Particle Physics 1

Lecture notes for the first year master course on the electroweak part of the Standard Model

Nikhef - autumn 2020

Marcel Merk, Ivo van Vulpen, Wouter Hulsbergen

Contents

Preliminaries

These are the lecture notes for the Particle Physics 1 (PP1) master course that is taught at Nikhef in the autumn semester. These notes contain 14 chapters, each corresponding to one lecture session. The topics discussed in this course are:

- Lecture 1 - 4: Electrodynamics of spinless particles
- Lecture 5 - 6: Electrodynamics of spin 1/2 particles
- Lecture 7: The weak interaction
- Lecture 8 - 10: Gauge symmetries and the electroweak theory
- Lecture 11-14: Electroweak symmetry breaking

Each lecture of 2×45 minutes is followed by a 1.5 hour problem solving session. The exercises are included in these notes, at the end of each chapter.

The notes mainly follow the material as discussed in the books of Halzen and Martin. The first ten chapters have been compiled by Marcel Merk in the period 2000-2011, and updated by Wouter Hulsbergen for the PP1 courses of 2012 and 2013. The last four chapters, written by Ivo van Vulpen, were added in 2014.

Literature

The following is a non-exhaustive list of course books on particle physics. (The comments reflect a personel opinion of your lecturers!)

Thomson: "Modern Particle Physics":
This is a new book (2013) that covers practically all the material in these lectures. If you do not have another particle physics book yet, then we recommend that you acquire this book.

Halzen & Martin: "Quarks & Leptons: an Introductory Course in Modern Particle Physics ":
This is the book that your lecturers used when they did their university studies. Though most of the theory is timeless, it is a bit outdated when it comes to experimental results. The book builds on earlier work of Aitchison (see below). Most of the course follows this book, but it is no longer in print.

Griffiths: “Introduction to Elementary Particle Physics”, second, revised ed.

The text is somewhat easier to read than H & M and is more up-to-date (2008) (e.g. neutrino oscillations) but on the other hand has a somewhat less robust treatment in deriving the equations. The introduction chapter of this book gives a very readable popular history of particle physics.

Aitchison & Hey: “Gauge Theories in Particle Physics”

Meanwhile in its 4th edition(2012), this 2-volume book provides a thorough theoretical introduction to particle physics, including field theory. It is excellent (notably its ’comments’ and appendices), but a bit more formal than needed for this course.

Perkins: “Introduction to High Energy Physics”, (1987) 3-rd ed., (2000) 4-th ed.

The first three editions were a standard text for all experimental particle physics. It is dated, but gives an excellent description of, in particular, the experiments. The fourth edition is updated with more modern results, while some older material is omitted.

Aitchison: “Relativistic Quantum Mechanics”

(1972) A classical, very good, but old book, often referred to by H & M.

Burcham & Jobes: “Nuclear & Particle Physics”

(1995) An extensive text on nuclear physics and particle physics. It contains more (modern) material than H & M. Formula’s are explained rather than derived and more text is spent to explain concepts.

Das & Ferbel: “Introduction to Nuclear and Particle Physics”

(2006) A book that is half on experimental techniques and half on theory. It is more suitable for a bachelor level course and does not contain a treatment of scattering theory for particles with spin.

Martin and Shaw: “Particle Physics ”, 2-nd ed.

(1997) A textbook that is somewhere inbetween Perkins and Das & Ferbel. In my opinion it has the level inbetween bachelor and master.

Particle Data Group: “Review of Particle Physics”

This book appears every two years in two versions: the book and the booklet. Both of them list all aspects of the known particles and forces. The book also contains concise, but excellent short reviews of theories, experiments, accellerators, analysis techniques, statistics etc. There is also a version on the web: http://pdg.lbl.gov

The Internet:

In particular Wikipedia contains a lot of information. However, one should note that Wikipedia does not contain original articles and they are certainly not reviewed! This means that they cannot be used for formal citations.

In addition, have a look at google books, where (parts of) books are online available.

About Nikhef

Nikhef is the Dutch institute for subatomic physics. (The name was originally an acronym for "Nationaal Instituut voor Kern en Hoge Energie Fysica".) The name Nikhef is used to indicate simultaneously two overlapping organisations:

- Nikhef is a national research lab ("institute") funded by the foundation FOM; the dutch foundation for fundamental research of matter.
- Nikhef is also a collaboration between the Nikhef institute and the particle physics departements of the UvA (A'dam), the VU (A'dam), the UU (Utrecht) and the RU (Nijmegen) contribute. In this collaboration all dutch activities in particle physics are coordinated.

In addition there is a collaboration between Nikhef and the Rijksuniversiteit Groningen (the former FOM nuclear physics institute KVI) and there are contacts with the Universities of Twente, Leiden and Eindhoven. For more information see the Nikhef web page: http://www.nikhef.nl.

The research at Nikhef includes both accelerator based particle physics and astro-particle physics. A strategic plan, describing the research programmes at Nikhef can be found on the web, from: www.nikhef.nl/fileadmin/Doc/Docs & pdf/StrategicPlan.pdf .

The accelerator physics research of Nikhef is currently focusing on the LHC experiments: Alice ("Quark gluon plasma"), Atlas ("Higgs") and LHCb ("CP violation"). Each of these experiments search answers for open issues in particle physics (the state of matter at high temperature, the origin of mass, the mechanism behind missing antimatter) and hope to discover new phenomena (eg supersymmetry, extra dimensions).

The LHC has started taking data in 2009 and the first LHC physics run has officially ended in winter 2013. In a joint meeting in summer 2012 the ATLAS and CMS experiments presented their biggest discovery so far, a new particle, called the Higgs particle, with a mass of approximately 125 GeV. The existence of this particle was a prediction of the Standard Model. Further research concentrates on signs of physics beyond the Standard Model, sometimes called "New Physics".

In preparation of these LHC experiments Nikhef has also been active at other labs: STAR (Brookhaven), D0 (Fermilab) and Babar (SLAC). Previous experiments that ended their activities are: L3 and Delphi at LEP, and Zeus, Hermes and HERA-B at Desy.

A more recent development is the research field of astroparticle physics. It includes Antares & KM3NeT ("cosmic neutrino sources"), Pierre Auger ("high energy cosmic rays"), Virgo & ET ("gravitational waves") and Xenon ("dark matter").

Nikhef houses a theory departement with research on quantum field theory and gravity, string theory, QCD (perturbative and lattice) and B-physics.

Driven by the massive computing challenge of the LHC, Nikhef also has a scientific computing departement: the Physics Data Processing group. They are active in the development of a worldwide computing network to analyze the large datastreams from the (LHC-) experiments ("The Grid").

