiny.

The conditions needed to obtain baryogenesis have been formulated long ago (1967) by Sakharov:

1. processes are needed that violate baryon number (i.e., the number of baryons minus the number of antibaryons; note that while this is not possible *perturbatively* in the Standard Model, it does allow for phase transitions where non-perturbative processes violate baryon number conservation);
2. processes are needed to violate C and CP symmetry;
3. and this needs to happen in an out-of-equilibrium situation (as otherwise, thermal equilibrium would dictate that reactions proceed both ways).

It turns out, however, that the observed asymmetry in the universe, even if small, is too large to be explained by the CP violation mechanisms as implemented in the Standard Model.

4.7 Neutrino Oscillations

4.7.1 Neutrino Puzzles

Solar neutrinos

The topic of neutrino oscillations has a long history. It has been known since Bethe's studies in 1938 that the origin of the energy radiated by the Sun is nuclear fusion, involving the net reaction

$$4\mathrm{p} \rightarrow {}^4\mathrm{He} + 2\mathrm{e}^+ + 2\nu_\mathrm{e} + \text{photons}.$$

This does not happen all at once, but instead proceeds via a number of intermediate reactions. This makes the computation of the relation between the solar neutrino spectrum and the energy radiated by the Sun in the form of electromagnetic radiation highly non-trivial. Nevertheless, the (now) *Standard Solar Model* does provide a definite prediction for the neutrino energy spectrum; it is shown in Fig. 4.20. This model has its origins in the early 1960's, when it was being developed by Bahcall and co-workers. Indications even then were that despite the very low ν_e interaction cross section at energies $\lesssim 10$ MeV, the high flux should allow for detection of solar neutrinos.

Davis set out to verify this experimentally using the same radio-chemical technique he used earlier to demonstrate the difference between ν and $\bar{\nu}$, with the reaction.

$$\nu_\mathrm{e} + {}^{37}\mathrm{Cl} \rightarrow \mathrm{e}^- + {}^{37}\mathrm{Ar},$$

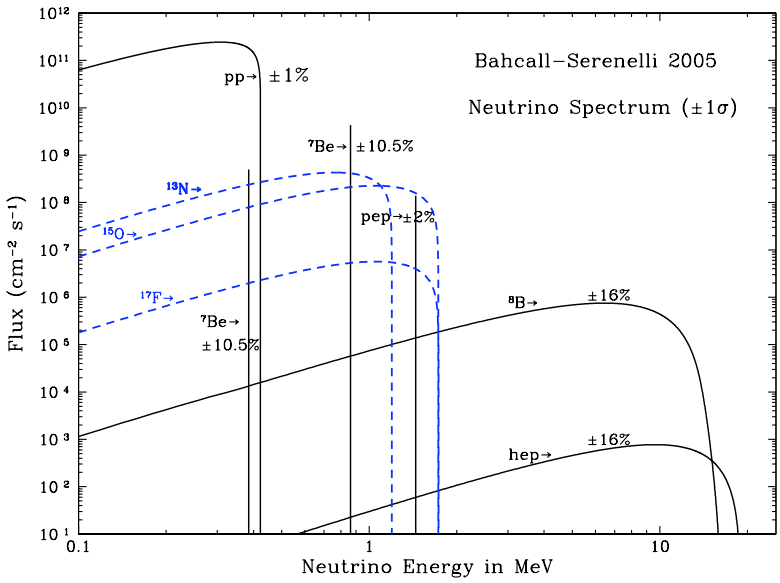

Figure 4.20: Solar neutrino energy spectrum as predicted by the Standard Solar Model. The sharp lines are due to monochromatic neutrinos originating from two-body processes. Figure taken from Ref. [9].

The experimental result, however, was that only about a third of the predicted number of solar neutrinos were observed. (Since the rate was so low – a few interactions per day, even in a volume as large as 400 m^3 – it was long thought that there was an error in the measurement; or in the flux prediction, given the complexity of the Standard Solar Model.) This is the *solar neutrino puzzle*.

This matter was settled definitively by the Sudbury Neutrino Observatory (SNO) experiment, in 2002. This ingenious experiment used *heavy* water, D_2O, instead of ordinary water, and as a consequence was sensitive to the "neutral current" interaction

$$\nu + n \rightarrow \nu + n.$$

The sensitivity was achieved by the fact that in this elastic scattering process, if the neutron concerned is part of a deuteron, may lead to deuteron dissociation. The free neutron may then recombine with another deuteron to yield an excited triton, which subsequently decays to its ground state under the emission of a relatively high-energy photon ($E_\gamma \approx 6.25$ MeV).

In addition to this neutral current interaction, the other relevant processes are:

- the "standard" charged current interaction $\nu_e + n \rightarrow e^- + p$, where only the Cherenkov radiation emitted by the electron is detected;
- the "quasi-elastic" process $\nu + e^- \rightarrow \nu + e^-$, where again only the Cherenkov radiation emitted by the electron is detected. This process is truly elastic if the incoming neutrino is

a ν_μ or ν_τ; however, in the case of ν_e there is a second diagram leading to a significantly increased cross section. As a result, the sensitivity to ν_e is significantly larger than that to $\nu_{\mu,\tau}$.

In the experiment, only energy and direction of photon radiation (whether Cherenkov radiation or gamma rays) were measured; nevertheless, it was possible to distinguish between the above three processes on a *statistical* basis (by considering a number of distributions of variables that are known to be distributed differently for the three processes).

The result is shown in Fig. 4.21. This demonstrates that

- the measured *total* solar neutrino flux is in good agreement with the predictions from the Standard Solar Model (combined with the modelling of the interaction cross sections and detector response);

- but a significant fraction of the ν_e have converted to ν_μ or ν_τ (the experiment is not able to tell these two apart).

