JACOB LINDER

Introduction to Particle Physics

JACOB LINDER

INTRODUCTION TO PARTICLE PHYSICS

CONTENTS

Preface

The aim of this book is to present fundamental concepts in particle physics. This includes topics such as the theories of quantum electrodynamics, quantum chromodynamics, weak interactions, Feynman diagrams and Feynman rules, important conservation laws and symmetries pertaining to particle dynamics, relativistic field theories, gauge theories, and more. In addition to explaining the underlying theories in a detailed manner, we shall also provide a number of examples that will illustrate the formalisms "in action".

This book is primarily based on my lecture notes from teaching this class to university students over several years, and the notes are in turn based on the excellent book "Introduction to Elementary Particles" by D. Griffiths, which they follow closely in structure and choice of examples. I have also included some additional topics and instructive examples which hopefully will allow the reader to obtain a more thorough physical understanding of the material. This book is suitable as material for a full-semester course in introductory particle physics.

It is my goal that students who study this book afterwards will find themselves well prepared to dig deeper into the remarkable world of theoretical physics at a more advanced level. I welcome feedback on the book (including any typos that you may find, although I have endeavored to eliminate as many of them as possible) and hope that you will have an exciting time reading it!

Jacob Linder (jacob.linder@ntnu.no)
Norwegian University of Science and Technology
Trondheim, Norway

About the author

J.L. holds a position as Professor of Physics at the Norwegian University of Science and Technology. His research is focused on theoretical quantum condensed matter physics and he has received several prizes for his Ph.D work on the interplay between superconductivity and magnetism. He has also received the American Physical Society "Outstanding Referee" award, selected among over 60.000 active referees. In teaching courses such as Quantum Mechanics, Classical Mechanics, and Particle Physics for both undergraduate and graduate students, he has invariably received high scores from the students for his pedagogical qualities and lectures. His webpage is found here. He has also written the books "Introduction to Lagrangian & Hamiltonian Mechanics" and "Intermediate Quantum Mechanics", which are both freely available to download on Bookboon.

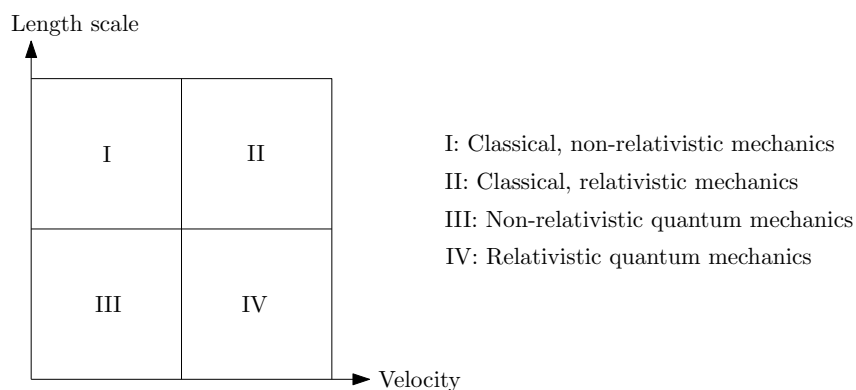## I. OVERVIEW OF ELEMENTARY PARTICLES AND THEIR INTERACTIONS

Learning goals. After reading this chapter, the student should:

- Be able to describe the four fundamental interactions, their mediating particles, and which particles that can participate for a given interaction.

- Understand how to read basic Feynman diagrams.

- Account qualitatively and quantitatively for neutrino oscillations.

In the so-called Standard Model of particle physics, matter is believed to consist of elementary particles. One often makes the distinction between matter and mediators (also known as force fields), so that although the following three groups are all elementary particles:

- Leptons

- Quarks

- Mediators

only the leptons and quarks (collectively known as *fermions*) comprise matter. In this book, you will get acquainted with all of these particles and how they interact with each other. We will begin by introducing the fundamental forces by which elementary particles interact and represent their interaction via *Feynman diagrams*. Our treatment will mostly be qualitative to begin with in order for the reader to get an overview of the situation, and then we will treat each scenario in detail mathematically. The appropriate physical theory to be used when considering physical phenomena in a system depends on which length scales and velocities we are considering. These quantities should be compared with the de Broglie wavelength λ of the particles and the velocity v of the particles, as illustrated in the figure below. In particular, λ roughly marks the transition for the length scales where a quantum mechanical description is required. Similarly, if a particle is moving with a velocity close to the speed of light c, a relativistic description is needed. In this way, region III describes length scales which are small or comparable to λ and velocities that are slow compared to c, whereas region IV describes length scales that are equally small, but where the velocity is comparable to c.

I: Classical, non-relativistic mechanics
II: Classical, relativistic mechanics
III: Non-relativistic quantum mechanics
IV: Relativistic quantum mechanics

A. The fundamental interactions

Each of the four fundamental interactions are believed to be mediated by the exchange of a particle: a so-called *mediator*. This is well-established except for the gravitational force, as there is no compelling evidence as of today for the existence of a graviton.

| Force	Theory	Mediator
Strong	Chromodynamics	Gluon
Electromagnetic	Electrodynamics	Photon
Weak	Flavordynamics	$W^{\pm}$ and Z^0
Gravitational	General theory of relativity	Graviton?

We will start by considering quantum electrodynamics (QED), which is the oldest, simplest, and most successful (in terms of comparison between theory and experiment) of all the above theories. The others are to a large extent in fact modelled on QED.

B. Quantum electrodynamics

All electromagnetic phenomena can ultimately be reduced to the following elementary process:
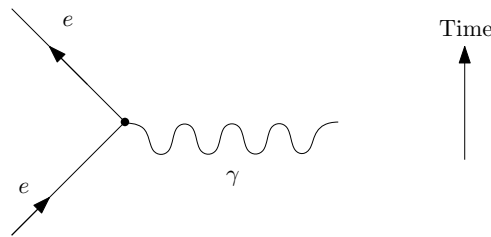
The diagram should be read as: a charge e enters, emits/absorbs a photon γ, and exits. The charged particle can be a lepton or a quark. For more complicated processes, one patches together several such so-called *primitive vertices*. For instance, the interaction between two electrons (Möller scattering) can be drawn as:
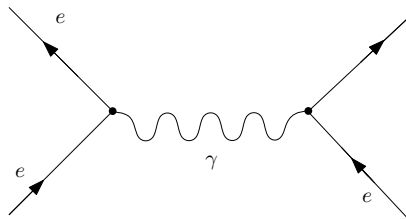
Diagrams of this type are known as Feynman diagrams. Note that a particle running backward in time (as indicated by the arrows) must be interpreted as the corresponding antiparticle moving forward in time. In fact, a universal feature of quantum field theory is:

For every kind of particle, there exists an antiparticle with equal mass and opposite charge.

There are six so-called *flavors* of leptons and these flavors form three *generations*: e^- and ν_e, μ^- and ν_μ, τ^- and ν_τ. Here, μ^- is a muon, τ^- is a tau particle, while ν_j are neutrinos of a particular type $j = \{e, \mu, \tau\}$. Notationwise, antiparticles are denoted with an overbar $.\bar{.}.$ or by specifying their charge explicitly: n for neutron, $\bar{n}$ for antineutron and e^- for electron, e^+ for positron. Note that since the neutron is not electrically charged, one could ask how its antiparticle is physically different. As we shall see, it turns out that particles such as neutrons carry other quantum numbers that also change sign for antiparticles.
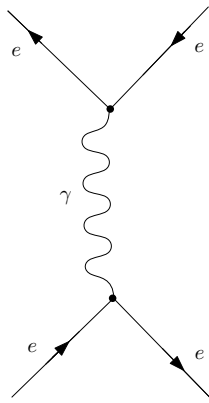
In the above figure, an electron and positron annihilate to form a photon and then produce and $e^- - e^+$ pair. The photon is its own antiparticle. The diagram then represents the interaction between two opposite charges, namely their Coulomb attraction. Another diagram also contributes to this process:
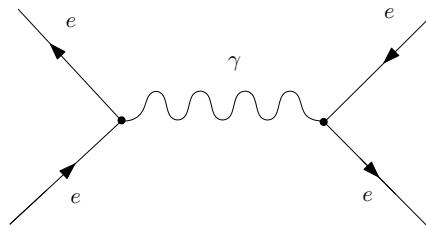
Diagrams with two vertices can also be used to represent several other processes:

Pair annihilation
$e^- + e^+ \rightarrow \gamma + \gamma$

Pair production
$\gamma + \gamma \rightarrow e^- + e^+$

Compton scattering
$e^- + \gamma \rightarrow e^- + \gamma$
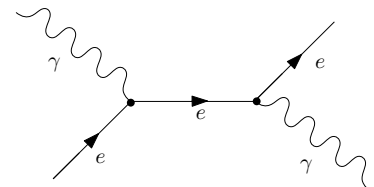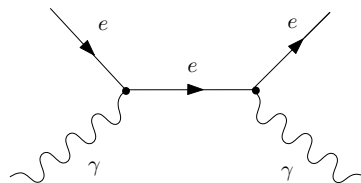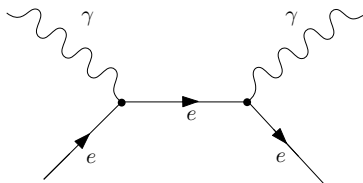
There is no need to explicitly specify the $+$ or $-$ charge on the external lines: the arrows suffice to indicate whether it is a particle or antiparticle. Note that all of the above processes are related via a so-called *crossing symmetry*. Suppose that $A + B \rightarrow C + D$ is allowed. Then:

$$A \rightarrow \bar{B} + C + D, \ A + \bar{C} \rightarrow \bar{B} + D, \ C + D \rightarrow A + B \tag{1.1}$$

are *in principle* also allowed. However, an important caveat is that conservation of energy must be fulfilled. For instance, if the masses satisfy $m_A < m_B + m_C + m_D$, then $A \rightarrow \bar{B} + C + D$ cannot occur. The reason is energy conservation. A massive particle has rest energy (we will have more to say about this in our treatment of the special theory of relativity later on) and thus m_A needs to be sufficiently large in order to enable the reaction. This is an example of a process which is not kinematically permissible, meaning that it does not satisfy conservation of for instance energy or momentum.

In terms of Feynman diagrams, processes which are related via crossing symmetry can be obtained by twisting or rotating the diagrams. We find additional possibilities if we allow for four vertices:
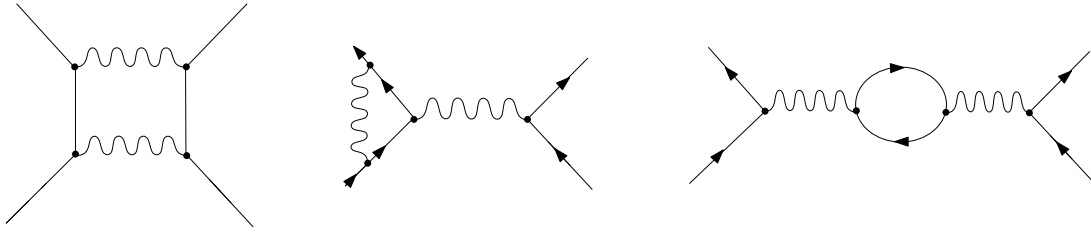
The common feature of all these diagrams is that two electrons enter and exit. The internal lines represent virtual particles that cannot be observed: they describe the precise mechanism of the interaction. It should be emphasized that diagrams are symbolic in the sense that they do not represent actual particle trajectories. Each Feynman diagram is actually associated with a particlar complex number $\mathcal{M}$ that may be computed by using *Feynman rules*. We shall say much more about this later when we develop our quantitative theory of particle interactions. Assume now that you want to analyze a particular physical process. The standard procedure is then:

- Draw all Feynman diagrams contributing to this process (with appropriate external lines) that have 2 vertices, then all with 4 vertices, and so on.

- Evaluate the contribution of each diagram via Feynman rules.

- Sum all these contributions: the result represents the actual physical process.

One aspect of the above recipe seems problematic: there are infinitely many Feynman diagrams, since we can construct higher and higher order contributions by introducing additionial vertices! We cannot consider all of these. The solution is that each vertex in a diagram introduces a factor $\alpha = e^2/\hbar c \simeq 1/137$: the fine structure constant. Therefore, higher order diagrams fortunately contribute less and less the more vertices they have. In practical calculations, it is often sufficient to include diagrams up to four vertices. Feynman rules include conservation of energy and momentum at each vertex (and thus automatically for the whole diagram).

The primitive QED vertex thus cannot represent a physical process: $e^- \to e^- + \gamma$ violates energy conservation since the electron does not have enough of the aforementioned *rest energy* to sustain the existence of both an electron and an additional photon γ. The same problem exists for $e^- + e^+ \to \gamma$ where momentum conservation cannot be fulfilled: in the CM system, e^- and e^+ enter with equal and oppsite velocities so that $\boldsymbol{p}_{\text{tot}} = 0$. However, the final momentum cannot be zero since the photon γ travels at the speed of light. It is important to note that virtual particles (internal lines in the diagrams) do not necessarily satisfy the same relation between energy, mass, and momentum as the corresponding physical particle. Instead, their mass is whatever the conservation laws at each vertex requires. In contrast, external lines represent real particles and must have the correct mass.

C. Quantum chromodynamics

In quantum chromodynamics (QCD), color plays the role of charge in QED. The fundamental process (primitive vertex) is quark $\rightarrow$ quark + gluon. There are six types of quarks, also referred to as *flavors*: up, down, strange, charm, top, and bottom (u, d, s, c, t, b).
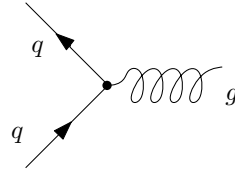
Leptons have no color and thus do not interact strongly. The force between quarks (responsible *e.g.* for binding quarks to make baryons) has the lowest-order diagram:
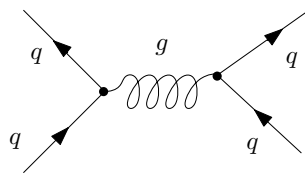
Note how the gluon is represented by a curly line whereas the photon is represented by a wavy line. A key difference from QED is that there are *three* kinds of color (instead of $\pm$ charge) and color may change in a process, although the total color must be conserved. In the example shown below, the gluon carries away one unit of blue and minus one unit of red:
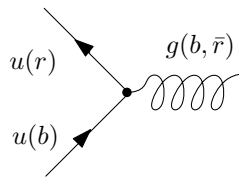
Note that the flavor of the quark may not change. A negative unit of color is denoted by an overbar, so that a negative unit of red color is $\bar{r}$. Gluons are therefore "bicolored": they carry one positive and one negative unit of color. This suggests that $3 \times 3 = 9$ possible gluons should exist. For reasons that we will come back to, there are nevertheless only 8. Moreover, individual quarks do not appear freely in nature, but quark-antiquark pairs called *mesons* exist. For now (we will later bring out a nuance in this statement), we may write that *all naturally occuring particles are colorless.* We define colorless as either the total amount of color being zero or that the three colors are present in equal amounts.

Since gluons carry color (unlike photons that do not carry charge), they may couple to other gluons:
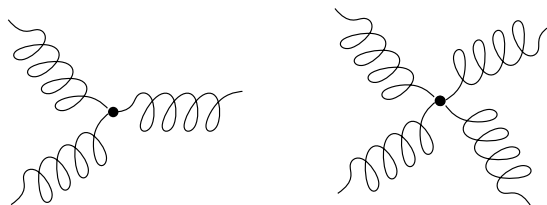
This $g - g$ coupling makes QCD richer and a lot more complicated than QED. Another essential difference is related to the size of the coupling constant. In QED, we stated that each vertex introduced a factor $\alpha \simeq 1/137$. In contrast, the coupling constant for strong forces α_S is (under some circumstances) greater than 1. It may be determined from *e.g.* the force between two protons. This leads to a problem: higher-order vertices and diagrams contribute more and more! The resolution to this, as we shall discuss in more detail later, is that α_S is in fact not

a true constant. It rather depends on the separation distance between the interacting particles, giving rise to the phenomenon of *asymptotic freedom*. For large distances, α_S becomes large. For small distances (less than the size of a proton), α_S becomes small. Therefore, within a proton or pion, quarks move around without interacting much.

The proton and neutron are comprised of three quarks, and such a composite particle is known as a *baryon*. A composite particle consisting of one quark and one antiquark (a $q\bar{q}$ pair) is known as a *meson*. Baryons and mesons are thus both quark-based composite particles which collectively are known as *hadrons*.

This distance-dependent effective coupling actually appears in electrodynamics as well in the form of charge screening.
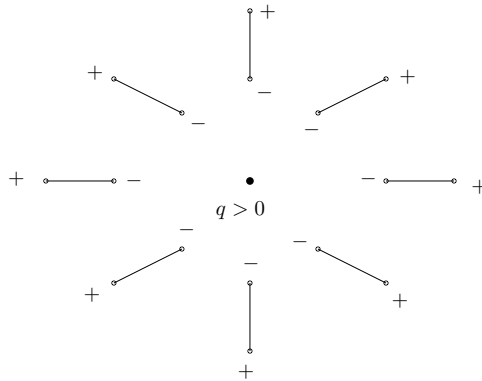
The "halo" of negative charge from molecular dipoles partially cancel the field from $q > 0$. Thus, the effective charge is reduced to $q_{\text{eff}} = q/\epsilon$ where ϵ is the dielectric constant of the medium. It measures how easily a substance becomes polarized by an electric field $\boldsymbol{E}$.
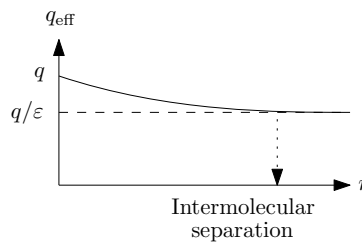
Now, in QED the vacuum actually acts as a dielectric due to the production of $e^- - e^+$ pairs that interacting in the following ways:
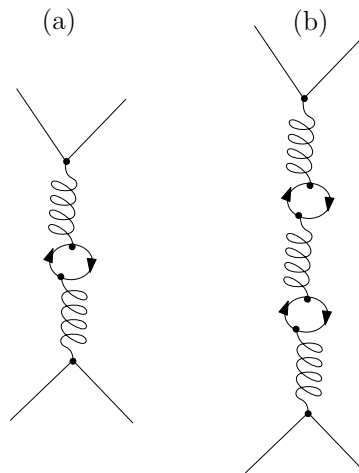
The key aspect to note in these diagrams is that the virtual electron in each bubble will be attracted to a positive charge q, whereas the positron is repelled. As a consequence of the screening occurring even in vacuum, what is measured in experiments is the screened effective charge. The same thing happens in QCD, but with an extra ingredient: gluon-gluon vertices. Therefore, in addition to diagrams of the type (a) and (b) above (where the photons should now be exchanged with gluons), we also have to include
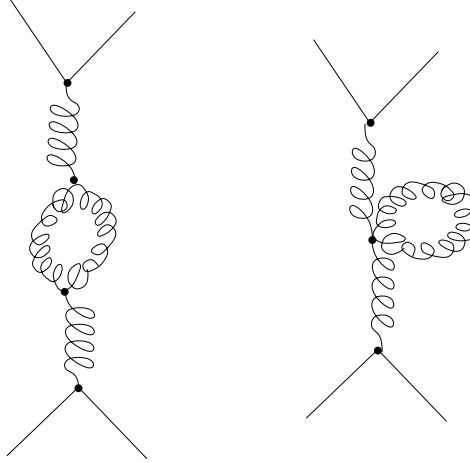
It is not obvious what the influence of these diagram will be, but it turns out that their effect is opposite: they drive down α_S at short distances, in contrast to quark polarization diagrams which enhance α_S at short distances. These effects then compete. The winner of the competition is determined by comparing the relative number of flavors (quarks) and colors (gluons). Detailed calculations beyond the scope of this textbook reveal that the critical parameter is $a = 2f - 11n$ where f is the number of flavors and n is the number of colors. In the Standard Model, $f = 6$ and $n = 3$, leading to $a = -21$. Therefore, the coupling decreases at short distances as stated previously.

Keep in mind that in contrast to charge, no naturally occurring particles carry color: quarks must be confined to colorless packages of two (mesons) and three (baryons). As a result, strong interactions between naturally occuring particles may be quite complicated. Consider for instance the strong force between two protons. One of the contributing diagrams is:

Effective π^0 exchange
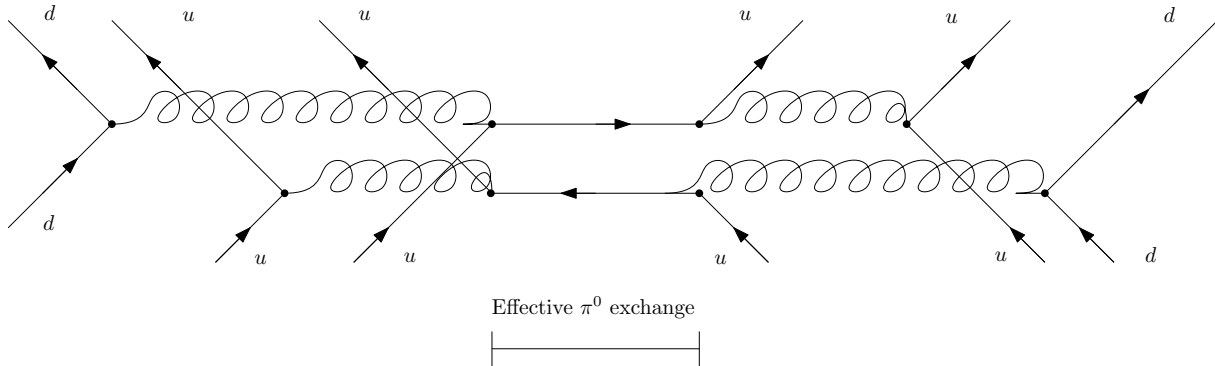
It is worth mentioning that we here see remnants of an early model for strong interactions introduced by the physicist Yukawa, who conjectured that the pion (and not the gluon) was the mediator of the strong force. In reality, the diagram shows that the interaction is much more complicated.

D. Weak interactions

Unfortunately, there is no particular name for the property that produces weak forces, but whatever you call it: quarks and leptons have it. Two types of weak interactions exist. Charged weak interactions are mediated by $W^{\pm}$ bosons while neutral weak interactions are mediated by the Z^0. For leptons, the fundamental charged vertex is:
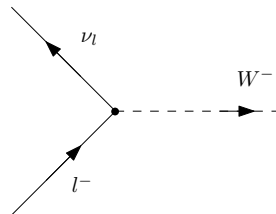
A negative lepton l^- converts into the corresponding neutrino and emits W^- (or absorbs W^+: the diagram can mean both things). As before, we combine primitive vertices for more complicated reactions:
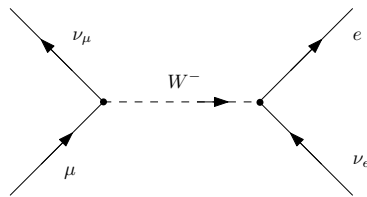
This type of neutrino-muon scattering event is hard to set up in the laboratory, but a slight twist gives us a muon-decay diagram that occurs all the time:
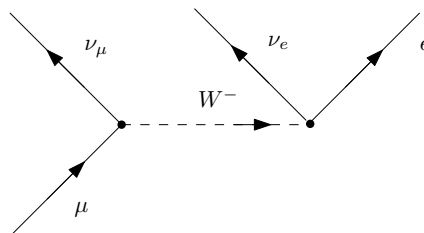
The fundamental neutral vertex for leptons is:
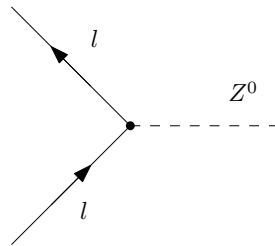
Here, l is any lepton including neutrinos. Neutral interactions were difficult to discover experimentally, being masked by much stronger electromagnetic interactions (we will see later why the electromagnetic ones usually dominate):
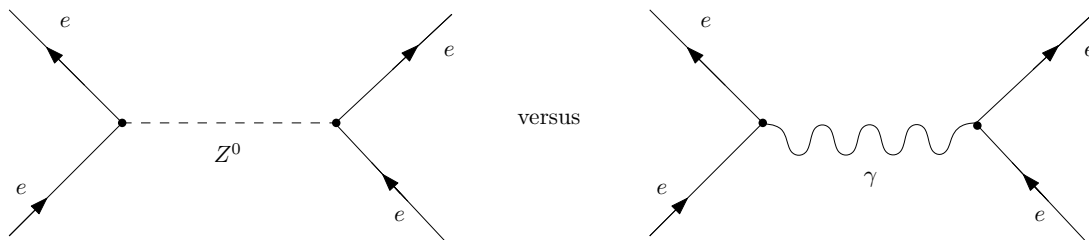
Therefore, to observe a *purely* weak interaction one has to consider neutrino scattering since there is no electromagnetic contribution to such a process. In terms of notation for the $W^{\pm}$ and Z^0 bosons in Feynman diagrams, it is worth remarking that another frequently encounted convention exists where these particles are drawn as wavy lines rather than dotted lines, similarly to the photon. Here, we stick with the dotted line convention to distinguish them more clearly from photon-mediated diagrams. Turning to how quarks interact weakly, recall first that leptonic weak vertices connect members of the same generation: e^- connects to ν_e, but never ν_μ, for instance. Thus, we have conservation of electron, muon, and tau numbers $\{L_e, L_\mu, L_\tau\}$ (see table below). Antileptons have all signs reversed.

Lepton l	Q	L_e	L_μ	L_τ
e	-1	1	0	0
ν_e	0	1	0	0
μ	-1	0	1	0
ν_μ	0	0	1	0
τ	-1	0	0	1
ν_τ	0	0	0	1

For quarks, the fundamental charged vertex would be natural to assume looks as follows:
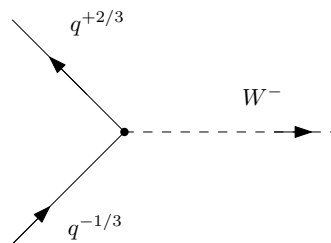
A quark with charge $-\frac{1}{3}$ (the d, s, b quarks) converts into the corresponding quark with charge $\frac{2}{3}$ (the u, c, t quarks) and emits a W^-. The outgoing quark carries the same color as the incoming one, but has a different flavor (type of quark). Importantly, the W^- does not carry the missing flavor. In fact, $W^{\pm}$ has no flavor since it must be able to couple to leptons. Therefore,

Flavor is not conserved in weak interactions.

The W^- in the above diagram can couple to leptons (a semileptonic process) or other quarks (a purely hadronic process). An example of a semileptonic process is the following:
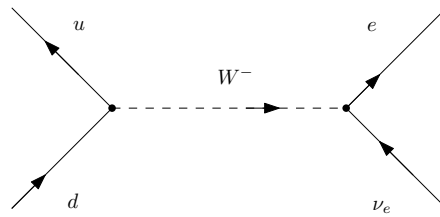
This process would never appear isolated in nature due to quark confinement, which we have discussed. However, if we turn it around to represent pion-decay, we obtain a feasible process as shown in (a).
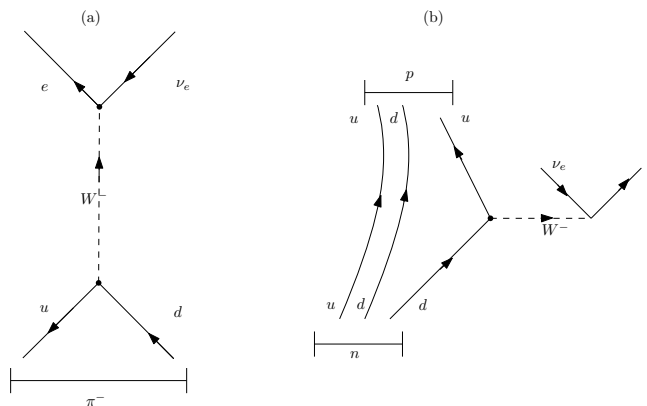
The same type of diagram holds for the beta decay of the neutron, as shown in (b). Since quarks interact both weakly and strongly, hadronic (quark-based) interactions can have both a weak and strong contribution. Consider for instance $\Delta^0 \to p^+ + \pi^-$:
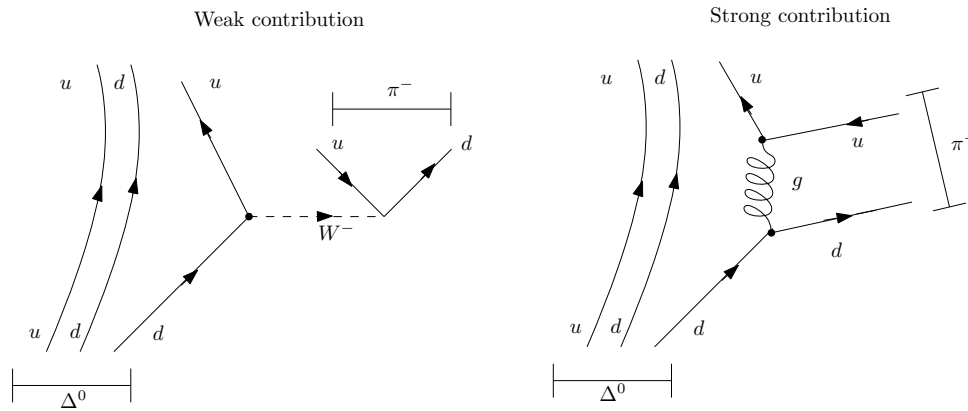
The strong contribution dominates completely in magnitude. As for the fundamental neutral vertex for quarks, it would be natural to suppose that it looks like:
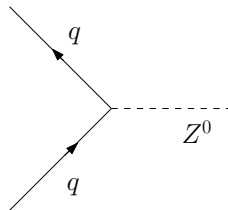
We may then use this to construct neutrino-scattering processes such as $\nu_\mu + p \to \nu_\mu + p$:
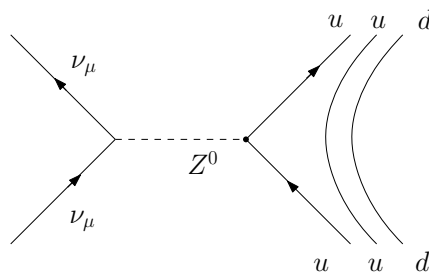
In essence, the above suggests that quarks mimic the leptons as far as weak interactions are concerned. However, we must make an *important modification to what we have stated*. To see why, consider the fact that if the fundamental quark vertex may only couple quarks within the same generation [similarly to leptons, the quarks also form three generations (u, d), (c, s), (t, b)], we cannot describe any strangeness-changing weak interactions such as changing a strange quark into an up quark. Yet such processes do exist in nature, with two examples shown below:

$$\Lambda \to p^+ + \pi^- \qquad\qquad \Omega^- \to \Lambda + K^-$$
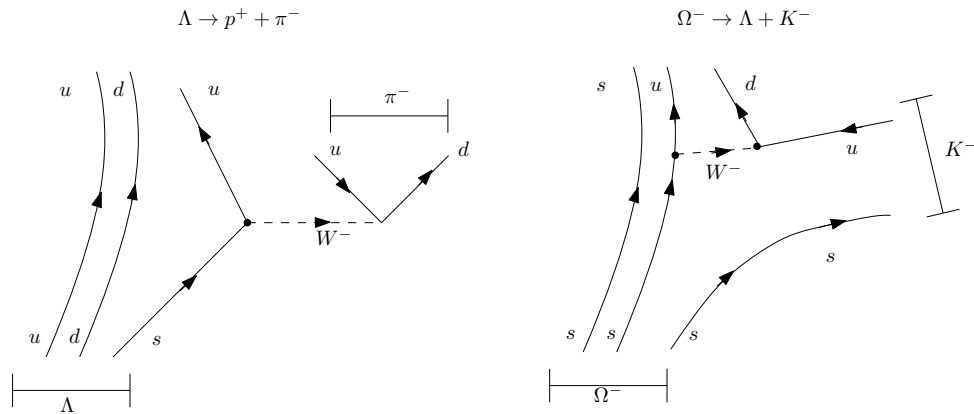
The solution to this dilemma was suggested by Cabibbo in 1963, and we shall treat this issue quantitatively in the chapter on QCD later on in this textbook. For now, we note that strangeness is a property (like charge or color) that is assigned to each particle. Strangeness is conserved in strong interactions, but not necessarily conserved in weak interactions.

Summarizing, the primitive vertices giving rise to physical processes via weak, strong, and electromagnetic interactions are the following:
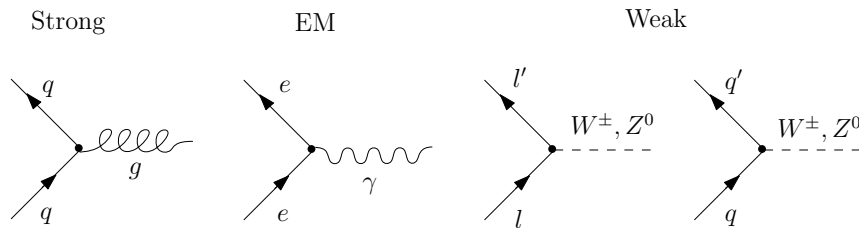
We also give a summary of how to characterize particles in terms of their spins in the table below

	Bosons (integer spin s)		Fermions (half-integer spin s)	
	Spin 0	*Spin 1*	*Spin $\frac{1}{2}$*	*Spin $\frac{3}{2}$*
Elementary particles $\longrightarrow$	Higgs particle	Mediators	Quarks, leptons	n/a
Composite particles $\longrightarrow$	Pseudoscalar mesons	Vector mesons	Baryon octet	Baryon decouplet

Weak and electromagnetic couplings to $W^\pm$ and Z^0

As a final note, just as there exist gluon-gluon couplings in QCD, the $W^\pm$ and Z^0 may couple to each other and the photon:
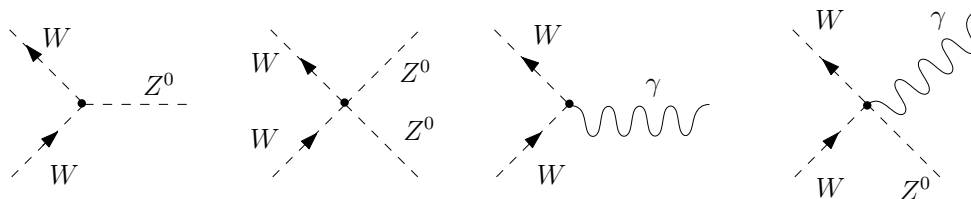
E. Briefly on neutrino oscillations

Neutrinos are notoriously difficult to observe, unfortunately, in particular since they are so light and electrically neutral. In the mid-1950s, Cowan and Reines performed an ingenious experiment where the idea was to look for a reaction triggered by neutrinos rather than trying to observe neutrinos as a byproduct of a reaction. In particular, they looked for inverse β-decay:

$$\bar{\nu} + p^+ \rightarrow n + e^+ \tag{1.2}$$

and successfully confirmed the existence of neutrinos. However, there turned out to be a problem associated with comparison between theory and several experiments. The experiments only saw about 1/3 of the theoretically predicted amount of neutrinos produced by the Sun, a mystery which was dubbed the solar neutrino problem. Pontecorvo suggested in 1968 a bold and elegant solution to the problem: what if the ν_e produced by the Sun were transformed into something else on their way to Earth which could not be detected by the experiments at hand, such as ν_μ (the experiments were specifically set up to measure ν_e)? This could be achieved via so-called *neutrino oscillations*.

To illustrate this idea, consider a scenario where we have two neutrino flavors ν_e and ν_μ. If indeed one may spontaneously convert to another, neither can be an eigenfunction of the Hamiltonian. Instead, the stationary states of the Hamiltonian describing the neutrinos should be a linear combination of them:

$$\nu_1(t) = \cos\theta \nu_\mu(t) - \sin\theta \nu_e(t), \ \nu_2(t) = \sin\theta \nu_\mu(t) + \cos\theta \nu_e(t). \tag{1.3}$$

This form ensures orthonormality of the states ν_1 and ν_2, which by virtue of being stationary have a simple time dependence:

$$\nu_i(t) = \nu_i(0)\mathrm{e}^{-\mathrm{i}E_i t/\hbar}, \ E_i = \sqrt{p_i^2 c^2 + m_i^2 c^4}. \tag{1.4}$$

Here, E_i and m_i is the energy and mass of neutrino type i. We note immediately that it follows that ν_e and ν_μ do not have well-defined masses since they are linear combinations of ν_1 and ν_2.