Nikhef program leaders/contact persons:

	Name	office	phone	email
Nikhef director	Stan Bentvelsen	H232	5001	z66@nikhef.nl
Theory	Robert Fleischer	H323	2172	r.fleischer@nikhef.nl
Atlas	Wouter Verkerke	H241	5134	w.verkerke@nikhef.nl
B-physics	Marcel Merk	N243	5107	marcel.merk@nikhef.nl
Alice	Raimond Snellings	N321	2178	r.snellings.nl
KM3Net	Paul de Jong	H253	2087	p.de.jong@nikhef.nl
Pierre Auger	Charles Timmermans			c.timmermans@hef.ru.nl
Virgo and ET	Frank Linde	H044	5140	f.linde@nikhef.nl
Xenon	Patrick Decowski	H349	2145	p.decowski@nikhef.nl
Detector R&D	Niels van Bakel	H045	2255	n.van.bakel@nikhef.nl
Scientific Computing	Jeff Templon	H158	2092	templon@nikhef.nl

A very brief history of particle physics

The book of Griffiths starts with a nice historical overview of particle physics in the previous century. This is a summary of key events:

Atomic Models

1897 *Thomson:* Discovery of Electron. The atom contains electrons as "plums in a pudding".

1911 *Rutherford:* The atom mainly consists of empty space with a hard and heavy, positively charged nucleus.

1913 *Bohr:* First quantum model of the atom in which electrons circled in stable orbits, quatized as: $L = \hbar \cdot n$

1932 *Chadwick:* Discovery of the neutron. The atomic nucleus contains both protons and neutrons. The role of the neutrons is associated with the binding force between the positively charged protons.

The Photon

1900 *Planck:* Description blackbody spectrum with quantized radiation. No interpretation.

1905 *Einstein:* Realization that electromagnetic radiation itself is fundamentally quantized, explaining the photoelectric effect. His theory received scepticism.

1916 *Millikan:* Measurement of the photo electric effect agrees with Einstein's theory.

1923 *Compton:* Scattering of photons on particles confirmed corpuscular character of light: the Compton wavelength.

Mesons

1934 *Yukawa:* Nuclear binding potential described with the exchange of a quantized field: the pi-meson or pion.

1937 *Anderson & Neddermeyer:* Search for the pion in cosmic rays but he finds a weakly interacting particle: the muon. (Rabi: "Who ordered that?")

1947 *Powell:* Finds both the pion and the muon in an analysis of cosmic radiation with photo emulsions.

Anti matter

1927 *Dirac* interprets negative energy solutions of Klein Gordon equation as energy levels of holes in an infinite electron sea: "positron".

1931 *Anderson* observes the positron.

1940-1950 *Feynman* and *Stückelberg* interpret negative energy solutions as the positive energy of the anti-particle: QED.

Neutrino's

1930 *Pauli* and *Fermi* propose neutrino's to be produced in β-decay ($m_\nu = 0$).

1958 *Cowan* and *Reines* observe inverse beta decay.

1962 *Lederman* and *Schwarz* showed that $\nu_e \neq \nu_\mu$. Conservation of lepton number.

Strangeness

1947 *Rochester* and *Butler* observe V^0 events: K^0 meson.

1950 *Anderson* observes V^0 events: Λ baryon.

The Eightfold Way

1961 *Gell-Mann* makes particle multiplets and predicts the Ω^-.

1964 Ω^- particle found.

The Quark Model

1964 *Gell-Mann* and *Zweig* postulate the existence of quarks

1968 Discovery of quarks in electron-proton collisions (SLAC).

1974 Discovery charm quark (J/ψ) in SLAC & Brookhaven.

1977 Discovery bottom quarks (Υ) in Fermilab.

1979 Discovery of the gluon in 3-jet events (Desy).

1995 Discovery of top quark (Fermilab).

Broken Symmetry

1956 *Lee and Yang* postulate parity violation in weak interaction.

1957 *Wu* et. al. observe parity violation in beta decay.

1964 *Christenson, Cronin, Fitch & Turlay* observe CP violation in neutral K meson decays.

The Standard Model

1978 *Glashow*, *Weinberg*, *Salam* formulate Standard Model for electroweak interactions

1983 W-boson has been found at CERN.

1984 Z-boson has been found at CERN.

1989-2000 LEP collider has verified Standard Model to high precision.

Introduction

i.1 Quantum fields

As far as we can tell all 'ordinary' matter is made of elementary spin-$\frac{1}{2}$ fermions, which we call quarks and leptons. These particles interact via four types of interactions, namely the electromagnetic force, the strong nuclear force, the weak nuclear force and gravity. Besides gravity, these particles and interactions can be well described by a relativistic quantum field theory (QFT). In this course we will only consider the electromagnetic and weak interaction, leaving quantum chromodynamics (QCD), the theory of the strong interaction, to the Particle Physics 2 course. The quantum field theory for electrodynamics is called *quantum electrodynamics* (QED). If the weak interaction is included, it is called the *electroweak theory.*

Most of our knowledge of the physics of elementary particles comes from scattering experiments, from decays, and from the spectroscopy of bound states. The theory of bound states in electrodynamics is essentially the theory of the hydrogen atom. Apart from a few subtle phenomena, it is well described by non-relativistic quantum mechanics (QM). We do not discuss the hydrogen atom: instead we concentrate on processes at high energies.

To understand when the classical quantum theory breaks down it is useful to look at typical distance (or energy) scales relevant for electromagnetic interactions. Consider Coulomb's law, the well-known expression for the electrostatic force between two electrons in vacuum,

$$F \;=\; \frac{e^2}{4\pi\epsilon_0}\frac{1}{r^2} \;=\; \frac{\alpha\,\hbar\,c}{r^2} \;\;, \tag{i.1}$$

where $-e$ is the electron charge, ϵ_0 is the vacuum permittivity, and

$$\alpha \;\equiv\; \frac{e^2}{4\pi\epsilon_0\hbar c} \tag{i.2}$$

is the fine structure constant. The latter is dimensionless and its value is approximately 1/137. It is because of the fact that $\alpha \ll 1$ that perturbation theory works so well in quantum electrodynamics.

The first typical distance scale is the Bohr radius, the distance at which an electron circles around an infinitely heavy object (a 'proton') of opposite charge. Using just classical mechanics and imposing quantization of angular momentum by requiring that $rp = \hbar$, you will find (try!) that this distance is given by

$$r_{\mathrm{Bohr}} \;=\; \frac{\hbar}{\alpha m_e c}\,. \tag{i.3}$$

(A proper treatment in QM tells you that the expectation value for the radius is not exactly the Bohr radius, but it comes close.) Hence, the velocity of the electron is

$$v_{\mathrm{Bohr}} \;=\; p/m \;=\; \alpha c\,, \tag{i.4}$$

which indeed makes the electron in the hydrogen atom notably non-relativistic.