As an aside, one would be tempted to argue that similar conclusions could perhaps been obtained from the study of only the charged-current and quasi-elastic processes (which would not need a heavy water experiment). This is true, in principle, but the fact that the sensitivity of the quasi-elastic scattering process to ν_e is so much larger than that to $\nu_{\mu,\tau}$ made it very difficult to exploit this, due to the measurement uncertainties.

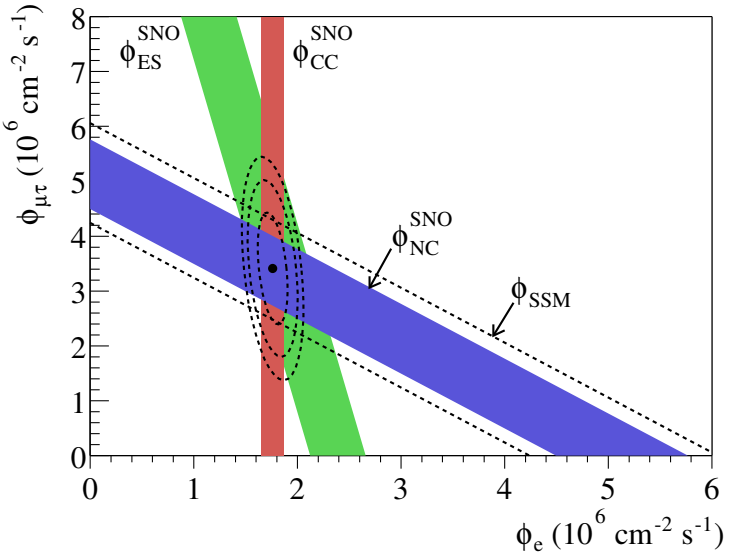

Figure 4.21: Constraints from the SNO experiment on the fluxes of solar ν_e and $\nu_{\mu,\tau}$. Figure taken from Ref. [10].

Atmospheric neutrinos

A second long-standing issue surrounding neutrinos is the *atmospheric neutrino puzzle*. This is related to the decays of charged pions produced in collisions of so-called "cosmic ray" particles with molecules in the atmosphere. As discussed above, such decays are almost exclusively to muons, $\pi^+ \to \mu^+ \nu_\mu$. However, the muons also decay, $\mu^+ \to \bar{\nu}_\mu e^+ \nu_e$, so twice as many ν_μ as ν_e are expected (roughly similar amounts of π^+ and π^- are produced in cosmic-ray interactions). The fact that not all muons decay before having reached ground level changes this conclusion only marginally.

The Super-Kamiokande experiment studied the ν_e and ν_μ fluxes in great detail, using a very large detector sensitive to charged-current interactions of both neutrino types for neutrino energies of order of at least a few hundred MeV (the atmospheric neutrino flux peaks at such low energies and then falls steeply with energy). It found that experimentally, the number of ν_μ and $\bar{\nu}_\mu$ is substantially smaller than expected, and their deficit exhibits a strong zenith angle dependence, as shown in Fig. 4.22. It is to be noticed that there is a strong correlation between the

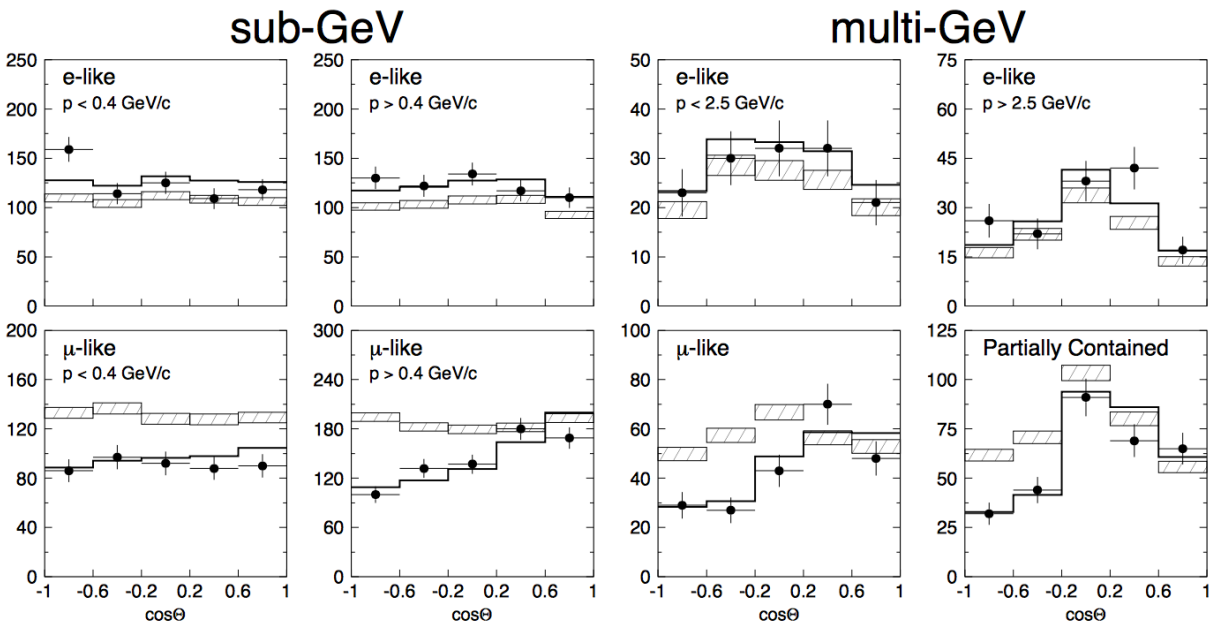

Figure 4.22: Comparisons between predicted and observed atmospheric interactions as observed by the Super-Kamiokande experiment. The hatched areas represent the predictions under the hypothesis that no neutrino oscillations occur; the solid line represents the best fit to a $\nu_\mu \to \nu_\tau$ oscillation hypothesis. Figure taken from Ref. [11].