To see how a flavor state evolves with time (*e.g.* from production in the Sun and propagating toward Earth), assume that we start with ν_e neutrinos:

$$\nu_e(t=0) = 1, \ \nu_\mu(t=0) = 0. \tag{1.5}$$

This gives

$$\nu_1(0) = -\sin\theta, \ \nu_2(0) = \cos\theta. \tag{1.6}$$

Now, rewrite Eq. (1.3) to express $\{\nu_e, \nu_\mu\}$ as a function of $\{\nu_1, \nu_2\}$ to obtain

$$\nu_\mu(t) = \cos\theta\nu_1(t) + \sin\theta\nu_2(t) = \sin\theta\cos\theta(\mathrm{e}^{-\mathrm{i}E_1t/\hbar} + \mathrm{e}^{-\mathrm{i}E_2t/\hbar}). \tag{1.7}$$

We then find the probability that ν_e has converted into ν_μ after a time t:

$$\begin{aligned}
|\nu_\mu(t)|^2 &= \frac{\sin^2 2\theta}{4}(1 - \mathrm{e}^{\mathrm{i}(E_2-E_1)t/\hbar} - \mathrm{e}^{-\mathrm{i}(E_2-E_1)t/\hbar} + 1) \\
&= \frac{\sin^2 2\theta}{4} 4\sin^2\left(\frac{E_2-E_1}{2\hbar}t\right) \\
&\to P_{\nu_e \to \nu_\mu} = \left[\sin 2\theta \sin\left(\frac{E_2-E_1}{2\hbar}t\right)\right]^2.
\end{aligned} \tag{1.8}$$

In this model, we have thus achieved precisely a conversion between ν_e and ν_μ which depends in an oscillatory manner on t (or equivalently the distance travelled). There are two necessary ingredients for neutrino oscillations to occur:

1. A finite mixing angle $\theta \neq 0$.

2. Different masses for the eigenstates ν_i (both cannot be zero).

Today, there is compelling experimental evidence (which was also awarded the Nobel Prize in physics in 2015) that neutrinos are not massless - in hindsight, one could even state that there actually is no fundamental reason for why they should be (unlike the photon, for which there exists a very good reason that its mass must be zero). In practical calculations where ν-oscillations are not of relevance, however, we will often approximate their mass to zero which still yields excellent results.

II. CONSERVATION LAWS AND SYMMETRIES

Learning goals. After reading this chapter, the student should:

- Understand the relation between continuous symmetries and conserved quantities.

- Be able to explain what a symmetry and a symmetry group is, and account for how isospin, parity, and charge conjugation symmetries work.

- Explain the difference between SU(2) and SO(3) and in which physical settings these groups are used.

A. Decays and conserved quantities

A general property of elementary particles is their tendency to decay:

> Every particle decays into lighter particles, unless prevented by some conservation law.

Neutrinos and photons are stable due to their low and vanishing mass, respectively: there is simply nothing lighter to decay into. Strictly speaking, however, there is a finite probability for a process of the type $\nu_i \to \nu_j + \gamma$ if the mass eigenstates for the neutrinos have masses that satisfy $m_{\nu_i} > m_{\nu_j}$. The electron is the lighest charged particle, so that conservation of charge makes it stable. The proton is the lightest baryon, so that conservation of baryon number saves it. Baryon number A is assigned so that baryons have $A = +1$ (neutrons, protons, et.c.) wheres antibaryons have $A = -1$ (non-baryons have $A = 0$). The positron and antiproton are also stable for the same reasons, but all other particles may spontaneously disintegrate. Even quarks decay: this is *e.g.* what happens for the weak decay $n \to p + e^- + \bar{\nu}_e$. A given decay is governed by one of the three fundamental forces we have discussed:

- $\Delta^{++} \to p^+ + \pi^+$ is a strong decay.

- $\pi^0 \to \gamma + \gamma$ is an electromagnetic decay.

- $\Sigma^- \to n + e^- + \bar{\nu}_e$ is a weak decay.

How can we know this? If a photon comes out, the process is certainly electromagnetic. If a neutrino emerges, the process is certainly weak. If neither γ or ν emerge, it is more tricky to determine the origin. For instance, $\Sigma^- \to n + \pi^-$ is weak, but $\Delta^- \to n + \pi^-$ is strong. The reason for this is that in the Σ^- decay strangeness is changed, which does not happen in strong interactions. Weak interactions *can* conserve strangeness, however, and in that case one can look at the decay time in order to distinguish the weak and strong interactions. This is in fact the experimentally most dramatic difference between the various types of decays:

Type of decay	Typical lifetime
Strong	$\sim 10^{-23}$ s
EM	$\sim 10^{-16}$ s
Weak	Ranges from 10^{-13} s (for τ) to several minutes (for n)

A decay of a given type in general proceeds more rapidly the larger the mass difference is between the original particle and the decay products (although exceptions exist). It is this kinematic effect that causes the great spread in the weak interaction lifetimes. The definition of the lifetime τ is related to the half-life $t_{1/2}$ via

$$t_{1/2} = (\ln 2)\tau \simeq 0.693\tau. \tag{2.1}$$

In turn, the half-life is the time required for half the particles in a large sample to decay.

Different types of conservation laws exist which dictate whether or not a reaction is possible. *The kinematic conservation laws* are concerned with energy, momentum and angular momentum. These apply to *all interactions*: strong, weak, and electromagnetic. In addition, there are additional conservation laws that apply to each vertex in a Feynman diagram. A quantity conserved at each vertex is automatically conserved for the total reaction.

1. **Conservation of charge.** Note that the $W^\pm$ bosons can carry away charge.

2. **Conservation of color.** Electromagnetic and weak interactions do not affect color, whereas gluons can carry away color.

3. **Conservation of baryon number.** As mentioned before, $A = +1$ for baryons, $A = -1$ for antibaryons, and $A = 0$ for everything else. The origin of this rule can be traced back to conservation of quark number.

4. **Conservation of lepton generation number.** Strong interactions do not influence leptons, whereas electromagnetic interactions only makes the particle emit a photon. Weak interactions mix leptons within the same generation (e, μ, τ).

5. **Conservation of flavor in strong and electromagnetic interactions.** Flavor is *not conserved* in weak interactions.

It is also worth mentioning that unlike leptons and baryons, *there is no conservation of mesons* (a quark-antiquark pair $q\bar{q}$).

B. The concept of symmetry and symmetry groups

Symmetry is an immensely powerful concept in physics and mathematics. Consider for instance the following graph.
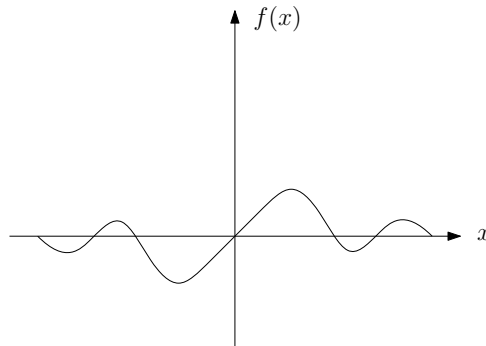
Even if we do not know the functional form of $f(x)$, we know a symmetry property it has, namely $f(x) = -f(-x)$. It then follows immediately that, for instance, $[f(-x)]^4 = [f(x)]^4$ and that $\frac{df}{dx}|_{x=2} = \frac{df}{dx}|_{x=-2}$. We actually know a lot about the mathematical properties of $f(x)$ simply by observing that it is antisymmetric. The concept of symmetry has profound implications in particle dynamics. A milestone was set in 1917 when Emmy Noether published her famous theorem:

> **Noether's theorem:** Continuous symmetry $\leftrightarrow$ Conservation law.

We will assume that the reader has been introduced to this theorem previously and not dwell further upon it, despite the fact that it is very interesting. If more details regarding its formal application to important scenarios such as translational or rotational symmetry is desirable, the reader can have a look here. Some examples are shown below.

Symmetry	Conservation law
Translation in time	Energy
Translation in space	Momentum
Rotation	Angular momentum
Gauge transformation	Charge (to be explained)

Now, what precisely is a symmetry? It is an operation you can perform on a system that leaves it invariant. Consider the following example.
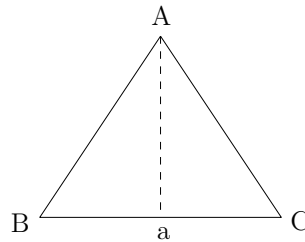
An equilateral triangle is carried into itself by clockwise rotation with $2\pi/3$ (an operation we denote R_+) or counter-clockwise rotation (R_-) with the same angle. Moreover, if we flip it about the axis Aa (R_a) or corresponding axis through B and C (R_b and R_c), it is also left invariant. A combination of these operations would also be a symmetry operation. Mathematically, a set of symmetry operations must have the following properties:

1. **Closure**. If R_i and R_j are in the set, then the product $R_i R_j$ is also in that set.

2. **Identity**. There is an element I such that $I R_i = R_i I = R_i$ for all elements R_i.

3. **Inverse**. For every element R_i, there is an inverse R_i^{-1} such that $R_i R_i^{-1} = R_i^{-1} R_i = I$.

4. **Associativity.** $R_i(R_j R_k) = (R_i R_j) R_k$.

These are the defining properties of a *group*. Note that commutativity of the elements is not required, so that $R_i R_j \neq R_j R_i$ is allowed, in which case we have a non-Abelian group.

In physics, most of the groups of interest are groups of matrices. Of particular importance is the U(n) group, meaning all unitary $n \times n$ matrices U (which thus satisfy $U^{-1} = U^\dagger$). Orthogonal matrices are a special case, namely unitary matrices with real entries.

- SU(n): unitary matrices with determinant +1.

- O(n): real unitary matrices (O stands for orthogonal).

- SO(n): real orthogonal matrices with determinant +1.

The SO(3) group, which we will treat in more detail in the next section, describes rotations in three dimensions and is mathematically quite similar to the SU(2) group. An important aspect of groups is that:

Every group can be represented by a group of matrices.

This means that for every group element a, there is a corresponding matrix M_a. This correspondence respects group multiplication: if $ab = c$, then $M_a M_b = M_c$. There may be several group elements represented by the same matrix. One then says that the group of matrices is homomorphic, but not necessarily isomorphic to G. If G itself is a group of matrices, such as SU(4), then it is a representation of itself (referred to as the *fundamental representation*). There can, however, be many other representations by matrices in various dimensions. For instance, a trivial example is that every element can be represented by the 1×1 unit matrix (although it is not a very interesting representation). The group SU(2) has representations of dimension 1 (the trivial one just mentioned), dimension 2 (the fundamental representation), dimensions 3, 4, 5, and so forth. One can always construct a new representation by combining two old in block-diagonal form:

$$M_a = \begin{bmatrix} M_a^{(1)} & 0 \\ 0 & M_a^{(2)} \end{bmatrix}. \qquad (2.2)$$

However, such representations are not counted separately: instead, one usually only lists *irreducible representations* which cannot be decomposed in this manner.

C. Introduction to group theory: SU(2) vs. SO(3)

These two groups are of particular importance since they are related to both rotations of spins (half-integer as well as integer) and to internal symmetries between particles and in Lagrangians describing relevant quantum fields. Let us start by introducing some terminology.

Isomorphism. A one-to-one correspondence between elements G of a group $\mathcal{G}$ and G' of a group $\mathcal{G}'$ such that if $G_i G_j = G_k$, then $G'_i G'_j = G'_k$.

Homomorphism. Let f be a mapping that maps the element G of $\mathcal{G}$ onto G' of $\mathcal{G}'$: $G' = f(G)$. If $f(G_i G_j) = f(G_i) f(G_j)$ holds for any two elements, f is a homomorphic mapping.

Note that homomorphism signifies a n-to-one correspondence between the elements of the group, in general. A homomorphism with $n = 1$ is thus an isomorphism.

Example 1. Equilateral triangle. The allowed symmetry operations constitute a group C_{3v}:

$$C_{3v} = \{E_1, C_3, C_3^{-1}, \sigma_1, \sigma_2, \sigma_3\}. \tag{2.3}$$

Here, C_3 means a rotation counterclockwise with an angle $2\pi/3$ around the center while σ_j means mirror reflection around symmetry axis. The multiplication table of a group provides information about how the product of two group elements is related to an element in the group. For this particular group, the multiplication table is shown in the figure.

(b)

(a)

Multiplication table for C_{3v} (entries $G_j \circ G_i$)

G_j \ G_i	E	C_3	C_3^{-1}	σ_1	σ_2	σ_3
E	E	C_3	C_3^{-1}	σ_1	σ_2	σ_3
C_3	C_3	C_3^{-1}	E	σ_3	σ_1	σ_2
C_3^{-1}	C_3^{-1}	E	C_3	σ_2	σ_3	σ_1
σ_1	σ_1	σ_2	σ_3	E	C_3	C_3^{-1}
σ_2	σ_2	σ_3	σ_1	C_3^{-1}	E	C_3
σ_3	σ_3	σ_1	σ_2	C_3	C_3^{-1}	E

(c)

Multiplication table for C_2

G_j \ G_i	E	C
E	E	C
C	C	E

Compare this group with $C_2 = \{E, C\}$ where C is a rotation with π so that $C^2 = E$. Let the following define the mapping f

$$\{E, C_3, C_3^{-1}\} \to E, \ \{\sigma_1, \sigma_2, \sigma_3\} \to C. \tag{2.4}$$

This fulfills the homomorphism criterion, as seen from the multiplication table in part (c) of the figure. As seen, there are three elements in C_{3v} corresponding to one element in C_2: this is a homomorphism, not an isomorphism.

The above considerations are related to *representations* of a group. Let $\mathcal{G}$ be a group with elements G_i and associate a square matrix with each element $\hat{D}(G_i)$. If the matrices satisfy $\hat{D}(G_j)\hat{D}(G_i) = \hat{D}(G_k)$ for the corresponding relation $G_j G_i = G_k$, then the set of matrices $\hat{D}(G_i)$ is a representation of $\mathcal{G}$. The mapping

$$\hat{D}: \ G_i \to \hat{D}(G_i) \tag{2.5}$$

is then homomorphic as the same matrix may be associated with several elements. As mentioned before, a group can have representations in several dimensions.

Example 2. 1D representation of C_{3v}**.** The following one-dimensional matrices (scalars) constitute a representation of C_{3v} and thus produce the same multiplication table as the above figure.

$$\hat{D}(E) = 1, \ \hat{D}(C_3) = 1, \ \hat{D}(C_3^{-1}) = 1,$$
$$\hat{D}(\sigma_1) = -1, \ \hat{D}(\sigma_2) = -1, \ \hat{D}(\sigma_3) = -1. \tag{2.6}$$

Consider now the rotation group in 3D, described by a matrix depending on three angles $\{\alpha, \beta, \gamma\}$:

$$\hat{R} = \begin{bmatrix} c_\alpha & -s_\alpha & 0 \\ s_\alpha & c_\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_\beta & 0 & s_\beta \\ 0 & 1 & 0 \\ -s_\beta & 0 & c_\beta \end{bmatrix} \begin{bmatrix} c_\gamma & -s_\gamma & 0 \\ s_\gamma & c_\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.7}$$

We introduced the short-hand notation $s_\alpha = \sin\alpha$, $c_\alpha = \cos\alpha$, and so forth. Since only proper rotations are considered (and thus not inversions), $\hat{R}$ is a 3×3 orthogonal matrix with $\det(\hat{R}) = +1$. Thus, the set of rotation matrices forms the group SO(3). The range of the parameters is important when discussing the relation of this group to SU(2), as we will now see. A possible representation of SO(3) in two dimensions is given by

$$\hat{D}^{(1/2)}[R(\alpha, \beta, \gamma)] = \begin{bmatrix} e^{-i(\alpha+\gamma)/2} \cos\frac{\beta}{2} & -e^{-i(\alpha-\gamma)/2} \sin\frac{\beta}{2} \\ e^{i(\alpha-\gamma)/2} \sin\frac{\beta}{2} & e^{i(\alpha+\gamma)/2} \cos\frac{\beta}{2} \end{bmatrix}. \tag{2.8}$$

We see that $\hat{D}^{(1/2)}$ is a special unitary 2×2 matrix. The word "special" means that the determinant is +1 rather than -1, and the superscript $^{(1/2)}$ indicates that the group SU(2) is of importance for rotation of spin 1/2 particles. Hence, the matrices given in the equation above form the group SU(2) *if* the parameter range of (α, β, γ) is such that all possible SU(2) matrices can be produced. To proceed, notice that $\hat{R}(0,0,0)$ and $\hat{R}(0,0,2\pi)$ are both corresponding to the same identity operation. However, $\hat{D}^{(1/2)}(0,0,0) = -\hat{D}^{(1/2)}(0,0,2\pi)$. In effect, both the matrix $\hat{D}^{(1/2)}(R)$ and $-\hat{D}^{(1/2)}(R)$ correspond to $\hat{R}$. This seems to indicate that SU(2) would be a double-valued representation of SO(3), meaning that there are two elements in SU(2) that can represent one element in SO(3). This is not allowed for a true representation as it provides an ambiguity. We can nevertheless resolve this issue by considering precisely the parameter range for the angles in $\hat{D}^{(1/2)}$. In order for these matrices to comprise SU(2), we need for instance $0 \leq \gamma \leq 4\pi$. However, by restricting the range to $0 \leq \gamma \leq 2\pi$, $\hat{D}^{(1/2)}$ can represent SO(3). Note that in this way, it is no longer the full group SU(2), but we have removed the double-valuedness by restricting the parameter range. We can also turn the argument around. Since there are two elements in SU(2) corresponding to SO(3), the groups are homomorphic:

$$\boxed{\text{SO(3) is a representation of SU(2).}}$$

This is not an isomorphism (one-to-one correspondence) and SU(2) is *not* a representation of SO(3).

We here summarize some general properties related to the groups SU(2) and SO(3).

- The representations of SU(2) may be classified by a number $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$ and the dimensionality of the representations is $2j + 1$.

- The spin-j representation of SU(2) is $\{\hat{D}^{(j)}\}$ where $\hat{D}^{(j)}$ are rotation matrices for angular momentum $= 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$

$j =$**half-integer:** the representation is faithful to SU(2), for instance like $j = \frac{1}{2}$ in our above treatment.

$j =$**integer:** the representation is also a representation for SO(3), and faithful as long as $j \neq 0$.

D. Isospin

The close similarity between the proton and neutron in terms of their mass ($m_p \simeq 938$ MeV/c^2 and $m_N \simeq 939$ MeV/c^2) caused Heisenberg to suggest that they could be regarded as two states of a single particle: the nucleon N. To explore such an idea, let us write

$$N = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \text{ where } p = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; n = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{2.9}$$

so that the proton and neutron can be characterized by their *isospin I*. This is in analogy to the treatment of spinors which have the same structure, only that their components are a part of spin-space. The physics behind this idea is that the strong interactions should be invariant under rotations in isospin space, just like electrical forces are invariant under rotations in ordinary space. Put differently, Heisenberg proposed that if the charge of the proton was somehow "turned off", it should be identical to the neutron. Noether's theorem dictates that isospin should then be conserved in all strong interactions, just like angular momentum is conserved in processes with rotational invariance.

Formulated in terms of our recently discussed group theory, strong interactions should then be invariant under an internal symmetry group SU(2) and nucleons belong to the two-dimensional representation with isospin $\frac{1}{2}$. Note that since m_p is not exactly equal to m_n, SU(2) isospin symmetry should not be expected to be an exact symmetry, but it should be very good. Isospin may also be used to classify other particles besides nucleons, such as the structure of the hadrons. The motivation for this is the same as for n and p: similar masses, but different charges. For the pions one assigns $I = 1$, so that the different pion particles are represented as different states in isospin space:

$$\pi^+ = |1, 1\rangle, \; \pi^0 = |1, 0\rangle, \; \pi^- = |1, -1\rangle. \tag{2.10}$$

The Λ has $I = 0$, so that $\Lambda = |0, 0\rangle$, whereas the Δ particles have $I = \frac{3}{2}$:

$$\Delta^{++} = |\frac{3}{2}, \frac{3}{2}\rangle, \; \Delta^+ = |\frac{3}{2}, \frac{1}{2}\rangle, \; \dots \Delta^- = |\frac{3}{2}, -\frac{3}{2}\rangle. \tag{2.11}$$

The number of particles r in a multiplet of hadrons is related to the isospin by $r = 2I + 1$. The third component I_3 is determined by the charge, so that $I_3 = I$ for the highest charge.

Isospin also has dynamical implications rather than simply being used a tool to classify particles. Consider the (isospin) angular momentum of two nucleons. Their total value can be $I = 0$ or $I = 1$, according to standard rules for how to add two angular momenta. In particular, we have three possible triplet states:

$$|1, 1\rangle = pp, \; |1, 0\rangle = \frac{1}{\sqrt{2}}(pn + np), \; |1, -1\rangle = nn \tag{2.12}$$

and one singlet state

$$|0, 0\rangle = \frac{1}{\sqrt{2}}(pn - np). \tag{2.13}$$

Now, the n and p form a bound state known as a deuteron (d). We can then immediately conclude that it has to be a singlet state in isospin space. Why? If it were a triplet state, then pp and nn should also occur in nature simply by rotating the isospin. However, such long-lived states are not known to exist (to be precise, nn was actually observed as an intermediate state in 2012, which shows that isospin is not a perfect symmetry). It is possible to extend the concept of isospin further to particles such as Σ's and Θ's, which have spin quantum number $s = 1/2$ and somewhat similar masses. Nevertheless, it is certainly a stretch to call all of these particles different states of one single particle. SU(2) isospin symmetry is a a good, but not exact, symmetry.

It is instructive to compare the isospin classification of π and K. There are three pions: π^+, π^0, π^-. All have similar mass and they fit nicely into a $I = 1$ multiplet. As for kaons, there are 4 of them: $K^+, K^-, K^0, \bar{K}^0$. Note

that K^0 and $\bar{K}^0$ are distinct (unlike π^0 which is its own antiparticle). Now, the question is: should the kaons be a $I = 3/2$ or $I = 1/2$ multiplet? In principle, we could assign them as follows (sticking with the notation $|I, I_3\rangle$):

$$K^+ = |\frac{3}{2}, \frac{3}{2}\rangle, \ K^0 = |\frac{3}{2}, \frac{1}{2}\rangle, \ \bar{K}^0 = |\frac{3}{2}, -\frac{1}{2}\rangle, \ K^- = |\frac{3}{2}, -\frac{3}{2}\rangle. \tag{2.14}$$

If we do, however, the charge is not descending with I_3 as it should. Moreover, there is a more serious problem: the kaons have different strangeness quantum number. The K^+ and K^0 have $S = -1$ as they consist of $u\bar{s}$ and $d\bar{s}$, respectively. The K^- and $\bar{K}^0$ have $S = +1$, as they consist respectively of $\bar{u}s$ and $\bar{d}s$. The whole idea with isospin is that the particles are supposed to interact strongly in the same way, and this cannot be the case above since strangeness is *conserved* in strong interactions. Therefore, we must assign the kaons into two isospin $I = 1/2$ doublets:
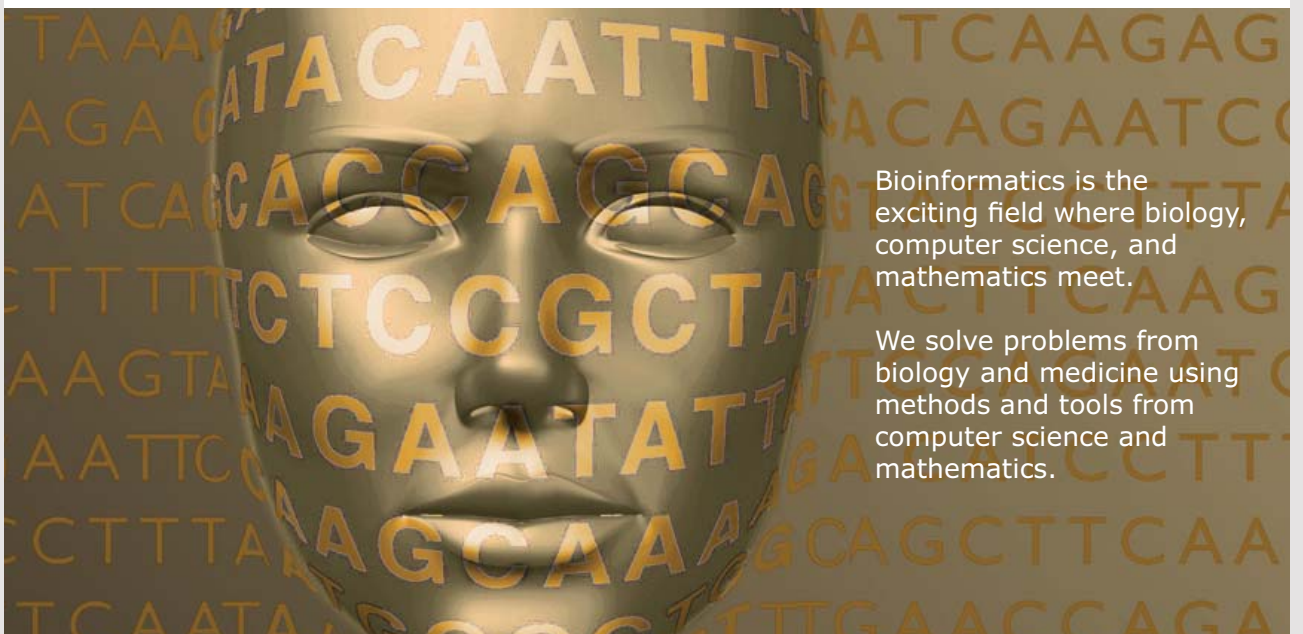
$$K^+ = |\frac{1}{2}, \frac{1}{2}\rangle, \ K^0 = |\frac{1}{2}, -\frac{1}{2}\rangle,$$
$$\bar{K}^0 = |\frac{1}{2}, \frac{1}{2}\rangle, \ K^- = |\frac{1}{2}, -\frac{1}{2}\rangle, \tag{2.15}$$

even if all have similar masses.

E. Charge conjugation and parity

Charge conjugation.
Classical electrodynamics is invariant under a sign change of charge (the forces remain the same). In particle physics, the generalization notion of this symmetry is the charge conjugation C:

$$C|p\rangle = |\bar{p}\rangle \qquad (2.16)$$

where $|p\rangle$ represents a general particle and not necessarily a proton. Therefore, the effect of the C operator is that it changes the sign of charge, baryon- and lepton-number, strangeness, and additional quantum numbers associated with quarks. It leaves mass, energy, momentum, and spin untouched (and thus also the helicity of a particle). We see that $C^2 = I$, which means that the eigenvalues of C are ± 1. Only particles that are their own antiparticles are eigenstates of C (such as the photon). However, a composite system such as spin-1/2 particle pair $p\bar{p}$ with relative orbital angular momentum l and total spin s is also an eigenstate of C with eigenvalue $(-1)^{l+s}$. Mesons are examples of this. C is a multiplicative quantum number and conserved in strong and electromagnetic interactions, whereas it is not conserved in weak interactions.

Example 3. Pion decay. The process $\pi^0 \rightarrow \gamma + \gamma$ is allowed since $C_{\pi^0} = +1$ and $C = (-1)^n$ for n photons. However, $\pi^0 \rightarrow \gamma + \gamma + \gamma$ would not be allowed via the electromagnetic interaction since C is not conserved in this process. A weak interaction could mediate this process, but its amplitude would be very small.

Can γ decay to multiple photons?
According to our discussion on charge conjugation, we see that $\gamma \rightarrow \gamma + \gamma$ violates C-symmetry. In fact, it is not even kinematically possible since the photon has no rest frame. Note that the process is not possible despite the fact that we can draw a Feynman diagram for it:
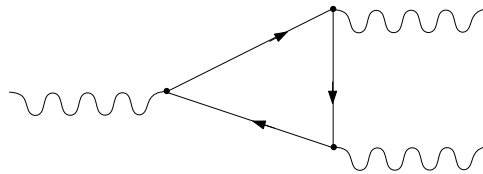
The existence of a Feynman diagram is thus no guarantee for the viability of a decay or scattering process, similarly to the primitive vertices. What about $\gamma \rightarrow \gamma + \gamma + \gamma$? It seems fine in terms of C-symmetry, but the only way that γ could decay into an odd number of photons is if all particles are moving colinearly in order to satisfy energy and momentum conservation. Computation of the actual Feynman diagram (which we will learn how to do later) nevertheless gives zero probability of this process to all orders perturbatively.

Parity.
Prior to 1956, physicists generally believed that nature had a parity symmetry: letting $\boldsymbol{r} \rightarrow (-\boldsymbol{r})$ in any physical process should also yield a physically permissible process. Lee and Yang found that there was strong experimental evidence of parity-invariance in the strong and electromagnetic processes, but not when it came to the weak interactions. Lee and Yang initiated a famous experiment, which was conducted by Wu and collaborators, that consisted of the setup shown in (a).
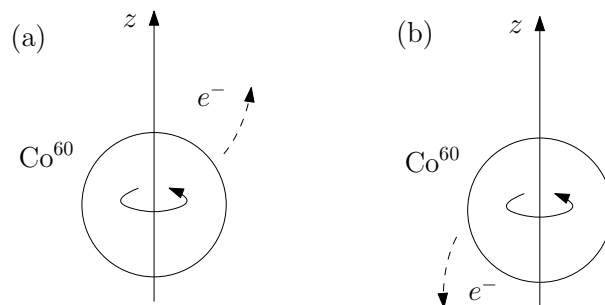
The spins of radioactive Co^{60} were aligned with the z-axis through application of a magnetic field. Upon decaying, the Co^{60} would emit electrons that turned out to escape mostly in one direction, as shown in (a). Based simply on this observation, the conclusion was that parity must be broken! If we let $r \to (-r)$, the parity-reversed process is obtained as shown in (b), but the experiment demonstrated that this does not occur. Recall that spin (angular momentum) is invariant under a parity transformation (for instance, $L = r \times p$ does not change since both r and p change sign under parity).

In fact, P violation appears to be prominent in weak interactions. This is exemplified via the neutrino. Define first the helicity of a particle as m_s/s where the z-axis is aligned with the direction of motion:

Positive helicity (right-handed)   Negative helicity (left-handed)

Earlier experiments showed that all neutrinos seemed to be left-handed while all anti-neutrinos were right-handed. Consider for instance the pion decay $\pi^- \to \mu^- + \bar{\nu}_\mu$. If π^- decays from rest (in the CM frame), the μ^- and $\bar{\nu}_\mu$ spins must be opposite since $s_{\pi^-} = 0$ and angular momentum is conserved:
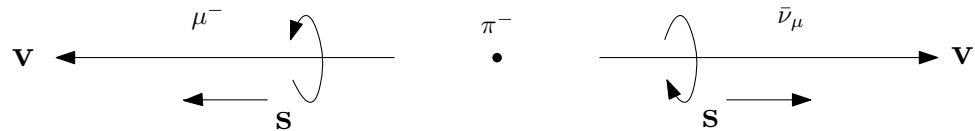
As seen, the helicities must then be the same, and experimentally one measures the muon to be right-handed in the π^- rest frame. Now, if the neutrino were truly massless, its helicity would have been the same in all inertial frames since then $v_\nu = c$. But if neutrinos do have mass (which we now know that they do), then $v_\nu < c$ and we could in principle find a reference frame S moving at a speed $v_\nu < v_S < c$. In this frame, the neutrino would be right-handed and parity symmetry would be restored. However, this has not yet been observed.

The parity operator P acts differently on different mathematical quantities. We have $Pr = -r$. However, $PL = L$ since $Pp = -p$. We thus have

- $Pv = -v$ for a vector (also known as polar vector).

- $Pa = a$ for a pseudovector (also known as axial vector).

Note that we can construct a *pseudoscalar* from the triple product of three polar vectors ($a \cdot (b \times c)$):

- $Ps = s$ for a scalar.

- $Pp = -p$ for a pseudoscalar.

Like C, P is multiplicative and conserved in strong and electromagnetic interactions. It is not conserved in weak interactions. The parity group has two elements: I and P (since $P^2 = I$). The eigenvalues are ± 1 and particle such as hadrons are eigenstates of P, so that they may be classified according to their eigenvalue. Quantum field theory dictates that the parity for fermions must be opposite to the corresponding antiparticle. For bosons, on the other hand, particle and antiparticle parity is the same. Spin-1/2 fermions have intrinsic positive parity (quarks, leptons, nucleons). In composite systems, parity is multiplicative. Excited states acquire an extra factor $(-1)^l$ where l is the angular momentum. The photon, being a vector particle, has parity -1. The terms "pseudo" and "vector" are in reference to a particle indicate the parity eigenvalue of the particle.

F. CP-violation and the TCP theorem

Although C and P are not invariant in terms of weak interactions, could it be that their product CP is? To investigate this, consider the process $K^0 \leftrightarrow \bar{K}^0$ which is possible via a second-order weak interaction:
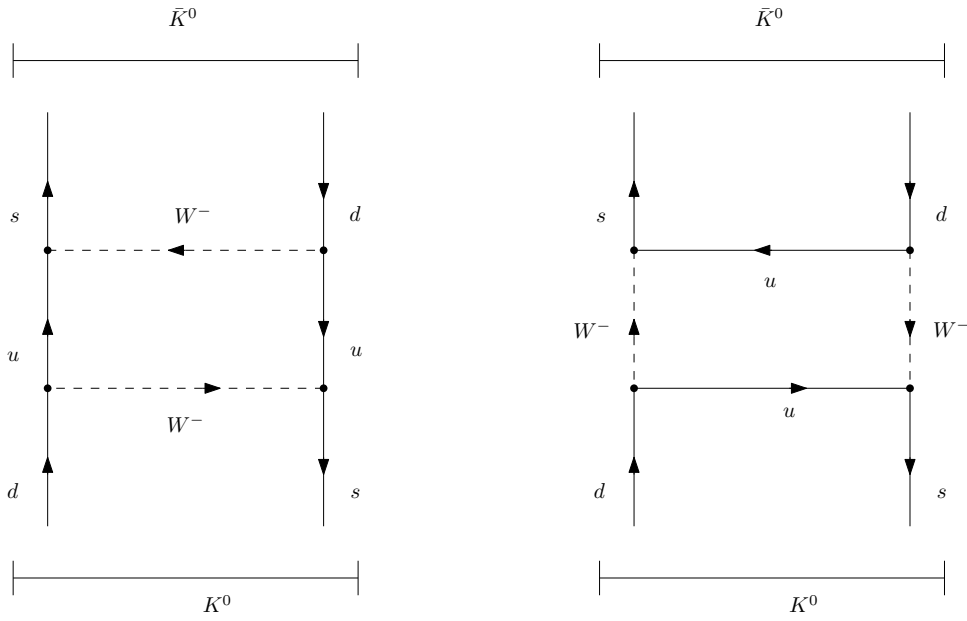
These are two of the contributing diagrams. Due to the interconversion, what is experimentally observed is usually some linear combination of K^0 and $\bar{K}^0$. We have

$$P|K^0\rangle = -|K^0\rangle, \ P|\bar{K}^0\rangle = -|\bar{K}^0\rangle, \tag{2.17}$$

since these are pseudoscalar particles. Moreover,

$$C|K^0\rangle = |\bar{K}]^0\rangle, \ C|\bar{K}^0\rangle = |K^0\rangle. \tag{2.18}$$

It follows that $CP|K^0\rangle = -|\bar{K}^0\rangle$ and $CP|\bar{K}^0\rangle = -|K^0\rangle$. Using these properties, we may then construct the normalized eigenstates of CP as:

$$CP|K_1\rangle = |K_1\rangle, \ CP|K_2\rangle = -|K_2\rangle, \tag{2.19}$$

where we defined the states

$$|K_1\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle - |\bar{K}^0\rangle), \ |K_2\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\bar{K}^0\rangle). \tag{2.20}$$

If CP is a conserved quantity, then K_1 (K_2) can only decay to a state with $CP = +1$ (-1). As a consequence, $K_1 \rightarrow 2\pi$ and $K_2 \rightarrow 3\pi$ are allowed, but $K_2 \rightarrow 2\pi$ should be impossible. Imagine then that we start the experiment with a beam of K^0 produced by an accelerator:

$$|K^0\rangle = \frac{1}{\sqrt{2}}(|K_1\rangle + |K_2\rangle). \tag{2.21}$$

The 2π decay is faster than the 3π since the released energy is much greater. Thus, for a long beam trajectory only 3π events should be observed at the end. However, an experiment by Cronin and Fitch in 1964 proved that 2π events did occur at the end of the beam, indicating violation of CP. The long-lived kaon state $|K_L\rangle$ should then read:

$$|K_L\rangle = \frac{1}{\sqrt{1 + |\epsilon|^2}}(|K_2\rangle + \epsilon|K_1\rangle), \tag{2.22}$$

since it apparently is not a perfect eigenstate of CP, where ϵ is a small number. Kaons are typically produced in strong interactions as eigenstates of strangeness, but they decay via the weak interaction. Once created, they are therefore better thought of as a superposition of weak eigenstates (K_1 and K_2).

What about time-reversal symmetry T? Should the laws of physics work equally well when reversing time? It seems not, due to the TCP-theorem from quantum field theory. It is derived only from general assumptions, such as Lorentz invariance, and states

The combined operation of T, C, and P is an exact symmetry of any interaction.

Therefore, if CP is violated, T should also be violated. If this theorem is correct, one can prove that every particle should have exactly the same mass and lifetime as its antiparticle. So far, no experiment has proven otherwise.

III. THEORY OF RELATIVISTIC KINEMATICS

Learning goals. After reading this chapter, the student should:

- Be able to perform calculations using 4-vectors and Lorentz-transformations.