The second distance scale is the Compton wavelength of the electron. Suppose that you study electrons by shooting photons at zero-velocity electrons. The smaller the wavelength of the photon, the more precise you look. However, at some point the energy of the photons becomes large enough that you can create a new electron. (In our real theory, you can only create pairs, but that factor 2 is not important now.) The energy at which this happens is when $\hbar\omega = m_e c^2$, or at a wavelength

$$\lambda_e \;=\; \frac{2\pi\hbar}{m_e c}\,. \tag{i.5}$$

Usually, we divide both sides by 2π and speak of the *reduced* Compton wavelength $\bar{\lambda}_e$, just like $\hbar$ is usually called the reduced Planck's constant. Note that $\bar{\lambda}_e = \alpha r_{\mathrm{Bohr}}$. In electromagnetic collisions at this energy, classical quantum mechanics no longer suffices: as soon as collisions involve the creation of new particles, one needs QFT.

Finally, consider the collisions of two electrons at even higher energy. If the electrons get close enough, the Coulomb energy is sufficient to create a new electron. (Again, ignore the factor two required for pair production.) Expressing the Coulomb potential as $V(r) = \alpha\hbar c/r$, and setting this equal to $m_e c^2$, one obtains for the distance

$$r_e \;=\; \frac{\alpha\hbar}{m_e c}\,. \tag{i.6}$$

Note that, taking into account the definition of α, this expression does not explicitly depend on $\hbar$: you do not need quantization to compute this distance, which is why it is usually called the *classical radius* of the electron. At energies this high lowest order perturbation theory may not be sufficient to compute a cross-section. The effect of 'screening' becomes important, amplitudes described by Feynman diagrams with loops contribute and QED needs *renormalization* to provide meaningful answers.

Fortunately for most of us, we will not discuss renormalization in this course. In fact, we will hardly discuss quantum field theory at all! Do not be disappointed: there are two pragmatic reasons for this. First, a proper treatment requires a proper QFT course with some non-trivial math. This would leave insufficient time for other things that we do need to address. Second, if you accept a little handwaving here and there, then we do not actually need QFT: starting from quantum mechanics and special relativity

	distance [m]	energy [MeV]
r_{Bohr}	5.3×10^{-11}	0.0037
$\bar{\lambda}_e$	3.9×10^{-13}	0.511
r_e	2.8×10^{-15}	70

Table i.1: Values for the Bohr radius, the reduced Compton wavelength of the electron and the classical radius of the electron, and the corresponding energy.

we can derive the 'Born level' — that is, 'leading order' — cross-sections, following a route that allows us to introduce new concepts in a somewhat historical, and hopefully enlightening, order.

However, before continuing and setting aside the field theory completely until chapter 8, it is worthwhile to briefly discuss some relevant features of QFT, in particular those that distinguish it from ordinary quantum mechanics. In QM particles are represented by waves, or wave packets. Quantization happens through the 'fundamental postulate' of quantum mechanics that says that the operators for space coordinates and momentum coordinates do not commute,

$$[\hat{x}, \hat{p}] = i\hbar \tag{i.7}$$

The dynamics of the waves is described by the Schrödinger equation. Scattering cross-sections are derived by solving, in perturbation theory, a Schrödinger equation with a Hamiltonian operator that includes terms for kinetic and potential energy. Usually we expand the solution around the solution for a 'free' particle and write the solution as a sum of plane waves. This is exactly what you have learned in your QM course and we will come back to this in Lecture 2.

In QFT particles are represented as 'excitations' (or 'quanta') of a field $q(x)$, a function of space-time coordinates x. There are only a finite number of fields, one for each type of particle, and one for each force carier. This solves one imminent problem, namely why all electrons are exactly *identical.* In its simplest form QED has only two fields: one for a spin-$\frac{1}{2}$ electron and one for the photon. The dynamics of these field are encoded in a Lagrangian density $\mathcal{L}$. Equations of motions are obtained with the principle of least action. Those for the free fields (in a Lagrangian without interaction terms) lead to wave equations, reminiscent of the Schrödinger equation, but now Lorentz covariant. Again, solutions are written as superpositions of plane waves. The fields are quantized by interpreting the fields as operators and imposing a quantization rule similar to that in ordinary quantum mechanics, namely

$$[q, p] = i\hbar \tag{i.8}$$

where the momentum $p = \partial\mathcal{L}/\partial\dot{q}$ is the so-called adjoint coordinate to q. (You may remember that you used similar notation to arrive at Hamilton's principle in your classical or quantum mechanics course.) The Fourier components of the quantized fields can be identified as operators that create or destruct field excitations, exactly what we need for a theory in which the number of particles is not conserved. The relation to classical QM can be made by identifying the result of a 'creation' operator acting on the vacuum as the QM wave in the Schrödinger equation.

That was a mouth full and you can forget most of it. One last thing, though: one very important aspect of quantum field theory is the role of *symmetries* in the Lagrangian. In fact, as we shall see in Lecture 8 and 9, the concept of phase invariance allows to define the standard model Lagrangian by specifying only the matter fields and the symmetries: once the symmetries are defined, the dynamics (the force carriers) come for free.

That said, we leave the formal theory of quantum fields alone. In the remainder of this chapter we briefly discuss some concepts and summarize our 'Standard Model' of elementary particles. In Lecture 1 we formulate a relativistic wave equation for a spin-0 particle. In Lecture 2 we discuss classical QM perturbation theory and Fermi's Golden rule, which allows us to formalize the computation of a cross-section. In Lecture 3, we show how the Maxwell equations take a very simple form when expressed in terms of a new spin-1 field, which we identify as the photon. In lecture 4 we apply the developed tools to compute the scattering of spin-0 particles. In Lectures 5 and 6 we turn to spin-$\frac{1}{2}$ fields, which are considerably more realistic given that all SM matter fields are indeed fermions. In Lectures 7 through 10, we introduce the weak interaction, gauge theory and electroweak unification. Finally, in Lectures 11-14 we look in more detail at electroweak symmetry breaking.

i.2 The Yukawa interaction

After Chadwick had discovered the neutron in 1932, the elementary constituents of matter were the *proton*, the *neutron* and the *electron*. The force responsible for interactions between charged particles was the electromagnetic force. A 'weak' interaction was responsible for nuclear decays. Moving charges emitted electromagnetic waves, which happened to be quantized in energy and were called *photons*. With these constituents the atomic elements could be described, as well as their chemistry.

However, there were already some signs that there were more elementary particles than just protons, neutrons, electrons and photons:

- Dirac had postulated in 1927 the existence of *anti-matter* as a consequence of his relativistic version of the Schrödinger equation in quantum mechanics. (We will come back to the Dirac theory later on.) The anti-matter partner of the electron, the positron, was discovered in 1932 by Anderson (see Fig. i.1).
- Pauli had postulated the existence of an invisible particle that was produced in nuclear beta decay: the *neutrino*. In a nuclear beta decay process $N_A \to N_B + e^-$ the energy of the emitted electron is determined by the mass difference of the nuclei N_A and N_B. It was observed that the kinetic energy of the electrons, however, showed a broad mass spectrum (see Fig. i.2), of which the maximum was equal to the expected kinetic energy. It was as if an additional invisible particle of low mass is produced in the same process: the (anti-) neutrino.