zenith angle and the path traversed by the neutrino produced in the decay processes in the atmosphere: neutrinos with $\cos\theta = 1$ are produced directly above the detector and have travelled only some 10 km before interacting in the detector, while neutrinos with $\cos\theta = -1$ have been produced on the other side of the Earth and have travelled about $12 \cdot 10^3$ km. The situation is

sketched in Fig. 4.23. This result (from 1998) was the first evidence for neutrino oscillations,

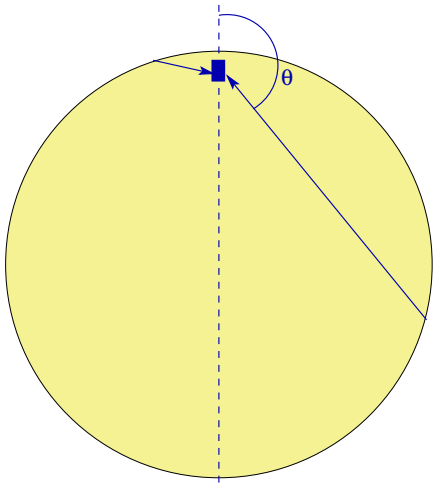

Figure 4.23: Correspondence between zenith angle and path length traversed by muon neutrinos detected in the Super-Kamiokande detector.

which we discuss next.

4.7.2 Neutrino Oscillation Formalism

So far, we have treated neutrinos as massless particles. However, it can be shown that if neutrinos *do* have a (small) mass, it is possible for neutrinos to be transformed from one flavour into another.

The first question that should be asked here is whether there is any direct experimental evidence for a nonzero neutrino mass. This is studied most easily using nuclear β processes, involving the underlying process $\mathrm{n} \to \mathrm{p} + \mathrm{e}^- + \bar{\nu}_\mathrm{e}$ or $\mathrm{p} \to \mathrm{n} + \mathrm{e}^+ + \nu_\mathrm{e}$ (depending on the nuclear binding energy for the nucleus in question). In either case, the decay is a three-body decay, leading to an electron (or positron) energy *spectrum*. This spectrum, and in particular its shape at the highest energies allowed kinematically, is sensitive to the mass of the (electron) neutrino. The most precise measurements to date have been done on tritium decay

$$^3\mathrm{H} \to ^3\mathrm{He} + \mathrm{e}^- + \bar{\nu}_\mathrm{e},$$

leading to an upper limit for a possible neutrino mass: $m(\nu_\mathrm{e}) < 2$ eV. Also cosmological measurements constrain neutrinos to be lighter than a few eV. So *if* neutrinos do have mass, their masses must at least be very small, and they are ultra-relativistic for all practical purposes.

The neutrino oscillation formalism is most easily demonstrated in a two-state system of, say, ν_e and ν_μ. We assume that, like in the case of the quarks, a *mixing* takes place between mass eigenstates and flavour eigenstates (the states that couple with charged leptons and W bosons). Denoting the former by ν_1 and ν_2 (with masses m_1 and m_2), and the latter by ν_e and ν_μ, we have:

$$\begin{pmatrix} \nu_\mathrm{e} \\ \nu_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}$$

Suppose now that at space-time point x_1, a neutrino is produced as a ν_e, and that its flavour is measured at point x_2. We have

$$\begin{aligned}
|\nu_e\rangle &= \cos\theta|\nu_1\rangle + \sin\theta|\nu_2\rangle, \\
|\nu_e(t)\rangle &= e^{-ip_1\cdot(x_2-x_1)}\cos\theta|\nu_1\rangle + e^{-ip_2\cdot(x_2-x_1)}\sin\theta|\nu_2\rangle \\
&= \left(e^{-ip_1\cdot(x_2-x_1)}\cos^2\theta + e^{-ip_2\cdot(x_2-x_1)}\sin^2\theta\right)|\nu_e\rangle \\
&+ \left(e^{-ip_2\cdot(x_2-x_1)} - e^{-ip_1\cdot(x_2-x_1)}\right)\sin\theta\cos\theta|\nu_\mu\rangle.
\end{aligned}$$

The oscillation probability therefore becomes

$$\begin{aligned}
P(\nu_e\to\nu_\mu) = |\langle\nu_\mu|\nu_e(t)\rangle|^2 &= \cos^2\theta\sin^2\theta\,(2-2\cos((p_2-p_1)\cdot(x_2-x_1))) \\
&= \sin^2(2\theta)\sin^2((p_2-p_1)\cdot(x_2-x_1)/2).
\end{aligned}$$

Clearly, since the two mass eigenstates propagate over macroscopic distances, the four-momenta p_1 and p_2 must be different. For the sake of simplicity, let us assume that they have the same $\vec{p}$ and differ only in their energy. The expression for the oscillation frequency then simplifies to

$$P(\nu_e\to\nu_\mu) = \sin^2(2\theta)\sin^2((E_2-E_1)(t_2-t_1)/2).$$

Finally, we exploit the fact that the neutrinos are ultra-relativistic, so that

$$E_i \approx |\vec{p}| + m_i^2/2|\vec{p}|,$$

so

$$P(\nu_e\to\nu_\mu) = \sin^2(2\theta)\sin^2(\Delta m^2 L/4E), \tag{4.24}$$

where $\Delta m^2 = m_2^2 - m_1^2$. In addition we have used $E \approx |\vec{p}|$ in the denominator, and $L = \Delta t$ (in typical experiments, distances rather than time differences are measured).