- Understand the basic principles of the special theory of relativity and how this influences particle decay and scattering.

We shall here establish the basic principles, notation, and terminology of relativistic kinematics. This will be crucial in order for us to develop the mathematical foundation for particle dynamics in later chapters of this book.

A. Lorentz transformations

We will assume here that the reader has been introduced the Einstein's special theory of relativity in previous courses, for instance in classical mechanics, and thus only briefly revisit some of the key insights here. For a more detailed introduction to this topic, the reader can have a look here. The special theory of relativity states that all laws of physics are equally valid in all inertial systems. An inertial system is defined as a system where Newton's first law is obeyed, namely that objects move along straight lines with constant speed unless acted upon by a force. Consider then two inertial frames S and S' where S' is moving at constant velocity $\boldsymbol{v}$ relative S:
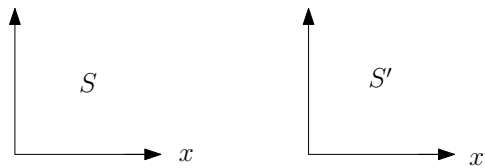
Assume that the origins of the frames coincide at $t = t' = 0$. A Lorentz-transformation establishes the relation between an event as seen in S at coordinates (x, y, z, t) and the same event as seen in S' at coordinates (x', y', z', t'):

$$x' = \gamma(x - vt), \ y' = y, \ z' = z, \ t' = \gamma(t - vx/c^2). \tag{3.1}$$

Here, we introduced

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}} \tag{3.2}$$

The inverse set of transformations (going from S' to S) are obtained by letting $\boldsymbol{r}' \rightarrow \boldsymbol{r}$ and $v \rightarrow (-v)$. The transformation rules can be derived by demanding that the speed of light c should be the same in both inertial frames. This set of of transformations has a number of consequences:

1. **The relativity of simultaneity.** If two events occur at the same time in S, but at different locations, then they do not occur at the same time in S'. Namely, if $t_A = t_B$, then $t'_A = t'_B + \frac{\gamma v}{c^2}(x_B - x_A)$.

2. **Lorentz contraction.** An object of length L' as measured in S' has a length $L = L'/\gamma$ when measured in S. The moving object is thus shortened by a factor γ, which applies to lengths along the direction of motion (dimensions $\perp$ the motion are unaffected).

3. **Time dilation**. A clock at the origin of S' ticking an interval T', will be seen to have ticked an interval $t = \gamma T'$ by an observer in S. Thus, one may state that moving clocks run slower. This is important with respect to particle physics since each unstable particle has their own "built-in" clock: a moving particle lasts longer than it would at rest. Note that tabulated lifetimes for particles always refer to the particles restframe. Due to time dilation, cosmic ray muons produced in the upper atmosphere make it to ground level even if their lifetime in the rest frame is not long enough to do so.

4. **Velocity addition.** Suppose that a particle is moving in the x-direction at speed u' with respect to S'. We then have $u' = \Delta x'/\Delta t'$. What is the speed of the particle with respect to S? It moves a distance $\Delta x = \gamma(\Delta x' + v\Delta t')$ in a time $\Delta t = \gamma[\Delta t' + (v/c^2)\Delta x']$, so that

$$u = \frac{\Delta x}{\Delta t} = \frac{(\Delta x'/\Delta t') + v}{1 + (v/c^2)(\Delta x'/\Delta t')} = \frac{u' + v}{1 + u'v/c^2}. \tag{3.3}$$

The classical, non-relativistic answer from a Galilei-transformation would be $u = u' + v$, so the correction from relativity is in the denominator. It only matters if u' and v are close to c. Note that $u' = c$ gives $u = c$.

B. 4-vectors

Define the position four-vector x^μ, $\mu = 0, 1, 2, 3, 4$:

$$x^0 = ct, \ x^1 = x, \ x^2 = y, \ x^3 = z \tag{3.4}$$

The Lorentz transformation may then be written compactly:

$$x^{\mu'} = \sum_{\nu=0}^{3} \Lambda^\mu_\nu x^\mu, \tag{3.5}$$

where Λ^μ_ν are the coefficients of the matrix Λ:

$$\Lambda = \begin{bmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.6}$$

so that for instance $\Lambda^0_0 = \gamma$, $\Lambda^1_0 = -\gamma\beta$, and so forth. We will from now on use Einstein's sum convention for greek indices, which means that for repeated indices a summation is implicitly understood. The Lorentz-transformation is then written simply as $x^{\mu'} = \Lambda^\mu_\nu x^\nu$. It can be shown that the particular combination:

$$I \equiv (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \tag{3.7}$$

has the same value in all inertial systems, so that I is also equal to $(x^{0'})^2 - (x^{1'})^2 - (x^{2'})^2 - (x^{3'})^2$. Therefore, I must be a *Lorentz-invariant* as it remains the same under a Lorentz-transformation, analogously to how $r^2 = x^2 + y^2 + z^2$ is invariant under rotations in three-dimensional space. To keep track of the different signs in I, we introduce the *metric* $g_{\mu\nu}$ which are the components of the matrix g:

$$g = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{3.8}$$

We may then express I as a double sum: $I = g_{\mu\nu}x^\mu x^\nu$. From this, we may define the *covariant four-vector* x_μ (index down):

$$x_\mu \equiv g_{\mu\nu}x^\nu. \tag{3.9}$$

It follows that $x_0 = x^0$ and $x_i = -x^i$, $i = 1, 2, 3$. For index up, we have a *contravariant four-vector* x^μ:

$$x^\mu = g^{\mu\nu}x_\nu. \tag{3.10}$$

We thus have:

> Index up (x^μ): contravariant vector. Index down (x_μ): covariant vector.

The metric g can thus be thought of as raising or lowering the index of a four-vector. One way to memorize this is the following: con*tra*variant vectors have index on *top*. Our invariant I can now be written even more compactly as

$$I = x_\mu x^\mu. \tag{3.11}$$

This type of notation also generalizes to non-Cartesian coordinate systems and curved spaces encountered in the theory of general relativity. We define a four-vector a^μ formally as an object that transforms in the same way as x^μ when going from one inertial-system to another:

$$a^{\mu'} = \Lambda^\mu_\nu a^\nu. \tag{3.12}$$

We also associate a covariant four-vector a_μ to each such contravariant vector:

$$a_\mu = g_{\mu\nu} a^\nu. \tag{3.13}$$

According to our definition above, we go from covariant to contravariant four-vectors via $a^\mu = g^{\mu\nu} a_\nu$. Here, $g^{\mu\nu}$ are the elements of the matrix g^{-1}. For our case, we see that $g = g^{-1}$ so that $g_{\mu\nu} = g^{\mu\nu}$. For any two four-vectors, a_μ and b^μ, the quantity $a_\mu b^\mu \equiv a \cdot b$ is invariant. We will sometimets also simply write $a_\mu b^\mu \equiv ab$ when there is no risk for confusion. Notation-wise, we may distinguish between time and spatial coordinates in this invariant since $a_\mu b^\mu = a^0 b^0 - \boldsymbol{a} \cdot \boldsymbol{b}$. Note that $a^2 = a_\mu a^\mu$ need not be positive:

- If $a^2 > 0$, then a^μ is said to be timelike.

- If $a^2 < 0$, then a^μ is said to be spacelike.

- If $a^2 = 0$, then a^μ is said to be lightlike.

Tensors transform in a generalized way compared to vectors: $s^{\mu\nu'} = \Lambda^\mu_\kappa \Lambda^\nu_\sigma s^\kappa \sigma$. Note that any tensor of rank $n + 2$ can be contracted to a tensor of rank n by performing a summation over upper and lower indices. For instance, s^μ_μ is a scalar while $t^{\mu\nu}_\nu$ is a four-vector, and so forth.

C. Energy and momentum

We previously stated that time dilation refers to the phenomenon that a moving clock measures an eigentime $d\tau$ whereas a stationary observer meaures dt. The two are related via

$$d\tau = dt/\gamma. \tag{3.14}$$

The quantity τ is an invariant: all observers agree on which time that is measured in the particle's restframe, even if the time measured in their own frame might differ. Just as 4-position vector gives an invariant $x^\mu x_\mu$, the 4-velocity:

$$\eta^\mu = dx^\mu/d\tau = \gamma(c, \boldsymbol{v}) \tag{3.15}$$

gives an invariant when contracted: $\eta_\mu \eta^\mu = \gamma^2(c^2 - \boldsymbol{v}^2) = c^2$. Note that if we had defined relativistic momentum as $m\boldsymbol{v}$, it would not have provided conservation of momentum in all inertial systems, assuming that it was valid in one such system. However, $m\eta_\mu$ assures that this holds. Thus, the appropriate 4-momentum vector reads:

$$p^\mu = (\gamma m c, \gamma m \boldsymbol{v}). \tag{3.16}$$

Let us define the relativistic energy as $E = \gamma m c^2$. Then, it follows that $p_\mu p^\mu = E^2/c^2 - \boldsymbol{p}^2 = E^2/c^2 - \gamma^2 m^2 v^2 = m^2 c^2$. We see that the relativistic momentum reduces to its classical counterpart when $v \ll c$ since $\gamma \to 1$ then. However, for the energy we obtain

$$E = mc^2\left(1 + \frac{1}{2}\frac{v^2}{c^2} + \frac{3}{8}\frac{v^4}{c^4} + \dots\right) \tag{3.17}$$

from which it is seen that an extra constant term mc^2 exists in the limit $v \ll c$. This term survives even in the absence of any kinetic energy, $v = 0$. It is the so-called *rest energy* of the particle. The remainder of the energy must then be associated with the kinetic content: $E_{\text{kinetic}} \equiv (\gamma - 1)mc^2$. An aspect which has no counterpart in non-relativistic mechanics is that a massless particle with non-zero momentum may exist: $E = |\boldsymbol{p}|c$. The fact that this is possible is demonstrated by the formulae

$$\boldsymbol{p} = m\boldsymbol{v}/\sqrt{1 - v^2/c^2} \text{ and } E = mc^2/\sqrt{1 - v^2/c^2}. \tag{3.18}$$

since $m \to 0$ still can give a finite $\boldsymbol{p}$ and E if simultaneously $v \to c$.

D. Collisions

Energy and momentum are, as in non-relativistic mechanics, conserved quantities in particle collisions. This may be compacty expressed via the 4-momentum vector: p^μ_{total} is a conserved quantity. Mass, however, is not necessarily conserved. The crucial concept here is that rest energy (and thus mass) may be converted into kinetic energy or vice versa. To obtain consistency with the non-relativistic limit, where kinetic energy is converted into some internal form of energy (such as heat), we need only realize that all such internal energies are reflected in the rest energy of the body. On a macroscopic scale, the rest energy is much greater than the internal energy, so that the added mass from internal energy is completely negligible. Strictly speaking, however, a hot potato weighs more than a cold potato. Imagine two objects that have identical masses when at the same temperature. If one of the objects is heated, so that its internal temperature increases, its weight will also become larger than the other object according to our reasoning above.

Example 4. Particles merging. Consider the following scenario where two particles merge into one via a head-on collision.
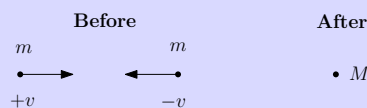
The task is to find the mass M of the final particle when $|v| = 3c/5$. Conservation of momentum gives $\boldsymbol{p}_1 = -\boldsymbol{p}_2$. For the energy, we obtain

$$Mc^2 = 2E_m = 2mc^2/\sqrt{1 - v^2/c^2} = 5mc^2/2. \tag{3.19}$$

Thus, we find $M = 5m/2 > 2m$, meaning that we have a so-called sticky collision where the mass has increased. Note that the reverse process, M decaying into two particles with mass m each, is only possible if $M > 2m$. A deuteron (bound state of a p and n) weigs less than $m_p + m_n$. Therefore, it will not decay unless energy is injected. This is actually a concrete example of how the (negative) binding energy is reflected in the total rest mass.

Example 5. Pion decay. A pion at rest decays into a muon and a neutrino. What is the speed of the muon? Let us start with the conservation laws at hand:

$$E_\pi = E_\mu + E_\nu, \ \boldsymbol{p}_\pi = \boldsymbol{p}_\mu + \boldsymbol{p}_\nu \rightarrow \boldsymbol{p}_\mu = -\boldsymbol{p}_\nu. \tag{3.20}$$

One solution strategy is to find the energy of a particle when you know its momentum, by using the invariant $E^2 - \boldsymbol{p}^2 c^2 = m^2 c^4$. When the energy has been identified in this way, use that $E = \gamma mc^2$ and $\boldsymbol{p} = \gamma m\boldsymbol{v}$ which gives $\boldsymbol{v} = \boldsymbol{p}c^2/E$. For this example, we have

$$E_\pi = m_\pi c^2, \ E_\mu = c\sqrt{m_\mu^2 c^2 + \boldsymbol{p}_\mu^2}, \ E_\nu = |\boldsymbol{p}_\nu|c = |\boldsymbol{p}_\mu|c. \tag{3.21}$$

This provides $|\boldsymbol{p}_\mu| = (m_\pi^2 - m_\mu^2)c/2m_\pi$ and $E_\mu = (m_\pi^2 + m_\mu^2)c^2/2m_\pi$. Plugging these expressions into $\boldsymbol{v} = \boldsymbol{p}c^2/E$ then provides the muon velocity.

We end this chapter by emphasizing the difference between a *conserved quantity* and a *invariant quantity*. Energy is conserved (same value before and after collision), but it is not an invariant: the energy can have different values in different inertial frames. On the other hand, mass is invariant (same value in all inertial systems), but it is not a conserved quantity in general (before and after a collision).

IV. QUANTUM ELECTRODYNAMICS AND FEYNMAN RULES

Learning goals. After reading this chapter, the student should:

- Be able to work with the Klein-Gordon and Dirac equation mathematically and understand what their solutions represent.

- Be able to compute the Feynman amplitude $\mathcal{M}$ for Feynman diagrams in quantum electrodynamics, and also be able to obtain decay rates and scattering cross sections from $\mathcal{M}$.

- Understand the principle behind renormalization and how it is related to higher-order Feynman diagrams.

This theory is a major cornerstone in particle physics and serves as the foundation for much of the material that we will look at later. We will here introduce the Dirac equation and its solution, Feynman rules and how they work specifically in QED, in addition to some pragmatic calculational tools.

A. The Dirac equation

In non-relativistic quantum mechanics (QM), it is the Schrödinger equation which is the fundamental equation that describes the behavior of physical systems. In relativistic QM, one has different fundamental equations depending on the spin of the particle.

- Spin-0: Klein-Gordon equation.

- Spin-$\frac{1}{2}$: Dirac equation.

- Spin-1: Proca equation.

We will later in this chapter establish a set of Feynman rules which will allow us to evaluate Feynman diagrams mathematically. Once this set of rules has been established, the underlying field equation for a particle with a given spin is no longer immediately needed. However, the notation of Feynman rules for spin-$\frac{1}{2}$ particles does require that we are familiar with the Dirac equation and we therefore treat it, and its solutions, here in detail.

It is instructive to first consider how the Schrödinger equation (SE) can be motivated. If one starts out with the classical, non-relativistic expression for energy, $E = \boldsymbol{p}^2/2m + V$, and then substitute the operators

$$\boldsymbol{p} \to \frac{\hbar}{\mathrm{i}}\nabla,\ E \to \mathrm{i}\hbar\partial_t \tag{4.1}$$

one ends up with precisely the time-dependent SE:

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi = \mathrm{i}\hbar\partial_t\Psi. \tag{4.2}$$

The same principle can be applied for the Klein-Gordon (KG) equation. The starting point is the classical, relativistic expression for energy, $E^2 - \boldsymbol{p}^2c^2 = m^2c^4$, or $p^\mu p_\mu - m^2c^2 = 0$. Consider a free particle to begin with such that $V = 0$. Now, perform the same substitution as in the SE case in order to introduce operators

$$p_\mu = \mathrm{i}\hbar\partial_\mu \tag{4.3}$$

which is simply a more compact way of writing Eq. (4.1). We introduced $\partial_\mu \equiv \frac{\partial}{\partial x^\mu}$. In effect, this means that

$$(E/c, -\boldsymbol{p}) \to \mathrm{i}\hbar(\partial_t/c, \nabla). \tag{4.4}$$

The KG equation now becomes:

$$-\hbar^2\partial^\mu\partial_\mu\Psi - m^2c^2\Psi = 0 \tag{4.5}$$

which alternatively can be written as:

$$-\frac{1}{c^2}\partial^2\Psi\partial t^2 + \nabla^2\Psi = \left(\frac{mc}{\hbar}\right)^2\Psi. \tag{4.6}$$

In the presence of an EM potential, we have to take into account the influence of the electric scalar potential and magnetic vector potential. This is done in the usual manner: the energy is changed to $E \rightarrow E - e\phi$ while the momentum is augmented by the magnetic vector potential via $\boldsymbol{p} \rightarrow \boldsymbol{p} - \frac{e\boldsymbol{A}}{c}$ in order for the physics to be gauge-invariant (the so-called minimal coupling). In this case, we obtain

$$(\mathrm{i}\hbar\partial_t - e\phi)^2\Psi = c^2\left(-\mathrm{i}\hbar\nabla - \frac{e\boldsymbol{A}}{c}\right)^2\Psi + m^2c^4\Psi. \tag{4.7}$$

Note that while the SE is $\mathcal{O}(t)$, the KG equation is $\mathcal{O}(t^2)$. As we now will turn our attention to the Dirac equation, it is also worth to emphasize that Ψ is a scalar field and that any solution to the Dirac equation is also a solution of the KG equation, as there is no reference to spin in the KG equation One can loosely think of the KG equation as describing the magnitude of the field, whereas the Dirac equation describes both the magnitude and the "direction" (spin) of the field in the case of spin-$\frac{1}{2}$ particlecs. Note that the opposite is not true: a solution of the KG equation is not necessarily a solution of the Dirac equation, since the KG equation has lost a degree of freedom (spin) compared to the Dirac equation.

Now, in order to obtain the Dirac equation, the strategy is to factor the $E - \boldsymbol{p}$ relation (often called a dispersion relation). If $\boldsymbol{p} = 0$, this is easy:

$$(p^0)^2 - m^2c^2 = 0 = (p^0 + mc)(p^0 - mc) \rightarrow (p^0 - mc) = 0 \text{ or } (p^0 + mc) = 0. \tag{4.8}$$

Both options guarantee that $p^\mu p_\mu - m^2c^2 = 0$. For the more interesting case $\boldsymbol{p} \neq 0$, we would need something like:

$$(p^\mu p_\mu - m^2c^2) = (\beta^\kappa p_\kappa + mc)(\gamma^\lambda p_\lambda - mc), \tag{4.9}$$

where β^κ and γ^λ are undetermined coefficients for now. To match the two sides of the equation, we obtain

$$\beta^\kappa\gamma^\lambda p_\kappa p_\lambda - mc(\beta^\kappa - \gamma^\kappa)p_\kappa - m^2c^2. \tag{4.10}$$

If $\beta^\kappa = \gamma^\kappa$, the linear terms in momentum are seen to cancel. Now, we must find γ^κ so that the second order terms also match:

$$p^\mu p_\mu = \gamma^\kappa\gamma^\lambda p_\kappa p_\lambda. \tag{4.11}$$

The problem with this equation is that it is impossible to satisfy for scalars γ^κ. However, it can be satisfied if the γ-quantities are *matrices* instead. Writing out Eq. (4.11) more explicitly, we obtain

$$(p^0)^2 - (p^1)^2 - (p^2)^2 - (p^3)^2 = (\gamma^0)^2(p^0)^2 + (\gamma^1)^2(p^1)^2 + (\gamma^2)^2(p^2)^2 + (\gamma^3)^2(p^3)^2$$
$$+ (\gamma^0\gamma^1 + \gamma^1\gamma^0)p_0p_1 + \dots(\gamma^2\gamma^3 + \gamma^3\gamma^2)p_2p_3. \tag{4.12}$$

We need a set of matrices to get rid off the cross terms, and we thus require that the matrices have the following properties in order to accomplish this:

$$(\gamma^0)^2 = 1, \ (\gamma^1)^2 = (\gamma^2)^2 = (\gamma^3)^2 = -1 \tag{4.13}$$

in addition to

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 0 \text{ for } \mu \neq \nu. \tag{4.14}$$

It is clear why scalars could not satisfy the above properties: scalars commute. The above requirements, can be written quite compactly as:

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}, \ g^{\mu\nu} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{4.15}$$

$g^{\mu\nu}$ is thus the Minkowski metric. There exists an infinite number of physically equivalent sets of γ-matrices that satisfy these properties. The smallest ones are 4×4 and take the explicit form:

$$\gamma^0 = \begin{bmatrix} \underline{1} & \underline{0} \\ \underline{0} & -\underline{1} \end{bmatrix}, \; \gamma^i = \begin{bmatrix} \underline{0} & \underline{\sigma^i} \\ -\underline{\sigma^i} & \underline{0} \end{bmatrix} \tag{4.16}$$

where $\underline{\dots}$ means a 2×2 matrix. Now, the Dirac equation factorizes:

$$(p^\mu p_\mu - m^2 c^2) = (\gamma^\kappa p_\kappa + mc)(\gamma^\lambda p_\lambda - mc) = 0. \tag{4.17}$$

The conventional choice is now to consider:

$$\gamma^\mu p_\mu - mc = 0. \tag{4.18}$$

If we let $p_\mu \rightarrow i\hbar\partial_\mu$ as for the other quantum equations (thus introducing operators), we obtain the final form of the Dirac equation:

$$\boxed{i\hbar\gamma^\mu\partial_\mu\psi - mc\psi = 0.}$$

Here, ψ is a Dirac spinor (which is *not a 4-vector*):

$$\psi = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{bmatrix} \tag{4.19}$$

Since it is not a 4-vector, it does not transform via a Lorentz transformation when changing inertial system.

We note in passing that the Pauli equation, which is essentially the SE with spin taken into account, is obtained by considering the Dirac equation in the low-energy (non-relativistic limit). We underline that spin is not a relativistic correction: it is a fundamental property of particles which existence does not rely on the speed at which the particles move. The Dirac equation thus describes both spin *and* relativistic effects.

Alternative derivation of the Dirac equation.
We saw above that we could fulfil the equation:

$$p^\mu p_\mu - m^2 c^2 = (\gamma^\kappa p_\kappa + mc)(\gamma^\lambda p_\lambda - mc) = 0 \tag{4.20}$$

with appropriate γ-matrices. We then took one of these factors, say $\gamma^\mu p_\mu - mc$, and set it to zero. Assigning operators and acting on a wavefunction, we got:

$$i\hbar\gamma^\mu \partial_\mu \psi - mc\psi = 0. \tag{4.21}$$

But $AB = 0$ does not imply that $A = 0$ or $B = 0$ when A and B are matrices. Therefore, we are strictly speaking not allowed to conclude that one of the factors in Eq. (4.20) is zero, even if it seems to give us the correct result. We therefore give here a more mathematically satisfying derivation which produces the same result.

Consider the KG equation:

$$\left(\nabla^2 - \frac{\partial_t^2}{c^2}\right)\psi = \left(\frac{mc}{\hbar}\right)^2 \psi. \tag{4.22}$$

If we factorize this equation (which is already in operator form), we get:

$$\left(\nabla^2 - \frac{\partial_t^2}{c^2}\right)\psi = (A\partial_x + B\partial_x + C\partial_z + iD\partial_t/c)(A\partial_x + B\partial_x + C\partial_z + iD\partial_t/c)\psi$$

$$= \left(\frac{mc}{\hbar}\right)^2 \psi \equiv \kappa^2 \psi. \tag{4.23}$$

This equation is satisfied if

$$(A\partial_x + B\partial_x + C\partial_z + iD\partial_t/c)\psi = \pm\kappa\psi \tag{4.24}$$

and if the terms $\{A, B, C, D\}$ are chosen so that the cross terms on the r.h.s. of the first line in Eq. (4.23) vanish. It turns out that this occurs when

$$\{A, B, C\} = i\gamma_k, \ D = \gamma^0. \tag{4.25}$$

so that Eq. (4.24) gives the Dirac equation:

$$i\hbar\gamma^\mu \partial_\mu \psi \pm mc\psi = 0. \tag{4.26}$$

This equation can also be brought to another form which is commonly used by multiplying it with γ^0:

$$\gamma^0(\gamma^0 p_0 - \boldsymbol{\gamma} \cdot \boldsymbol{p} - mc)\psi \rightarrow i\partial_t \psi - \gamma^0 \boldsymbol{\gamma} \cdot \boldsymbol{p}\psi - m\gamma^0 \psi = 0. \tag{4.27}$$

This may be written as:

$$i\partial_t \psi = H_D \psi \tag{4.28}$$

where we defined the Hamiltonian

$$H_D = \boldsymbol{\alpha} \cdot \boldsymbol{p} + \beta m \tag{4.29}$$

and $\boldsymbol{\alpha} = \gamma^0 \boldsymbol{\gamma}$ and $\beta = \gamma^0$. This form of the equation describing a relativistic spin-1/2 particle also makes it clear how to include a potential energy $V(\boldsymbol{r})$, namely $H_D \rightarrow H_D + V(\boldsymbol{r})$ as usual.

Solutions to the Dirac equation.
Consider first the simple case where ψ is independent on $\boldsymbol{r}$: $\partial\psi/\partial x_j = 0$ $(j = x, y, z)$. This should be the case for $\boldsymbol{p} = 0$, since $p_\mu \rightarrow i\hbar\partial_\mu$. The Dirac equation then reads

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{\partial\psi_A}{\partial t} \\ \frac{\partial\psi_B}{\partial t} \end{bmatrix} = -i\frac{mc^2}{\hbar} \begin{bmatrix} \psi_A \\ \psi_B \end{bmatrix} \tag{4.30}$$

where $\psi_A = [\psi_1, \psi_2]^T$ and $\psi_B = [\psi_3, \psi_4]^T$. The superscript ^{T} denotes matrix transposition. The equations for ψ_A and ψ_B are decoupled, and the solutions are obtained as:

$$\psi_A(t) = e^{-i(mc^2/\hbar)t}\psi_A(0), \ \psi_B(t) = e^{+i(mc^2/\hbar)t}\psi_B(0). \tag{4.31}$$

For a particle at rest, $E = mc^2$, so ψ_A is seen to carry the usual factor $e^{-iEt/\hbar}$. The negative sign in the exponent of ψ_B represents *antiparticles* with positive energy. For instance, ψ_A may represent an electron, in which case ψ_B represents a positron.

The particle and antiparticle (p and $\bar{p}$) parts are each 2×1 spinors since they have $s = \frac{1}{2}$. For $\boldsymbol{p} = 0$, we then have four independent solutions

- Electron spin-↑: $\psi^{(1)} = e^{-i(mc^2/\hbar)t}[1, 0, 0, 0]^T$.

- Electron spin-↓: $\psi^{(2)} = e^{-i(mc^2/\hbar)t}[0, 1, 0, 0]^T$.

- Positron spin-↑: $\psi^{(3)} = e^{+i(mc^2/\hbar)t}[0, 0, 1, 0]^T$.

- Positron spin-↓: $\psi^{(4)} = e^{+i(mc^2/\hbar)t}[0, 0, 0, 1]^T$.

We now turn to the more difficult, but much more interesting, case of $\boldsymbol{p} \neq 0$. We look for solutions to the Dirac equation in a plane-wave form:

$$\psi(\boldsymbol{r}, t) = ae^{-i(Et - \boldsymbol{p} \cdot \boldsymbol{r})/\hbar}u(E, \boldsymbol{p}) \tag{4.32}$$

which in 4-vector notation can be written (more compactly, as usual): $\psi(x) = ae^{-ix \cdot p/\hbar}u(p)$. Here, a is a normalization constant and we must determine $u(p)$. Inserting our ansatz into the Dirac equation produces: $(\gamma^\mu p_\mu - mc)u = 0$. It is simpler than the original equation because there are no derivatives present. Now, we see that since

$$\gamma^\mu p_\mu = \frac{E}{c}\begin{bmatrix} \underline{1} & 0 \\ 0 & -\underline{1} \end{bmatrix} - \boldsymbol{p} \cdot \begin{bmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{bmatrix}, \tag{4.33}$$

it follows that

$$(\gamma^\mu p_\mu - mc)u = \begin{bmatrix} (E/c - mc)u_A - \boldsymbol{p} \cdot \boldsymbol{\sigma}u_B \\ \boldsymbol{p} \cdot \boldsymbol{\sigma}u_A - (E/c + mc)u_B \end{bmatrix}. \tag{4.34}$$

The $(E/c \pm mc)$ terms are implicitly understood to have an $\underline{1}$ structure as dimensionality requires, and we follow this convention in what follows (i.e. omitting identity matrices where it is clear from the context that such a matrix should be present). The solution to the above equation is:

$$u_A = \frac{c}{E - mc^2}(\boldsymbol{p} \cdot \boldsymbol{\sigma})u_B, \ u_B = \frac{c}{E + mc^2}(\boldsymbol{p} \cdot \boldsymbol{\sigma})u_A. \tag{4.35}$$

Combining these equations, we find

$$u_A = \frac{c^2(\boldsymbol{p} \cdot \boldsymbol{\sigma})^2}{E^2 - m^2c^4}u_A, \tag{4.36}$$

which is seen to be consistent as $(\boldsymbol{p} \cdot \boldsymbol{\sigma})^2 = \boldsymbol{p}^2 = (E^2 - m^2c^4)/c^2$. The wavefunctions are then determined up to a normalization constant:

1. $u_A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, which gives $u_B = \frac{c}{E+mc^2}\begin{bmatrix} p_z \\ p_x + ip_y \end{bmatrix}$.

2. $u_A = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, which gives $u_B = \frac{c}{E+mc^2}\begin{bmatrix} p_x - ip_y \\ -p_z \end{bmatrix}$.

3. $u_B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, which gives $u_A = \frac{c}{E-mc^2} \begin{bmatrix} p_z \\ p_x + \mathrm{i}p_y \end{bmatrix}$.

4. $u_B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, which gives $u_A = \frac{c}{E-mc^2} \begin{bmatrix} p_x - \mathrm{i}p_y \\ -p_z \end{bmatrix}$.

In the first two cases, we have to use $E = +\sqrt{\boldsymbol{p}^2 c^2 + m^2 c^4}$ in order to avoid u_B diverging when $\boldsymbol{p} \to 0$. These are the *particle solutions*. In the two last cases, the opposite is true: we must use $E = -\sqrt{\boldsymbol{p}^2 c^2 + m^2 c^4}$ to avoid divergences. These are the *antiparticle* solutions. Regarding the normalization, there exists different conventions. A common convention that we shall stick with here is to normalize the above spinors so that $u^\dagger u = 2|E|/c$. The † means conjugate transpose, so that

$$u = \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} \to u^\dagger = [\alpha^*, \beta^*, \gamma^*, \delta^*]. \tag{4.37}$$

Note that all of the choices 1-4 above do not correspond immediately to spin-$\uparrow, \downarrow$ electrons or positrons as they have more than one non-zero entry in the spinor. Neither of them are in fact eigenstates of the spin operator

$$\boldsymbol{S} = \frac{\hbar}{2}\boldsymbol{\Sigma} = \frac{\hbar}{2}\begin{bmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{bmatrix}. \tag{4.38}$$

Only for the special case of $p_x = p_y = 0$ will the spinors 1-4 be eigenspinors of S_z and $\boldsymbol{S}^2$. We noted that $E = -\sqrt{\boldsymbol{p}^2 c^2 + m^2 c^4}$ corresponds to antiparticles with positive energy (since free particles must have positive energy). It is customary to relabel the antiparticle states to v and flip the sign of the energy and momentum, so that we can use $E = \sqrt{\boldsymbol{p}^2 c^2 + m^2 c^4}$ everywhere without having to worry about whether the energy belongs to a particle or antiparticle state. The four independent solutions valid for finite momentum $\boldsymbol{p} \neq 0$ then read:

$$u^{(1)} = N \begin{bmatrix} 1 \\ 0 \\ \frac{cp_z}{E+mc^2} \\ \frac{cp_+}{E+mc^2} \end{bmatrix}, \; u^{(2)} = N \begin{bmatrix} 0 \\ 1 \\ \frac{cp_-}{E+mc^2} \\ -\frac{cp_z}{E+mc^2} \end{bmatrix}, \; v^{(1)} = N \begin{bmatrix} \frac{cp_-}{E+mc^2} \\ -\frac{cp_z}{E+mc^2} \\ 0 \\ 1 \end{bmatrix}, \; v^{(2)} = -N \begin{bmatrix} \frac{cp_z}{E+mc^2} \\ \frac{cp_+}{E+mc^2} \\ 1 \\ 0 \end{bmatrix}, \tag{4.39}$$

where we defined $N = \sqrt{(|E| + mc^2)/c}$ and $p_\pm = p_x \pm \mathrm{i}p_y$. Thus, $u^{(i)}$ are two spin states for an electron with energy E and momentum $\boldsymbol{p}$ while $v^{(i)}$ are two spin states for a positron with energy E and momentum $\boldsymbol{p}$. Notice that the u-spinors satisfy $(\gamma^\mu p_\mu - mc)u = 0$ while the v-spinors satisfy $(\gamma^\mu p_\mu + mc)v = 0$ since we have reversed the sign of E and $\boldsymbol{p}$ by convention.

Plane-wave solutions are particularly useful since they describe particles with specific energies and momenta, both quantities that are usually possible to control quite well in experiments.

Transformation of Dirac spinors.
We mentioned previously that a Dirac spinor is not a 4-vector, meaning that it does not transform via the Lorentz transformation under a change of inertial system. How does it transform then? The full proof is not given here and the reader is referred to *e.g.* the textbook by Bjorken & Drell. The reader is instead encouraged to try to derive the transformation rule via the following strategy. We want to discover how the solution ψ of the Dirac equation in one frame, which satisfies

$$\mathrm{i}\hbar\gamma^\mu\partial_\mu\psi - mc\psi = 0 \tag{4.40}$$

looks in a different frame, $\psi' = S\psi$, where the Dirac equation reads

$$\mathrm{i}\hbar\gamma^\mu\partial'_\mu\psi' - mc\psi' = 0. \tag{4.41}$$

To help us, we know how the differential operators transform as follows under a change of inertial system:

$$\partial'_\mu = \frac{\partial}{\partial x^{\mu'}} = \frac{\partial x^\nu}{\partial x^{\mu'}}\partial_\nu. \tag{4.42}$$

The result is that:

$$\psi \to \psi' = S\psi \tag{4.43}$$

where the $'$ system is moving with speed v in the x-direction and

$$S = \begin{bmatrix} a_+\underline{1} & a_-\underline{\sigma}_1 \\ a_-\underline{\sigma}_1 & a_+\underline{1} \end{bmatrix}. \tag{4.44}$$

Besides $\gamma = 1/\sqrt{1 - (v/c)^2}$ as usual, we defined

$$a_\pm = \pm\sqrt{\frac{1}{2}(\gamma \pm 1)}. \tag{4.45}$$

Now, $\psi^\dagger\psi$ does not transform as a scalar under a change of inertial system because $(\psi^\dagger\psi)' \neq \psi^\dagger\psi$:

$$(\psi^\dagger\psi)' = (\psi^\dagger)'\psi' = \psi^\dagger S^\dagger S\psi \tag{4.46}$$

and

$$S^\dagger S = \gamma \begin{bmatrix} \underline{1} & -\frac{v}{c}\underline{\sigma}_1 \\ -\frac{v}{c}\underline{\sigma}_1 & \underline{1} \end{bmatrix} \neq 1. \tag{4.47}$$

Similarly to a 4-vector, only the proper metric (relative sign between the components x^μ) will make a quantity transform like a scalar. In this case, it turns out to be

$$\bar{\psi}\psi = \psi^\dagger \gamma^0 \psi = |\psi_1|^2 + |\psi_2|^2 - |\psi_3|^2 - |\psi_4|^2 \tag{4.48}$$

which is a relativistic invariant. Here, we introduced the adjoint spinor:

$$\bar{\psi} \equiv \psi^\dagger \gamma^0. \tag{4.49}$$

It is also of interest to consider the transformation properties of ψ under a parity transformation. Recall how scalars and pseudoscalars are distinguished based on how they transform under a parity operation. We have seen that $\bar{\psi}\psi$ is Lorentz-invariant (transforms as a scalar under a change of inertial system), but how does it transform under parity? Using a similar strategy as above (working out the parity transformation effect on the differential operators), one finds that parity has the effect:

$$\psi \rightarrow \psi' = \gamma^0 \psi. \tag{4.50}$$

It follows that these relations hold for a parity transformation:

- $\bar{\psi}\psi = $ scalar.

- $\bar{\psi}\gamma^5\psi = $ pesudoscalar.

- $\bar{\psi}\gamma^\mu\psi = $ vector.

- $\bar{\psi}\gamma^\mu\gamma^5\psi = $ pseudovector.