Furthermore, though the constituents of atoms were fairly well established, there was something puzzling about atoms: What was keeping the nucleus together? It clearly

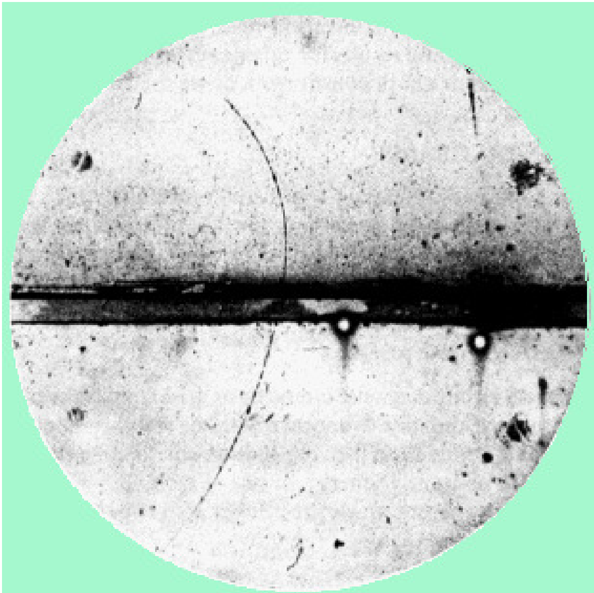


Figure i.1: The discovery of the positron as reported by Anderson in 1932. Knowing the direction of the B field Anderson deduced that the trace was originating from an anti electron. *Question: how?*

had to be a new force, something beyond electromagnetism. Rutherford's scattering experiments had given an estimate of the size of the nucleus, of about 1 fm. With protons packed this close, the new force had to be very strong to overcome the repulsive coulomb interaction of the protons. (Being imaginative, physicists simply called it the *strong* nuclear force.) Yet, to explain scattering experiments, the *range* of the force had to be small, bound just to the nucleus itself.

In an attempt to solve this problem Japanese physicist Yukawa published in 1935 a fundamentally new view of interactions. His idea was that *forces*, like the electromagnetic force and the nuclear force, could be described by the exchange of *virtual particles*, as illustrated in Fig. i.3. These particles (or rather, their *field*) would follow a relativistic wave-equation, just like the electromagnetic field.

In this picture, the massless photon was the carrier of the electromagnetic field. As we will see in exercise 1.4.3 the relativistic wave equation for a massless particle leads to an electrostatic potential of the form (in natural units, $\hbar = \overline{c} = 1$)

$$V(r) \;=\; -\alpha\,\frac{1}{r}. \tag{i.9}$$

Because of its $1/r$ dependence, the force is said to be of 'infinite range'.

By contrast, in Yukawa's proposal the strong force were to be carried by a massive particle, later called the pion. A massive force carrier leads to a potential of the form

$$U(r) = -g^2\,\frac{e^{-r/R}}{r} \tag{i.10}$$

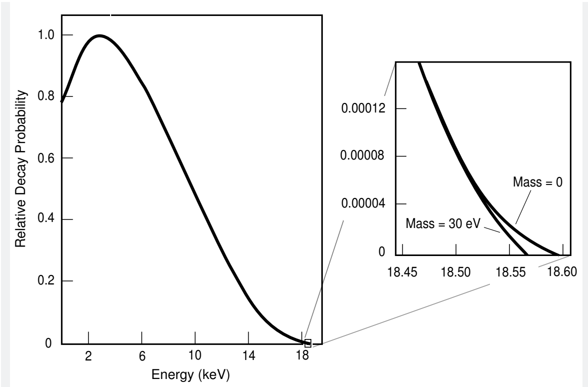


Figure 1. The Beta Decay Spectrum for Molecular Tritium
The plot on the left shows the probability that the emerging electron has a particular energy. If the electron were neutral, the spectrum would peak at higher energy and would be centered roughly on that peak. But because the electron is negatively charged, the positively charged nucleus exerts a drag on it, pulling the peak to a lower energy and generating a lopsided spectrum. A close-up of the endpoint (plot on the right) shows the subtle difference between the expected spectra for a massless neutrino and for a neutrino with a mass of 30 electron volts.

Figure i.2: The beta spectrum as observed in tritium decay to helium. The endpoint of the spectrum can be used to set a limit of the neutrino mass. *Question: how?*

which is called the *one-pion-exchange-potential*. Since it falls of exponentially, it has a finite range. The range R is inversely proportional to the mass of the force carrier and for a massless carrier the expression reduces to that for the electrostatic potential.

In the exercise you will derive the relation between mass and range properly, but it can also be obtained with a heuristic argument, following the Heisenberg uncertainty principle. (As you can read in Griffiths, whenever a physicists refers to the uncertainty principle to explain something, take all results with a grain of salt.) In some interpretation, the principle states that we can borrow the energy $\Delta E = mc^2$ to create a virtual particle from the vacuum, as long as we give it back within a time $\Delta t \approx \hbar/\Delta E$. With the particle traveling at the speed of light, this leads to a range $R = c\Delta t = c\hbar/mc^2$.

From the size of the nucleus, Yukawa estimated the mass of the force carrier to be approximately $100\,\mathrm{MeV}/c^2$. He called the particle a *meson*, since its mass was somewhere in between the mass of the electron and the nucleon.

In 1937 Anderson and Neddermeyer, as well as Street and Stevenson, found that cosmic rays indeed consist of such a middle weight particle. However, in the years after, it became clear that this particle could not be Yukawa's meson, since it did not interact strongly, which was very strange for a carrier of the strong force. In fact this particle turned out to be the *muon*, the heavier brother of the electron.