From the above, it should be evident that if the oscillation probability can be measured as a function of L/E, both the value of $\sin^2(2\theta)$ and that of $|\Delta m^2|$ can be determined (and moreover, proves that the apparent lack of neutrinos in the above-mentioned puzzles is really due to oscillations).

4.7.3 Oscillation Signals

It is beyond the scope of this lecture course to give an exhaustive overview of all the experiments that have been conducted since the solar neutrino puzzle first arose. It should be pointed out, however, that both the SNO and Super-Kamiokande measurements fit into a coherent picture.

The oscillation signals themselves have now been established unambiguously. The solar neutrino oscillation signal has been measured by the KamLAND experiment, situated in Japan and sensitive to the $\bar{\nu}_e$ emitted by fission reactors at distances up to 180 km. This oscillation signal corresponds to $|\Delta m^2| = (8.0\pm0.3)\cdot10^{-5}$ eV^2 and $\sin^2(2\theta) = 0.86\pm0.03$. The atmospheric neutrino oscillation signal has been measured using the Super-Kamiokande experiment.

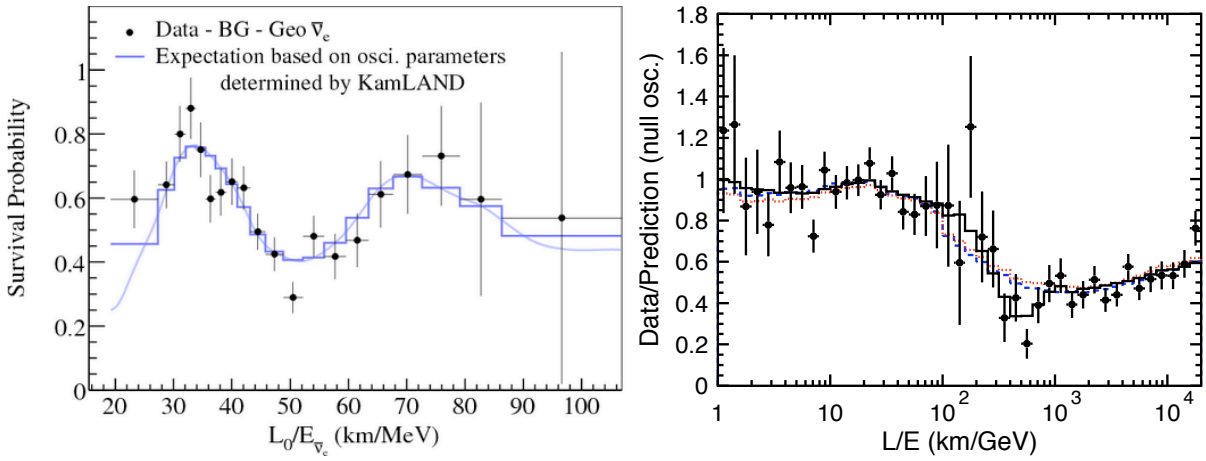

Figure 4.24: Neutrino oscillation signals as measured by the KamLAND experiment (left) and the Super-Kamiokande experiment (right). Notice the difference in scales between the two signals.

This measurement makes use of the correlation between zenith angle and path length. This signal corresponds to $|\Delta m^2| \approx 2.5 \cdot 10^{-3}$ eV^2 and $\sin^2(2\theta) > 0.92$. The oscillation signals are shown in Fig. 4.24. In addition, the Tokai-to-Kamioka (T2K) experiment (which actually uses the Super-Kamiokande experiment but uses it to detect (anti)neutrinos created as higher energy ν_μ ($E_\nu \sim 600$ MeV) resulting from decays of charged pions produced in a high energy beam. This resulted in an *appearance* measurement, demonstrating that $\nu_\mu \rightarrow \nu_e$ oscillations do take place at the "atmospheric" frequency (even if they are a small effect).

Finally, it should be added that there are quite a few more experiments that have not been mentioned above, but that all help to provide a solid picture of neutrino oscillations. Even with the experiments described above, however, it is clear that the simplistic two-neutrino oscillation formalism discussed in Section 4.7.2 cannot be adequate. It is therefore necessary to extend to the full picture.

4.7.4 The PMNS Matrix

Of course, we know that there are three, rather than two, neutrino species. So the two-state result of Eqn. 4.24 is incomplete. Like the CKM matrix, the Pontecorvo-Maki-Nakagawa-Sakata or PMNS matrix (actually conceived before the CKM matrix) features four parameters: three rotation angles and one complex phase. It can be parametrized as

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \cdot \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \cdot \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{4.25}$$

where we use shorthands $c_{23} \equiv \cos\theta_{23}$, $s_{23} \equiv \sin\theta_{23}$ for the terms in the first matrix, and similarly for the sine and cosine terms related to the other two rotation angles.

Proper three-family analyses of the available data have been carried out, leading to

$$\begin{aligned} \sin^2\theta_{12} &= 0.308 \pm 0.017, \\ \sin^2\theta_{23} &= 0.390 \pm 0.033, \\ \sin^2\theta_{13} &= 0.0242 \pm 0.0026. \end{aligned}$$

While the precision on these measurements is still being improved, attention is now focusing on the measurement of the phase δ. This is relevant especially as a nonzero phase implies CP violation in the lepton sector, and in turn this could offer alternative explanations of the matter-antimatter asymmetry.

4.7.5 Neutrinos and the Standard Model

The neutrino oscillation measurements provide information about the difference between squared neutrino masses. However, they do not indicate what these absolute masses should be – at most that at least two of them must be nonzero.