B. The photon

Classical electrodynamics is described by Maxwell's equation (using a Gaussian units convention, but the procedure below is equally valid for SI units):

$$(i)\ \nabla \cdot \boldsymbol{E} = 4\pi\rho,\ (ii)\ \nabla \cdot \boldsymbol{B} = 0,$$
$$(iii)\ \nabla \times \boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} = 0,\ (iv)\ \nabla \times \boldsymbol{B} - \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} = \frac{4\pi}{c}\boldsymbol{J}. \tag{4.51}$$

Here, ρ is the charge density and $\boldsymbol{J}$ is the current density. In relativistic theory notation, one introduces the antisymmetric field tensor $F_{\mu\nu}$:

$$F_{\mu\nu} = \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{bmatrix} \tag{4.52}$$

and the 4-vector $J^\mu = (c\rho, \boldsymbol{J})$. The first and fourth of Maxwell's equations listed above can now be compactly written:

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c}J^\nu. \tag{4.53}$$

Since $F^{\mu\nu}$ is an antisymmetric tensor, $F^{\mu\nu} = -F^{\nu\mu}$, one finds that $\partial_\mu J^\mu = 0$. This can be written as

$$\nabla \cdot \boldsymbol{J} = -\frac{\partial \rho}{\partial t}, \tag{4.54}$$

which is simply the charge continuity equation. On the other hand, the homogeneous equations (the second and third of Maxwell's equations listed above) may be reexpressed via the scalar potential ϕ and vector potential $\boldsymbol{A}$ as follows:

$$\boldsymbol{B} = \nabla \times \boldsymbol{A},\ \boldsymbol{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t} \tag{4.55}$$

which can be compactly expressed through:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{4.56}$$

The fields $\boldsymbol{E}$ and $\boldsymbol{B}$ are the physical observables, whereas the potentials ϕ and $\boldsymbol{A}$ [or simply $A^\mu \equiv (\phi, \boldsymbol{A})$] are not uniquely determined:

$$A'_\mu = A_\mu + \partial_\mu \lambda \tag{4.57}$$

gives the same fields since $\partial^\mu A^{\nu\prime} - \partial^\nu A^{\mu\prime} = \partial^\mu A^\nu - \partial^\nu A^\mu$. Here, λ is any differentiable function of position and time. Eq. (4.57) is an example of a *gauge-transformation*, because we can choose any particular gauge (λ) that we like. For instance, $\partial_\mu A^\mu = 0$ is the Lorentz condition, which means that we restrict λ to satisfy $\Box \lambda = 0$ where $\Box \equiv \partial^\mu \partial_\mu$. In this gauge, the inhomogeneous Maxwell equations simplify to

$$\Box A^\mu = 4\pi J^\mu / c. \tag{4.58}$$

For empty space with $J^\mu = 0$, we may choose a gauge with $A^0 = 0$ such that the Lorentz condition takes the form $\nabla \cdot \boldsymbol{A} = 0$. This special case is known as the *Coulomb gauge*.

Note that by selecting $A^0 = 0$, we restrict ourselves to one particular inertial system, and thus break so-called *Lorentz covariance* (which is the property that the equations are written in a form which is valid in any inertial system). Alternatively, we must perform a gauge transformation along with every Lorentz transformation in order to keep the Coulomb gauge intact in a new inertial system.

In QED, A^μ can be thought of as the "wavefunction" of the photon. For a free photon ($J^\mu = 0$), the equation of motion for this wavefunction in our chosen gauge takes the simple form: $\Box A^\mu = 0$. Closer inspection reveals that this is in fact equivalent to the KG equation for a massless field ($m = 0$). It has plane-wave solutions of the type:

$$A^\mu(x) = a e^{-ip\cdot x/\hbar} \epsilon^\mu(p), \tag{4.59}$$

where ϵ^μ is the polarization vector characterizing the *spin* of the photon. Inserting this solution into $\Box A^\mu = 0$ [which is valid only under the Lorentz condition gauge], we find $p^\mu p_\mu = 0$. This means that indeed $m = 0$ and $E = |p|c$. Now, the Lorentz condition dictates that $p^\mu \epsilon_\mu = 0$. In the Coulomb gauge, we further have that

$$\epsilon^0, \ \boldsymbol{\epsilon} \cdot \boldsymbol{p} = 0. \tag{4.60}$$

Therefore, the polarization vector $\boldsymbol{\epsilon}$ is $\perp$ to $\boldsymbol{p}$: a free photon is transversely polarized in the Coulomb gauge. For a given $\boldsymbol{p}$, there are thus two linearly independent 3-vectors perpendicular to $\boldsymbol{p}$, for instance $\boldsymbol{\epsilon}_{(1)} = (1, 0, 0)$ and $\boldsymbol{\epsilon}_{(2)} = (0, 1, 0)$ for $\boldsymbol{p} \propto (0, 0, 1)$.

It might seem puzzling that there are only two polarization states since the photon is a spin-1 particle ($s = 1$). Should there not be three available spin states? We might expect this based on how *massive* particles of spin s behave: in this case, we are used to there being $2s + 1$ different spin orientations. However, this changes for *massless* particles of spin s. For such particles, there are only 2 different spin orientations regardless of s except for $s = 0$ in which case there is only one spin orientation. This fact is related to the fact that the photon has no rest frame as it moves with the speed of light.

With our gauge-choice, we explicitly eliminated the $m_s = 0$ solution. However, the same physics would transpire if we didn't specify the gauge. In that case, longitudinal "ghost" photons decoupled from everything else would appear (which would thus not be of any physical consequence since they would not interact with anything).

C. Feynman calculus: application to decays and scattering

In order to develop a quantitative formulation of elementary particle dynamics, such as decay rates Γ and scattering cross sections σ, we need two main ingredients:

- Evaluate Feynman diagrams to find the amplitude $\mathcal{M}$ for a given process.

- Use Fermi's Golden rule to compute Γ or σ from $\mathcal{M}$.

Decay rate and scattering cross section.
The *lifetime* of a particle implicitly refers to a particle at rest, since otherwise we always have to take into account time dilation. Even so, we can only compute the average lifetime of particles from a large sample. The decay rate Γ is the probability per unit time that a muon will disintegrate. For a large sample $N(t)$, we have

$$dN = -\Gamma N dt \rightarrow N(t) = N_0 e^{-\Gamma t}. \tag{4.61}$$

The mean lifetime is then defined as $\tau = \Gamma^{-1}$. If a particle has several decay routes Γ_i (for instance, a π^+ can decay to $\mu^+ + \nu_\mu, e^+ + \nu_e + \gamma$, et.c.), the total decay rate is

$$\Gamma_{\text{tot}} = \sum_{i=1}^n \Gamma_i \tag{4.62}$$

The branching ratio is defined as $\Gamma_i/\Gamma_{\text{tot}}$.

Concerning scattering, it gives information about how particles quantitatively interact with each other. Imagine firing a stream of arrows against a target as an analogue of particles scattering.

- Unlike a physical object, the arrows (particles in our case) do not simply miss or hit. Instead, their deflection depends on the distance to the target. In this way, the effective scattering cross section $\sigma_{\text{eff}} \neq A$ where A is the physical area of the object in general.

- The scattering cross section σ also depends on the arrows themselves. Particles scatter differently due to different interactions, such as electromagnetic or weak ones.

- Finally, there is also a dependence on the final state of the arrows, such as whether or not the scattering is elastic or inelastic.

Each process has a specific scattering cross section. It depends on the energy of the particles as well: so-called resonances in the $\sigma - E$ diagram indicate bound states such as short-lived particles.
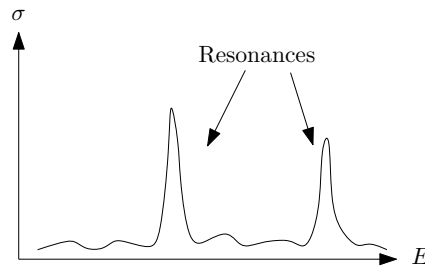
The scattering cross section is a measure for how likely a particle reaction is to occur. In fact, we can measure the particle scattering cross section σ for a given reaction via the rate W at which it occurs. This is because $W \propto \sigma$. When there are similar routes to a final end-product one performs (similarly to the decay rate Γ):

$$\sigma_{\text{tot}} = \sum_{i=1}^{n} \sigma_i \tag{4.63}$$

where n is the total number of processes leading to the end-product. For a particle scattering off some potential center, we may envision the situation as follows:

where b is the impact parameter and θ is the scattering angle. The reader is assumed to be familiar with the classical treatment of a scattering cross section and we will not give a detailed treatment of classical scattering theory - if required, the reader can find a comprehensive treatment here. We denote the differential scattering cross section as $d\sigma/d\Omega$.
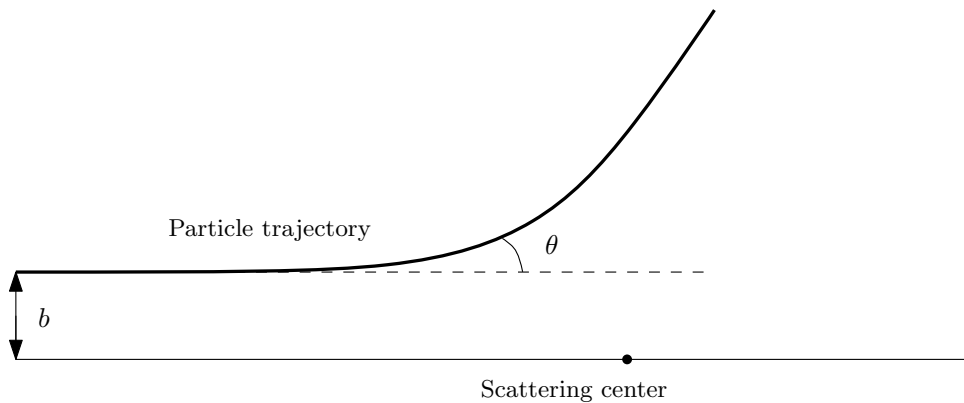
The Golden rule.
The basic ingredients for calculating decay rates and scattering cross sections quantitatively will be the amplitude $\mathcal{M}$ of the Feynman diagram for a given process (which contains information about how the interactions give rise to the process, how they depend on energy and momentum, and so forth) and the available phase-space. More phase-space indicates a more likely process, in general.

Example 6. Heavy particle decay. When a particle with a large mass decays into several light secondary particles, a large phase-space is involved. This means that there are many ways to distribute the available energy and causes (generally, but there are exceptions) heavy particles to decay faster than light particles. In contrast, a process such as $n \to p + e^- + \bar{\nu}_e$ has very little extra mass to spare since m_n and m_p are almost equal, and thus very small phase-space. As a result, the neutron has a very long lifetime (compared to other particles that decay via weak interactions).

Before we work out a quantitative theory concerning decays and scattering cross sections, we state Fermi's Golden rule in a qualitative form to promote the physical understanding:

$$\text{Transition rate} = \frac{2\pi}{\hbar}|\mathcal{M}|^2 \times \text{phase-space.}$$

We will now look at two particular cases of interest where we quantify what phase-space means mathematically.

Golden rule for decays.
The decay rate for a process $1 \rightarrow 2 + 3 + 4 + \ldots + n$ (which is thus a very general process) is given by

$$d\Gamma = |\mathcal{M}|^2 \frac{\mathcal{S}}{2\hbar m_1} \Big[\Big(\frac{cd^3p_2}{(2\pi)^3 2E_2} \Big) \Big(\frac{cd^3p_3}{(2\pi)^3 2E_3} \Big) \cdots \Big(\frac{cd^3p_n}{(2\pi)^3 2E_n} \Big) \Big] \times (2\pi)^4 \delta^4(p_1 - p_2 - p_3 - \ldots p_n).$$

It is assumed here that particle 1 is at rest, so that $p_1 = (m_1c, \mathbf{0})$. In general, $p_i = (E_i/c, \mathbf{p}_i)$ is the 4-momentum for particle i with mass m_i ($E_i^2 - \mathbf{p}_i^2 c^2 = m_i^2 c^4$). The detailed derivation of the expression for $d\Gamma$ is beyond the scope of this textbook, but we shall in fact later outline briefly how it is derived using quantum field theory. The δ-function enforces conservation of E and $\mathbf{p}$. $\mathcal{S}$ is a product of statistical factors: it contains a factor $1/j!$ for each group of j identical particles in the final state. The expression $d\Gamma$ is to be understood as the differential rate of decay where the momentum of particle 2 ends up in the range d^3p_2 around $\mathbf{p}_2$, and so forth for the remaining of the produced particles. The total decay rate is obtained via integration. For instance, with two particles in the final state:

$$\Gamma = \frac{\mathcal{S}}{\hbar m_1} \Big(\frac{c}{4\pi} \Big)^2 \frac{1}{2} \int \frac{|\mathcal{M}|^2}{E_2 E_3} \delta^4(p_1 - p_2 - p_3) d^3p_2 d^3p_3. \tag{4.64}$$

In general, the amplitude will depend on the momenta: $\mathcal{M} = \mathcal{M}(\mathbf{p}_2, \mathbf{p}_3)$.
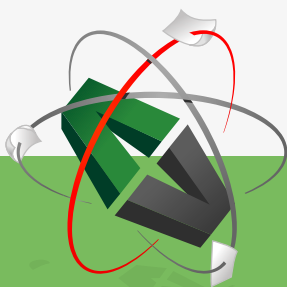
Example 7. Two-body decay. Consider the process $m_1 \to m_2 + m_3$. Assume that $\mathcal{M}$ is a known quantity and compute Γ. Since $E_j = \sqrt{m_j^2 c^4 + \boldsymbol{p}_j^2 c^2}$, we have

$$\Gamma = \frac{\mathcal{S}}{2(4\pi)^2 \hbar m_1} \int \frac{|\mathcal{M}|^2 \delta(m_1 c - \sqrt{m_2^2 c^2 \boldsymbol{p}_2^2} - \sqrt{m_3^2 c^2 + \boldsymbol{p}_2^2})}{\sqrt{m_2^2 c^2 \boldsymbol{p}_2^2}\sqrt{m_3^2 c^2 \boldsymbol{p}_3^2}} d^3 p_2. \tag{4.65}$$

We have carried out the integral over $d^3 p_3$ by noting it only gives a contribution for $\boldsymbol{p}_2 = -\boldsymbol{p}_3$:

$$\delta^4(p_1 - p_2 - p_3) = \delta(m_1 c - E_2/c - E_3/c)\delta^3(\boldsymbol{0} - \boldsymbol{p}_2 - \boldsymbol{p}_3). \tag{4.66}$$

Now, $|\mathcal{M}|^2$ only depends on $\boldsymbol{p}_2$ since we have set $\boldsymbol{p}_3 = -\boldsymbol{p}_2$. In fact, it should only depend on $|\boldsymbol{p}_2|$ since it is a scalar (only the combination $\boldsymbol{p}_2 \cdot \boldsymbol{p}_2$ represents a scalar). Introducing spherical coordinates and performing the angular integration gives:

$$\Gamma = \frac{\mathcal{S}}{8\pi \hbar m_1} \int_0^\infty \frac{|\mathcal{M}|^2 \delta(m_1 c - \sqrt{m_2^2 c^2 \boldsymbol{p}_2^2} - \sqrt{m_3^2 c^2 + \boldsymbol{p}_2^2})}{\sqrt{m_2^2 c^2 \boldsymbol{p}_2^2}\sqrt{m_3^2 c^2 \boldsymbol{p}_3^2}} \rho^2 d\rho. \tag{4.67}$$

Here, we defined $\rho \equiv |\boldsymbol{p}_2|$. To solve this integral, introduce the total energy of the final particles:

$$E \equiv c\left[\sqrt{m_2^2 c^2 + \rho^2} + \sqrt{m_3^2 c^2 + \rho^2}\right]. \tag{4.68}$$

This yields:

$$\Gamma = \frac{\mathcal{S}}{8\pi \hbar m_1} \int_{(m_2 + m_3)c^2}^\infty |\mathcal{M}|^2 \frac{\rho}{E} \delta(m_1 c - E/c) dE. \tag{4.69}$$

Finally, using that $\delta(m_1 c - E/c) = c\delta(E - m_1 c^2)$, we obtain

$$\Gamma = \mathcal{S}|\mathcal{M}|^2 \rho_0 / 8\pi \hbar m_1^2 c \tag{4.70}$$

where ρ_0 is the value of ρ which gives $E = m_1 c^2$. Explicitly, it reads:

$$\rho_0 = c\sqrt{m_1^4 + m_2^4 + m_3^4 - 2m_1^2 m_2^2 - 2m_1^2 m_3^2 - 2m_2^2 m_3^2}/2m_1. \tag{4.71}$$

Also, $|\mathcal{M}|$ is evaluated at the momentum consistent with energy conservation and momentum conservation. Note that $m_1 > m_2 + m_3$ in order to perform the integration. This makes physical sense, since otherwise particle 1 does not have enough rest energy to produce the masses of particles 2 and 3, let alone provide them with any kinetic energy.

Example 8. Decay rate of $\pi^0 \to \gamma + \gamma$. Use the formula derived in the previous example and set $m_2 = m_3 = 0$, leading to $\rho_0 = m_1 c/2$. Moreover, we have to set $\mathcal{S} = 1/2! = 1/2$ since we are producing two identical particles. Note that both of these examples were done without knowing the explicit form of $\mathcal{M}$. This is not possible for decays into three particles.

Golden rule for scattering
Suppose that 1+2 collide and produce a number of particles via the process $1 + 2 \to 3 + 4 + \ldots + n$. The scattering cross section is then:

$$d\sigma = |\mathcal{M}|^2 \frac{\hbar^2 S}{4\sqrt{(p_1 p_2)^2 - (m_1 m_2 c^2)^2}} \left[\left(\frac{c d^3 p_3}{(2\pi)^3 2 E_3} \right) \cdots \left(\frac{c d^3 p_n}{(2\pi)^3 2 E_n} \right) \right] \times (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - \ldots p_n).$$

Typically, one studies at which angle one of the final particles emerge (say, particle 3): the integration required to find the cross section for this process is then performed over all the other momenta $(\boldsymbol{p}_4, \boldsymbol{p}_5, \ldots, \boldsymbol{p}_n)$ and the magnitude of $\boldsymbol{p}_3$. We are then left with the differential scattering cross section $d\sigma/d\Omega$ for scattering particle 3 into solid angle $d\Omega$.

Example 9. Scattering $1 + 2 \to 3 + 4$ **in the CM frame.** Assume that $\mathcal{M}$ is known and calculate $d\sigma/d\Omega$.
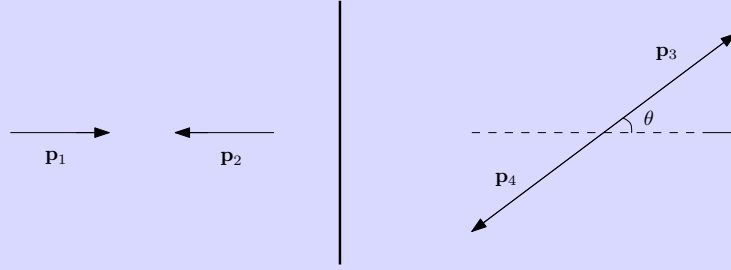
In the CM frame, we have $\boldsymbol{p}_1 = -\boldsymbol{p}_2$ which leads to $p_1 \cdot p_2 = E_1 E_2/c^2 + \boldsymbol{p}_1^2$. Therefore:

$$d\sigma = \left(\frac{\hbar c}{8\pi} \right)^2 \frac{S|\mathcal{M}|^2 c}{(E_1 + E_2)|\boldsymbol{p}_1|} \frac{d^3 p_3 d^3 p_4}{E_3 E_4} \delta[(E_1 + E_2 - E_3 - E_4)/c] \times \delta^3(-\boldsymbol{p}_3 - \boldsymbol{p}_4). \tag{4.72}$$

Now, express the energies with $\boldsymbol{p}_3$ and $\boldsymbol{p}_4$ and perform $\int d^3 p_4$ so that $\boldsymbol{p}_4 = -\boldsymbol{p}_3$ everywhere:

$$d\sigma = \left(\frac{\hbar}{8\pi} \right)^2 \frac{S|\mathcal{M}|^2 c}{(E_1 + E_2)|\boldsymbol{p}_1|} \frac{\delta[(E_1 + E_2)/c - \sqrt{m_3^2 c^2 + \boldsymbol{p}_3^2} - \sqrt{m_4^2 c^2 + \boldsymbol{p}_3^2}]}{\sqrt{m_3^2 c^2 + \boldsymbol{p}_3^2}\sqrt{m_4^2 c^2 + \boldsymbol{p}_3^2}} d^3 p_3. \tag{4.73}$$

This time, $|\mathcal{M}|^2$ can in general depend not only on $|\boldsymbol{p}_3|$, but also its direction since it depends on all momenta: due to the assumption of working in the CM system, it depends on $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$, so that the direction may play a role from terms such as $\boldsymbol{p}_1 \cdot \boldsymbol{p}_3$. We cannot carry out the integration over $d\Omega$ unless $|\mathcal{M}|^2$ is specified, but *we can integrate over* $|\boldsymbol{p}_3|$:

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar}{8\pi} \right)^2 \frac{S c}{(E_1 + E_2)|\boldsymbol{p}_1|} \frac{|\mathcal{M}|^2 \delta[(E_1 + E_2)/c - \sqrt{m_3^2 c^2 + \boldsymbol{p}_3^2} - \sqrt{m_4^2 c^2 + \boldsymbol{p}_3^2}]}{\sqrt{m_3^2 c^2 + \boldsymbol{p}_3^2}\sqrt{m_4^2 c^2 + \boldsymbol{p}_3^2}} \rho^2 d\rho, \tag{4.74}$$

where we introduced $\rho \equiv |\boldsymbol{p}_3|$. This is the same integral as in our previous example of a decay if we perform the substitutions $m_2 \to m_4$ and $m_1 \to (E_1 + E_2)/c^2$. The result is:

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar}{8\pi} \right)^2 \frac{S|\mathcal{M}|^2 c}{(E_1 + E_2)^2} \frac{|\boldsymbol{p}_f|}{|\boldsymbol{p}_i|}. \tag{4.75}$$

Here, $|\boldsymbol{p}_f|$ is the magnitude of either final momentum (does not matter since we work in the CM frame) and $|\boldsymbol{p}_i|$ is the incoming momentum. As before, it is implicitly understood that $|\mathcal{M}|^2$ in the final expression is evaluated at these momenta.

Let us also say a few words about the dimensions of the quantities we have considered. We have $[\Gamma] = \mathrm{s}^{-1}$ and $[\sigma] = \mathrm{m}^2$, although in practice the cross section is more commonly given in cm^2 or even barn where 1 barn $= 10^{-24} \, \mathrm{cm}^2$. The dimension of the Feynman amplitude is $[\mathcal{M}] = (mc)^{4-n}$, in effect momentum raised to the power $4 - n$ where n is the number of incoming plus outgoing particles.

How does one derive the formula for dΓ and dσ?

We here outline how the formulae we have given for $d\Gamma$ and $d\sigma$ are derived, following the 1988 textbook by Mandl and Shaw. Both decay and scattering have in common that one starts with an initial state $|i\rangle$ and ends up in a final state $|f\rangle$. The rate or likelihood of this process is determined by matrix elements of the type $|\langle f|\Phi(\infty)\rangle|^2$ where $|\Phi(\infty)\rangle$ is the state into which $|i\rangle$ evolves as $t \to \infty$. It includes all possible states that $|i\rangle$ can evolve into, and the matrix element thus measures the overlap with one specific final state $|f\rangle$ that we are interested in. Now, to find $|\Phi(\infty)\rangle$ we use that $|\Phi(\infty)\rangle = S|i\rangle$ where S is the S-matrix determining the time-evolution of $|i\rangle$. S may be obtained via the time-dependent standard QM equation of motion (Schrödinger equation):

$$i\frac{d|\Phi(t)\rangle}{dt} = H_I(t)|\Phi(t)\rangle \tag{4.76}$$

in an iterative fashion. Here, $H_I(t)$ contains the interactions that govern the decay/scattering process. One finds that:

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2 \dots \mathcal{T}\{H_I(t_1)H_I(t_2)\dots H_I(t_n)\}, \tag{4.77}$$

where $\mathcal{T}$ is the time-ordering operator. Essentially, we thus need to compute $S_{fi} = \langle f|S|i\rangle$ for the specified states $|i\rangle$ and $|f\rangle$ for a given interaction H_I. When computing this matrix element, one typically ends up with a result of the type:

$$S_{fi} = \delta_{fi} + (2\pi)^4\delta^4\left(\sum p_f - \sum p_i\right)\prod_i \left(\frac{1}{2VE_i}\right)^{1/2} \times \prod_f \left(\frac{1}{2VE_f}\right)^{1/2}C\mathcal{M} \tag{4.78}$$

where C is a constant and $\mathcal{M}$ contains info about external lines, i.e. the states $|i\rangle$ and $|f\rangle$, and the relevant interactions H_I. After this, we just have to multiply with an appropriate factor for the number of states available [typically $\prod_f \frac{V d^3 p_f}{(2\pi)^3}$], and we are left with $d\Gamma$ or $d\sigma$.

Feynman rules for QED

Turning now to QED specifically, let us preface our discussion of the Feynman rules with a summary of the properties for electron and positron wavefunctions and there properties. By electron and positron, we really mean particle and antiparticle in general, but for concreteness we focus on those particular cases. The wavefunctions for free electrons and positrons of momentum $p = (E/c, \boldsymbol{p})$ and $E = \sqrt{m^2c^4 + \boldsymbol{p}^2c^2}$ have the following form.

Electrons: $\psi(x) = a\mathrm{e}^{-\mathrm{i}px/\hbar}u^{(s)}(p)$ where $s = 1, 2$ denotes two spin states. The spinor satisfies $(\gamma^\mu p_\mu - mc)u = 0$ and its adjoint $\bar{u} = u^\dagger\gamma^0$ satisfies $\bar{u}(\gamma^\mu p_\mu - mc) = 0$. Moreover, the spinors are orthogonal, normalized, and complete:

$$\bar{u}^{(1)}u^{(2)} = 0, \ \bar{u}u = 2mc, \ \sum_{s=1,2} u^{(s)}\bar{u}^{(s)} = (\gamma^\mu p_\mu + mc). \tag{4.79}$$

Positrons: $\psi(x) = a\mathrm{e}^{\mathrm{i}px/\hbar}v^{(s)}(p)$ where $s = 1, 2$ denotes two spin states. The spinor satisfies $(\gamma^\mu p_\mu + mc)v = 0$ and its adjoint $\bar{v} = v^\dagger\gamma^0$ satisfies $\bar{v}(\gamma^\mu p_\mu + mc) = 0$. Moreover, the spinors are orthogonal, normalized, and complete:

$$\bar{v}^{(1)}v^{(2)} = 0, \ \bar{v}v = -2mc, \ \sum_{s=1,2} v^{(s)}\bar{v}^{(s)} = (\gamma^\mu p_\mu - mc). \tag{4.80}$$

The completeness relations are important as one normally averages over electron and positron spins (since these are often random in experiments), so that one requires the complete set. Note that all of the above relations are equally valid for *e.g.* μ^- and μ^+, quarks and antiquarks, and so forth (since all are spin-1/2 particles).

Photons: The wavefunction reads $A^\mu(x) = a\mathrm{e}^{-\mathrm{i}px/\hbar}\epsilon^\mu_{(s)}$ with $s = 1, 2$ being the two spin polarization states. The polarization vectors satisfy $\epsilon^\mu p_\mu = 0$ (Lorentz condition) and are orthogonal and normalized:

$$(\epsilon^\mu_{(1)})^*\epsilon_{\mu(2)} = 0, \ (\epsilon^\mu_{(i)})^*\epsilon_{\mu(i)} = 1. \tag{4.81}$$

In the Coulomb gauge, $\epsilon^0 = 0$ and $\epsilon \cdot \boldsymbol{p} = 0$, and one has

$$\sum_{s=1,2} (\epsilon_{(s)})_i(\epsilon^*_{(s)})_j = \delta_{ij} - \hat{p}_i\hat{p}_j \tag{4.82}$$

for the 3-vectors. An explicit pair is, as previously mentioned, $\boldsymbol{\epsilon}_{(1)} = (1, 0, 0)$ and $\boldsymbol{\epsilon}_{(2)} = (0, 1, 0)$.

With the above notation in mind, we now explain *how to compute $\mathcal{M}$ for a given Feynman diagram*:

1. **Notation**. Label the incoming and outgoing momenta and spin with $p_1, p_2, \ldots$ and $s_1, s_2, \ldots$ Internal momenta are labelled q_n. Arrows on external lines indicate if it is an e^- or e^+, while arrows on internal lines are conventionally assigned so that the direction of flow is preserved (one arrow in, one arrow out). External photons have arrows that point forward:

2. **External lines**.
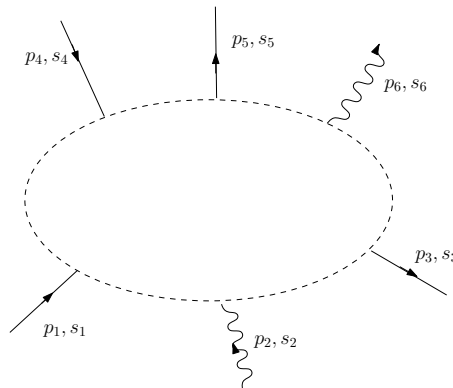
	Electrons		Positrons		Photons	
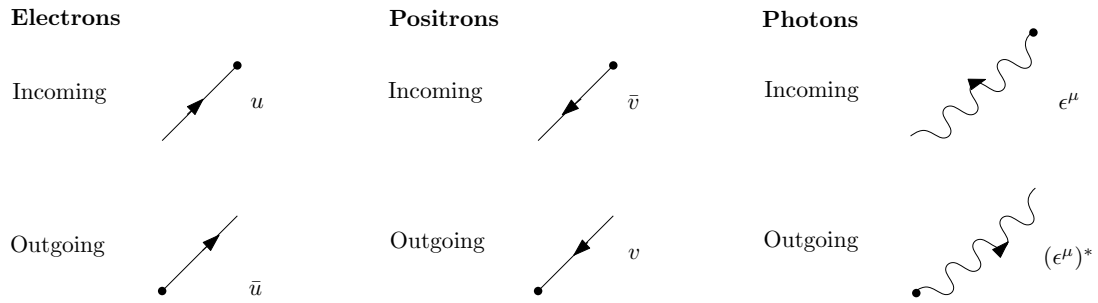
3. **Vertex factors**. Each vertex contributes $ig_e\gamma^\mu$ where $g_e = \sqrt{4\pi\alpha}$ is a dimensionless coupling constant. For quarks where $|q| \neq e$, one should use $g = -q\sqrt{4\pi/\hbar c}$.

4. **Propagators.** Internal lines contribute with

$$\frac{i(\gamma^\mu q_\mu + mc)}{q^2 - m^2 c^2} \tag{4.83}$$

for electrons and positrons, while photons contribute with

$$-\frac{ig_{\mu\nu}}{q^2}. \tag{4.84}$$

5. **Conservation of E and p.** For each vertex, write $(2\pi)^4 \delta^4(k_1 + k_2 + k_3)$. Positive k for arrow into the vertex and negative k for arrow out of the vertex, *except* for external positrons.

6. **Integrate over internal momenta**. For each internal q, write $\int d^4q/(2\pi)^4$.

7. **Cancel the δ-function**. Remove $(2\pi)^4 \delta^4(p_1 + p_2 + \ldots - p_n)$, and what remains is $-i\mathcal{M}$. It is worth remarking that if one chooses $C = +i$ in Eq. (4.78), one ends up with $+i\mathcal{M}$ which is a commonly used convention. However, the overall sign does not matter since it is $|\mathcal{M}|^2$ that is used to compute quantities such as decays and scattering cross sections. We shall stick with the convention that one ends up with $-i\mathcal{M}$ after cancelling the δ-function.

The total amplitude $\mathcal{M}$ is then the sum of all the amplitudes for each contributing diagram while taking into account:

Antisymmetrization. Include a relative minus sign between diagrams that differ only by either (1) interchanging two incoming (or outgoing) electrons or positrons (2) or that differ by exchanging an incoming electron with an outgoing positron or vice versa. This rule is taken into account in order to incorporate the Pauli principle for fermionic wavefunctions.

Later, we will comment explicitly on how to deal with *fermion loops* which are of relevance in so-called vacuum polarization diagrams. To test our skills in using the Feynman rules, we now work out $\mathcal{M}$ in detail for some of these processes.

Example 10. $e - \mu$ **scattering.** The lowest-order diagram looks as follows.
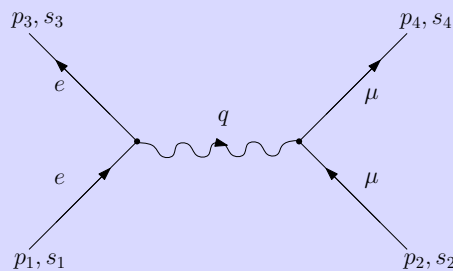
Using the Feynman rules, we obtain the following complex scalar:

$$(2\pi)^4 \int [\bar{u}^{(s_1)}(p_3)(\mathrm{i}g_e\gamma^\mu)u^{(s_1)}(p_1)]\left(-\frac{\mathrm{i}g_{\mu\nu}}{q^2}\right)[\bar{u}^{(s_4)}(p_4)\mathrm{i}g_e\gamma^\nu u^{(s_2)}(p_2)]\delta^4(p_1-p_3-q)\delta^4(p_2+q-p_4)d^4q.$$

(4.85)

After performing the q-integration and dropping the δ-function, we are left with

$$\mathcal{M} = -\frac{g_e^2}{(p_1-p_3)^2}[\bar{u}^{(s_3)}(p_3)\gamma^\mu u^{(s_1)}(p_1)][\bar{u}^{(s_4)}(p_4)\gamma_\mu u^{(s_2)}(p_2)].$$

(4.86)

We emphasize again that $\mathcal{M}$ is a *number*.

Example 11. $e - e$ **scattering.** The diagram is similar to the above example, but there are now two contributing diagrams since electrons are identical particles and cannot be distinguished. Therefore, we have to account for an additional diagram where $(p_3, s_3) \leftrightarrow (p_4, s_4)$ since we cannot tell which electron that is coming out where. As a result, the total amplitude is obtained by adding the two contributions with a relative minus sign due to the aforementioned antisymmetrization rule:

$$\mathcal{M} = -\frac{g_e^2}{(p_1-p_3)^2}[\bar{u}(3)\gamma^\mu u(1)][\bar{u}(4)\gamma_\mu u(2)] + \frac{g_e^2}{(p_1-p_4)^2}[\bar{u}(4)\gamma^\mu u(1)][\bar{u}(3)\gamma_\mu u(2)].$$

(4.87)

We were able to simply write down this $\mathcal{M}$ by using our result from the $e - \mu$ scattering example. For brevity, we introduced the notation $u(i) \equiv u^{(s_i)}(p_i)$.

Example 12. $e^- - e^+$ **scattering.** Again, there are two lowest-order contributions to this process:
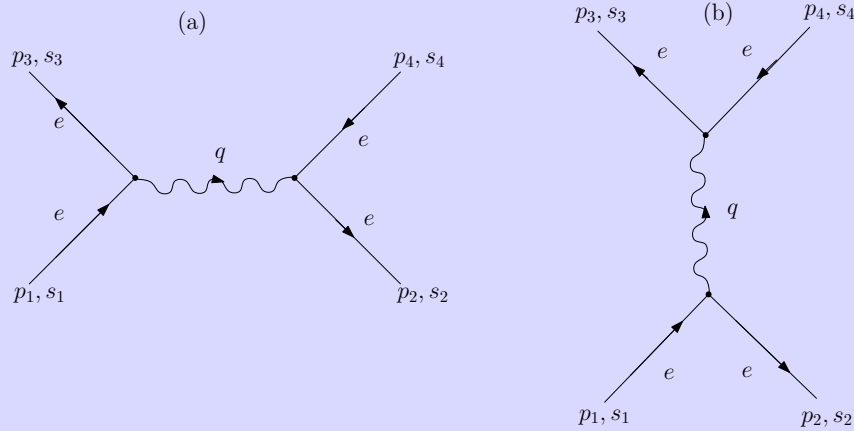
The first contribution [(a)] is evaluated, according to the Feynman rules, to the following complex scalar:

$$(2\pi)^4 \int [\bar{u}(3)(ig_e\gamma^\mu)u(1)]\left(-\frac{ig_{\mu\nu}}{q^2}\right)[\bar{v}(2)(ig_e\gamma^\nu)v(4)]\delta^4(p_1-p_3-q)\delta^4(p_2+q-p_4)d^4q. \tag{4.88}$$

Note how the order is always adjoint spinor/gamma matrix/spinor when moving along a fermion line, in order for the dimensionality of the factors to match. Moving backwards (along the arrow) of an antiparticle line is equivalent to moving forward in time. The second contribution [(b)] provides:

$$(2\pi)^4 \int [\bar{u}(3)(ig_e\gamma^\mu)v(4)]\left(-\frac{ig_{\mu\nu}}{q^2}\right)[\bar{v}(2)(ig_e\gamma^\nu)u(1)]\delta^4(q-p_3-p_4)\delta^4(p_1+p_2-q)d^4q. \tag{4.89}$$

By interchanging the incoming e^+ and outgoing e^-, we obtain
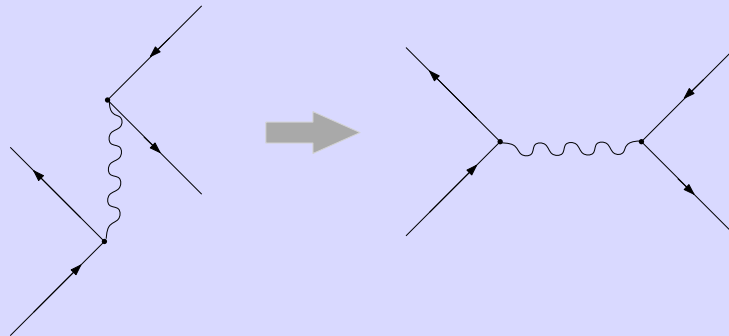
which is equivalent to the first contributing diagram. Therefore, due to the rule of antisymmetrization, the two amplitudes must add with a relative minus sign.