Only in 1947 Powell and Perkins found Yukawa's pion in cosmic rays. They took photographic emulsion plates to mountain tops to study the contents of cosmic rays (see Fig. i.4). In a cosmic ray event a cosmic proton scatters with high energy on an atmospheric

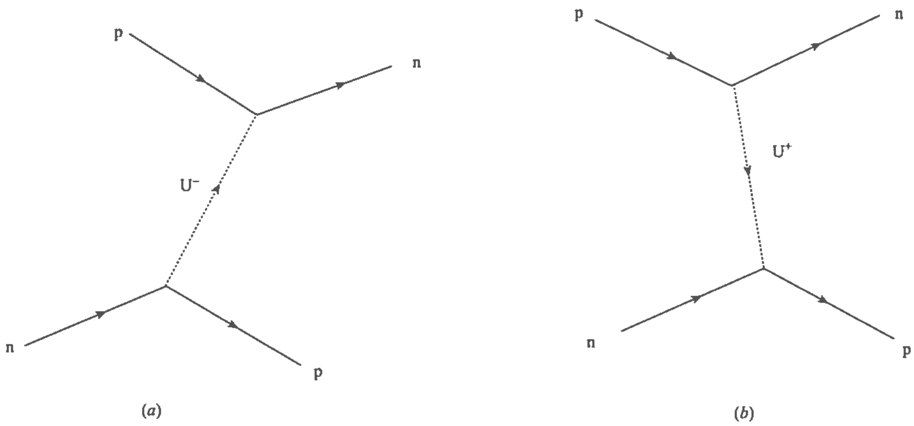


Figure i.3: Illustration of the interaction between protons and neutrons by charged pion exchange. (From Aichison and Hey.)

nucleon and produces many secondary particles. Pions produced in the atmosphere decay long before they reach sea level, which is why they had not been observed before.

As a carrier of the strong force Yukawa's meson did not stand the test of time. We now know that the pion is a composite particle and that the true carrier for the strong force is the massless gluon. The range of the strong force is small, not because the force carrier is massive, but because gluons carry a strong interaction charge themselves. However, even if Yukawa's original meson model did not survive, his interpretation of forces as the exchange of virtual particles is still central to the description of particle interactions in quantum field theory.

i.3 Feynman diagrams

Figure i.3 is an example of a Feynman diagram. You have probably seen Feynman diagrams before and already know that they are not just pictures that help us to 'visualize' a scattering process: they can be translated efficiently into mathematical expressions for the computation of quantum mechanical transition amplitudes.

In this course, we will always draw Feynman diagrams such that *time* runs from left to right. This is just a convention: the diagrams in Fig. i.3 are equally valid if time runs from right to left, or from top to bottom, etc.

With this convention, the two diagrams in Fig. i.3 represent two different ways of scattering a proton and a neutron via pion exchange: In case (a) a negative virtual pion is first emitted by the neutron and then absorbed by the proton, while in the case (b) a positive virtual pion is first emitted by the proton and then absorbed by the neutron. As usual in quantum mechanics these two contributions are represented by two complex amplitudes, which need to be added in order to obtain the total amplitude. It turns out that only if both amplitudes are taken into account, Lorentz-covariant results can be obtained in a quantum theory.

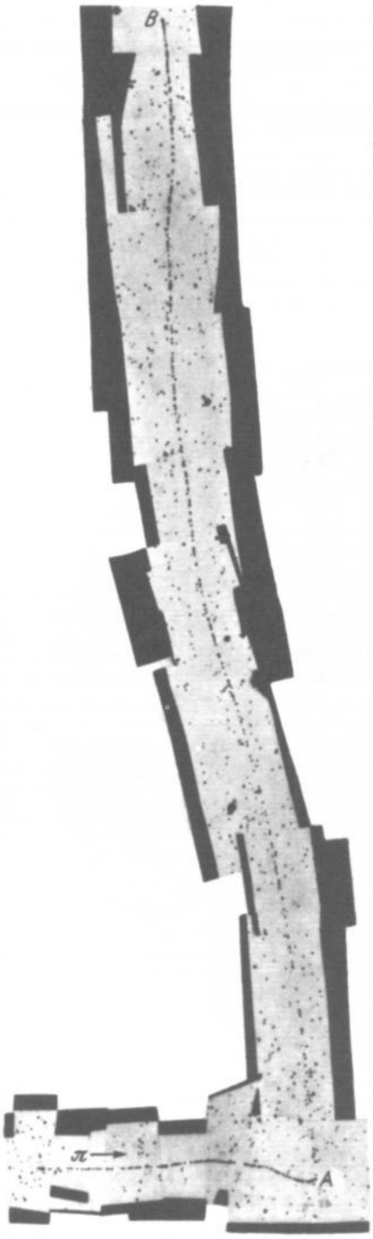


Figure 1.4 One of Powell's earliest pictures showing the track of a pion in a photographic emulsion exposed to cosmic rays at high altitude. The pion (entering from the left) decays into a muon and a neutrino (the latter is electrically neutral, and leaves no track). Reprinted by permission from C. F. Powell, P. H. Fowler, and D. H. Perkins, *The Study of Elementary Particles by the Photographic Method* (New York: Pergamon, 1959). First published in *Nature* **159**, 694 (1947).

Figure i.4: A pion entering from the left decays into a muon and an invisible neutrino.

However, now that we know that both amplitudes must be taken into account, we no longer need to draw both of them! In fact, in the remainder of this course, we will always draw only one diagram, with the line that represents the pion exchange drawn vertically. By convention, Feynman diagrams always present all possible time orderings for the 'internal' lines, the virtual particles.

i.4 The Standard Model

In the Standard Model (SM) of particle physics all matter particles are spin-$\frac{1}{2}$ fermions and all force carriers are spin-1 bosons. The fermions are the quarks and leptons, organized in three families (table i.2). The force carriers are the photon for the electromagnetic interaction, the Z and W for the weak interaction and the gluon for the strong interaction (table i.3). In the SM the forces originate from a symmetry by a mechanism called *local gauge invariance*, discussed later on in the course.

charge	**Quarks**		
$\frac{2}{3}$	u (up) 1.5–4 MeV	c (charm) 1.15–1.35 GeV	t (top) (174.3 ± 5.1) GeV
$-\frac{1}{3}$	d (down) 4–8 MeV	s (strange) 80–130 MeV	b (bottom) 4.1–4.4 GeV
charge	**Leptons**		
0	ν_e (e neutrino) < 3 eV	ν_μ (μ neutrino) < 0.19 MeV	ν_τ (τ neutrino) < 18.2 MeV
-1	e (electron) 0.511 MeV	μ (muon) 106 MeV	τ (tau) 1.78 GeV

Table i.2: Matter particles in the Standard Model, with their approximate mass.

force	vector boson	coupling strength at 1 GeV
strong	g (8 gluons)	$\alpha_s \sim \mathcal{O}(1)$
electromagnetic	γ (photon)	$\alpha \sim \mathcal{O}(10^{-3})$
weak	Z ,$W^{\pm}$ (weak bosons)	$\alpha_W \sim \mathcal{O}(10^{-8})$

Table i.3: Standard Model forces, the mediating bosons, and the associated strength of the coupling at an energy of about 1 GeV. (The latter are taken from Thomson, 2013.)

Apart from the charged weak interaction (mediated by the W), all interactions conserve quark and lepton flavour: They do not change one type of fundamental fermion into another type. On the other hand, the charged weak interaction allows for transitions between an up-type quark and a down-type quark, and between charged leptons and neutrinos. Some of the fundamental diagrams are shown in figure i.5.