This is the last bit of information relevant to Table 1.1. Now that it has been established that neutrinos do have mass, it makes sense to introduce right-handed neutrino states in addition to right-handed charged lepton states. So why the question marks in Table 1.1? The thing is that such right-handed neutrinos carry no conserved quantum numbers. This makes it possible, by a construction due to Majorana, that they are in fact their own anti-particles! This possibility is being investigated in so-called *neutrinoless double beta decay* experiments, in which double beta decays are searched for that do *not* involve the emission of additional neutrinos. Such processes can only occur if neutrinos are indeed Majorana particles, and would be sensitive to the neutrino masses themselves rather than (squared) mass differences.

Appendix A

Clebsch-Gordan coefficients

36. CLEBSCH-GORDAN COEFFICIENTS, SPHERICAL HARMONICS, AND *d* FUNCTIONS

Note: A square-root sign is to be understood over *every* coefficient, *e.g.*, for $-8/15$ read $-\sqrt{8/15}$.

Notation:

J	J	...
M	M	...
m_1	m_2	Coefficients
m_1	m_2	
.	.	
.	.	
.	.	

$1/2\times1/2$

	1		
+1		1	0
+1/2	+1/2	1	0 0
+1/2	−1/2	1/2 1/2	1
−1/2	+1/2	1/2 −1/2	−1
	−1/2	−1/2	1

$$Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\theta$$

$$Y_1^1 = -\sqrt{\frac{3}{8\pi}}\sin\theta\, e^{i\phi}$$

$$Y_2^0 = \sqrt{\frac{5}{4\pi}}\left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)$$

$$Y_2^1 = -\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta\, e^{i\phi}$$

$$Y_2^2 = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\sin^2\theta\, e^{2i\phi}$$

$2\times1/2$

	5/2		
+2	+1/2	1	3/2 3/2
+2	−1/2	1/5 4/5	5/2 3/2
+1	+1/2	4/5 −1/5	1/2 1/2
+1	−1/2	2/5 3/5	5/2 3/2
0	+1/2	3/5 −2/5	−1/2 −1/2
0	−1/2	3/5 2/5	5/2 3/2
−1	+1/2	2/5 −3/5	−3/2 −3/2
−1	−1/2	4/5 1/5	5/2
−2	+1/2	1/5 −4/5	−5/2
−2	−1/2		1

$1\times1/2$

	3/2		
+1	+3/2	1	3/2 1/2
+1	−1/2	1/3 2/3	3/2 1/2
0	+1/2	2/3 −1/3	−1/2 −1/2
0	−1/2	2/3 1/3	3/2
−1	+1/2	1/3 −2/3	−3/2
	−1	−1/2	1

$3/2\times1/2$

	2		
+2	+3/2	1	2 1
+2	+1/2	1/4 3/4	2 1
+1	+1/2	3/4 −1/4	0 0
+1	−1/2	1/2 1/2	2 1
0	+1/2	1/2 −1/2	−1 −1
0	−1/2	3/4 1/4	2
−1	+1/2	1/4 −3/4	−2
−1	−1/2		1

2×1

	3		
+2	+1	1	3 2
+2	0	1/3 2/3	3 2 1
+1	+1	2/3 −1/3	+1 +1 +1
+2	−1	1/15 1/3	3/5
+1	0	8/15 1/6	−3/10
0	+1	2/5 −1/2	1/10

$3/2\times1$

	5/2		
+3/2	+1	1	5/2 3/2
+3/2	0	2/5 3/5	5/2 3/2 1/2
+1/2	+1	3/5 −2/5	+1/2 +1/2 +1/2
+3/2	−1	1/10 2/5	1/2
+1/2	0	3/5 1/15	−1/3
−1/2	+1	3/10 −8/15	1/6

1×1

	2		
+1	+1	1	2 1
+1	0	1/2 1/2	2 1 0
0	+1	1/2 −1/2	0 0 0
+1	−1	1/6 1/2	1/3
0	0	2/3	0 −1/3
−1	+1	1/6 −1/2	1/3
		0 −1	1/2 1/2 2
		−1 0	1/2 −1/2 −2
			−1 −1 1

		3	2 1
		0	0 0
+1	−1	1/5	1/2 3/10
0	0	3/5	0 −2/5
−1	+1	1/5	−1/2 3/10

		3	2 1
		−1	−1 −1
0	−1	2/5	1/2 1/10
−1	0	8/15	−1/6 −3/10
−2	+1	1/15	−1/3 3/5

		3	2
		−2	−2
−1	−1	2/3	1/3
−2	0	1/3	−2/3
		−2	−1 1

		5/2	3/2 1/2
		−1/2	−1/2 −1/2
+1/2	−1	3/10	8/15 1/6
−1/2	0	3/5	−1/15 −1/3
−3/2	+1	1/10	−2/5 1/2

		5/2	3/2
		−3/2	−3/2
−1/2	−1	3/5	2/5
−3/2	0	2/5	−3/5
		−3/2	−1 1

$$Y_\ell^{-m} = (-1)^m Y_\ell^{m*}$$

$$d^\ell_{m,0} = \sqrt{\frac{4\pi}{2\ell+1}}\, Y_\ell^m\, e^{-im\phi}$$

$$\langle j_1 j_2 m_1 m_2 | j_1 j_2 JM\rangle = (-1)^{J-j_1-j_2}\langle j_2 j_1 m_2 m_1 | j_2 j_1 JM\rangle$$

$$d^j_{m',m} = (-1)^{m-m'} d^j_{m,m'} = d^j_{-m,-m'}$$

$3/2\times3/2$

	3		
+3/2	+3/2	1	3 2
+3/2	+1/2	1/2 1/2	3 2 1
+1/2	+3/2	1/2 −1/2	+1 +1 +1
+3/2	−1/2	1/5 1/2	3/10
+1/2	+1/2	3/5 0	−2/5
−1/2	+3/2	1/5 −1/2	3/10