D. Casimir's trick and trace theorems

If all spins and polarizations are known *a priori* in the experiment, the appropriate spinors and polarization vectors are simply inserted into the expression for $\mathcal{M}$. However, it is more often the case that the spins are not known, but random. Then, the cross section is given by the average over all initial spin configurations and the sum over all final spin configurations. Thus, $\langle|\mathcal{M}|^2\rangle \equiv$ average over initial spins and sum over final spins. Let us introduce some convenient notation known as Feynman *slash* notation: $\not{a} \equiv a^\mu\gamma_\mu$, $\not{a}^* \equiv \gamma^\mu a_\mu^*$, and $\bar{\Gamma} \equiv \gamma^0\Gamma^\dagger\gamma^0$ for a

matrix Γ. Consider now $e - \mu$ scattering. From our previous result for $\mathcal{M}$ for this process, we find:

$$|\mathcal{M}|^2 = \frac{g_e^4}{(p_1 - p_3)^4}[\bar{u}(3)\gamma^\mu u(1)][\bar{u}(4)\gamma_\mu u(2)][\bar{u}(3)\gamma^\nu u(1)]^*[\bar{u}(4)\gamma_\nu u(2)]^*. \tag{4.90}$$

This mathematical object is built up from the schematic structure: $G \equiv [\bar{u}(a)\Gamma_1 u(b)][\bar{u}(a)\Gamma_2 u(b)]^*$, where $\Gamma_{1,2}$ are 4×4 matrices. For a scalar (1×1 matrix) the operation * is equivalent to †, which means that

$$[\bar{u}(a)\Gamma_2 u(b)]^* = [u^\dagger(a)\gamma^0 \Gamma_2 u(b)]^\dagger = u^\dagger(b)\Gamma_2^\dagger \gamma^{0\dagger} u(a) = \bar{u}(b)\bar{\Gamma}_2 u(a) \tag{4.91}$$

where we used that $\gamma^{0\dagger} = \gamma^0$ and $(\gamma^0)^2 = 1$. We now perform a summation over the spin orientations of particle b:

$$\sum_{b \text{ spins}} G = \bar{u}(a)\Gamma_1 \left\{ \sum_{s_b=1,2} u^{(s_b)}(p_b)\bar{u}^{(s_b)}(p_b) \right\} \bar{\Gamma}_2 u(a) = \bar{u}(a)\Gamma_1(\not{p}_b + m_b)\bar{\Gamma}_2 u(a) = \bar{u}(a)Qu(a), \tag{4.92}$$

where $Q \equiv \Gamma_1(\not{p}_b + m_b c)\bar{\Gamma}_2$. Similarly, for a summation over the a spins we get:

$$\sum_{a \text{ spins}} \sum_{b \text{ spins}} \bar{u}_i^{(s_a)} Q_{ij} u_j^{(s_a)} = Q_{ij} \left\{ \sum_{s_a=1,2} u^{(s_a)}\bar{u}^{(s_a)} \right\}_{ji} = Q_{ij}(\not{p}_a + m_a c)_{ji} = \text{Tr}\{Q(\not{p}_a + m_a c)\}. \tag{4.93}$$

Here, $\text{Tr}(A) = \sum_i A_{ii}$. We also used that u and $\bar{u}$ are 4×1 and 1×4 spinors, respectively. In total, we thus have

$$\sum_{\text{all spins}} [\bar{u}(a)\Gamma_1 u(b)][\bar{u}(a)\Gamma_1 u(b)]^* = \text{Tr}\{\Gamma_1(\not{p}_b + m_b c)\bar{\Gamma}_2(\not{p}_a + m_a c)\}. \tag{4.94}$$

We see that there are no spinors left after the summation over spins: only matrix multiplication and a trace at the end. This procedure is known as *Casimir's trick*. Note that if u is replaced with v, the corresponding mass should change sign.

For $e - \mu$ scattering, we then have $\Gamma_2 = \gamma^\nu$ and $\bar{\Gamma}_2 = \gamma^0\gamma^{\nu\dagger}\gamma^0 = \gamma^\nu$. We find:

$$\langle|\mathcal{M}|^2\rangle = \frac{g_4^4}{4(p_1 - p_3)^4}\text{Tr}\{\gamma^\mu(\not{p}_1 + mc)\gamma^\nu(\not{p}_3 + mc)\}\text{Tr}\{\gamma_\mu(\not{p}_2 + Mc)\gamma_n u(\not{p}_4 + Mc)\}, \tag{4.95}$$

where $m = m_e$, $M = m_\mu$, and $\frac{1}{4}$ is included to average over initial spins (2 particles with 2 configurations each = 4 configurations). To evaluate the traces appearing in $\langle|\mathcal{M}|^2\rangle$, there is a set of trace theorems that may be derived by using the fundamental mathematical properties of the Tr operation. Some useful relations, which we will draw upon in later calculations, are:

- $\text{Tr}\{A + B\} = \text{Tr}\{A\} + \text{Tr}\{B\}$.
- $\text{Tr}\{\alpha A\} = \alpha \text{Tr}\{A\}$.
- $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$.

- $\text{Tr}\{AB\} = \text{Tr}\{BA\}$.
- $g_{\mu\nu}g^{\mu\nu} = 4$.
- $\not{a}\not{b} + \not{b}\not{a} = 2ab$.

From this follows, for instance, that $\text{Tr}\{\gamma^\mu\gamma^\nu\} = 4g^{\mu\nu}$ and $\text{Tr}\{\not{a}\not{b}\} = 4ab$. Also, the trace over an odd number of γ-matrices is equal to zero (try to prove this yourself, by using that $\text{Tr}\{\gamma^5\} = 0$ and $\{\gamma^5, \gamma^\mu\} = 0$).

E. Treatment of bubble diagrams

In some Feynman diagrams, fermion loops (so-called bubbles) appear. The general rule for treating these is:

> Include a factor -1, take the Tr, and follow the fermion lines along the arrows.

Taking the Tr physically corresponds to including all possible spin orientations of the fermions as it connects with itself in the end, in the same spirit as the integration over all internal momenta.

Example 13. Evaluating a fermion bubble. Consider the following contribution to a Feynman diagram.
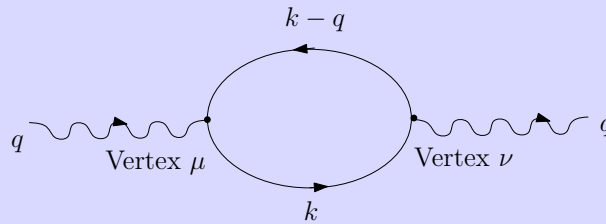
Using the rule stated above, we obtain

$$-\text{Tr} \int \frac{d^4k}{(2\pi)^4} \frac{\{\gamma_\mu(\not{k}+m)\gamma_\nu(\not{k}-\not{q}+m)\}}{[(k-q)^2-m^2c^2][k^2-m^2c^2]}. \tag{4.96}$$

However, is the direction in which we follow the arrows important? Could we instead follow the arrows in the opposite direction? The answer is that for a loop with two (in general, an *even*) number of fermions, we can. The proof goes as follows. Consider the trace $\text{Tr}\{\gamma_\mu\not{p}\gamma_\nu\not{q}\}$ obtained by traversing a fermion loop in a given direction. This can be written as:

$$\text{Tr}\{\gamma_\mu\not{p}\gamma_\nu\not{q}\} = p^\lambda q^\sigma \text{Tr}\{\gamma_\mu\gamma_\lambda\gamma_\nu\gamma_\sigma\} = 4p^\lambda p^\sigma (g_{\mu\lambda}g_{\nu\sigma} - g_{\mu\sigma}g_{\lambda\sigma} + g_{\mu\sigma}g_{\lambda\nu}) = 4(p_\mu q_\nu - g_{\mu\nu}p^\lambda q_\lambda + q_\mu p_\nu). \tag{4.97}$$

If we instead had traversed the fermion loop in the opposite direction, we would have obtained $\text{Tr}\{\gamma_\mu\not{q}\gamma_\nu\not{p}\}$. But Eq. (4.97) is invariant upon exchanging $p \leftrightarrow q$. Therefore, the two traces must be identical and the direction in which we move around the loop does not matter. This proof can then be generalized to any even number of fermions, where one makes use of the general formula $\text{Tr}\{\gamma_{\mu_1}\gamma_{\mu_2}\ldots\gamma_{\mu_n}\} = \text{Tr}\{\gamma_{\mu_n}\ldots\gamma_{\mu_2}\gamma_{\mu_1}\}$.

F. Computation of various cross sections and lifetimes

We are now in a position to explicitly evaluate $d\sigma/d\Omega$ for *e.g.* Mott and Rutherford scattering. If m_e scatters off a much heavier muon $M \gg m_e$, the task is to find $d\sigma/d\Omega$ in the lab frame where M is at rest when neglecting the recoil of M. Using the rules for cross-section related to $\mathcal{M}$ given previously, one finds

$$\frac{d\sigma}{d\Omega} = \left(\frac{\hbar}{8\pi Mc}\right)\langle|\mathcal{M}|^2\rangle. \tag{4.98}$$

Before   **After**

$E, \mathbf{p}_3$

θ

$E, \mathbf{p}_1$
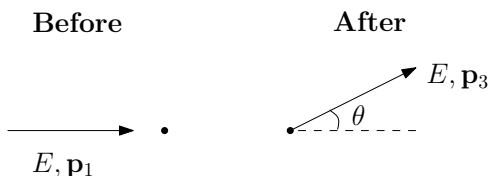
We now want to compute $\mathcal{M}$ explicitly. We have

$$p_1 = (E/c, \boldsymbol{p}_1), \ p_2 = (Mc, 0), \ p_3 = (E/c, \boldsymbol{p}_3), \ p_4 = (Mc, 0). \tag{4.99}$$

where $|\boldsymbol{p}_1| = |\boldsymbol{p}_3|$ due to our assumptions. The averaged amplitude is given by (using Casimir's trick):

$$\langle|\mathcal{M}|^2\rangle = \frac{8g_e^4}{(p_1 - p_3)^4}[(p_1p_2)(p_3p_4) + (p_1p_4)(p_2p_3) - (p_1p_3)M^2c^2 - (p_2p_4)m^2c^2 + 2(mMc^2)^2] \tag{4.100}$$

by evaluating the traces that arise. Inserting the above 4-momenta gives:

$$\langle|\mathcal{M}|^2\rangle = \left(\frac{g_e^2 Mc}{\boldsymbol{p}^2 \sin^2(\theta/2)}\right)^2 [(mc)^2 + \boldsymbol{p}^2 \cos^2(\theta/2)]. \tag{4.101}$$

Inserted into our expression for the differential scattering cross section, we obtain the Mott formula:

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{lab}} = \left(\frac{\alpha\hbar}{2\boldsymbol{p}^2 \sin^2(\theta/2)}\right)^2 [(mc)^2 + \boldsymbol{p}^2 \cos^2(\theta/2)]. \tag{4.102}$$

In the non-relativistic limit $\boldsymbol{p}^2 \ll (mc)^2$, one arrives at the Rutherford formula:

$$\frac{d\sigma}{d\Omega} = \left(\frac{e^2}{2mv^2 \sin^2(\theta/2)}\right)^2. \tag{4.103}$$

Whereas the above illustrates an example of how the differential scattering cross section may be evaluated fully, including the calculation of $\mathcal{M}$, we should also consider an example of a decay. But a moment of reflection reveals an interesting fact: *there are no decays in pure QED*. There is no mechanism that can convert *e.g.* μ^- to e^- since fermion lines cannot simply terminate in a diagram. If we allow for neutrino oscillations, such a process can become possible, but then we have introduced a type of interaction that is not contained in pure QED. On the other hand, annihilation events such as $e^- + e^+ \rightarrow \gamma + \gamma$ are fully possible in pure QED, but this is conventionally viewed as a scattering event rather than a decay.

As an example, let us in fact analyze this annihilation in the positronium rest frame (the CM-frame for the e^-e^+ pair). We assume that the bound-state is in the singlet spin configuration and at rest to begin with. There are two contributing diagrams, as seen in the figure.

p_3, s_3 p_4, s_4 p_4, s_4 p_3, s_3

q q

p_1, s_1 p_2, s_2 p_1, s_1 p_2, s_2
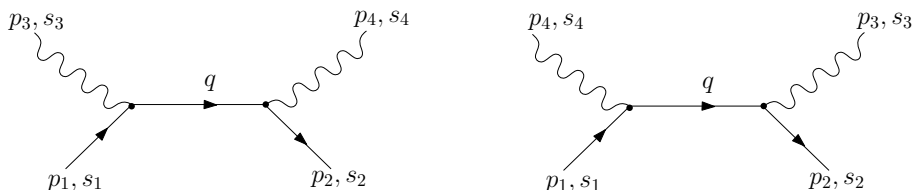
The corresponding amplitudes are

$$\mathcal{M}_1 = \frac{g_e^2}{(p_1 - p_3)^2 - m^2 c^2} \bar{v}(2) \not{\epsilon}_4 (\not{p}_1 - \not{p}_3 + mc) \not{\epsilon}_3 u(1), \quad \mathcal{M}_2 = \frac{g_e^2}{(p_1 - p_4)^2 - m^2 c^2} \bar{v}(2) \not{\epsilon}_3 (\not{p}_1 - \not{p}_4 + mc) \not{\epsilon}_4 u(1).$$

$$(4.104)$$

We have omitted the complex conjugation sign on the polarization vectors for brevity of notation - however, we will reinstate this in the final result. The total amplitude reads $\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2$ and according to our statements about the frame and starting point we have the following 4-vectors:

$$p_1 = mc(1,0,0,0), \; p_2 = mc(1,0,0,0), \; p_3 = mc(1,0,0,1), \; p_4 = mc(1,0,0,-1). \qquad (4.105)$$

It follows that $(p_1 - p_3)^2 - m^2 c^2 = (p_1 - p_4)^2 - m^2 c^2 = -2(mc)^2$. Now use the rule

$$\not{p}_1 \not{\epsilon}_3 + \not{\epsilon}_3 \not{p}_1 = 2 p_1 \epsilon_3 \qquad (4.106)$$

and invoke the Coulomb gauge $\epsilon^0 = 0$. Since $\boldsymbol{p}_1 = 0$, this gives $\not{p}_1 \not{\epsilon}_3 = -\not{\epsilon}_3 \not{p}_1$. Moreover, $\not{p}_3 \not{\epsilon}_3 = -\not{\epsilon}_3 \not{p}_3$ since $p_3 \epsilon_3 = 0$ due to the Lorentz condition. We finally also use that $(\not{p}_1 - mc)u(1) = 0$ [since $u(1)$ solves the Dirac equation] leads to

$$(\not{p}_1 - \not{p}_3 + mc) \not{\epsilon}_3 u(1) = \not{\epsilon}_3 \not{p}_3 u(1). \qquad (4.107)$$

Inserting all of this into $\mathcal{M}$ produces

$$\mathcal{M} = -\frac{g_e^2}{2(mc)^2} \bar{v}(2) [\not{\epsilon}_4 \not{\epsilon}_3 \not{p}_3 + \not{\epsilon}_3 \not{\epsilon}_4 \not{p}_4] u(1). \qquad (4.108)$$

Using the properties of the γ-matrices the contraction rules, we find that the above can be rewritten as

$$\mathcal{M} = \frac{g_e^2}{(mc)} \bar{v}(2) [\boldsymbol{\epsilon}_3 \cdot \boldsymbol{\epsilon}_4 \gamma^0 + \mathrm{i}(\boldsymbol{\epsilon}_3 \times \boldsymbol{\epsilon}_4) \cdot \boldsymbol{\Sigma} \gamma^3] u(1) \qquad (4.109)$$

where we introduced

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\sigma} & 0 \\ \underline{0} & \boldsymbol{\sigma} \end{bmatrix}. \qquad (4.110)$$

Now, take into account the spin-singlet symmetry so that the final amplitude in the singlet state reads:

$$\mathcal{M}_{\text{singlet}} = (\mathcal{M}_{\uparrow\downarrow} - \mathcal{M}_{\downarrow\uparrow})/\sqrt{2}. \qquad (4.111)$$

We have obtained an expression for $\mathcal{M}$ in Eq. (4.109), so to evaluate $\mathcal{M}_{\uparrow\downarrow}$ we simply have to use a spin-up spinor for the electron and spin-down spinor for the positron:

$$u(1) = \sqrt{2mc} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \; \bar{v}(2) = \sqrt{2mc}[0,0,1,0]. \qquad (4.112)$$

This gives $\bar{v}(2) \gamma^0 u(1) = 0$ and $\bar{v}(2) \boldsymbol{\Sigma} \gamma^3 u(1) = -2mc\hat{\boldsymbol{z}}$, so that

$$\mathcal{M}_{\uparrow\downarrow} = -2\mathrm{i} g e^2 (\boldsymbol{\epsilon}_3 \times \boldsymbol{\epsilon}_4)_z. \qquad (4.113)$$

The same procedure for the opposite spin configuration yields

$$\mathcal{M}_{\downarrow\uparrow} = 2\mathrm{i} g e^2 (\boldsymbol{\epsilon}_3 \times \boldsymbol{\epsilon}_4)_z = -\mathcal{M}_{\uparrow\downarrow}. \qquad (4.114)$$

This means that the total amplitude, according to Eq. (4.111) is:

$$\mathcal{M}_{\text{singlet}} = -2\sqrt{2}\mathrm{i} g_e^2 (\boldsymbol{\epsilon}_3 \times \boldsymbol{\epsilon}_4)_z \qquad (4.115)$$

for annihilation of a stationary $e^- e^+$ pair in a spin-singlet configuration into two photons which emerge in the directions $\pm \hat{\boldsymbol{z}}$.

Note that the amplitude for this process in a triplet configuration $(\uparrow\downarrow + \downarrow\uparrow)/\sqrt{2}$ gives 0, since $\mathcal{M}_{\uparrow\downarrow} = -\mathcal{M}_{\downarrow\uparrow}$. The reason is conservation of charge conjugation in EM processes which only allow 2γ, and not 3γ, from $e^- e^+$ in the singlet configuration. On the other hand, in a triplet configuration the 3γ production is allowed from a charge conjugation point of view.

G. Renormalization

We now consider specifically higher-order diagrams which require *renormalization*. Consider $e - \mu$ scattering.
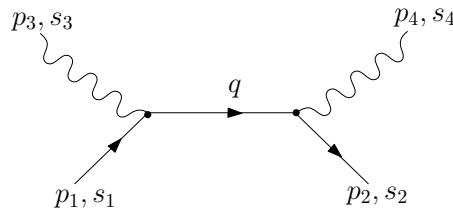
With $q = p_1 - p_3$, the amplitude for this diagram is:

$$\mathcal{M} = -g_e^2 [\bar{u}(p_3)\gamma^\mu u(p_1)]\frac{g_{\mu\nu}}{q^2}[\bar{u}(p_4)\gamma^\nu u(p_2)]. \tag{4.116}$$

A higher-order correction (4th order, to be specific) to this process is the vacuum-polarization diagram shown below.

The origin of this name is that the temporary e^-e^+ pair modifies the effective electric charge of the electron (or alternatively put, the effective coupling constant of the electron). One may derive the amplitude by using the extra Feynman rule for ferm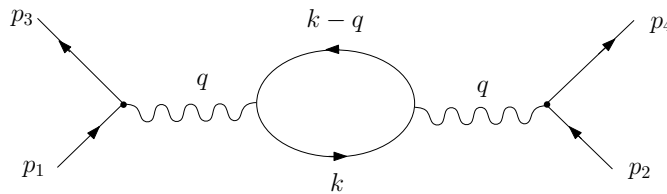ion loops that we treated earlier: for each closed fermion loop, include an overall factor -1 and take the trace. This gives:

$$\mathcal{M} = -\frac{\mathrm{i}g_e^4}{q^4}[\bar{u}(p_3)\gamma^\mu u(p_1)]\left\{\frac{d^4k}{(2\pi)^4}\frac{\mathrm{Tr}\{\gamma_\mu(\not{k}+mc)\gamma_\nu(\not{k}-\not{q}+mc)\}}{(k^2-m^2c^2)[(k-q)^2-m^2c^2]}\right\}[\bar{u}(p_4)\gamma^\nu u(p_2)] \tag{4.117}$$

Including this contribution, one observes the higher-order diagram can in fact be incorporated as an effective modification of the photon propagator:

$$\frac{g_{\mu\nu}}{q^2} \to \frac{g_{\mu\nu}}{q^2} - \frac{\mathrm{i}}{q^4}I_{\mu\nu} \text{ where } I_{\mu\nu} = -g_e^2\left\{\int\frac{d^4k}{(2\pi)^4}\cdots\right\} \tag{4.118}$$

The problem is the integral $I_{\mu\nu}$ is logarithmically divergent upon evaluation. Our strategy for dealing with this problem is representative for the main idea behind renormalization: we will start by absorbing this "infinity" into a new mass and coupling constant. This might seem horrendous from a mathematical point of view, but let us see how it works out in the end. The general form of $I_{\mu\nu}$ after integration (only q remaining as a 4-vector) is $g_{\mu\nu}(\ldots)+q_\mu q_\nu(\ldots)$. We thus write $I_{\mu\nu} = -\mathrm{i}g_{\mu\nu}I(q)+q_\mu q_\nu J(q)$. The second term actually makes no contribution to the final amplitude, which can be verified as follows: q_μ contracts with γ^μ in the expression for $\mathcal{M}$ and thus gives

$$[\bar{u}(3)\not{q}u(p_1)] = \bar{u}(p_3)(\not{p}_1 - \not{p}_3)u(p_1) = 0 \tag{4.119}$$

as seen from the basic Dirac equations for u and $\bar{u}$. The first term, on the other hand, may be rewritten in the following manner (the calculation is beyond the level of this book):

$$I(q) = \frac{g_e}{12\pi^2}\Big[\int_{m^2}^\infty\frac{dx}{x} - 6\int_0^1 z(1-z)\ln\Big(1-\frac{q^2}{m^2c^2}z(1-z)\Big)dz. \tag{4.120}$$

We thus manage to contain the logarithmic divergence in the first term of this quantity. If we temporarily introduce a cut-off Λ instead of the infinite limit of the integral, and define

$$f(x) \equiv 6\int_0^1 z(1-z)\ln[1+xz(1-z)]dz \tag{4.121}$$

with $x \equiv -q^2/(m^2c^2)$, we can evaluate the integral determining $I(q)$ in Eq. (4.120) as:

$$I(q) = \ln\Big(\frac{\lambda^2}{m^2}\Big)\frac{g_e^2}{12\pi^2} - \frac{g_e^2}{12\pi^2}f\Big(\frac{-q^2}{m^2c^2}\Big). \tag{4.122}$$

Two particular limiting cases of $f(x)$ allow for a simpler expression:

$$\lim_{x\ll 1}f(x) \simeq x/5, \ \lim_{x\gg 1}f(x) \simeq \ln(x). \tag{4.123}$$

We are now ready to write down the total amplitude for $e - \mu$ scattering, *including the fourth order vacuum-polarization diagram as a correction*, as follows:

$$\mathcal{M} = -g_e^2[\bar{u}(3)\gamma^\mu u(p_1)]\frac{g_{\mu\nu}}{q^2}\Big[1 - \frac{g_e^2}{12\pi^2}\Big(\ln\Big(\frac{\Lambda^2}{m^2}\Big) - f\Big(\frac{-q^2}{m^2c^2}\Big)\Big)\Big][\bar{u}(p_4)\gamma^\nu u(p_2)]. \tag{4.124}$$

Here is the critical step: let us now introduce the *renormalized* coupling constant

$$g_R \equiv g_e\sqrt{1 - \frac{g_e^2}{12\pi^2}\ln(\Lambda^2/m^2)}, \tag{4.125}$$

so that up to $\mathcal{O}(g_e^2)$ [we make the approximation $g_e^4/g_R^2 = g_R^2$ for the last term] we obtain

$$\mathcal{M} = -g_R^2[\bar{u}(p_3)\gamma^\mu u(p_1)]\frac{g_{\mu\nu}}{q^2}\Big[1 + \frac{g_R^2}{12\pi^2}f\Big(\frac{-q^2}{m^2c^2}\Big)\Big][\bar{u}(p_4)\gamma^\nu u(p_2)]. \tag{4.126}$$

Remarkably, this has an identical form as the lowest order diagram except for two differences:

1. We have a new effective coupling constant $g_e \rightarrow g_R$. It depends on the cut-off Λ, which is physically a spurious effect since we have introduced Λ by hand as a way to avoid a divergence. Nevertheless, this is not a problem in practice because g_R can be experimentally measured and has been determined to be finite. Therefore, the divergent term is apparently cancelled out by compensating infinities from other higher-order diagrams. The main point of the renormalization procedure is that the bare quantity g_e is not the physical quantity (or equivalently the bare electron charge e): higher-order corrections to the bare quantity gives the effective quantity g_R (or equivalently e_{eff}) which is physically measurable. Therefore, we can work with the expression for $\mathcal{M}$ obtained from the lowest-order diagram so long as we keep in mind that the masses and coupling constants we use must be the *renormalized ones* (whose values can, in some cases, be determined experimentally).

2. There is also a finite correction from the term $\propto \frac{g_R^2}{12\pi^2} f\left(\frac{-q^2}{m^2 c^2}\right)$. This can also be absorbed into the effective coupling constant, in which case we obtain a so-called *running* coupling constant which means that it now depends on energy and momentum (through q):

$$g_R(q^2) = g_R(0) \sqrt{1 + \frac{g_R^2}{12\pi^2} f\left(\frac{-q^2}{m^2 c^2}\right)}. \tag{4.127}$$

High momentum q is equivalent to a closer approach between the particles, so that the effective coupling constant depends on the distance to other particles (which is physically reasonable due to vacuum polarization and screening effects). For non-relativistic situations, this is nevertheless usually a small effect. The fact that a coupling constant is running has crucial physical consequences as it leads to *e.g.* asymptotic freedom in QCD, which we shall say more about in a little while.

V. WEAK INTERACTIONS AND ELECTROWEAK THEORY

Learning goals. After reading this chapter, the student should:

- Be able to compute the Feynman amplitude $\mathcal{M}$ for weak interaction Feynman diagrams, and also be able to obtain decay rates and scattering cross sections from $\mathcal{M}$.

- Understand which particles interact via charged and neutral weak interactions and how these interactions are mediated.

- Be able to describe how electromagnetic and weak interactions can be combined in a unified framework via chiral fermion states.

We will here establish Feynman rules for charged $(W^{\pm})$ and neutral (Z^0) weak interactions related to leptons and quarks, and treat some important processes in detail. Finally, we gather EM and weak vertices under the same umbrella, namely Glashow-Weinberg-Salam electroweak theory, by using chiral fermion states.

A. Charged leptonic weak interactions

The mediators of weak interactions are the charged $W^{\pm}$ bosons and the neutral Z^0 boson. Their masses are

$$M_W = 80.385 \pm 0.015 \text{ GeV}/c^2, \ M_z = 91.188 \pm 0.002 \text{ GeV}/c^2. \tag{5.1}$$

Since these are massive spin-1 bosons, there are three available spin polarization states ($m_s = -1, 0, 1$). Their propagator has the form

$$\frac{-\mathrm{i}[g_{\mu\nu} - q_\mu q_\nu/(Mc)^2]}{q^2 - M^2 c^2}. \tag{5.2}$$

Since the boson masses are large, it is often experimentally the case that the momentum transfer satisfies $q^2 \ll (Mc)^2$. In that case, we may simplify the propagator to simply $\mathrm{i}g_{\mu\nu}/(Mc)^2$. The fundamental leptonic vertex and the reverse process are shown below. Note that W^+ and W^- are antiparticles of each other.
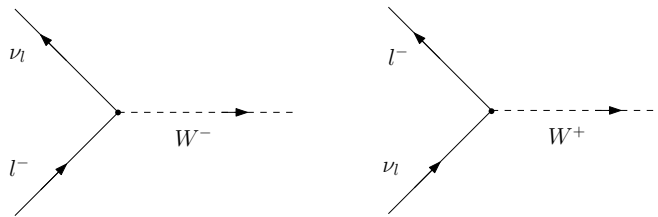
As for the Feynman rules, they are the same as in QED except for two things:

1. The propagator expression is modified since it is massive.

2. Since the interaction is different from QED, the weak vertex factor reads

$$-\frac{\mathrm{i}g_W}{2\sqrt{2}}\gamma^\mu(1 - \gamma^5). \tag{5.3}$$

The extra factor $(1 - \gamma^5)$ causes parity violation. γ^μ on its own gives a vector coupling (in QED), while $\gamma^\mu\gamma^5$ gives an axial vector coupling, as described in the previous chapter.

Decay of the muon.
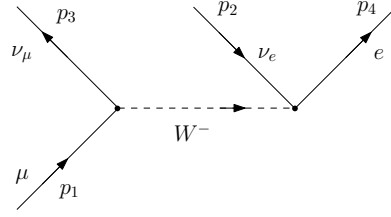
The amplitude is obtained as

$$\mathcal{M} = \frac{g_W^2}{8(M_W c)^2}[\bar{u}(3)\gamma^\mu(1-\gamma^5)u(1)][\bar{u}(4)\gamma_\mu(1-\gamma^5)v(2)]. \tag{5.4}$$

Using Casimir's trick, we find that

$$\langle|\mathcal{M}|^2\rangle = 2\Big(\frac{g_W}{M_W c}\Big)^4 (p_1 p_2)(p_3 p_4). \tag{5.5}$$

Analyzing it in the muon rest frame where $p_1 = (m_\mu c, \mathbf{0})$ we obtain $p_1 p_2 = m_\mu E_2$. Moreover, $p_1 = p_2 + p_3 + p_4$ from which it follows that

$$(p_3 + p_4)^2 = m_e^2 c^2 + 2p_3 p_4 = (p_1 - p_2)^2 + m_\mu^2 c^2 - 2p_1 p_2 \tag{5.6}$$

which in turn leads to

$$p_3 p_4 = \frac{(m_\mu^2 - m_e^2)c^2}{2} - m_\mu E_2. \tag{5.7}$$

We have set the neutrino mass $m_\nu = 0$ and since $m_\mu \gg m_e$, we can set $m_e = 0$ as well as an approximation. Inserting our 4-vectors into the expression for $\mathcal{M}$ gives:

$$\langle|\mathcal{M}|^2\rangle = \Big(\frac{g_W}{M_W c}\Big)^4 m_\mu^2 E_2(m_\mu c^2 - 2E_2). \tag{5.8}$$

We use the Golden rule to calculate the decay rate:

$$d\Gamma = \frac{\langle|\mathcal{M}|^2\rangle}{2\hbar m_\mu}\Big(\prod_{i=2,3,4} \frac{cd^3 p_i}{(2\pi)^3 2E_i}\Big)(2\pi)^4\delta^4(p_1 - p_2 - p_3 - p_4). \tag{5.9}$$

Here, we have $E_i = |\boldsymbol{p}_i|c$. The $\boldsymbol{p}_3$ integral is performed:

$$d\Gamma = \frac{\langle|\mathcal{M}|^2\rangle c^3}{16(2\pi)^5 \hbar m_\mu} \frac{(d^3 p_2)(d^3 p_4)}{E_2 E_3 E_4}\delta(m_\mu c - E_2/c - E_3/c - E_4/c), \tag{5.10}$$

where $E_3 = |\boldsymbol{p}_2 + \boldsymbol{p}_4|c$. We may then continue with the $\boldsymbol{p}_2$ integral. Let $\boldsymbol{p}_4 \parallel \hat{\boldsymbol{z}}$ so that

$$(E_3/c)^2 = |\boldsymbol{p}_2 + \boldsymbol{p}_4|^2 = (E_2^2 + E_4^2 + 2E_2 E_4 \cos\theta)/c^2. \tag{5.11}$$

Moreover, $d^3 p_2 = (\frac{E_2}{c})^2 \frac{dE_2}{c}\sin\theta d\theta d\phi$ and since there is no ϕ dependence, $\int d\phi = 2\pi$. To do the θ-integration, let $x = E_3/c$ so that

$$dx = -\frac{E_2 E_4 \sin\theta d\theta}{cE_3}. \tag{5.12}$$

We then have

$$\int_0^\pi \frac{\sin\theta d\theta}{E_3}\delta[m_\mu c - (E_2 + E_3 + E_4)/c] = \frac{c}{E_2 E_4}\int_{x_-}^{x_+} \delta(m_\mu c - x - E_2/c - E_4/c)dx$$

$$= \begin{cases} \frac{c}{E_2 E_4} & \text{if } x_- < m_\mu c - E_2/c - E_4/c < x_+ \\ 0 & \text{otherwise.} \end{cases} \tag{5.13}$$

Here, we defined $x_\pm = |E_2 \pm E_4|/c$. We can write the above inequality in a physically more transparent way by adding $(E_2 + E_4)$ and dividing on 2:

$$\frac{1}{2}[|E_2 - E_4| + E_2 + E_4] < \frac{1}{2}m_\mu c^2 < E_2 + E_4. \tag{5.14}$$

This is in fact equivalent to three inequalities: $E_2 < m_\mu c^2/2$, $E_4 < m_\mu c^2/2$, and $(E_2 + E_4) > m_\mu c^2/2$. If we reflect upon these equations for a moment, we realize that they make sense physically. The maximum energy for particle is obtained if it emerges *opposite* to the two others. Conservation of momentum then dictates that the particle picks up half of the available energy. Therefore, a pair of particles (*e.g.* 2 and 4) must always have at least an energy of $m_\mu c^2/2$ If there is an angle between the particles, one will always acquire less energy.
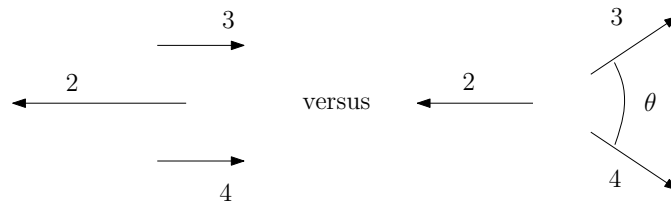
The inequalities dictate the limits for the dE_2 integral:

$$d\Gamma = \left(\frac{g_W}{4\pi M_W c}\right)^4 \frac{m_\mu c}{\hbar} \frac{d^3 p_4}{E_4^2} \int_{m_\mu c^2/2 - E_4}^{m_\mu c^2/2} E_2(m_\mu c^2 - 2E_2)dE_2 \quad = \left(\frac{g_W}{4\pi M_W c}\right)^4 \frac{m_\mu c}{\hbar} \left(\frac{m_\mu c^2}{2} - \frac{2E_4}{3}\right)d^3 p_4. \tag{5.15}$$

Finally, setting $E \equiv E_4$ and then $d^3 p_4 = 4\pi (E_4/c)^2 dE_4/c$ gives us:

$$\frac{d\Gamma}{dE} = \left(\frac{g_W}{M_W c}\right)^4 \frac{m_\mu^2 E^2}{2\hbar(4\pi)^3}\left(1 - \frac{4E}{3m_\mu c^2}\right). \tag{5.16}$$

This is the energy distribution of the electrons emitted in muon decay, and gives a very good experimental match. The total decay rate can be obtained as:

$$\Gamma = \left(\frac{g_W}{M_W c}\right)^4 \frac{m_\mu^2}{2\hbar(4\pi)^3} \int_0^{m_\mu c^2/2} E^2\left(1 - \frac{4E}{3m_\mu c^2}\right)dE \tag{5.17}$$

which in turn gives the lifetime $\tau = \Gamma^{-1}$. Comparing with the experimentally determined value of τ, one obtains $g_W = 0.66$. This gives a weak fine structure constant

$$\alpha_W = \frac{g_W^2}{4\pi} \simeq \frac{1}{29} \tag{5.18}$$

which is roughly five times α_{QED}. The intrinsic coupling in weak interactions is then large, but the interactions still remain feeble at low energies because the propagator is so massive $[\sim (q^2 - M_W^2 c^2)^{-1}]$. For high energies (momentum transfer), $q^2 \sim M_W^2 c^2$, the weak interactions can dominate over the electromagnetic ones.