The fundamental particles carry a charge (essentially a quantum number) to characterize how they couple to the force carriers. For the electromagnetic interaction this charge

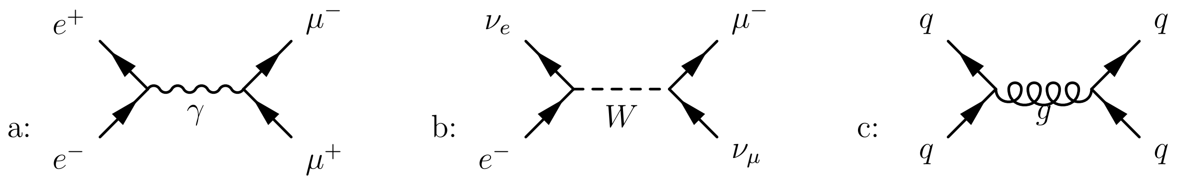


Figure i.5: Feynman diagrams of fundamental lowest order perturbation theory processes in a: electromagnetic, b: weak and c: strong interaction.

is the electric charge, for the weak interaction it is called *weak isospin* and for the strong interaction it is called *colour*. The strength of an interaction is determined by the charges of the particles involved, by a coupling strength that is a characteristic of the interaction (see table i.3), and by the mass of the vector boson.

There is an important difference between the electromagnetic force on one hand, and the weak and strong force on the other hand. The photon does not carry electric charge and, therefore, does not interact with itself. The gluons, however, carry colour and do interact amongst each other. Also, the weak vector bosons carry weak isospin charge and undergo this so-called *self-coupling*.

The coupling strengths are not actually constant, but vary as a function of energy, which is called the *running* of the coupling constants. At a momentum transfer of 10^{15} GeV the couplings of electromagnetic, weak and strong interaction all obtain approximately the same value. (See figure i.6.) *Grand unification* refers to the hypothesis that at high energy there is actually only a single force, originating from a single gauge symmetry with a single coupling constant.

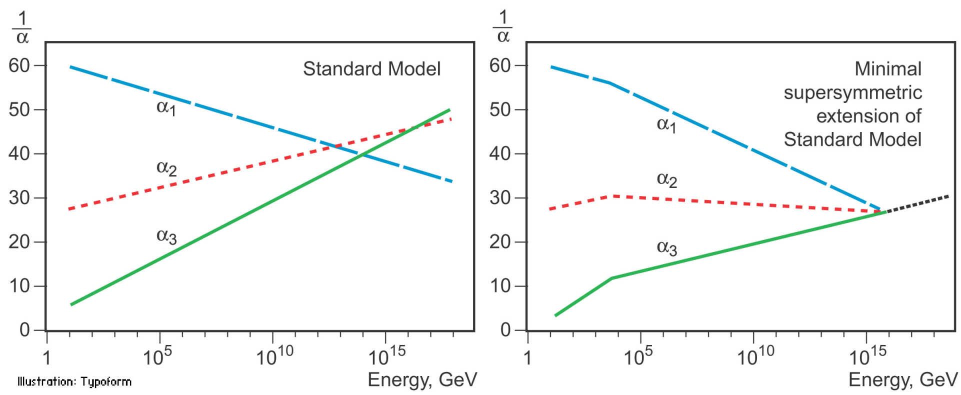


Figure i.6: Running of the coupling constants and possible unification point. On the left: Standard Model. On the right: Supersymmetric Standard Model. (Source: `https://www.nobelprize.org/prizes/physics/2004/popular-information/`.)

Due to the self-coupling of the force carriers the running of the coupling constants of the weak and strong interaction are opposite to that of electromagnetism. Electromagnetism becomes weaker at low momentum (i.e. at large distance), the weak and the strong force become stronger at low momentum or large distance. The strong interaction coupling becomes so large at momenta less than a few 100 MeV that perturbation theory is no longer applicable. (The coupling constant is larger than 1.) Although this is not rigorously proven, it is assumed that the self-coupling of the gluons is also responsible for *confinement*: the existence of free coloured objects (i.e. objects with net strong charge) is forbidden.

Confinement means that free quarks do not exist, at least, not at time-scales longer than that corresponding to the range of the strong interaction. Quarks always appear in bound states, either as combinations of three quarks (*baryons*) or as combinations of a quark an an anti-quark (*mesons*). Together these are called *hadrons*. In the *quark model* the various species of hadrons are organized by exploiting *quark flavour symmetry*, the fact that equally charged quarks of different families are indistinguisable except for their mass. Due to lack of time, we will not discuss the quark model in this course. For reference table i.4 gives a list of common hadrons, some of which we encounter in examples in the lectures.

name	nickname	symbol	quark content	mass / MeV
proton		p	uud	938.3
neutron		n	udd	939.6
charged pion	pion	π^+	$u\bar{d}$	139.6
neutral pion	pi-zero	π^0	$(u\bar{u} - d\bar{d})/\sqrt{2}$	135.0
charged kaon	kaon	K^+	$u\bar{s}$	493.7
neutral kaon	K-zero	K^0	$d\bar{s}$	497.6
charged charmed meson	D-plus	D^+	$c\bar{d}$	1869.6
neutral charmed meson	D-zero	D^0	$c\bar{u}$	1864.8
strange charmed meson	D-sub-s	D_s^+	$c\bar{s}$	1968.3
charged bottom meson	B-plus	B^+	$u\bar{b}$	5279.3
neutral bottom meson	B-zero	B^0	$d\bar{b}$	5279.6
strange bottom meson	B-sub-s	B_s^0	$s\bar{b}$	5366.8

Table i.4: Name, quark content and approximate mass of common baryons and mesons. The complete list of all known hadrons, together with a lot of experimental data, can be found in the particle data book, `http://pdglive.lbl.gov`.

Finally, the Standard Model includes a scalar boson field, the Higgs field, which provides mass to the vector bosons and fermions in the Brout-Englert-Higgs mechanism. The motivation for the Higgs particle and corresponding precision tests of the SM are the subject of the last four lectures of this course.

Despite the success of the standard model in describing all physics at 'low' energy scale, there are still many open questions, such as:

- why are the masses of the particles what they are?
- why are there 3 generations of fermions?
- are quarks and leptons truly fundamental?
- is there really only one Higgs particle?
- why is the charge of the electron *exactly* opposite to that of the proton? Or phrased differently: why is the total charge of leptons and quarks in one generation exactly

zero?

- is a neutrino its own anti-particle?
- can all forces be described by a single gauge symmetry (unification)?
- why is there no anti-matter in the universe?
- what is the source of dark matter?
- what is the source of dark energy?