		3	2 1 0
		0	0 0 0
+3/2	−3/2	1/20	1/4 9/20 1/4
+1/2	−1/2	9/20	1/4 −1/20 −1/4
−1/2	+1/2	9/20	−1/4 −1/20 1/4
−3/2	+3/2	1/20	−1/4 9/20 −1/4

$2\times3/2$

	7/2		
+2	+3/2	1	7/2 5/2
+2	+1/2	3/7 4/7	7/2 5/2 3/2
+1	+3/2	4/7 −3/7	+3/2 +3/2 +3/2
+2	−1/2	1/7 16/35	2/5
+1	+1/2	4/7 1/35	−2/5
0	+3/2	2/7 −18/35	1/5

		7/2	5/2 3/2 1/2
		+1/2	+1/2 +1/2 +1/2
+2	−3/2	1/35	6/35 2/5 2/5
+1	−1/2	12/35	5/14 0 −3/10
0	+1/2	18/35	−3/35 −1/5 1/5
−1	+3/2	4/35	−27/70 2/5 −1/10

		7/2	5/2 3/2 1/2
		−1/2	−1/2 −1/2 −1/2
+1	−3/2	4/35	27/70 2/5 1/10
0	−1/2	18/35	3/35 −1/5 −1/5
−1	+1/2	12/35	−5/14 0 3/10
−2	+3/2	1/35	−6/35 2/5 −2/5

		7/2	5/2 3/2
		−3/2	−3/2 −3/2
0	−3/2	2/7	18/35 1/5
−1	−1/2	4/7	−1/35 −2/5
−2	+1/2	1/7	−16/35 2/5

		7/2	5/2
		−5/2	−5/2
−1	−3/2	4/7	3/7 7/2
−2	−1/2	3/7	−4/7 −7/2
		−2	−3/2 1

2×2

	4		
+2	+4	1	4 3
+2	+2	1/2 1/2	+3 +3
+2	+1	1/2 −1/2	4 3 2
+2	0	3/14 1/2	+2 +2 +2
+1	+1	4/7 0	2/7
0	+2	3/14 −1/2	2/7

		4	3 2 1
		+1	+1 +1 +1
+2	−1	1/14	3/10 3/7 1/5
+1	0	3/7	1/5 −1/14 −3/10
0	+1	3/7	−1/5 −1/14 3/10
−1	+2	1/14	−3/10 3/7 −1/5

		4	3 2 1 0
		0	0 0 0 0
+2	−2	1/70	1/10 2/7 2/5 1/5
+1	−1	8/35	2/5 1/14 −1/10 −1/5
0	0	18/35	0 −2/7 0 1/5
−1	+1	8/35	−2/5 1/14 1/10 −1/5
−2	+2	1/70	−1/10 2/7 −2/5 1/5

		4	3 2 1
		−1	−1 −1 −1
+1	−2	1/14	3/10 3/7 1/5
0	−1	3/7	1/5 −1/14 −3/10
−1	0	3/7	−1/5 −1/14 3/10
−2	+1	1/14	−3/10 3/7 −1/5

		4	3 2
		−2	−2 −2
0	−2	3/14	1/2 2/7
−1	−1	4/7	0 −3/7
−2	0	3/14	−1/2 2/7

		4	3
		−3	−3
−1	−2	1/2	1/2 4
−2	−1	1/2	−1/2 −4
		−2	−2 1

$$d^{3/2}_{3/2,3/2} = \frac{1+\cos\theta}{2}\cos\frac{\theta}{2}$$

$$d^{3/2}_{3/2,1/2} = -\sqrt{3}\frac{1+\cos\theta}{2}\sin\frac{\theta}{2}$$

$$d^{3/2}_{3/2,-1/2} = \sqrt{3}\frac{1-\cos\theta}{2}\cos\frac{\theta}{2}$$

$$d^{3/2}_{3/2,-3/2} = -\frac{1-\cos\theta}{2}\sin\frac{\theta}{2}$$

$$d^{3/2}_{1/2,1/2} = \frac{3\cos\theta-1}{2}\cos\frac{\theta}{2}$$

$$d^{3/2}_{1/2,-1/2} = -\frac{3\cos\theta+1}{2}\sin\frac{\theta}{2}$$

$$d^2_{2,2} = \left(\frac{1+\cos\theta}{2}\right)^2$$

$$d^2_{2,1} = -\frac{1+\cos\theta}{2}\sin\theta$$

$$d^2_{2,0} = \frac{\sqrt{6}}{4}\sin^2\theta$$

$$d^2_{2,-1} = -\frac{1-\cos\theta}{2}\sin\theta$$

$$d^2_{2,-2} = \left(\frac{1-\cos\theta}{2}\right)^2$$

$$d^2_{1,1} = \frac{1+\cos\theta}{2}(2\cos\theta-1)$$

$$d^2_{1,0} = -\sqrt{\frac{3}{2}}\sin\theta\cos\theta$$

$$d^2_{1,-1} = \frac{1-\cos\theta}{2}(2\cos\theta+1)$$

$$d^2_{0,0} = \left(\frac{3}{2}\cos^2\theta - \frac{1}{2}\right)$$

$$d^{1/2}_{0,0} = \cos\theta \qquad d^{1/2}_{1/2,1/2} = \cos\frac{\theta}{2} \qquad d^{1/2}_{1/2,-1/2} = -\sin\frac{\theta}{2}$$

$$d^1_{1,1} = \frac{1+\cos\theta}{2} \qquad d^1_{1,0} = -\frac{\sin\theta}{\sqrt{2}} \qquad d^1_{1,-1} = \frac{1-\cos\theta}{2}$$

Figure 36.1: The sign convention is that of Wigner (*Group Theory*, Academic Press, New York, 1959), also used by Condon and Shortley (*The Theory of Atomic Spectra*, Cambridge Univ. Press, New York, 1953), Rose (*Elementary Theory of Angular Momentum*, Wiley, New York, 1957), and Cohen (*Tables of the Clebsch-Gordan Coefficients*, North American Rockwell Science Center, Thousand Oaks, Calif., 1974).