Decay of the neutron.
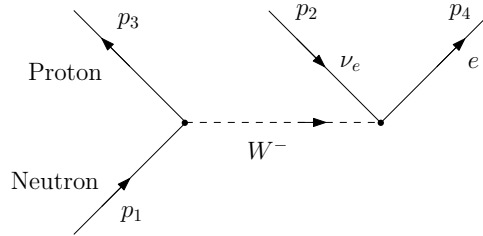
Neutrons and protons are composite particles since they consist of quarks, but perhaps we can expect that treating them as point Dirac particles will give a reasonable approximation? The calculation proceeds in essentially the same way as for μ-decay, and the final result is:

$$\Gamma = \frac{1}{4\pi^3\hbar}\left(\frac{g_W}{2M_W c}\right)^4 (m_e c^2)^5 [\frac{1}{15}(2a^4 - 9a^2 - 8)\sqrt{a^2 - 1} + a\ln(a + \sqrt{a^2 - 1})], \tag{5.19}$$

where $a \equiv (m_n - m_p)/m_e$. Note that we cannot here neglect the electron mass, since the rest energy of the electron is comparable to the released energy of the reaction, $(m_n - m_p - m_e)c^2$, due to the small mass difference between the neutron and proton masses. This is not the case for μ-decay, where the released energy is $(m_\mu - m_e)c^2$. Putting in the numbers, we obtain $\tau = \frac{1}{\Gamma} = 1316$ s. The experimentally value, on the other hand, is $\tau = 898 \pm 16$ s. The order of magnitude is thus correct, but there is still a deviation. Considering that weak decay processes range from 15 minutes (the neutron) to 10^{-13} s, our theoretical estimate is not that bad. But what is the reason for the discrepancy? We did assume that p and n are point particles (neglecting their internal quark structure) and assumed an interaction with W in the same way as leptons interact with it. At the same time, we also know that the Mott formula works very well for $e - p$ scattering mediated by γ, where p is treated as a point particle. The crucial question becomes: what is the net coupling strength of the proton and the neutron to $W^\pm$?

In electrodynamics, all internal complications do not matter because electric charge is conserved. However, we do not know that the same is necessarily true for weak interactions. For instance, a gluon splitting to a $q\bar{q}$ pair might make a finite contribution to the effective weak coupling vertex since quarks interact weakly. To account for this in the $n \to p + W$ vertex, we make the substitution: $(1 - \gamma^5) \to (c_V - c_A\gamma^5)$ where c_V is the correction to the vector "weak charge" while c_A is the correction to the axial vector "weak charge". Experimentally, one finds $c_V \simeq 1.00$ and $c_A \simeq 1.26$. The corrected theoretical estimate for the lifetime then comes into much closer agreement with the experimental one: $\tau = 914$ s.

Decay of the pion.

The process $\pi^- \to l^- + \bar{\nu}_l$ is actually a scattering event of bound quarks [see Fig. (a)], and can in a sense be viewed analogously to positronium decay ($e^+ + e^- \to \gamma + \gamma$).
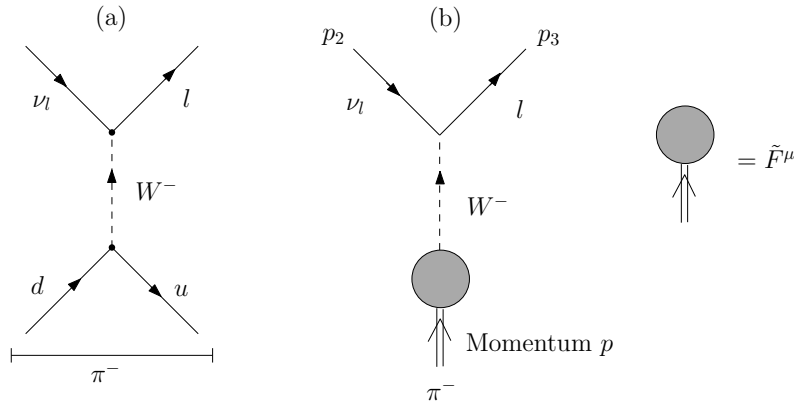
Analyzing this decay in the framework of weak interactions, we may represent the scattering as in (b). The dark blob represents the (unknown) coupling of π^- to W^-. Let $\tilde{F}^\mu$ describe the blob. The Feynman amplitude takes the form:

$$\mathcal{M} = \bar{u}(3)\Big(\frac{-\mathrm{i}g_W}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)\Big)v(2)\frac{\mathrm{i}g_{\mu\nu}}{(M_W c)^2}\tilde{F}^\mu$$

$$= \frac{g_W^2}{8(M_W c)^2}[\bar{u}(3)\gamma_\mu(1-\gamma^5)v(2)]F^\mu. \tag{5.20}$$

We here absorbed a constant into $\tilde{F}^\mu \to F^\mu$. We know it must be a 4-vector in order to contract γ_μ, so that $\mathcal{M}$ ends up being a scalar. Since π^- has spin 0, F^μ can only depend on p^μ. Therefore, $F^\mu = f_\pi p^\mu$ where f_π is a scalar. We now perform a summation over outgoing spins:

$$\langle |\mathcal{M}|^2 \rangle = \Big[\frac{f_\pi}{8}\Big(\frac{g_W}{M_W c}\Big)^2\Big]^2 p_\mu p_\nu \mathrm{Tr}\{\gamma^\mu(1-\gamma^5)\not{p}_2\gamma^\nu(1-\gamma^5)(\not{p}_3+m_l c)\}$$

$$= \frac{1}{8}\Big[\frac{f_\pi}{8}\Big(\frac{g_W}{M_W c}\Big)^2\Big]^2 [2(pp_2)(pp_3) - p^2(p_2 p_3)]. \tag{5.21}$$

Since $p = p_2 + p_3$, we have $2p_2 p_3 = (m_\pi^2 - m_l^2)/c^2$. As a result, we obtain

$$\langle |M|^2 \rangle = \Big(\frac{g_W}{2M_W}\Big)^4 f_\pi^2 m_l^2 (m_\pi^2 - m_l^2). \tag{5.22}$$

This is as far as we get without specifying what f_π is. However, we can calculate *branching ratios*:

$$\frac{\Gamma(\pi^- \to e^- + \bar{\nu}_e)}{\Gamma(\pi^- \to \mu^- + \bar{\nu}_\mu)} = \frac{m_e^2(m_\pi^2 - m_e^2)^2}{m_\mu^2(m_\pi^2 - m_\mu^2)^2} = 1.28 \times 10^{-4}. \tag{5.23}$$

Experimental measurements give $(1.23 \pm 0.02) \times 10^{-4}$, so it is a good fit.

However, let us consider if this result is physically reasonable. It appears that π seems to prefer to decay into μ in spite of the fact that $m_\mu \gg m_e$. This seems to contradict the general rule that decays occur faster the larger the mass difference, since there is more phase space available for decay into a lighter particle. There exists, however, an explanation. Consider the hypothetical case where $m_e = 0$. If so, then $\pi^- \to e^- + \bar{\nu}_e$ would be completely forbidden for the following reason. π^- has $s = 0$, so that e and $\bar{\nu}_e$ have to emerge with opposite spins. This means that they have equal helicity.
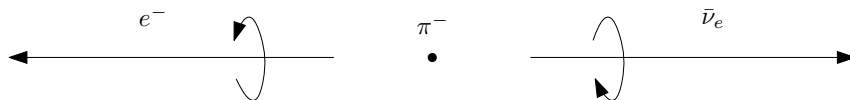
Since $\bar{\nu}_e$ is considered to be right-handed in all cases treated here (although nowadays it is known that neutrinos have a small, but finite mass, meaning that left-handed antineutrinos should exist as well), e should also be that here. However, if $m_e = 0$, then $(1 - \gamma^5)$ in the weak vertex would only couple to left-handed electrons (just as it only couples to left-handed neutrinos). Therefore, the decay to e is strongly suppressed since m_e is so small compared to the other energy scales.

B. Charged weak interactions of quarks

For leptons, a coupling to $W^{\pm}$ occurs only within a particular *generation*:

$$\begin{bmatrix} \nu_e \\ e \end{bmatrix}, \quad \begin{bmatrix} \nu_\mu \\ \mu \end{bmatrix}, \quad \begin{bmatrix} \nu_\tau \\ \tau \end{bmatrix}. \tag{5.24}$$

No cross-terms of the type $e^- \to \nu_\mu + W^-$ occur. For quarks, the generation structure is similar:

$$\begin{bmatrix} u \\ d \end{bmatrix}, \quad \begin{bmatrix} c \\ s \end{bmatrix}, \quad \begin{bmatrix} t \\ b \end{bmatrix}, \tag{5.25}$$

but now cross-generational coupling is allowed. For instance, $s \to u + W^-$ underlies the decay $\Lambda \to p + e + \bar{\nu}_e$. In 1963, when only the u, d, s quarks were known, Cabibbo suggested that $d \to u + W^-$ vertex carries an extra factor $\cos\theta_C$, whereas $s \to u + W^-$ carries $\sin\theta_C$.

$$-\frac{ig_W}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)\cos\theta_C \qquad\qquad -\frac{ig_W}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)\sin\theta_C$$
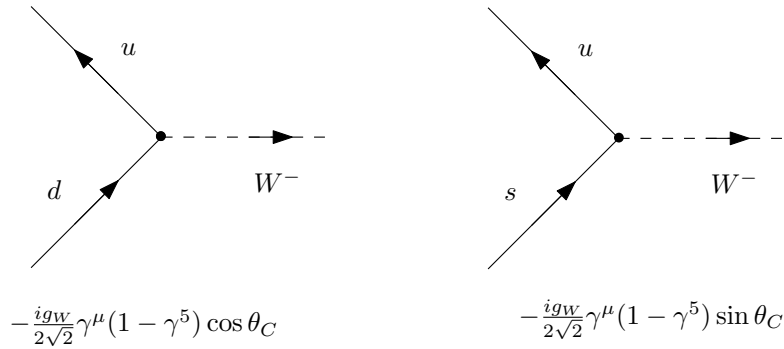
Experimentally, one finds $\theta_C = 13.1°$, so that weak interactions *almost* respect quark generations. Various decays may be classified as follows:

- Leptonic decays ($K^- \to l^- + \bar{\nu}_l$)

- Semileptonic decays ($\pi^- \to \pi^0 + e^- + \bar{\nu}_e$)

- Non-leptonic decays ($K^- \to \pi^0 + \pi^-$).

Cabibbo's idea was very successful for many decay rates, but a problem arose when considering $K^0 \to \mu^+ + \mu^-$. The amplitude should be proportional to $\sin\theta_C \cos\theta_C$, since the diagram has the following form:
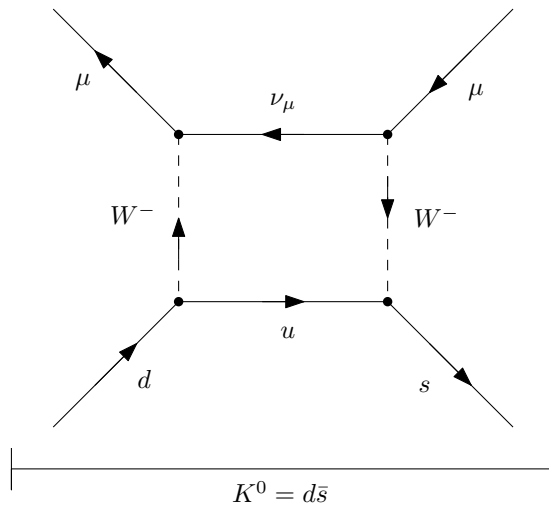
However, this did not agree with the experimentally measured rate, which was much lower than the prediction. The paradox was resolved in 1970 by Glashow, Iliopoulos, and Maiani who introduced a fourth quark c (even prior to its experimental discovery. The new quark would interact in the following manner:

$$-\frac{ig_W}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)(-\sin\theta_C) \qquad\qquad -\frac{ig_W}{2\sqrt{2}}\gamma^\mu(1-\gamma^5)\cos\theta_C$$
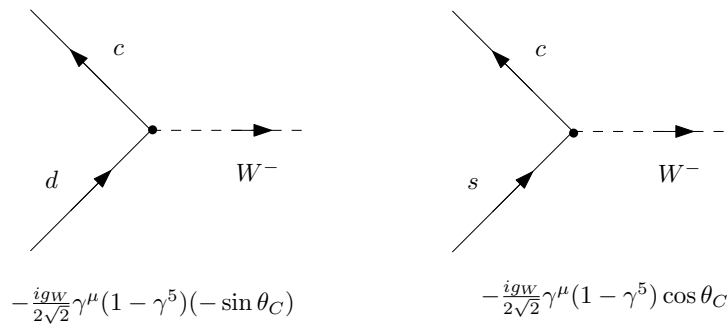
Now, the diagram for $K^0 \to \mu^+ + \mu^-$ was almost completely cancelled by the corresponding one where u was replaced with c, since that one was proportional to $(-\sin\theta_C \cos\theta_C)$. The cancellation was not perfect due to the mass mismatch $m_c \neq m_u$, leaving behind a small net amplitude.

The Cabibbo-GIM scheme then suggests the following: in weak interactions, one should replace the physical quarks d and s with d' and s':

$$d' = d\cos\theta_C + s\sin\theta_C, \ \ s' = -d\sin\theta_C + s\cos\theta_C. \tag{5.26}$$

The $W^{\pm}$'s then couple to Cabibbo-rotated states

$$\begin{bmatrix} u \\ d' \end{bmatrix} = \begin{bmatrix} u \\ d\cos\theta_C + s\sin\theta_C \end{bmatrix}, \ \begin{bmatrix} c \\ s' \end{bmatrix} = \begin{bmatrix} c \\ -d\sin\theta_C + s\cos\theta_C \end{bmatrix}. \tag{5.27}$$

In this way, $d \to u + W^-$ carries a factor $\cos\theta_C$, $s \to u + W^-$ carries $\sin\theta_C$, and so forth.

Kobayashi and Maskawa (KM) generalied this to three "weak interaction generations" of quarks:

$$\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = \begin{bmatrix} U_{ud} & U_{us} & U_{ub} \\ U_{cd} & U_{cs} & U_{cb} \\ U_{td} & U_{ts} & U_{tb} \end{bmatrix} \begin{bmatrix} d \\ s \\ b \end{bmatrix}. \tag{5.28}$$

where U_{ud} represents the coupling of u to d ($d \to u + W^-$). Not all entries are independent, because U must be a unitary matrix.

C. Neutral weak interactions

The fundamental vertex for these interactions looks as (a).
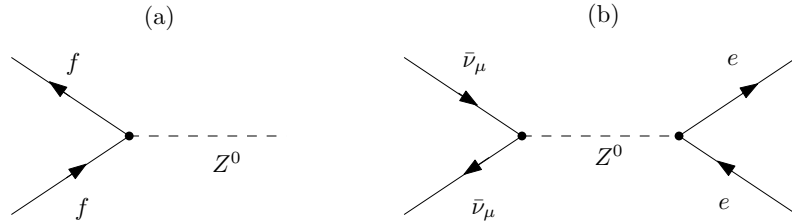
Here, f is any lepton or quark. There is no fermion-mixing, though. The first experimental indication of neutral weak interactions was provided in 1973 through the process $\bar{\nu}_\mu + e \to \bar{\nu}_\mu + e$ shown in (b). We know that the coupling of quarks and leptons to $W^{\pm}$ is $-(\mathrm{i}g_W/2\sqrt{2})\gamma^\mu(1 - \gamma^5)$. This is slightly modified when coupling to composite particles, like the proton, but that is due to a contamination from strong interaction. The coupling to Z^0 is similarly given by the vertex factor

$$-\mathrm{i}g_z\gamma^\mu(c_V^f - c_A^f \gamma^5)/2. \tag{5.29}$$

Here, g_z is the neutral coupling constant while $c_{V,A}^f$ are coefficients depending on the fermion type f. All of these numbers are determined by the *weak mixing angle*, also known as the Weinberg angle, θ_W. These relations are summarized as follows.

f	c_V	c_A
ν_e, ν_μ, ν_τ	$\frac{1}{2}$	$\frac{1}{2}$
e^-, μ^-, τ^-	$-\frac{1}{2} + 2\sin^2\theta_W$	$-\frac{1}{2}$
u, c, t	$\frac{1}{2} - \frac{4}{3}\sin^2\theta_W$	$\frac{1}{2}$
d, s, b	$-\frac{1}{2} + \frac{2}{3}\sin^2\theta_W$	$-\frac{1}{2}$

This will be motivated later on, as we unify QED and weak interactions into electroweak theory. There is no way to compute θ_W theoretically in the Standard Model, but its value can be inferred from experiments: $\theta_W \simeq 28.7°$. Finally, the propagator is:

$$-\mathrm{i}(g_{\mu\nu} - q_\mu q_\nu/M_Z^2 c^2)/(q^2 - M_Z^2 c^2) \tag{5.30}$$

with $M_W = M_Z \cos\theta_W$.

Example 14. Elastic $\nu_e - e$ scattering.
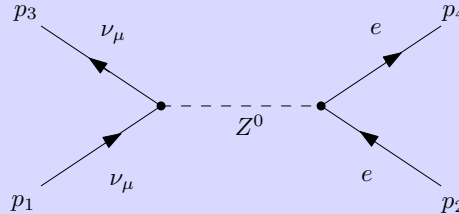
The amplitude is obtained using the Feynman rules:

$$\mathcal{M} = \frac{g_Z^2}{8(M_Z c)^2}[\bar{u}(3)\gamma^\mu(1-\gamma^5)u(1)][\bar{u}(4)\gamma_\mu(c_V - c_A\gamma^5)u(2)]. \tag{5.31}$$

As before, $\{c_V, c_A\}$ are the neutral weak couplings for electrons. If we go to the CM frame and assume very high energy scattering so that we may neglect the electron mass (its rest energy), we obtain:

$$\langle|\mathcal{M}|^2\rangle = 2\Big(\frac{g_Z E}{M_Z c^2}\Big)^4[(c_V + c_A)^2 + (c_V - c_A)^2\cos^4(\theta/2)] \tag{5.32}$$

where E is the electron energy (or equivalently the neutrino energy, since we have neglected m_e and operate in the CM frame) and θ is the scattering angle.

Before   After

Example 15.
The differential scattering cross section for this situation can be worked out according to our previous treatment of how to relate $d\sigma/d\Omega$ to $\mathcal{M}$ in for two-particle scattering in the CM frame:

$$\frac{d\sigma}{d\Omega} = 2\Big(\frac{\hbar c}{\pi}\Big)^2 \Big(\frac{g_z}{4M_z c}\Big)^4 E^2 [(c_V + c_A)^2 + (c_V - c_A)^2 \cos^4(\theta/2)] \tag{5.33}$$

so that the total σ becomes:

$$\sigma = \frac{2}{3\pi}(\hbar c)^2 \Big(\frac{g_z}{2M_z c^2}\Big)^4 E^2 (c_V^2 + c_A^2 + c_V c_A). \tag{5.34}$$

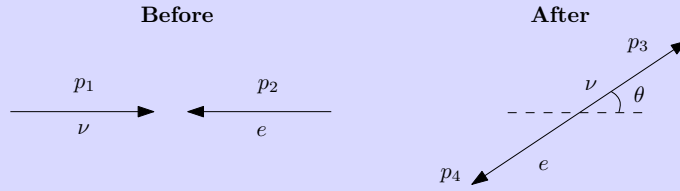It is important to note that most (but not all) neutral processes are masked by competing electromagnetic ones. For instance, $e^- + e^+ \to \mu^- + \mu^+$ can occur both via exchange of a virtual Z^0 or γ. Conversely, there is *a weak interaction contamination in **every** electromagnetic process* since Z^0 couples to everything that γ couples to (and more). Even if the effect is small, its smoking gun signature, when it is observable, is parity violation. To access weak interactions alone, one has to use neutrinos which do not have any electromagnetic coupling. Alternatively, if one is able to access energies so high that $q \sim M_Z c$, the denominator of the Z^0 propagator becomes very small and leads to a large interaction that dominates even the γ contribution. Let us consider an example of such a situation.

Example 16. $e^- e^+$ **scattering near the Z^0 pole.** We are considering the process $e^- + e^+ \to f + \bar{f}$ where f is a quark or a lepton.
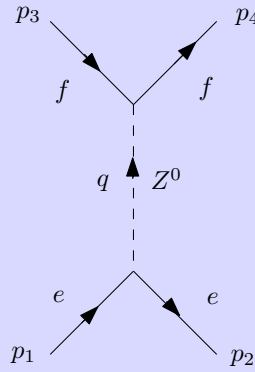
Asume that $m_f \ll M_Z$, but let us keep the exact form of the Z^0 propagator since we are interested in large momenta transfers $q \sim M_Z c$. By the way, note that Z^0 is its own antiparticle. We obtain:

$$\mathcal{M} = -\frac{g_Z^2}{4[q^2 - (M_Z c)^2]}[\bar{u}(4)\gamma^\mu (c_V^f - c_A^f \gamma^5)v(3)][g_{\mu\nu} - q_\mu q_\nu/(M_z c)^2][\bar{v}(2)\gamma^\nu (c_V^e - c_A^e \gamma^5)u(1)] \tag{5.35}$$

with $q = p_1 + p_2 = p_3 + p_4$. Since we will consider energies near the M_Z mass of 90 GeV, we may safely neglect all masses. From the second term, we find that q_μ contracts with γ^μ to yield

$$\bar{u}(4)\slashed{q}(c_V - c_A\gamma^5)v(3). \tag{5.36}$$

Moreover, since $\displaystyle{\not{q}} = {\not{p}}_3 + {\not{p}}_4$ and $\bar{u}(4){\not{p}}_4 = 0$ (this is the Dirac equation for $m = 0$) in addition to:

$$\not{p}_3(c_V - c_A\gamma^5)v(3) = (c_V + c_A\gamma^5)\not{p}_3 v(3) = 0, \tag{5.37}$$

we see that the term dependent on $q_\mu q_\nu$ gives no contribution to $\mathcal{M}$. Performing now the spin summation (Casimir's trick):

$$\langle|\mathcal{M}|^2\rangle = \left[\frac{g_Z^2}{8(q^2 - M_Z^2 c^2)}\right]^2 \text{Tr}\{\gamma^\mu(c_V^f - c_A^f\gamma^5)\not{p}_3\gamma^\nu(c_V^f - c_A^f\gamma^5)\not{p}_4\} \times \text{Tr}\{\gamma^\mu(c_V^e - c_A^e\gamma^5)\not{p}_1\gamma_\nu(c_V^e - c_A^e\gamma^5)\not{p}_2\}. \tag{5.38}$$

Performing the traces (using rules for γ-matrices) and integrating over the scattering angle, we find:

$$\sigma = \frac{1}{3\pi}\left[\frac{\hbar c g_Z^2 E}{4[(2E)^2 - (M_z c^2)^2]}\right]^2 [(c_V^f)^2 + (c_A^f)^2][(c_V^e)^2 + (c_A^e)^2] \tag{5.39}$$

in the CM frame. We now see that as the total energy $2E \to M_Z c^2$, the cross section σ diverges! To counter this (a true mathematical divergence can never be observed physically), take into account the finite lifetime τ_Z of Z^0 which modifies the propagator to:

$$\frac{1}{q^2 - (M_Z c)^2} \to \frac{1}{q^2 - (M_Z c)^2 + i\hbar M_Z \Gamma_Z} \tag{5.40}$$

where $\Gamma_Z = \tau_Z^{-1}$. To derive why the finite lifetime leads to the imaginary term in the denominator is beyond the scope of this course, but as a sidenote it is interesting to be aware of the fact that the same thing happens when describing finite lifetimes of quasiparticles in condensed matter systems. In any case, the presence of the term $\propto \Gamma_Z$ is that it "smears" out the effective mass of the propagator so that the divergence right at the bare mass M_Z is avoided and one obtains

$$\sigma = \frac{(\hbar c g_Z^2 E)^2}{48\pi}\frac{[(c_V^f)^2 + (c_A^f)^2][(c_V^e)^2 + (c_A^e)^2]}{[(2E)^2 - (M_Z c^2)^2]^2 + (\hbar M_Z c^2 \Gamma_Z)^2}. \tag{5.41}$$

The correction from the finite lifetime is negligible *except* if $2E \simeq M_Z c^2$, in which case it is crucial since it prevents a divergence. Now, the same process mediated by a photon gives:

$$\sigma = \frac{(\hbar c g_e^2)^2 (Q^f)^2}{48\pi E^2} \tag{5.42}$$

where Q^f is the charge of f in units of e. We can actually compare γ- and Z^0-mediated scattering directly via the ratio:

$$\frac{\sigma(e^+ e^- \to Z^0 \to \mu^+\mu^-)}{\sigma(e^+ e^- \to \gamma \to \mu^+\mu^-)} \simeq \frac{2E^4}{[(2E)^2 - (M_Z c^2)^2]^2 + (\hbar \Gamma_Z M_Z c^2)^2} \tag{5.43}$$

when inserting for θ_W in $c_{V,A}^f$ and $c_{V,A}^e$. Two particular limiting cases are of interest:

$$\lim_{2E \ll M_Z c^2} \frac{\sigma_Z}{\sigma_\gamma} \simeq 2\left(\frac{E}{M_Z c^2}\right)^4 \ll 1 \tag{5.44}$$

whereas near the Z^0 pole

$$\lim_{2E \to M_Z c^2} \frac{\sigma_Z}{\sigma_\gamma} \simeq \frac{1}{8}\left(\frac{M_Z c^2}{\hbar \Gamma_z}\right)^2 \gg 1. \tag{5.45}$$

We have used the value $\hbar \Gamma_Z = 2.5$ GeV. Hence, the weak mechanism is strongly favored near Z^0 pole.

D. Electroweak unification

Chiral fermion states.
We would now like to explore where the GWS-parameters, such as c_V and c_A, and their dependence on the Weinberg-angle θ_W come from. First, note that Glashow's original aim was to unify weak and electromagnetic interactions as manifestations of one fundamental "electroweak" interaction. An immediate problem arises: if it is indeed the same underlying interaction, why is γ massless while $W^\pm$ and Z^0 so heavy? The solution turns out to be the *Higgs mechanism*, but we shall delay a detailed treatment of this to a later chapter in this book.

There is another issue which also must be resolved in the quest to merge weak and electromagnetic interactions: the structural difference between the vertices, namely γ^μ versus $\gamma^\mu(1-\gamma^5)$. One way to fix this is to *absorb* $(1-\gamma^5)$ into the particle spinor itself:

$$u_L(p) \equiv (1-\gamma^5)u(p)/2, \tag{5.46}$$

so that the vertex structure becomes the same in QED and weak interactions. In general, u_L is not an eigenstate of helicity operator in spite of "L" representing "left-handed", i.e. helicity -1. We have:

$$\gamma^5 u(p) = \begin{bmatrix} \frac{c(\boldsymbol{p}\cdot\boldsymbol{\sigma})}{E+mc^2} & 0 \\ 0 & \frac{c(\boldsymbol{p}\cdot\boldsymbol{\sigma})}{E-mc^2} \end{bmatrix} u(p) \tag{5.47}$$

which is clearly seen to not be proportional to $u(p)$. We can prove the above equation as follows. First, note that

$$\gamma^5 u = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_A \\ u_B \end{bmatrix}. \tag{5.48}$$
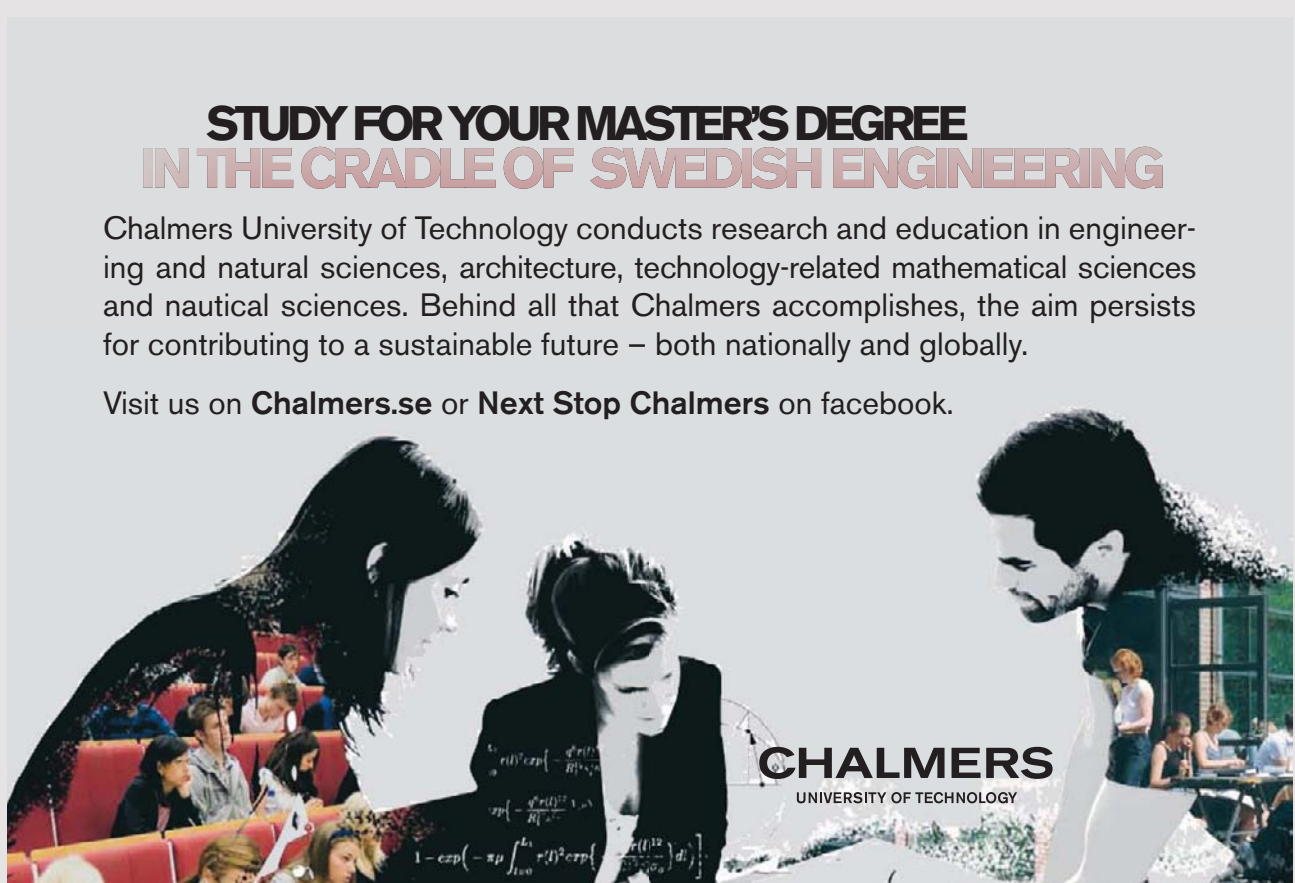
Now, use that $(\not{p}-mc)u = 0$ gives us two equations (the upper and lower components):

$$u_A = \frac{c}{E-mc^2}(\boldsymbol{p}\cdot\boldsymbol{\sigma})u_B, \; u_B = \frac{c}{E+mc^2}(\boldsymbol{p}\cdot\boldsymbol{\sigma})u_A. \tag{5.49}$$

Inserted into Eq. (5.48), we then obtain the desired result Eq. (5.47). Now, if $m = 0$, then we obtain

$$\gamma^5 u(p) = (\hat{\boldsymbol{p}} \cdot \boldsymbol{\Sigma}) u(p) \tag{5.50}$$

where $\hat{\boldsymbol{p}} \cdot \boldsymbol{\Sigma}$ is the helicity operator with eigenvalues ± 1. Thus, for mass $m = 0$ one has:

$$\frac{1}{2}(1 - \gamma^5) u(p) = \begin{cases} 0 \text{ if } u(p) \text{ has helicity } +1 \\ u(p) \text{ if } u(p) \text{ has helicity } -1 \end{cases} \tag{5.51}$$

It is important to emphasize that this holds exactly only for $m = 0$, but u_L is nevertheless always referred to as a left-handed particle. We can think of $\frac{1}{2}(1 - \gamma^5)$ as a projection operator that picks out the -1 helicity component from $u(p)$ for $m = 0$.

For an antiparticle:

$$v_L(p) \equiv \frac{(1 + \gamma^5)}{2} v(p). \tag{5.52}$$

For a right-handed counterpart of a particle or antiparticle, let $\gamma^5 \to (-\gamma^5)$. The states $u_{L,R}$ and $v_{L,R}$ are *chiral fermion states*, where the word "chiral" is Greek for "hand" (in reference to their left/right nature). In this way, weak and electromagnetic interactions can be expressed in a more unified form. Consider for instance the process:
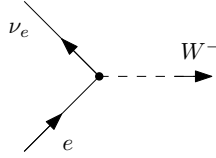
which contributes with a factor

$$j_\mu^- = \bar{\nu} \gamma_\mu \left(\frac{1 - \gamma^5}{2} \right) e \tag{5.53}$$

to the amplitude $\mathcal{M}$. Here, $\bar{\nu}$ and e represents the spinors. The quantity j_μ^- is the *weak current*, and can be viewed as an analogue to the electric current in QED. Now, anticommutation of $\{\gamma^\mu, \gamma^5\} = \{\gamma_\mu, \gamma^5\} = 0$ gives that $\gamma_\mu(1 - \gamma^5)/2 = (1 + \gamma^5)\gamma_\mu/2$. Also, we have $[(1 - \gamma^5)/2]^2 = (1 - \gamma^5)/2$. Therefore, the relation

$$\gamma_\mu \left(\frac{1 - \gamma^5}{2} \right) = \left(\frac{1 + \gamma^5}{2} \right) \gamma_\mu \left(\frac{1 - \gamma^5}{2} \right). \tag{5.54}$$

This allows us to write the weak current in the manner:

$$j_\mu^- = \bar{\nu}_L \gamma_\mu e_L \tag{5.55}$$

which is a purely vectorial vertex, but only couples left-handed ν's to left-handed e's. We can also accomplish the same thing in QED (writing the current in terms of chiral fermion states) even if the vertex is already vectorial there:

$$u = \left(\frac{1 - \gamma^5}{2} \right) u + \left(\frac{1 + \gamma^5}{2} \right) u = u_L + u_R, \tag{5.56}$$

so that for instance

$$j_\mu^{\text{QED}} = -\bar{e} \gamma_\mu e = -\bar{e}_L \gamma_\mu e_L - \bar{e}_R \gamma_\mu e_R. \tag{5.57}$$

Cross-terms can be shown to vanish, and we built in a factor (-1) due to the negative electric charge.

The main merit of the above formulation is then that the actual vertex may be turned into a purely vectorial one for both QED and the weak interactions by allowing the $(1 - \gamma^5)$ factor to characterize the particles instead. Again, we emphasize that the "L" and "R" notation only represents *true* handedness for mass $m = 0$ (and approximately for $E \gg mc^2$). In general, it should be viewed as convenient notation.

Weak isospin and hypercharge.
The negatively charged weak current and positive equivalent look as follows.
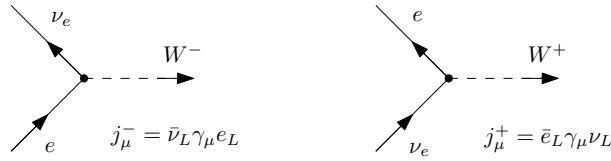
They may be expressed compactly by introducing the notation

$$\chi_L = \begin{bmatrix} \nu_e \\ e \end{bmatrix}_L , \ \tau^+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} , \ \tau^- = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{5.58}$$

so that we may write:

$$j_\mu^\pm = \bar{\chi}_L \gamma_\mu \tau^\pm \chi_L. \tag{5.59}$$

We see that $\tau^\pm = (\tau^1 \pm i\tau^2)/2$ where $\tau^{1,2}$ are the first two Pauli matrices. This is similar to our earlier treatment of isospin. In fact, we could envision a full "weak isospin" SU(2) symmetry (meaning that the Lagrangian describing these interactions would be invariant under an SU(2) transformation of the fermion spinors: more on this in the final chapter) if a third current existed, namely

$$j_\mu^3 = \bar{\chi}_L \gamma_\mu \frac{1}{2} \tau^3 \chi_L. \tag{5.60}$$

This is precisely the neutral weak current, which we already do know exists. But what about right-handed couplings via γ, which should also be allowed? To incorporate these couplings and the electromagnetic current into this unified framework, we introduce the *weak hypercharge* current

$$j_\mu^Y = 2j_\mu^{\text{QED}} - 2j_\mu^3. \tag{5.61}$$

which includes right-handed electrons [to see this, revisit the definition of j_μ^{QED} in Eq. (5.57)]. The quantity Y equals $2Q - 2I^3$ where Q is the electric charge in units of e while I^3 is the third component of the weak isospin. Note that j_μ^Y is invariant with respect to rotations in weak isospin space (such as $e_L \leftrightarrow \nu_L$), since such transformations do not influence right-handed states.