Particle physicists try to address these questions with scattering experiments in the laboratory and by studying high energy phenomena in the cosmos.

i.5 Units in particle physics

In particle physics we often make use of *natural units* to simplify expressions. In this system of units the action is expressed in units of Planck's constant

$$\hbar \approx 1.055 \times 10^{-34}\mathrm{Js} \tag{i.11}$$

and velocity is expressed in units of the speed of light in vacuum

$$c \approx 2.998 \times 10^{8}\mathrm{m/s}. \tag{i.12}$$

such that all factors $\hbar$ and c can be omitted. As a consequence (see textbooks), there is only one basic unit for length (L), time (T), mass (M), energy and momentum. In high energy physics this basic unit is often chosen to be the energy in MeV or GeV, where 1 eV is the kinetic energy an electron obtains when it is accelerated over an electrostatic potential of 1 V. Momentum and mass then get units of energy, while length and time get units of inverse energy.

To confront the result of calculation with experiments the factors $\hbar$ and c usually need to be reintroduced. There are two ways to do this. First one can take the final expressions in natural units and then use the table i.5 to convert the quantities for space, time, mass, energy and momentum back to their original counterparts. (For the positron charge, see below.) Alternatively, one can express all results in GeV, then use the table i.6 with conversion factors to translate it into SI units.

Where it concerns electromagnetic interactions, there is also freedom in choosing the unit of electric charge. Consider again Coulomb's law,

$$F = \frac{e^2}{4\pi\epsilon_0}\frac{1}{r^2}\ , \tag{i.13}$$

where ϵ_0 is the vacuum permittivity. The dimension of the factor e^2/ϵ_0 is fixed — it is $[\mathrm{L}^3\mathrm{M}/\mathrm{T}^2]$ — but this still leaves a choice of what to put in the charges and what in the vacuum permittivity.

quantity	symbol in natural units	equivalent symbol in ordinary units
space	x	$x/\hbar c$
time	t	$t/\hbar$
mass	m	mc^2
momentum	p	pc
energy	E	E
positron charge	e	$e\sqrt{\hbar c/\epsilon_0}$

Table i.5: Conversion of basic quantities between natural and ordinary units.

quantity	conversion factor	natural unit	normal unit
mass	$1\ \mathrm{kg} = 5.61 \times 10^{26}\,\mathrm{GeV}$	GeV	GeV/c^2
length	$1\ \mathrm{m} = 5.07 \times 10^{15}\,\mathrm{GeV}^{-1}$	GeV^{-1}	$\hbar c/\,\mathrm{GeV}$
time	$1\ \mathrm{s} = 1.52 \times 10^{24}\,\mathrm{GeV}^{-1}$	GeV^{-1}	$\hbar/\,\mathrm{GeV}$

Table i.6: Conversion factors from natural units to ordinary units.

In the SI system the unit of charge is the Coulomb. Until 2019 it was defined via the Ampére, which in turn was defined as the current leading to a particular force between two current-carrying wires. However, since 2019 it is defined in terms of the charge of the proton (or minus the charge of the electron), such that that charge expressed in Coulombs is exactly

$$e \;=\; 1.602176634 \times 10^{-19}\mathrm{C}\ . \tag{i.14}$$

The vacuum permittivity is approximately

$$\epsilon_0 \approx 8.854 \times 10^{-12}\mathrm{C}^2\mathrm{s}^2\mathrm{kg}^{-1}\mathrm{m}^{-3}. \tag{i.15}$$

As we shall see in Lecture 3 the Maxwell equations look much more neat if, in addition to $c = 1$, we choose $\epsilon_0 = 1$. This is called the *Heaviside-Lorentz* system. Obviously, this choice affects the numerical value of e. However, the coupling constant α, defined in equation i.2, is dimensionless and hence independent of the system of units. In this course we will often write e^2, when in fact we mean α.

Finally, it is customary to express scattering cross-sections in *barn*: one barn is equal to $10^{-24}\mathrm{cm}^2$.

i.6 Four-vector notation

We define the coordinate four-vector x^μ as

$$x^\mu = (x^0, x^1, x^2, x^3) \tag{i.16}$$

where the first component $x^0 = ct$ is the time coordinate and the latter three components are the spatial coordinates $(x^1, x^2, x^3) = \boldsymbol{x}$. Under a Lorentz transformation along the

x^1 axis with velocity $\beta = v/c$, x^μ transforms as

$$
\begin{aligned}
x^{0\prime} &= \gamma(x^0 - \beta x^1) \\
x^{1\prime} &= \gamma(x^1 - \beta x^0) \\
x^{2\prime} &= x^2 \\
x^{3\prime} &= x^3
\end{aligned}
\tag{i.17}
$$

where $\gamma = 1/\sqrt{1-\beta^2}$.

A 'contravariant four-vector' is defined to be any set of four quantities $A^\mu = (A^0, A^1, A^2, A^3) = (A^0, \boldsymbol{A})$ that transforms under Lorentz transformations exactly as the corresponding components of the coordinate four-vector x^μ. It is the *transformation property* that defines what a 'contravariant' vector is.

Lorentz transformations leave the quantity

$$|A|^2 = A^{0^2} - |\boldsymbol{A}|^2 \tag{i.18}$$

invariant. This expression may be regarded as the scalar product of A^μ with a related 'covariant vector' $A_\mu = (A^0, -\boldsymbol{A})$, such that

$$A \cdot A \equiv |A|^2 = \sum_\mu A^\mu A_\mu. \tag{i.19}$$

From now on we omit the summation sign and implicitly sum over any index that appears twice. Defining the metric tensor

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{i.20}$$

we have $A_\mu = g_{\mu\nu} A^\nu$ and $A^\mu = g^{\mu\nu} A_\nu$. A scalar product of two four-vectors A^μ and B^μ can then be written as

$$A \cdot B = A_\mu B^\mu = g_{\mu\nu} A^\mu B^\nu. \tag{i.21}$$

One can show that such a scalar product is indeed also a Lorentz invariant.

You will show in exercise 1.6 that if the contravariant and covariant four-vectors for the coordinates are defined as above, then the four-vectors of their derivatives are given by

$$\partial^\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\boldsymbol{\nabla}\right) \quad \text{and} \quad \partial_\mu = \left(\frac{1}{c}\frac{\partial}{\partial t}, \boldsymbol{\nabla}\right). \tag{i.22}$$

The position of the minus sign is opposite to that of the coordinate four-vector itself.

Lecture 1

Wave Equations and Anti-Particles

1.1 Particle-wave duality

Ever since Maxwell we know that electromagnetic fields propagating in a vacuum are described by a wave equation

$$\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right]\phi(\boldsymbol{x},t) = 0 \; . \tag{1.1}$$

The solution to this equation is given by plane waves of the form

$$\phi(\boldsymbol{x},t) = e^{i\boldsymbol{k}\cdot\boldsymbol{x} - i\omega t} \; , \tag{1.2}$$

where the wave-vector $\boldsymbol{k}$ and the angular frequency ω are related by the *dispersion relation*

$$\omega = c|\boldsymbol{k}| \; . \tag{1.3}$$

Since the equation above is real, we can restrict ourselves to real solutions. In fact, the photon field is described by just the real component of the wave. However, it is often more convenient to work with complex waves. The real solutions can always be identified with one component of the complex wave.