Appendix B

Basics of Group Theory

The mathematical description of symmetries, of eminent importance in particle physics, is provided by the branch of mathematics called *group theory*. The aim of this appendix is to collect some basics of group theory that are made use of in the lectures.

B.1 Axioms and definitions

In an abstract fashion, groups are sets of *elements*, together with an object called the *group operator* $\cdot$, which (starting from two elements) can produce another element. Calling the set G and its elements a, b, c, ..., one speaks of a group if the following requirements (called the *group axioms*) are satisfied:

closure: for any elements a and b of G, also $a \cdot b$ is an element of G;

associativity: for any elements a, b, and c, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;

identity: there exists an element (let's call it i) satisfying $a \cdot i = i \cdot a = a$ for all a. This element is called the *identity* element;

inverse: for every a there exists an element b, satisfying $a \cdot b = b \cdot a = i$. This element is called the *inverse* of a and is usually denoted as a^{-1}.

An easy example of a group is the set of operations that can be applied to a regular polygon (take for example a square) in such a way that all sides of the "transformed" polygon coincide with sides of the original polygon. For the case of the square, one can think of its center being at a position $(0,0)$. Then it is not very hard to see that the relevant group (called the *dihedral* group D_4) is given by the following set of operations:

- a rotation over 0, 90, 180, or 270 degrees around the centre (the first is the identity element);
- and reflecting the square in a line either joining two opposing points or bisecting two opposing sides.

In addition to the above, two group elements a and b are said to *commute* if $a \cdot b = b \cdot a$. If *all* elements of a group commute, then the group is called *abelian*. This would *e.g.* be the case for our square if we weren't allowed to apply the reflections mentioned above and could only use the rotations.

A *group homomorphism* is a function that preserves the group structure. That is, if we call this function f, then it implies that for any a and b, $f(a) \cdot f(b) = f(a \cdot b)$. This naturally implies that f maps G onto another group (let's call this H); if a reverse map g also exists (*i.e.*, $g(f(a)) = a$ for all a), the function f is called a *group isomorphism*.

All of the above can be formulated in a very abstract way. More concrete implementations are often given in terms of *group representations*, describing linear transformations of vector spaces. This naturally leads to group elements being represented by *matrix multiplications*. It will not come as a surprise that such representations are of eminent importance in a description of symmetries in a quantum mechanical context.

B.2 Continuous symmetries

The above examples of regular polygons all refer to *finite* groups, *i.e.*, groups containing a finite number of elements. But these are not the only existing groups. Instead of the square above, take *e.g.* the circle. It is then evident that a rotation (about its centre) over *any* angle transforms this circle onto itself; in other words, this group contains infinitely many elements.

B.2.1 SO(2)

This group is called SO(2), for *special orthogonal* transformations in two dimensions:

- *orthogonal* means that the norm of vectors is left invariant; and
- *special* restricts the possible operations to rotations only (reflections across any given axis would otherwise be possible).

A straightforward representation of this group is given by the 2×2 rotation matrices describing counter-clockwise rotations about angles α:

$$M_\alpha = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}. \tag{B.1}$$

Now this group is abelian, and hence this single rotation can be thought of *e.g.* as the product of n rotations over angles α/n: $M_\alpha = (M_{\alpha/n})^n$. This property can be exploited by letting $n \to \infty$, expanding the sine and cosine terms in Taylor series, and retaining only their first terms:

$$M_\alpha = \lim_{n\to\infty} \begin{pmatrix} 1 & \alpha/n \\ -\alpha/n & 1 \end{pmatrix}^n = \lim_{n\to\infty} (\mathbb{1} - i\frac{\alpha}{n} J)^n, \quad \text{with}^{1} \; J \equiv \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \tag{B.2}$$

and $\mathbb{1}$ representing the 2×2 identity matrix. Finally one can make use of the identity $e^A = \lim_{n\to\infty}(1+A/n)^n$ to obtain

$$M_\alpha = e^{-i\alpha J}. \tag{B.3}$$

This means that *all* of these rotations can be generated starting from the matrix J: J is called the *generator* of this group.

The above can in fact be concluded in a much easier way: these SO(2) transformations can be considered as equivalent to rotations in the *complex plane*, which are given as multiplications by the number $e^{i\alpha}$ for some (real) value of α. This is formally another group, U(1), meaning the group of *unitary* transformations in one (complex) dimension, *i.e.*, transformations leaving the norm of complex numbers invariant. From the above, it is straightforward to see that U(1) and SO(2) are isomorphic to each other.