Note that we are here considering the original leptonic version of the electroweak model, where the right-handed neutrino did not couple to any gauge fields as it was thought to not exist, and could thus simply be dropped. However, today we know that neutrinos have a small mass (as discussed in our earlier treatment of neutrino oscillations), so that the right-handed neutrino should in fact be kept.

The total underlying symmetry group of the (original) electroweak theory is $\text{SU}(2)_L \otimes \text{U}(1)$ where U(1) is associated with the weak hypercharge quantum number. The latter involves both chiralities. The meaning of the above statement regarding the symmetry group is that the Lagrangian describing these interactions is invariant under both SU(2) transformations (in the left-handed sector) and U(1) transformations of the fermion spinors. This formalism can also be extended to other lepton/quark doublets, such as

$$\chi_L \rightarrow \begin{bmatrix} \nu_e \\ e \end{bmatrix}_L , \ \begin{bmatrix} \nu_\mu \\ \mu \end{bmatrix}_L , \ \begin{bmatrix} u \\ d' \end{bmatrix}_L , \ \begin{bmatrix} c \\ s' \end{bmatrix}_L , \tag{5.62}$$

and so forth. We then formally define three weak isospin currents $\boldsymbol{j}_\mu$ and a weak hypercharge current j_μ^Y:

$$\boldsymbol{j}_\mu = \frac{1}{2} \bar{\chi}_L \gamma_\mu \boldsymbol{\tau} \chi_L, \ j_\mu^Y = 2j_\mu^{\text{QED}} - 2j_\mu^3, \tag{5.63}$$

where

$$j_\mu^{\text{QED}} = \sum_{i=1}^{2} Q_i(\bar{u}_{iL}\gamma_\mu u_{iL} + \bar{u}_{iR}\gamma_\mu u_{iR}) \tag{5.64}$$

and the summation goes over particles i in the doublet.

E. Electroweak mixing

The GWS model states the the weak isospin currents couple with strength g_W to a weak isotriplet of intermediate vector bosons $\boldsymbol{W}^\mu$ (a vector in weak isospin space) whereas j_μ^Y couples with strength $g'/2$ to an isosinglet intermediate vector boson B^μ. This may be mathematically expressed via the following structure, which could be included in a Lagrangian describing such interactions:

$$-\mathrm{i}\Big[g_W \boldsymbol{j}_\mu \cdot \boldsymbol{W}^\mu + \frac{g'}{2}j_\mu^Y B^\mu\Big] \tag{5.65}$$

We shall in fact examine Lagrangians describing particle interactions in much more detail in the last chapter of this book. Eq. (5.65) contains all the electrodynamic and weak interactions. We may express the first term via the charged currents:

$$\boldsymbol{j}_\mu \cdot \boldsymbol{W}^\mu = \frac{1}{\sqrt{2}}(j_\mu^+ W^{\mu+} + j_\mu^- W^{\mu-}) + j_\mu^3 W^{\mu 3} \tag{5.66}$$

where $W^{\mu\pm} \equiv \frac{1}{\sqrt{2}}(W^{\mu 1} \mp \mathrm{i}W^{\mu 2})$ are the "wavefunctions" for the $W^\pm$ particles. The precise couplings to the $W^\pm$ particles can now be read out from the coefficients of Eq. (5.65). Let us look at this in more detail.

Example 17. Coupling to intermediate vector bosons. Consider $e^- \to \nu_e + W^-$. This process is described by the negative weak current:

$$j_\mu^- = \bar{\nu}_L \gamma_\mu e_L = \bar{\nu}\gamma_\mu \frac{1-\gamma^5}{2}e. \tag{5.67}$$

Inserted into Eq. (5.65), we obtain that

$$-\mathrm{i}g_W \frac{1}{\sqrt{2}}j_\mu^- W^{\mu-} = -\frac{\mathrm{i}g_W}{2\sqrt{2}}[\bar{\nu}\gamma_\mu(1-\gamma^5)e]W^{\mu-}. \tag{5.68}$$

Thus, we recover the correct vertex factor $-\mathrm{i}g_W/(2\sqrt{2})\gamma_\mu(1-\gamma^5)$.

What happens in electroweak theory is that the $\mathrm{SU}(2)_L\otimes \mathrm{U}(1)$ symmetry is reduced to an $\mathrm{U}(1)$ symmetry (but not in the hypercharge sector, although this is not crucial for this argument) via the *Higgs mechanism* that occurs through *spontaneous symmetry breaking*. We shall have much more to say about both of these phenomena in the next chapter, but for now we only need to know that the implication of this is that the two neutral intermediate vector bosons in our theory so far (W^3 and B) mix and produce one massless linear combination (the photon) and one massive combination (the Z^0) according to:

$$A_\mu = B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W,$$
$$Z_\mu = -B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W. \tag{5.69}$$

If we write the neutral part portion of the electroweak interaction with the physical A_μ and Z_μ states, we obtain

$$-\mathrm{i}[g_W j_\mu^3 W^{\mu 3} + \frac{g'}{2}j_\mu^Y B^\mu] = -\mathrm{i}\Big([g_W \sin\theta_W j_\mu^3 + (g'/2)\cos\theta_W j_\mu^Y]A^\mu + [g_W \cos\theta_W j_\mu^3 - (g'/2)\sin\theta_W j_\mu^Y]\Big) \tag{5.70}$$

so that the previous symmetry in isospin space [SU(2)], made evident by the coupling to the Pauli matrix vector $\boldsymbol{\tau}$, is now gone. Instead, by comparing this with the known electromagnetic coupling $-\mathrm{i}g_e j_\mu^{\mathrm{QED}} A^\mu$, and using that $j_\mu^{\mathrm{QED}} = j_\mu^3 + \frac{1}{2}j_\mu^Y$, we obtain consistency only if

$$g_W \sin\theta_W = g'\cos\theta_W = g_e. \tag{5.71}$$

This is the origin of the dependence of the interaction parameters on the mixing angle θ_W, and shows that the weak and electromagnetic coupling constants are not independent. A similar procedure for Z^0 yields $g_Z = g_E / \sin \theta_W \cos \theta_W$ as previously announced. We can also read out the vector and axial couplings $\{c_V, c_A\}$ for neutral weak processes, as we show in the example below.

Example 18. Reading out coupling constants. Consider the process $\nu_e \to \nu_e + Z^0$. We can only have a contribution from j_μ^3 and not j_μ^{QED} since neutrinos only interact weakly. The total coupling to Z^μ can therefore be written as:

$$-\mathrm{i}g_Z (j_\mu^3 - \overbrace{\sin^2 \theta_W j_\mu^{\mathrm{QED}}}^{=0}) Z^\mu \tag{5.72}$$

By inserting $j_\mu^3 = \bar{\nu}_L \gamma_\mu \nu_L$, we obtain from the above equation

$$-\frac{\mathrm{i}g_Z}{2} \left[\bar{\nu}\gamma_\mu \left(\frac{1 - \gamma^5}{2} \right) \nu \right] Z^\mu. \tag{5.73}$$

From this expression, we then read out that $c_V^\nu = c_A^\nu = \frac{1}{2}$ by comparing with Eq. (5.29).

Although we have made a lot of progress here in terms of constructing a unified framework for weak and electromagnetic interactions, we still have not provided a detailed explanation of an important step of this task, namely how the gauge fields acquire mass. This will be remedied in the last chapter in this book which treats gauge theories and spontaneous symmetry breaking.

VI. QUANTUM CHROMODYNAMICS

Learning goals. After reading this chapter, the student should:

- Be able to compute the Feynman amplitude $\mathcal{M}$ for Feynman diagrams in quantum chromodynamics, and also be able to obtain decay rates and scattering cross sections from $\mathcal{M}$.

- Understand how to account for the composite structure of protons and neutrons when treating scattering processes via Feynman diagrams.

- Understand how the color quantum number affects particles in strong interactions and how it is accounted for quantitatively.

All that we have stated so far about electrons in QED applies to quarks as well after substituting $(-e)$ to either $2e/3$ or $(-e/3)$, depending on which quark we are considering. The problem which complicates our observation of how quarks behave is that they, in contrast to electrons, never appear freely: we must infer information indirectly through hadrons. We therefore start by considering two central examples of hadron-production via $e^- - e^+$ and $e - p$ scattering. Then, we proceed to develop the framework of QCD: the theory of how *colored* particles interact. We discuss Feynman rules, color factors, pair annihilation, and finally asymptotic freedom.

A. Hadron production via $e^- - e^+$ collisions

Consider the process $e^- + e^+ \to q + q^-$.
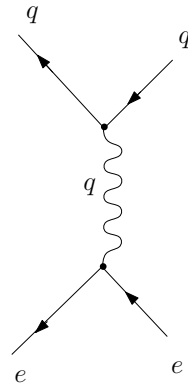
Briefly, the quarks escape as "free particles", but when their separation reaches $\sim 10^{-15}$ m the strong interaction is so great that it triggers a cascade of quark-antiquark pairs. The result is jet formation of hadrons.
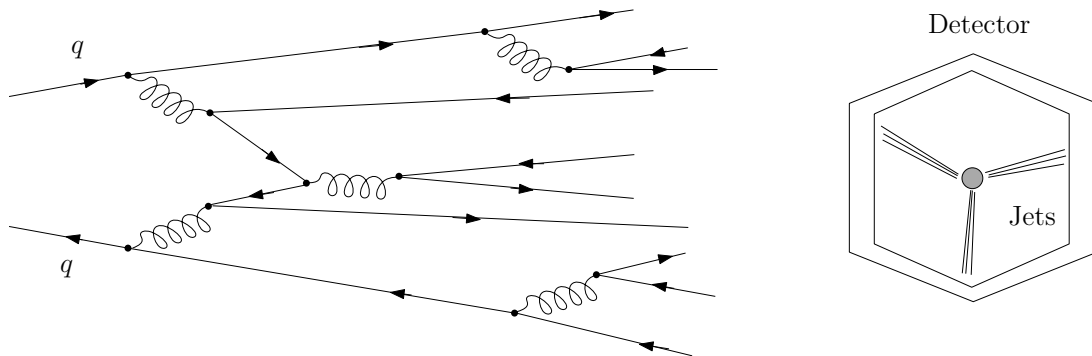
What is observed is then $e^- + e^+ \to$ hadrons. However, the fingerprint of quarks in this process is that the hadrons typically emerge in two back-to-back jets along the direction of q and $\bar{q}$, respectively. One also experimentally observes three-jet events, which is indicative of an emitted gluon carrying a portion of the energy. The observation of three-jet events is generally regarded as the most direct evidence for the existence of gluons.

Note that the first stage in the hadronization and jet formation ($e^- + e^+ \to \gamma \to q + \bar{q}$) is ordinary QED, so we obtain (evaluated in the CM frame):

$$\langle |\mathcal{M}|^2 \rangle = Q^2 g_e^4 \Big[1 + \Big(\frac{mc^2}{E} \Big)^2 + \Big(\frac{Mc^2}{E} \Big)^2 + \Big[1 - \Big(\frac{mc^2}{E} \Big)^2 \Big] \Big[\Big(\frac{Mc^2}{E} \Big)^2 \Big] \cos^2 \theta \Big] \tag{6.1}$$

which leads to a total scattering cross section

$$\sigma = \frac{\pi Q^2}{3} \Big(\frac{\hbar c \alpha}{E} \Big)^2 \sqrt{\frac{1 - (Mc^2/E)^2}{1 - (mc^2/E)^2}} \Big[1 + \frac{1}{2} \Big(\frac{Mc^2}{E} \Big)^2 \Big] \Big[1 + \frac{1}{2} \Big(\frac{mc^2}{E} \Big)^2 \Big]. \tag{6.2}$$

Here, Q is the quark charge in units of e (2/3 for u, c, t, and so forth) while M is the mass of the quark.

An interesting point is that σ becomes imaginary for $E < Mc^2$. Take a moment to reflect on why this happens - clearly, it must be nonphysical. The reason is that the process becomes kinematically forbidden for energies below Mc^2: there is not enough energy available to even produce the rest masses of the quarks. On the other hand, for high energies $E > Mc^2 \gg mc^2$, we obtain the simple expression

$$\sigma = \frac{\pi}{3} \Big(\frac{\hbar Q c \alpha}{E} \Big)^2. \tag{6.3}$$

There exists a number of thresholds for the energy as we increase it: first one is able to produce muons, light quarks, then (at $\simeq 1300$ MeV) the c quark, τ at 1777 MeV, b quark at 4500 MeV, and finally the t quark. A prediction for experiments would then be to consider the ratio

$$R = \frac{\sigma(e^- e^+ \to \text{ hadrons})}{\sigma(e^- e^+ \to \mu^- \mu^+)}. \tag{6.4}$$

Above each threshold, we then have according to our expression derived for σ above: $R(E) = 3 \sum_i Q_i^2$ where the sum is over all quark flavors with threshold below E. The factor 3 is due to the three possible colors for each flavor. According to this, we should have what resembles a "staircase" graph with a step up every time a new threshold is exceeded.
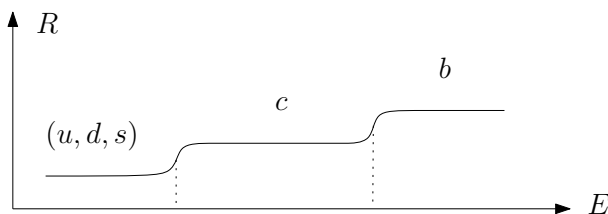
Comparison between experiment and theory is pretty good, except right next to the actual thresholds. One thing missed by our approximations is that the $q\bar{q}$ pairs are not truly free particles, although we have described them as if they were. Therefore, the hadronization process cannot in reality be split into two artificial steps $e^- e^+ \to q\bar{q}$ and then $q\bar{q} \to$ hadrons. For instance, it is possible to produce a bound state (such as $\phi = s\bar{s}$ or $\psi = c\bar{c}$) where the quarks are strongly interacting and our procedure fails. Such events show up as resonant peaks in σ. Nevertheless, it is noteworthy that the color factor of 3 is crucial for agreement between experiment and theory, and thus constitutes compelling evidence of the color quantum number.

B. Elastic $e - p$ scattering

Let us now see how we can probe the internal structure of the proton. *If p had no internal structure, we should be able to treat it as a point-particle and copy/paste our QED treatment of $e - \mu$ scattering.* However, since p is not a simple point charge, we need a more flexible formalism to account for such scattering. To lowest order in QED, we may represent the process as follows.
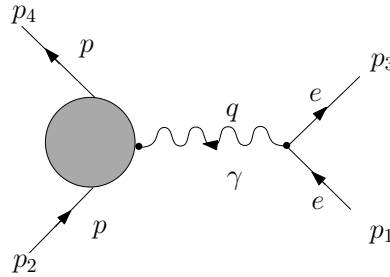
The dark blob represents the unknown exact structure of the photon-proton vertex interaction. What we do know, however, is that the $e - e$ vertex and γ propagator are the same as in QED. Therefore, we can write generally:

$$\langle |\mathcal{M}|^2 \rangle = \frac{g_e^4}{q^4} L_{\text{electron}}^{\mu\nu} K_{\mu\nu,\text{proton}} \tag{6.5}$$

where the tensor $K_{\mu\nu}$ is, so far, unknown. Since $K_{\mu\nu}$ is a second-rank tensor, it can only depend on three quantities: p_2, p_4 and q. But as $q = p_4 - p_2$, only two of these are independent. We choose q and p_2 and drop the subscript so that $p \equiv p_2$ is the initial proton momentum. With this in mind, the most general form of $K_{\mu\nu}$ reads:

$$K^{\mu\nu} = -K_1 g^{\mu\nu} + \frac{K_2}{(Mc)^2} p^\mu p^\nu + \frac{K_4}{(Mc)^2} q^\mu q^\nu + \frac{K_5}{(Mc)^2} (p^\mu q^\nu + p^\nu q^\mu). \tag{6.6}$$

In this way, all scalars K_i have the same dimension and are functions of the only scalar variable in the problem: q^2. Note that $p^2 = M^2 c^2$ and $p_4^2 = M^2 c^2 = (q + p)^2 \rightarrow q \cdot p = -q^2/2$. The K_i functions are nevertheless not independent of each other. One may show that $q_\mu K^{\mu\nu} = 0$ which in turn leads to
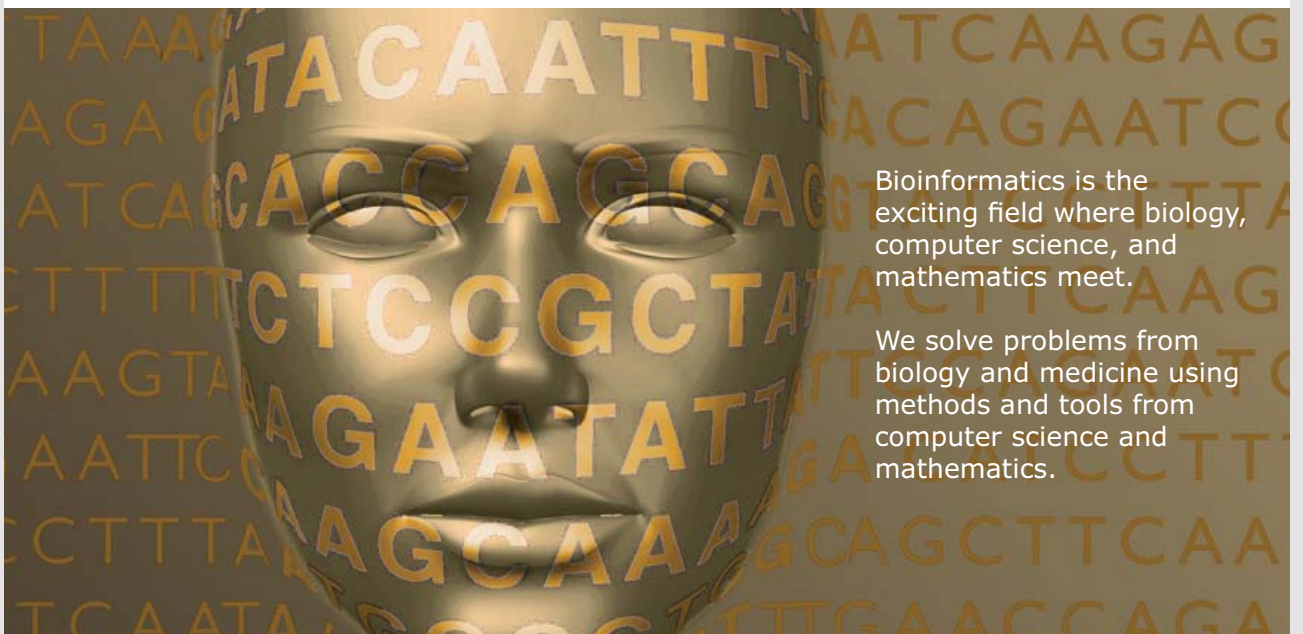
$$K_4 = (Mc)^2 K_1/q^2 + K_2/4, \ K_5 = K_2/2. \tag{6.7}$$

We have then managed to express $K^{\mu\nu}$ with two so-called *form factors* $K_1 = K_1(q^2)$ and $K_2 = K_2(q^2)$:

$$K^{\mu\nu} = K_1\Big(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\Big) + \frac{K_2}{(Mc)^2}(p^\mu + q^\mu/2)(p^\nu + q^\nu/2). \tag{6.8}$$

With this in hand, we may now evaluate $\langle|\mathcal{M}|^2\rangle$ in the standard way:

$$\langle|\mathcal{M}|^2\rangle = \Big(\frac{2g_e^2}{q^2}\Big)^2\Big[K_1[(p_1 p_3) - 2(mc)^2] + K_2[(p_1 p)(p_3 p)/(Mc)^2 + q^2/4]\Big]. \tag{6.9}$$

Working in the lab frame with the proton at rest, one obtains

$$\frac{d\sigma}{d\Omega} = \Big(\frac{\alpha\hbar}{4ME\sin^2(\theta/2)}\Big)^2\frac{E'}{E}[2K_1\sin^2(\theta/2) + K_2\cos^2(\theta/2)] \tag{6.10}$$

where E is the incident electron energy, E' is the outgoing electron energy, while θ is the scattered angle of the electron. The point is then that by experimentally counting the number of electrons scattered into a given direction for a specific range of energies, one determines $K_1(q^2)$ and $K_2(q^2)$. This gives us a way to model the internal structure of the proton. However, a complete theory should be able to *calculate* what K_1 and K_2 are. In the simplest model (proton as a point charge), one obtains $K_1 = -q^2$ and $K_2 = (2Mc)^2$. This is fine at low energies where e never gets close enough to see the inside of the proton. It fails dramatically at high energies, since the proton has a rich internal structure which comes into play at close enough distances.

C. Feynman rules for QCD

QCD is the interaction between colored particles via gluons. The strong coupling constant is $g_S = \sqrt{4\pi\alpha_S}$, so that α_S sets the strength of the force. Put simply, we may think of g_S as the fundamental unit of color. Now, to specify a quark state requires both a Dirac spinor $u^{(s)}(p)$ and a there-element column vector c which provides the color state. We have $c = [1,0,0]^T$ for red, $[0,1,0]^T$ for blue, and $c = [0,0,1]^T$ for green. Let c_i, $i = 1,2,3$ run over quark color indices. We can then have processes such as:
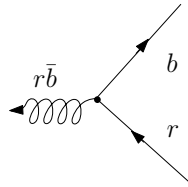
where a red quark turns into a blue quark and a gluon carries off the missing color. Initially, we may thus expect nine types of gluons since there are 3×3 combinations of color and anticolor ($r\bar{r}, r\bar{b}, \ldots$). These states can be represented as a color octet and a color singlet, which is a compatible representation with regard to the SU(3) symmetry that QCD is based on (more on this later). Specifically, the octet states read

$$|1\rangle = (r\bar{b} + b\bar{r})/\sqrt{2},\ |2\rangle = -\mathrm{i}(r\bar{b} - b\bar{r})/\sqrt{2},\ |3\rangle = (r\bar{r} - b\bar{b})/\sqrt{2},\ |4\rangle = (r\bar{g} + g\bar{r})/\sqrt{2},$$
$$|5\rangle = -\mathrm{i}(r\bar{g} - g\bar{r})/\sqrt{2},\ |6\rangle = (b\bar{g} + g\bar{b})/\sqrt{2},\ |7\rangle = -\mathrm{i}(b\bar{g} - g\bar{b})/\sqrt{2},\ |8\rangle = (r\bar{r} + b\bar{b} - 2g\bar{g})/\sqrt{6}. \tag{6.11}$$

while the singlet state is

$$|9\rangle = (r\bar{r} + b\bar{b} + g\bar{g})/\sqrt{3}. \tag{6.12}$$

All types of gluon combinations can be obtained using these nine states. For instance, a $r\bar{b}$ gluon is obtained by $(|1\rangle + \mathrm{i}|2\rangle)/\sqrt{2}$. Note that the state $|9\rangle$ is invariant under SU(3) transformations. That is why it is called a singlet, analogously to a $S_z = 0, S = 0$ spin-singlet state which is invariant under SU(2) spin rotations. We must now make a minor, but important modification compared to what we have stated initially in this book:

All naturally occurring particles are ~~colorless~~ color singlets.

A color singlet is not the same as what is meant by a colorless particle. For instance, $|3\rangle$ and $|8\rangle$ are colorless in the sense that they have equally much of color and the corresponding anticolor. However, they are not singlets. To see this, consider an analogy from spin. $S_z = 0$ does not imply $S = 0$: for instance, the $m = 0$ triplet state $(\uparrow\downarrow + \downarrow\uparrow)/\sqrt{2}$ has $S_z = 0$, but $S = 1$. On the other hand, $S = 0$ does necessarily imply $S_z = 0$. One can think of $|9\rangle$ as an invariant under color rotations in the same way as $\boldsymbol{r}^2 = x^2 + y^2 + z^2$ is invariant under spatial rotations.

Therefore, octet gluons do not appear as free particles, because they are not singlets. But if $|9\rangle$ existed, it should appear not only as a mediator but also as a *free particle*. Since gluons are massless, the existence of a free singlet gluon would imply that a long-ranged strong force exists. However, this has never been observed and we may conclude that only eight gluons exist. Strictly speaking, there are additional details regarding the gluon mass in relation to the so-called trace anomaly in QCD, which we do not go into here. The conclusion nevertheless remains that a free singlet gluon has not been experimentally observed.

In order to quantitatively describe gluons, let us start by noting that they are massless spin-1 particles, so that we can set the polarization vector ϵ^μ orthogonal to the momentum p_μ. Let us, as before, use the Coulomb gauge: $\epsilon^0 = 0$, $\boldsymbol{\epsilon} \cdot \boldsymbol{p} = 0$. Note that selecting a particular gauge spoils Lorentz covariance: doing a Lorentz transformation now demands that we also do a belonging gauge transformation to restore the Coulomb gauge. As always, the gauge-choice does not ultimately affect the physics: it is just a matter of mathematical convenience.

To describe the color of the gluon, we use an eight-element vector a:

$$a = [1, 0, 0, 0, 0, 0, 0, 0]^T \text{ for } |1\rangle, a = [0, 0, 0, 0, 0, 0, 1, 0]^T \text{ for } |7\rangle, \text{ and so forth.} \tag{6.13}$$

Gluons couple to each other since they carry color (unlike photons that do not carry charge).

Before stating the Feynman rules for QCD, we need to establish some notation (precisely as we also did for QED before stating the rules). The first thing is to introduce the Gell-Mann matrices. These are to SU(3) what the Pauli matrices are to SU(2):

$$\lambda^1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ \lambda^2 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ \lambda^3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ \lambda^4 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

$$\lambda^5 = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}, \ \lambda^6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \ \lambda^7 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \ \lambda^8 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}. \tag{6.14}$$

We also need the commutators of λ-matrices which define the structure constants $f^{\alpha\beta\gamma}$ of SU(3) via the relation:

$$[\lambda^\alpha, \lambda^\beta] = 2i f^{\alpha\beta\gamma} \lambda^\gamma \tag{6.15}$$

where summation over γ from 1 to 8 is implied due to the repeated index. The structure constants are antisymmetric, so that $f^{\alpha\beta\gamma} = -f^{\alpha\gamma\beta} = -f^{\alpha\beta\gamma}$. With notation in hand, we can now write down the Feynman rules of QCD.

External lines. A quark or antiquark with momentum p, spin s, and color c is represented as follows.

Quark		Antiquark	
Incoming $\longrightarrow\bullet$	$u^{(s)}(p)c$	Incoming $\longleftarrow\bullet$	$\bar{v}^{(s)}(p)c^\dagger$
Outgoing $\bullet\longrightarrow$	$\bar{u}^{(s)}(p)c^\dagger$	Outgoing $\bullet\longleftarrow$	$v^{(s)}(p)c$
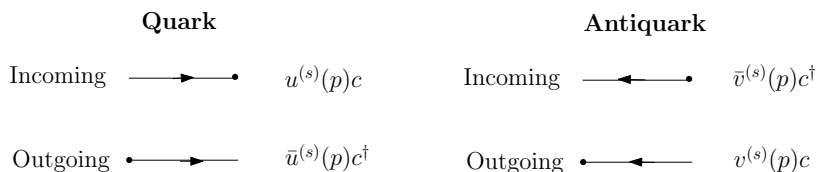
A gluon with momentum p, polarization ϵ, and color a is written as follows.

Incoming $\quad$ α, μ $\quad\quad$ $\epsilon_\mu(p)a^\alpha$

Outgoing $\quad$ α, μ $\quad\quad$ $\epsilon_\mu^*(p)(a^\alpha)^*$

Propagators.

Quarks & Antiquarks $\quad\quad q \quad\quad$ $\dfrac{i(q^\mu \gamma_\mu + mc)}{q^2 - m^2 c^2}$

Gluons $\quad\quad q \quad\quad$ $-\dfrac{ig_{\mu\nu}\delta^{\alpha\beta}}{q^2}$

$\alpha, \mu \quad\quad\quad\quad \beta, \nu$

Vertices. Each vertex introduces a factor as follows.

Quark-gluon

$-\frac{ig_s}{2}\lambda^\alpha\gamma^\mu$

Three-gluon

$-g_S f^{\alpha\beta\gamma}[g_{\mu\nu}(k_1-k_2)_\lambda + g_{\nu\lambda}(k_2-k_3)_\mu + g_{\lambda\mu}(k_3-k_1)_\nu]$

(notice how gluon momenta point into the vertex)

Four-gluon

$-ig_S^2[f^{\alpha\beta\eta}f^{\gamma\delta\eta}(g_{\mu\lambda}g_{\nu\rho} - g_{\mu\rho}g_{\nu\lambda}) + f^{\alpha\delta\eta}f^{\beta\gamma\eta}$
$\times(g_{\mu\nu}g_{\lambda\rho} - g_{\mu\lambda}g_{\nu\rho}) + f^{\alpha\gamma\eta}f^{\delta\beta\eta}(g_{\mu\rho}g_{\nu\lambda} - g_{\mu\nu}g_{\lambda\rho})]$
(sum over η implied)
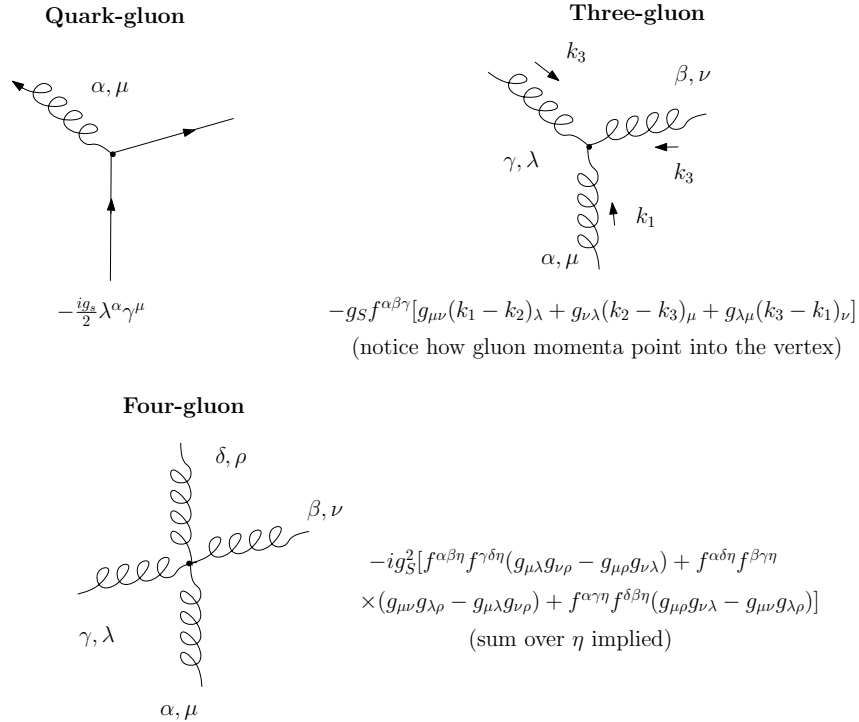
Apart from these rules, the same applies as in QED. In what follows, we shall have a look at some concrete examples of this framework in action.

D. Color factors

Consider the interaction between two quarks to lowest-order QCD. The focus will be on saying something on the effective potential between quarks, analogously to the QED Coulomb potential. Keep in mind that we are effectively doing a perturbation theory calculation, assuming that α_S is small. Because of this, however, we cannot hope to obtain the result of *asymptotic freedom* at lowest order since we should only be able to describe the short-range potential behavior where the strong interaction is much weaker in magnitude compared to large separation distances between the quarks. We will discover a central result (to be shown quantitatively below):

> Quarks attract most strongly when they are in a color singlet configuration.

Quark & Antiquark.
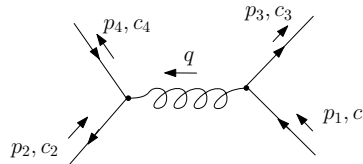Consider then $q\bar{q}$ scattering with different flavors (such as $u + \bar{d} \to u + \bar{d}$). The lowest-order diagram looks as follows (the small arrows emphasize the direction of momentum flow).

The amplitude is obtained according to the rules:

$$\mathcal{M} = i[\bar{u}(3)c_3^\dagger][-ig_S\lambda^\alpha\gamma^\mu/2][u(1)c_1][-ig_{\mu\nu}\delta^{\alpha\beta}/q^2][\bar{v}(2)c_2^\dagger][-ig_s\lambda^\beta\gamma^\nu/2][v(4)c_4]$$

$$= -\frac{g_S^2}{4q^2}[\bar{u}(3)\gamma^\mu u(1)][\bar{v}(2)\gamma_\mu v(4)](c_3^\dagger\lambda^\alpha c_1)(c_2^\dagger\lambda^\alpha c_4). \tag{6.16}$$

Now, the structure is identical to $e - \bar{e}$ scattering except for (1) $g_e \rightarrow g_S$ and (2) the additional color factor $f = \frac{1}{4}(c_3^\dagger \lambda^\alpha c_1)(c_2^\dagger \lambda^\alpha c_4)$. Thus, we can state that the potential describing $q\bar{q}$ interactions is the same as in QED for opposite charges when replacing α with $f\alpha_S$, so that we can at least phenomenologically introduce the potential:

$$V_{q\bar{q}}(r) = -\frac{f\alpha_S\hbar c}{r}. \tag{6.17}$$

The factor f depends on the color state of the interacting quarks: we can create a singlet state or one of the color octets. In the latter case, all octets actually turn out to yield the same f. Let us compute this one to begin with and choose the combination $r\bar{b}$ to be concrete. Color is conserved overall, and thus the incoming q has the same color as the outgoing q, namely red. Conversely, the incoming and outgoing $\bar{q}$ have antiblue color. Therefore,

$$c_1 = c_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ c_2 = c_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \tag{6.18}$$

In turn, this produces the following color factor:

$$f = \frac{1}{4}\Big([1,0,0]\lambda^\alpha[1,0,0]^T\Big)\Big([0,1,0]\lambda^\alpha\Big) = \frac{1}{4}\lambda_{11}^\alpha\lambda_{22}^\alpha. \tag{6.19}$$

Only the λ^3 and λ^8 matrices have non-zero entries in the (11) and (22) positions, so that $f = -\frac{1}{6}$.

We can now repeat the same procedure for the singlet state $\frac{1}{\sqrt{3}}(r\bar{r} + b\bar{b} + g\bar{g})$, which gives $f = \frac{4}{3}$. We have then obtained:

$$V_{q\bar{q}}(r) = \begin{cases} -\frac{4}{3}\frac{\alpha_S\hbar c}{r} & \text{for color singlet} \\ \frac{1}{6}\frac{\alpha_S\hbar c}{r} & \text{for color octet.} \end{cases} \tag{6.20}$$

It is clear that the force is attractive for the singlet state, which helps to explain why $q\bar{q}$ bindings (mesons) occur as color singlets but not octets (which would have produced colored mesons).

Quark & quark.
Consider again different flavors (*e.g.* $u + d \rightarrow u + d$). The lowest-order amplitude for such scattering is:

$$\mathcal{M} = -\frac{g_S^2}{4q^2}[\bar{u}(3)\gamma^\mu u(1)][\bar{u}(4)\gamma_\mu u(2)](c_3^\dagger \lambda^\alpha c_1)(c_4^\dagger \lambda^\alpha c_2). \tag{6.21}$$

The procedure is then similar to $e - \mu$ scattering except for (1) $g_e \rightarrow g_S$ and (2) the additional color factor $f = (c_3^\dagger \lambda^\alpha c_1)(c_4^\dagger \lambda^\alpha c_2)$. The effective potential becomes formally equivalent to that of like charges in electrodynamics:

$$V_{qq}(r) = \frac{f\alpha_S\hbar c}{r}. \tag{6.22}$$

However, the color configuration for two quarks cannot be singlet. Instead, we have

$$\text{Triplet (antisymmetric)} : \begin{cases} (rb - br)/\sqrt{2} \\ (bg - gb)/\sqrt{2} \\ (gr - rg)/\sqrt{2} \end{cases} \tag{6.23}$$

$$\text{Sextet (symmetric)} : \begin{cases} rr, bb, gg \\ (rb + br)/\sqrt{2} \\ (bg + gb)/\sqrt{2} \\ (gr + rg)/\sqrt{2} \end{cases}. \tag{6.24}$$

Performing the calculation as before for the color factors, we find

$$V_qq(r) = -\frac{2}{3}\frac{\alpha_S\hbar c}{r} \text{ for the triplet state, } V_qq(r) = \frac{1}{3}\frac{\alpha_S\hbar c}{r} \text{ for the sextet state.} \tag{6.25}$$

The different sign is obvious, but at the same time we know that neither occurs in nature in spite of the attractive interaction for the triplet state. To understand this, we must realize that the short-range attraction does not *prove* that binding occurs: for that, we would have to know about the long-range behavior for the same color configuration. The above result nevertheless has important consequences for the binding of three quarks where you can show that complete mutual attraction between three quarks is obtained only in a singlet configuration.