B.2.2 SO(3)

The situation gets more interesting when considering *e.g.* SO(3), describing rotations in three (real) dimensions. Clearly one now has rotations about three possible axes. Taking these to be the three axes $\hat{x}$, $\hat{y}$, and $\hat{z}$ of a usual cartesian system, one finds that the corresponding rotation matrices are given by

$$R_{\alpha;\hat{x}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{pmatrix},\ R_{\beta;\hat{y}} = \begin{pmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{pmatrix},\ R_{\gamma;\hat{z}} = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It is straightforward to derive again the generators for these matrices:

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & i \\ 0 & -i & 0 \end{pmatrix},\ J_2 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix},\ J_3 = \begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{B.4}$$

and that an arbitrary rotation can be described as

$$R = e^{-i\vec{\alpha}\cdot\vec{J}}. \tag{B.5}$$

It is well known that arbitrary SO(3) rotations do not commute (contrary to SO(2)): this is a *non-abelian* group. This can be expressed by nontrivial commutation relations between the three generators (as can be verified explicitly using Eqn. B.4:

$$[J_k, J_l] = 2i\sum_m \varepsilon_{klm} J_m, \tag{B.6}$$

where the symbol ε_{klm} is the completely antisymmetric unit tensor of rank 3 ($\varepsilon_{123} = -\varepsilon_{213} = \ldots = 1$; and a null value results if any of the three indices are identical). These commutation relations constitute the so-called *Lie algebra* of SO(3), which determine to a large extent the group properties (as can be deduced from Eqn. B.5). Equations B.6 constitute the *Lie algebra* of the SO(3) symmetry group.

B.2.3 SU(2)

The group SU(2) deals with all unitary matrices in two dimensions, with the "special" requirement (the "S"in the group name) that the matrices have determinant equal to unity. An Ansatz can be made for such matrices: suppose that they have the form

$$A = \begin{pmatrix} r_0 - ir_3 & -r_2 - ir_1 \\ r_2 - ir_1 & r_0 + ir_3 \end{pmatrix}.$$

A priori, these four degrees of freedom (the r_i are real) are one too many. This can be resolved by realising that the unitarity condition needs to be verified explicitly. It can be seen to be satisfied if

$$r_0^2 + r_1^2 + r_2^2 + r_3^2 = 1,$$

and under these conditions also the unit determinant is obtained immediately. With the above condition, we can now write

$$r_0 = \sqrt{1 - (r_1^2 + r_2^2 + r_3^2)}. \tag{B.7}$$

Expanding for small $\mathrm{d}r_1, \mathrm{d}r_2, \mathrm{d}r_3$, Eqn. B.7 implies that to first order, $r_0 \approx 1$, and A can be written as

$$A = \mathbb{1} - i\mathrm{d}\vec{r} \cdot \vec{\sigma},$$

where the σ_i are the (hermitian) Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

But we know the commutation relations for the Pauli matrices: $[\sigma_k, \sigma_l] = 2i\sum_m \varepsilon_{klm}\sigma_m$. So the Lie algebra of SO(3) is identical to that involving the Pauli matrices! This means that there is a close connection between SO(3) on the one hand and SU(2) on the other: they are locally equivalent. Globally they are different, as can be realised from the fact that in Eqn. B.7, we could also have chosen a minus sign for r_0. We won't delve into the details here, but suffice it to say that the description of half-integer spins requires the use of SU(2).

B.2.4 SU(3)

The group SU(3) is the exact equivalent of SU(2), but in three complex dimensions rather than two. The addition of this last dimension makes the discussion a bit more complicated: a general SU(n) matrix has $n^2 - 1$ free parameters (obtained as $2n^2$ for a general $n \times n$ complex matrix, minus n^2 from the unitarity constraint, and minus one from the requirement that the matrix's determinant be equal to unity). So, for $n = 3$ we find 8 independent degrees of freedom; so there are also 8 *generators* of the group, in analogy with the 3 Pauli matrices of SU(2). These are often

taken to be the Gell-Mann matrices:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \qquad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$
$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \qquad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

With this convention, the Lie algebra of SU(3) becomes

$$[\frac{\lambda_k}{2}, \frac{\lambda_l}{2}] = i \sum_m f_{klm} \frac{\lambda_m}{2},$$

with *structure constants* f_{klm}, which can be explicitly evaluated to be:

$$f_{123} = 1, \quad f_{458} = f_{678} = \sqrt{3}/2, \quad f_{147} = f_{165} = f_{246} = f_{257} = f_{345} = f_{376} = \frac{1}{2}$$

while the remaining constants can be obtained by using the fact that the f_{klm} must be antisymmetric under the exchange of any two indices. Note that the normalisation of the λ_i is chosen such that

$$\mathrm{Tr}(\lambda_i \lambda_j) = 2\delta_{ij}$$

(the same normalisation holds for the Pauli matrices of SU(2)).

Bibliography

[1] C. Itzykson and J.-B. Zuber, Quantum Field Theory, (McGraw Hill, 1985)

[2] M.E. Peskin and D.V. Schroeder, An Introduction to Quantum Field Theory, (Westview, 1995)

[3] C.D. Anderson, Phys. Rev. **43** (1933) 491

[4] J.D. Jackson, Classical Electrodynamics, (John Wiley and Sons, second edition, 1975)

[5] M. Thomson, Modern Particle Physics, (Cambridge University Press, 2013)

[6] R.W. McAllister and R. Hofstadter, Phys. Rev. **102** (1956) 851

[7] J.D. Bjorken, Phys. Rev. **179** (1969) 1547

[8] F. Halzen and A.D. Martin, Quarks & Leptons: An Introductory Course in Modern Particle Physics, (Wiley, 1984)

[9] Bahcall, John N. and Serenelli, Aldo M. and Basu, Sarbani, Astrophys.J. **621** (2005) L85–L88

[10] Ahmad, Q.R. and others, Phys.Rev.Lett. **89** (2002) 011301

[11] Fukuda, Y. and others, Phys.Rev.Lett. **81** (1998) 1562–1567