E. Briefly on asymptotic freedom

In the QED chapter of this book, we found that loop diagrams of the type caused the effective charge of the electron to be a function of the momentum transfer q:

$$\alpha(q) = \alpha(0)\Big[1 + \frac{\alpha(0)}{3\pi}\ln\Big(\frac{|q|^2}{(mc)^2}\Big)\Big] \tag{6.26}$$

for $|q|^2 \gg (mc)^2$. The physics in this case is that as the charges get closer to each other (larger $|q|^2$), the coupling strength increases due to the vacuum polarization. The screening is less effective at shorter distances. We recall that this type of diagram also introduces a divergent term that we "soak up" in an effective charge (which is what is experimentally measured). Now, the above equation holds to order $\mathcal{O}\{[\alpha(0)]^2\}$. The contribution from higher-order diagrams such as:
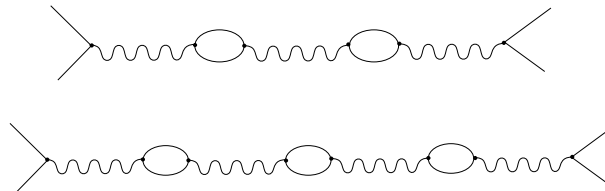

can be performed a summation over explicitly and provides the result:

$$\alpha(q) = \frac{\alpha(0)}{1 - [\alpha(0)/3\pi]\ln\Big(\frac{|q|^2}{(mc)^2}\Big)}. \tag{6.27}$$

We can understand why the expression looks like this when taking into account higher-order loop diagrams since we obtain a series of the type $1 + x + x^2 + \ldots = \frac{1}{1-x}$ where x represents a bubble.

Now, much of the same thing happens in QCD as well where $q - \bar{q}$ bubbles screen quark color and gives (modulo color factors f) the same as Eq. (6.26). However, QCD has a twist that is not present in QED. There are also *virtual gluon bubbles* due to the gluon-gluon coupling:
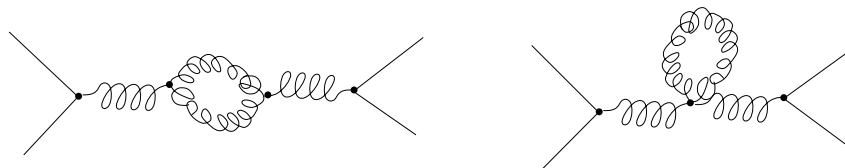
The gluon contribution actually works in the opposite direction as the $q - \bar{q}$ bubbles and thus produces antiscreening or "camouflage". The formula for the running coupling constant in QCD reads:

$$\alpha_S(|q|^2) = \frac{\alpha_S(\mu^2)}{1 + [\alpha_S(\mu^2)/12\pi](11n - 2f)\ln\Big(\frac{|q|^2}{\mu^2}\Big)}, \ |q|^2 \gg \mu^2. \tag{6.28}$$

Here, n is the number of colors and f is the number of flavors. In the Standard Model, we know that $n = 3$ and $f = 6$. As a result, $11n > 2f$ and the coupling constant α_S decreases as $|q|^2$ increases. This means that short

distances, the strong interaction in fact becomes quite weak in magnitude. This is an important consequence of asymptotic freedom $\rightarrow$ our license to use perturbation theory (Feynman diagrams) for interquark potentials.

But we have not specified what μ is yet. In electrodynamics it is natural to define the charge of a particle as its long-range (fully screened) value. We could, however, use the effective charge at any $|q|^2$ as the reference value so long as $\alpha(|q|^2)$ is small there (so that we are allowed to use Feynman perturbation theory). The problem in QCD is that α_S is large when $q^2 \rightarrow 0$, so we cannot use that as our reference point. Instead, we use $\alpha_S(\mu^2) < 1$ as the "bare" strength of the coupling constant which we base our perturbation expansion on. The quantity μ is therefore chosen so that α_S satisfies this. Note that $\alpha_S(|q|^2)$ varies substantially over the experimentally available energy range whereas $\alpha(|q|^2)$ varies much less.

VII. RELATIVISTIC FIELD THEORY AND GAUGE THEORIES

Learning goals. After reading this chapter, the student should:

- Be able to derive the equation of motion for a relativistic field Lagrangian $\mathcal{L}$ and interpret the terms present in $\mathcal{L}$ physically.

- Be able to identify the Klein-Gordon, Dirac, and Proca Lagrangians.

- Understand the concepts of local gauge invariance, spontaneous symmetry breaking, and how these are related to the Higgs mechanism.

We here assume familiarity with the Lagrangian formulation of classical particle mechanics (see *e.g.* the free textbook "Introduction to Lagrangian and Hamiltonian Mechanics" available to download from Bookboon), and develop Lagrangian field theory. Then, we proceed to introduce the fundamental concepts of local gauge invari-anec, spontaneous symmetry breaking, and the Higgs mechanism.

A. Lagrangians in relativistic field theory

The first issue to address is fundamental: what is a field? When we discuss particles, we think of localized entities. In classical physics, one is typically interested in identifying the position $x = x(t)$ of a particle as a function of time t. On the other hand, a field occupies some region of space and time, and we are typically interested in identfiying the value of the field $\phi_i = \phi_(x, y, t)$ at a given position in space and time. Temperature or electric potential are classical examples of fields. Here, we shall use fields to describe relativistic particles.

In field theory, one starts with a Lagrangian density $\mathcal{L}$ which is a function of ϕ_i and its derivatives

$$\partial_\mu \phi_i \equiv \frac{\partial \phi_i}{\partial x^\mu} \tag{7.1}$$

The Euler-Lagrange equation reads

$$\partial_\mu \Big(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \Big) = \frac{\partial \mathcal{L}}{\partial \phi_i}, \ i = 1, 2, 3, \dots \tag{7.2}$$

These equations resemble their equivalents in classical, non-relativistic mechanics, but a key difference here is that space and time coordinates are treated equally due to the special theory of relativity.

Example 19. Klein-Gordon Lagrangian. This $\mathcal{L}$ describes a scalar spin-0 field:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2}(mc/\hbar)^2 \phi^2. \tag{7.3}$$

To compute the corresponding equation of motion for the field ϕ, i.e. Eq. (7.2), we first observe that

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} = \partial^\mu \phi. \tag{7.4}$$

To see this, note that we can write the Lagrangian as

$$\mathcal{L} = \frac{1}{2}(\partial_0 \phi \partial_0 \phi - \partial_1 \phi_1 \partial_1 \phi_1 - \dots) \tag{7.5}$$

from which it becomes clear that

$$\frac{\partial}{\partial(\partial_0\phi)} = \partial_0\phi = \partial^0\phi, \quad \frac{\partial}{\partial(\partial_i\phi)} = -\partial_i\phi = \partial^i\phi. \tag{7.6}$$

Moreover, we note that

$$\frac{\partial\mathcal{L}}{\partial\phi} = -\frac{mc^2}{\hbar}\phi. \tag{7.7}$$

The total equation then takes the form:

$$\partial_\mu\partial^\mu\phi + (mc/\hbar)^2\phi = 0 \tag{7.8}$$

which is precisely the Klein-Gordon equation we have discussed previously in this book.

In the same way as shown in the above example, one can show that from the Dirac Lagrangian which describes a spin-1/2 field:

$$\mathcal{L} = i(\hbar c)\bar{\psi}\gamma^\mu\partial_\mu\psi - (mc^2)\bar{\psi}\psi \tag{7.9}$$

that the Euler-Lagrange equation provides the Dirac equation:

$$i\gamma^\mu\partial_\mu\psi - (mc/\hbar)\psi = 0. \tag{7.10}$$

Here, ψ and $\bar{\psi}$ are treated as independent fields. Moreover, the Proca Lagrangian describes a massive vector (spin-1) field:

$$\mathcal{L} = -\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{8\pi}\left(\frac{mc^2}{\hbar}\right)A^\nu A_\nu \tag{7.11}$$

and leads to the Proca equation

$$\partial_\mu F^{\mu\nu} + (mc^2/\hbar)A^\nu = 0. \tag{7.12}$$

In the special limit of a massless field $m = 0$, this reduces to Maxwell's equations for empty space. A *source term*, representing charge and current densities, can be added by writing

$$\mathcal{L} = -\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} - \frac{1}{c}J^\mu A_\mu \tag{7.13}$$

which provides the equation

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c}J^\nu. \tag{7.14}$$

Note that it follows that $\partial_\nu J^\nu = 0$: the current inserted must satisfy the continuity equation.

Now, where do all the above Lagrangians come from? We have simply written them down without any justification. In fact, they were all chosen so that they would provide the correct field equations for their corresponding fields. Therefore, in contrast to deriving the classical, non-relativistic $L = T - U$ where T is kinetic energy and U is potential energy, $\mathcal{L}$ is axiomatic in relativistic field theory. Similarly to classical, non-relativistic Lagrangians, $\mathcal{L}$ is not unique. We can add a simple constant or a divergence $\partial_\mu M^\mu$ without affecting the equation of motion Eq. (7.2) for the field.

B. Local gauge invariance

The Dirac Lagrangian is invariant under the transformation $\psi \rightarrow \mathrm{e}^{\mathrm{i}\theta}\psi$. This is a global transformation as θ is independent on position and time. Moreover, θ is a real number here. Therefore, since $\bar{\psi} \rightarrow \mathrm{e}^{-\mathrm{i}\theta}\bar{\psi}$, all exponential factors cancel due to the fact that it is the combination $\bar{\psi}\psi$ that appears in $\mathcal{L}_{\mathrm{Dirac}}$. When $\theta = \theta(x)$, the above is a local gauge transformation. The phase-factor is usually thought of as a mere convention: as known from quantum mechanics, the phase of a wavefunction is unobservable, and only the difference in phases between wavefunctions may be observed. A local gauge symmetry just means that we should be able to freely change the convention (reference point) of the phase at any point in space-time without changing the physics, and not only change the reference point globally (i.e. by equally much at each point in space-time). However, there seems to be a problem: $\mathcal{L}_{\mathrm{Dirac}}$ is *not invariant* under a local transformation $\psi \rightarrow \psi \mathrm{e}^{\mathrm{i}\theta(x)}$. Instead, it transforms as $\mathcal{L} \rightarrow \mathcal{L} - \hbar c (\partial_\mu \theta)\bar{\psi}\gamma^\mu \psi$. We define $\lambda(x) \equiv -(\hbar c/q)\theta(x)$, so that when $\psi \rightarrow \mathrm{e}^{-\mathrm{i}q\lambda(x)/\hbar c}\psi$, we obtain

$$\mathcal{L} \rightarrow \mathcal{L} + (q\bar{\psi}\gamma^\mu \psi)\partial_\mu \lambda. \tag{7.15}$$

We see that there is a way in which to reobtain local gauge invariance, namely if we add a term to $\mathcal{L}$ which cancels the extra contribution obtained during the local gauge transformation. Assume that

$$\mathcal{L} = [\mathrm{i}\hbar c\bar{\psi}\gamma^\mu \partial_\mu \psi - mc^2 \bar{\psi}\psi] - (q\bar{\psi}\gamma^\mu \psi)A_\mu, \tag{7.16}$$

where A_μ is the so-called gauge field. If the fields now transform like: $\psi \rightarrow \mathrm{e}^{-\mathrm{i}q\lambda(x)/\hbar c}\psi$ and $A_\mu \rightarrow A_\mu + \partial_\mu \lambda$, then $\mathcal{L}$ is invariant under such a local gauge transformation. Whereas this mathematically restores the gauge invariance, we still have to physically justify this procedure. This can be done by realizing that the term $q\bar{\psi}\gamma^\mu \psi A_\mu$ describes the coupling between the vector field A_μ and the fermion field ψ. If A_μ is present, however, there should in all fairness also exist a "free" term that describes A_μ in $\mathcal{L}$ irrespective of whether or not a fermion field ψ is present that it can couple to. This can be accomplished via the Proca Lagrangian with the mass term set to zero:

$$\mathcal{L} = -F^{\mu\nu}F_{\mu\nu}/16\pi \tag{7.17}$$

since $F^{\mu\nu}$ is invariant under the gauge transformation in question (whereas a mass term $\propto A^\nu A_\nu$ is not). As a result, we arrive at the important conclusion that:

> Imposing local gauge-invariance on the Dirac $\mathcal{L}$, we must introduce a massless vector field A_μ.

The full Lagrangian describing a spin-1/2 fermion field coupled to a massless spin-1 boson field then reads:

$$\mathcal{L} = [\mathrm{i}\hbar c\bar{\psi}\gamma^\mu \partial_\mu \psi - mc^2\bar{\psi}\psi] + \left[-\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} \right] - [q\bar{\psi}\gamma^\mu\psi A_\mu].$$

We identify A_μ as the electromagnetic vector potential since:

1. $A_\mu \to A_\mu + \partial_\mu \lambda$ leaves $F^{\mu\nu}$ invariant.

2. The two last terms in the above $\mathcal{L}$ give the Maxwell Lagrangian with a source term $J^\mu = cq(\bar{\psi}\gamma^\mu\psi)$, and this is precisely the current produced by Dirac particles.

We have therefore identified a Lagrangian which describes electrodynamics through the coupling between Dirac fermions and photons, in particular the current J^μ produced by the former. Remarkably, the *origin* of this coupling was that we demanded *local gauge invariance*.

The main difference, mathematically, between global and local gauge transformations arises due to derivatives of the field. Let

$$\partial_\mu \to D_\mu \equiv \partial_\mu + \mathrm{i}\frac{q}{\hbar c}A_\mu \tag{7.18}$$

to convert a global gauge invariance to a local one. The above relation is known as *minimal coupling*. Note that the gauge field here must be massless. In terms of symmetry, we have a local U(1) gauge symmetry since $\mathrm{e}^{\mathrm{i}\theta}$ belongs to U(1). The mathematical structure here may be extended to a Lagrangian with two spin-1/2 fields ψ_1 and ψ_2, in which case we obtain a local SU(2) gauge symmetry. This is an example for the so-called Yang-Mills theories which we shall treat in more detail later.

Before moving on to discussing how various terms in $\mathcal{L}$ can be interpreted physically in general, let us formally define precisely what is meant by a *gauge theory*:

A gauge theory is a field theory described by a Lagrangian $\mathcal{L}$ that is invariant under a group of local continuous transformations. The mathematical procedure which adjusts the degrees of freedom related to the invariance of $\mathcal{L}$ is referred to as the gauge.

C. Interpreting $\mathcal{L}$: importance of the mass term

In general, $\mathcal{L}$ consists of two kinds of terms.

- Free Lagrangian for each field ($\mathcal{L}_0$).
- Interaction terms for the fields ($\mathcal{L}_{\mathrm{int}}$).

$\mathcal{L}_{\mathrm{int}}$ may be obtained by invoking local gauge invariance, as we showed in the above treatment of the Dirac and Proca Lagrangians. Local gauge invariance thus provides a way to identify the vertex couplings in a theory, and it works beautifully for strong and electromagnetic interactions. However, in the weak interactions the gauge field is far from massless. It is nevertheless still possible to create a gauge theory with massive gauge fields via spontaneous symmetry breaking and the Higgs mechanism. Before exploring these phenomena, we consider how to identify a mass term in $\mathcal{L}$ in the first place.

Example 20. Identifying the mass term. What is the mass term in the following Lagrangian?

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) + \mathrm{e}^{-(\alpha\phi)^2}. \tag{7.19}$$
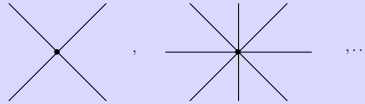
Let us expand $\mathcal{L}$ around the point $\phi = 0$:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) + 1 - \alpha^2 \phi^2 + \frac{1}{2}\alpha^4 \phi^4 - \ldots \tag{7.20}$$

Since constants do not affect the equations of motion for the field (although constants do contribute to the so-called cosmological constant), we discard them and thus identify the mass term from the $\mathcal{O}(\phi^2)$-term:

$$m = \frac{\sqrt{2}\alpha\hbar}{c}. \tag{7.21}$$

The higher-order terms [$\mathcal{O}(\phi^4)$ and higher] represent couplings of the type:

However, we cannot always simply extract the mass from the term which is second order in the field. An exception is for instance:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) + \frac{1}{2}\mu^2 \phi^2 - \frac{1}{4}\lambda^2 \phi^4. \tag{7.22}$$

The sign is the problem here: it looks like the mass should be imaginary, since the sign is opposite to the previous examples we have encountered.

To resolve this problem, we have realize that Feynman calculus using field Lagrangians is really a *perturbation theory* starting from the ground-state. The higher order terms ($\phi^4, \phi^6, \ldots$) then represent higher order corrections to the ground-state. But this is because, so far, the field configuration that gives the minimum energy has been $\phi = 0$. In Eq. (7.22), this is no longer the case. We thus have to rewrite $\mathcal{L}$ in terms of deviations from the ground-state. To do so, we first have to identify the value of ϕ which gives the ground-state. We can do so by conjecturing that $\mathcal{L} = \mathcal{T} - \mathcal{U}$ (just like in classical, non-relativistic Lagrangian theory), where

$$\mathcal{T} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) \tag{7.23}$$

is the kinetic energy while

$$\mathcal{U} = -\frac{1}{2}\mu^2 \phi^2 + \frac{1}{4}\lambda^2 \phi^4 \tag{7.24}$$

is the potential energy. We then see that the minimum of $\mathcal{U}(\phi)$ occurs at $\phi = \phi_m = \pm\mu/\lambda$. We introduce a new field η as the deviation from the minimum value:
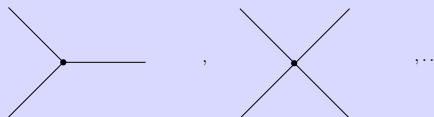
$$\eta \equiv \phi \pm \mu/\lambda \tag{7.25}$$

and then express $\mathcal{L}$ in terms of the new field η:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2 \eta^2 \pm \mu\lambda\eta^3 - \frac{1}{4}\lambda^2 \eta^4 + \frac{1}{4}(\mu^2/\lambda)^2. \tag{7.26}$$

Now, we can identify the mass term since the second order term in η has the appropriate sign:

$$m = \frac{\sqrt{2}\mu\hbar}{c} \tag{7.27}$$

and the other terms in $\mathcal{L}$ are couplings of the type:

Note how the $\mathcal{L}$'s expressed with ϕ and η represent exactly the same physical system - we have only changed the notation. However, the key insight here is that the ϕ-version is not suitable for Feynman calculus since a perturbation series in ϕ would not converge (as $\phi = 0$ is not the ground-state).

The conclusion is then that in order to identify the mass term, we must do the following:

1. Locate the ground-state.

2. Express $\mathcal{L}$ as a function of the deviation η from the ground-state.

3. Expand $\mathcal{L}$ in powers of η: the mass term comes from the η^2 term.

What about linear terms in $\mathcal{L}$?
Assume that we have a Lagrangian of the form:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - a\phi^2 - b\phi^3 - c\phi. \tag{7.28}$$

Introduce the field $\eta \equiv \phi - \phi_0$ and write

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - A\eta^2 - B\eta^3. \tag{7.29}$$

Now identify $\{A, B, \phi_0\}$ so that

$$a\phi^2 + b\phi^3 + c\phi = A\eta^2 + B\eta^3. \tag{7.30}$$

This gives us three equations: one for each order of ϕ. Therefore, a linear term can simply be absorbed into a new field (η above) which describes the deviation from some constant value of the original field (ϕ in the above example). However, note that $\eta = 0$ is not necessarily the ground-state of the field: in fact, in general it will not be. So even though we can drop the linear term without loss of generality when considering a generic form of $\mathcal{L}$, we must pay attention to how it modifies the constants for a *specific theory* so that we may correctly identify the ground-state around which the field is expanded (*e.g.* in the potential energy).

D. Spontaneous symmetry breaking

This phenomenon refers to a situation where the Lagrangian (or equations of motion) have a given symmetry, whereas the ground-state does not share the same symmetry. Let us borrow an example from condensed matter theory, namely a ferromagnetic material. In the simplest model, the Lagrangian (or Hamiltonian, if you prefer) only depends on the magnitude of the magnetization $M = |\boldsymbol{M}|$. Therefore, it does not matter energywise in which direction the magnetization points in such a material: all states are equivalent. However, as soon as a material becomes magnetic, the magnetization *has to point in some direction*. In other words, the ground-state has chosen *one particular solution* out of all the available ones that have equal energy. In this way, the ground-state has lowered the symmetry that the original Lagrangian had. The word "breaking" could therefore in some sense be substituted by "choosing": spontaneous symmetry breaking occurs when choosing one particular ground-state out of many available.

Example 21. Mexican hat potential. Consider the Lagrangian for a two-component field (with components ϕ_1 and ϕ_2):

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi_1)(\partial^\mu \phi_1) + \frac{1}{2}(\partial_\mu \phi_2)(\partial^\mu \phi_2) + \frac{1}{2}\mu^2(\phi_1^2 + \phi_2^2) - \frac{1}{4}\lambda^2(\phi_1^2 + \phi_2^2)^2. \tag{7.31}$$

This is invariant under rotations in (ϕ_1, ϕ_2)-space. This means that we have an SO(2) symmetry: $\mathcal{L}$ is invariant under the transformation

$$\begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} \rightarrow \begin{bmatrix} \phi_1' \\ \phi_2' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix}. \tag{7.32}$$

We recognize the rotation matrix in 2D, which belongs to the SO(2) group. The minimum of the potential in $\mathcal{L}$ is given by the equation

$$\phi_{m,1}^2 + \phi_{m,2}^2 = \mu^2/\lambda^2. \tag{7.33}$$

Circle of potential minima
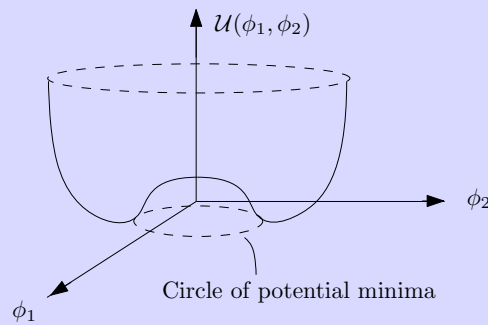
Thus, we in fact have an entire circle of minima. In order for us to be allowed to use Feynman calculus, we thus expand around the ground-state. This means that we must *choose* one particular ground-state on this circle. Note that the ground-state is often referred to simply as *vacuum* for a given model. The physics should not depend on which ground-state we choose, so we might as well take

$$\phi_{m,1} = \mu/\lambda, \ \phi_{m,2} = 0. \tag{7.34}$$

The fluctuation fields are then η and ξ where:

$$\eta \equiv \phi_1 - \mu/\lambda, \ \xi = \phi_2. \tag{7.35}$$

We rewrite $\mathcal{L}$ in terms of these to obtain:

$$\mathcal{L} = [\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2 \eta^2] + [\frac{1}{2}(\partial_\mu \xi)(\partial^\mu \xi)]$$
$$+ [\mu\lambda(\eta^3 + \eta\xi^3) - \frac{\lambda^2}{4}(\eta^4 + \xi^4 + 2\eta^2\xi^2)] + \frac{\mu^4}{4\lambda^2}. \tag{7.36}$$

The terms can now be interpreted as follows.

- 1st term: free K-G Lagrangian with mass $m_\eta = \sqrt{2}\mu\hbar/c$.

- 2nd term: free K-G Lagrangian with mass $m_\xi = 0$.

- 3rd term: five coupling terms of the type
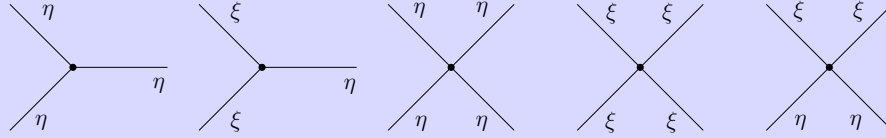
- 4th term: irrelevant constant.

The Lagrangian is no longer symmetric in any of the fields since we have chosen to express in terms of one particular solution. One of the new fields, ξ, is actually massless.

The above is an example of a more fundamental theorem, namely:

> **Goldstone's theorem:** Spontaneous breaking of a continuous global symmetry always generates one or more massless scalar (spin-0) particles, called Goldstone bosons

This does not seem very helpful for our purpose - we were looking for a way to generate massive gauge fields, not massless ones. As we shall now see, the gauge fields become massive when we apply spontaneous symmetry breaking to a *local* gauge invariance rather than global gauge invariance.

E. Higgs mechanism

Let us first make the notation a bit more convenient

$$\phi \equiv \phi_1 + i\phi_2 \rightarrow \phi^*\phi = \phi_1^2 + \phi_2^2. \tag{7.37}$$

Then, the Lagrangian we have considered takes the form:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu\phi)^*(\partial^\mu\phi) + \frac{1}{2}\mu^2(\phi^*\phi) - \frac{1}{4}\lambda^2(\phi^*\phi)^2. \tag{7.38}$$

This $\mathcal{L}$ is invariant under U(1) phase transformations $\phi \rightarrow e^{i\theta}\phi$. Note that there is no contrast between this result and our previous statement that the Lagrangian had an SO(2) symmetry. The reason for this is that the groups SO(2) and U(1) are *isomorphic* (see our previous treatment of symmetry groups), so we have not changed anything but notation.

We can now make $\mathcal{L}$ invariant under a local U(1) gauge, by introducing a massless A^μ field and a minimal coupling

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + i\frac{q}{\hbar c}A_\mu. \tag{7.39}$$

Then, the following $\mathcal{L}$ has a local gauge symmetry:

$$\mathcal{L} = \frac{1}{2}\left[\left(\partial_\mu - \frac{iq}{\hbar c}A_\mu\right)\phi^*\right]\left[\left(\partial^\mu + \frac{iq}{\hbar c}A^\mu\right)\phi\right] + \frac{1}{2}\mu(\phi^*\phi) - \frac{1}{4}\lambda^2(\phi^*\phi)^2 - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}. \tag{7.40}$$

If we again define the fluctuation fields $\eta = \phi_1 - \mu/\lambda$ and $\xi = \phi_2$, we obtain

$$\mathcal{L} = \left[\frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \mu^2\eta^2\right] + \left[\frac{1}{2}(\partial_\mu\xi)(\partial^\mu\xi)\right] + \left[-\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}\left(\frac{q\mu}{\hbar c\lambda}\right)A_\mu A^\mu\right] + \text{ interaction terms.} \tag{7.41}$$

The gauge-field A_μ is now massive: $m_A = 2\sqrt{\pi}(\frac{q\mu}{\lambda c^2})^2$. However, we still have the massless Goldstone boson ξ present. Moreover, there are strange interaction terms $\propto (\partial_\mu \xi) A^\mu$, suggesting a conversion between ξ and A_μ. We see that both of these problems are related to the $\xi = \phi_2$ field. Interestingly, we can remove this field completely by exploiting local gauge invariance.

To see this, let

$$\phi \to \phi' = (\phi_1 c - \phi_2 s) + i(\phi_1 s + \phi_2 c) \tag{7.42}$$

so that if we choose $\theta = -\text{atan}(\phi_2/\phi_1)$, then $\phi_2' = 0$. Since $\mathcal{L}$ has the same form expressed with (ϕ', A_μ') as it does with (ϕ, A_μ) (that is what gauge invariant means per definition), we now have obtained $\xi = 0$ in this particular gauge and $\mathcal{L}$ takes the form (where it is assumed that A^μ has been gauge-transformed accordingly):

$$\mathcal{L} = [\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2 \eta^2] + \left[-\frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} + \frac{1}{2}\left(\frac{q\mu}{\hbar c \lambda}\right)^2 A_\mu A^\mu \right]$$
$$+ \frac{\mu}{\lambda}\left(\frac{q}{\hbar c}\right)^2 \eta(A_\mu A^\mu) + \frac{1}{2}\left(\frac{q}{\hbar c}\right)^2 \eta^2 A_\mu A^\mu - \lambda\mu\eta^3 - \frac{1}{4}\lambda^2 \eta^4 + \left(\frac{\mu^2}{2\lambda}\right)^2. \tag{7.43}$$

Again, the $\mathcal{L}$'s in these two gauges describe *exactly* the same physics, but the new choice of gauge gives us an $\mathcal{L}$ [Eq. (7.43)] that is easier to interpret. We are left with a single massive scalar (the Higgs particle) and a massive gauge field A^μ. Physicists like to say that A^μ "ate" the Goldstone boson ξ and thus acquired mass. This can also be thought of as A^μ acquiring a longitudinal polarization (since it now massive), which is a new degree of freedom.

The Higgs mechanism has then allowed for the gauge boson A^μ to become massive and this is what happens in the Standard Model where a spontaneous symmetry breaking of the Higgs field gives mass to $W^\pm$ and Z^0, while leaving the photon massless. The spontaneous symmetry breaking of the Higgs field gives mass not only to the gauge fields, but also to the fermion fields. The latter couple to the Higgs field via a so-called Yukawa coupling.

This is necessary in order for the fermion fields to acquire mass, because in the electroweak Lagrangian there can be no bare mass-terms for the fermions of the type $m\bar{\psi}\psi = m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L)$, since these would explicitly break the SU(2)$_L$ symmetry. The spontaneous symmetry breaking of the Higgs field demotes the electroweak symmetry group SU(2)$_L\otimes$ U(1)$_Y$ to a U(1)$_{EM}$ symmetry (which can be seen by writing out the new Lagrangian after the symmetry breaking), so that one is left with only one massless gauge field (the photon) in agreement with experiments. In effect, there is good reason to believe that fundamental interactions (such as weak, strong, and electromagnetic interactions) can be described by local gauge theories.

Note that "symmetry breaking" is often used in the literature without specifying if one refers to explicit or spontaneous symmetry breaking - it is supposed to be clear from the context. This might not always be so and for clarity we state what the difference is here.

Explicit symmetry breaking: A Lagrangian (or Hamiltonian) contains terms that exclude a certain symmetry operation. For instance, a magnetic field applied in the $\hat{z}$-direction breaks spin rotation symmetry since spin pointing in the $\hat{z}$ direction now have a different energy than *e.g.* spins pointing along the $\hat{x}$ direction.

Spontaneous symmetry breaking: A Lagrangian (or Hamiltonian) is invariant under a certain symmetry operation, but the ground-state in which the same system resides does not share that symmetry. For instance, in the absence of a magnetic field the energy of a ferromagnetic material is independent on the direction of the magnetization. However, once the material turns ferromagnetic the magnetization points in a given direction and thus has broken the original symmetry by selecting one particular realization of the ground-state.

F. Yang-Mills theory

Instead of looking at the details of the full electroweak Lagrangian, we give an example of a simpler SU(2) invariant Lagrangian and its belonging gauge fields. This will allow us to illustrate similar physics of the electroweak theory, but in a technically more transparent way. In particular, we will recover the coupling structure used in GWS theory ($\boldsymbol{j}_\mu \cdot \boldsymbol{W}^\mu$) and see that three gauge fields are required (unlike the single gauge field required in QED). Yang-Mills theory is also known as non-Abelian gauge theory because its Lagrangian has an SU(2) symmetry. In turn, SU(2) is a non-Abelian group since matrices do not in general commute.

Suppose that we have two spin-1/2 fields, ψ_1 and ψ_2. This is just like in the weak isospin doublet case, $\chi_L = \begin{bmatrix} \nu_e \\ e \end{bmatrix}_L$. The free Lagrangian takes the form

$$\mathcal{L} = [\mathrm{i}\hbar c\bar{\psi}_1\gamma^\mu\partial_\mu\psi_1 - m_1c^2\bar{\psi}_1\psi_1] + (1 \to 2). \tag{7.44}$$

Consider for simplicity the case of equal masses, so that we can make the notation more compact upon introducing

$$\psi = \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix}, \ \bar{\psi} = [\bar{\psi}_1, \bar{\psi}_2]. \tag{7.45}$$

which leads to:

$$\mathcal{L} = \mathrm{i}\hbar c\bar{\psi}\gamma^\mu\partial_\mu\psi - mc^2\bar{\psi}\psi. \tag{7.46}$$

This $\mathcal{L}$ is invariant under the transformation $\psi \to U\psi$ where U is any 2×2 unitary matrix. The general form of a SU(2) matrix is

$$U = \mathrm{e}^{\mathrm{i}\boldsymbol{\tau}\cdot\boldsymbol{a}} \tag{7.47}$$

where $\boldsymbol{\tau}$ is the Pauli matrix vector. Therefore, $\mathcal{L}$ is invariant under global SU(2) transformations. If we now insist on the SU(2) symmetry being local, we have to introduce gauge fields as before. For a local transformation, we let $\boldsymbol{a} = \boldsymbol{a}(x^\mu)$. Let us write this as:

$$\psi \to S\psi, \ S = \mathrm{e}^{-\mathrm{i}q\boldsymbol{\tau}\cdot\boldsymbol{\lambda}(x)/\hbar c} \tag{7.48}$$

which is a local SU(2) transformations. The original $\mathcal{L}$ is not invariant under this transformation, but we have already laid out a strategy for fixing this: we replace ∂_μ with the covariant derivative D_μ:

$$D_\mu \equiv \partial_\mu + \mathrm{i}\frac{q}{\hbar c}\boldsymbol{\tau} \cdot \boldsymbol{A}_\mu. \tag{7.49}$$

The remaining task is to identify a transformation rule for the gauge fields $\boldsymbol{A}_\mu$ so that

$$D_\mu\psi \to S(D_\mu\psi), \tag{7.50}$$

in which case we have an $\mathcal{L}$ that is invariant under local SU(2) transformations as can be verified by direct insertion of the above transformation. It turns out that (try to show this!) the following transformation rule for $\boldsymbol{A}_\mu$ does the trick:

$$\boldsymbol{\tau} \cdot \boldsymbol{A}'_\mu = S(\boldsymbol{\tau} \cdot \boldsymbol{A}_\mu)S^{-1} + \mathrm{i}(\hbar c/q)(\partial_\mu S)S^{-1}. \tag{7.51}$$

It is particularly instructive to consider the special case of infinitesimal transformations where $|\boldsymbol{\lambda}|$ is small. In that case, an expansion of S yields:

$$S \simeq 1 - \frac{\mathrm{i}q}{\hbar c}\boldsymbol{\tau} \cdot \boldsymbol{\lambda},\ \partial_\mu S \simeq -\frac{\mathrm{i}q}{\hbar c}\boldsymbol{\tau} \cdot \partial_\mu\boldsymbol{\lambda}. \tag{7.52}$$

Inserted into Eq. (7.51), we obtain the transformation rule for the vector of gauge fields:

$$\boldsymbol{A}'_\mu \simeq \boldsymbol{A}_\mu + \partial_\mu\boldsymbol{\lambda} + \frac{2q}{\hbar c}(\boldsymbol{\lambda} \times \boldsymbol{A}_\mu). \tag{7.53}$$

With these transformation rules, the following Lagrangian:

$$\begin{aligned}\mathcal{L} &= \mathrm{i}\hbar c\bar{\psi}\gamma^\mu D_\mu\psi - mc^2\bar{\psi}\psi \\ &= \mathrm{i}\hbar c\bar{\psi}\gamma^\mu\partial_\mu\psi - mc^2\bar{\psi}\psi - (q\bar{\psi}\gamma^\mu\boldsymbol{\tau}\psi) \cdot \boldsymbol{A}_\mu\end{aligned} \tag{7.54}$$

is invariant under local SU(2) transformations. We should also include the free Lagrangian of the three new vector fields as well:

$$\mathcal{L}_A = -\frac{1}{16\pi}\sum_j F_j^{\mu\nu}F_{\mu\nu,j} = -\frac{1}{16\pi}\boldsymbol{F}^{\mu\nu} \cdot \boldsymbol{F}_{\mu\nu}. \tag{7.55}$$

These must be massless gauge fields, since at term $m_A^2\boldsymbol{A}^\nu\boldsymbol{A}_\nu$ would break SU(2) invariance. Also, we must revise the structure of $\boldsymbol{F}^{\mu\nu}$ compared to our previously used $F^{\mu\nu}$ since the $\boldsymbol{A}$ field now has an extra term in the transformation rule Eq. (7.53):

$$\boldsymbol{F}^{\mu\nu} = \partial^\mu\boldsymbol{A}^\nu - \partial^\nu\boldsymbol{A}^\mu - \frac{2q}{\hbar c}(\boldsymbol{A}^\mu \times \boldsymbol{A}^\nu) \tag{7.56}$$

In conclusion, the full Yang-Mills Lagrangian describing two equal mass Dirac fields interacting with three massless gauge fields is:

$$\mathcal{L} = \mathrm{i}\hbar c\bar{\psi}\gamma^\mu\partial_\mu\psi - mc^2\bar{\psi}\psi - (q\bar{\psi}\gamma^\mu\boldsymbol{\tau}\psi) \cdot \boldsymbol{A}_\mu - \frac{1}{16\pi}\boldsymbol{F}^{\mu\nu} \cdot \boldsymbol{F}_{\mu\nu}. \tag{7.57}$$

Notice the form of the coupling $\boldsymbol{J}^\mu \cdot \boldsymbol{A}_\mu$, exactly like in electroweak theory.