Maxwell identified propagating electromagnetic fields with light, and thereby firmly established what everybody already knew: light behaves as a wave. However, to explain the photo-electric effect Einstein hypothesized in 1904 that light is also a particle with zero mass. For a given frequency, lights comes in packets ('quanta') with a fixed energy. The energy of a quantum is related to the frequency by

$$E \;=\; h\nu \;=\; \hbar\omega \; , \tag{1.4}$$

while its momentum is related to the wave-number

$$\boldsymbol{p} \;=\; \hbar\boldsymbol{k} \; . \tag{1.5}$$

In terms of energy and momentum the dispersion relation takes the familiar form $E = pc$. The idea of light as a particle was received with much skepticism and only generally accepted after Compton showed in 1923 that photons scattering of electrons behave as one would expect from colliding particles.

So, by 1923 light was a wave and a particle: it satisfied a wave equation, yet it only came about in packets of discrete energy. That lead De Broglie in 1924 to make another bold preposition: if light is both a wave and a particle, then why wouldn't matter particles be waves as well? It took another few years before physicists established the wave-like character of electrons in diffraction experiments, but well before that people took De Broglie hypothesis seriously and started looking for a suitable wave-equation for massive particles.

The crucial element is to establish the dispersion relation for the wave. Schrödinger started with the relativistic equation for the total energy

$$E^2 = m^2c^4 + \boldsymbol{p}^2c^2 \, , \tag{1.6}$$

but abandoned the idea, for reasons we will discuss later. He then continued with the equation for the kinetic energy in the non-relativistic limit

$$E \;=\; \frac{\boldsymbol{p}^2}{2m} \, , \tag{1.7}$$

which, as we shall see now, led to his famous equation.

1.2 The Schrödinger equation

One pragmatic way to quantize a classical theory is to take the classical equations of motion and substitute energy and momentum by their operators in the coordinate representation,

$$E \;\rightarrow\; \hat{E} = i\hbar\frac{\partial}{\partial t} \qquad \text{and} \qquad \boldsymbol{p} \;\rightarrow\; \hat{\boldsymbol{p}} = -i\hbar\boldsymbol{\nabla} \, . \tag{1.8}$$

Inserting these operators in Eq. (1.7), leads to the Schrödinger equation for a free particle,

$$i\hbar\frac{\partial}{\partial t}\,\psi \;=\; -\frac{\hbar^2}{2m}\,\nabla^2\psi \, . \tag{1.9}$$

In quantum mechanics we interprete the square of the wave function as a probability density. The probability to find a particle at time t in a box of finite size V is given by the volume integral

$$P(\text{particle in volume } V, t) \;=\; \int_V \rho(\boldsymbol{x}, t)\,\mathrm{d}^3x \, , \tag{1.10}$$

where the density is

$$\rho(\boldsymbol{x}, t) = |\psi(\boldsymbol{x}, t)|^2 \, . \tag{1.11}$$

Since total probability is conserved, the density must satisfy a so-called continuity equation

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{j} = 0 \tag{1.12}$$

where j is the *density current* or *flux*. When considering charged particles you can think of ρ as the charge per volume and j as the charge times velocity per volume. The continuity equation can then be stated in words as "The change of charge in a given volume equals the current through the surrounding surface".

What is the current corresponding to a quantum mechanical wave ψ? It is straightforward to obtain this current from the continuity equation by writing $\partial\rho/\partial t = \psi\partial\psi^*/\partial t + \psi^*\partial\psi/\partial t$ and inserting the Schrödinger equation. However, because this is useful later on, we follow a slightly different approach. First, rewrite the Schrödinger equation as

$$\frac{\partial}{\partial t}\psi = \frac{i\hbar}{2m}\nabla^2\psi.$$

Now multiply both sides on the left by ψ^* and add the expression to its complex conjugate

$$\begin{array}{rcl} \psi^* \dfrac{\partial \psi}{\partial t} & = & \psi^* \left(\dfrac{i\hbar}{2m}\right) \nabla^2 \psi \\ \psi \, \dfrac{\partial \psi^*}{\partial t} & = & \psi \, \left(\dfrac{-i\hbar}{2m}\right) \nabla^2 \psi^* \\ \hline \dfrac{\partial}{\partial t} \underbrace{(\psi^* \, \psi)}_{\rho} & = & -\boldsymbol{\nabla} \cdot \underbrace{\left[\dfrac{i\hbar}{2m} \, (\psi \, \boldsymbol{\nabla}\psi^* - \psi^* \, \boldsymbol{\nabla}\psi)\right]}_{\boldsymbol{j}} \end{array} + \tag{1.13}$$

where in the last step we have used that $\boldsymbol{\nabla}\cdot(\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*) = \psi^*\nabla^2\psi - \psi\nabla^2\psi^*$. In the result we can recognize the continuity equation if we interpret the density and current as indicated.

Plane waves of the form

$$\psi \;=\; N\, e^{i(\boldsymbol{p}\cdot\boldsymbol{x} - Et)/\hbar} \tag{1.14}$$

with $E = p^2/2m$ are solutions to the free Schrödinger equation. (In fact, starting from the idea of particle-wave duality, Schrödinger took the plane wave form above and 'derived' his equation as the equation that described its time evolution.) We will leave the definition of the normalization constant N for the next lecture: as the plane wave is not localized in space (it has precise momentum, and infinitely imprecise position!), it can only be normalized on a finite volume.

To get rid of the inconvenient factor $\hbar$ in the exponent, we can express energy and momentum in terms of the wave vector and angular frequency defined above or work in natural units. Since it is sometimes useful to verify expressions with a dimensional analysis, we shall keep the factors $\hbar$ for now.

For the density of the plane wave we obtain

$$\rho \equiv \psi^* \psi = |N|^2 \tag{1.15}$$

$$\boldsymbol{j} \equiv \frac{i\hbar}{2m} \left(\psi \, \boldsymbol{\nabla}\psi^* - \psi^* \, \boldsymbol{\nabla}\psi\right) = \frac{|N|^2}{m} \, \boldsymbol{p} \tag{1.16}$$

As expected, the density current is equal to the density times the non-relativistic velocity $v = p/m$.

Any solution to the free Schrödinger can be written as a superposition of plane waves. Ignoring boundary conditions (which usually limit the energy to quantized values), the decomposition is written as the convolution integral

$$\psi(\boldsymbol{x}, t) = \left(\hbar\sqrt{2\pi}\right)^{-3} \int \psi(\boldsymbol{p}) e^{i(\boldsymbol{p}\cdot\boldsymbol{x} - Et)/\hbar} \mathrm{d}^3\boldsymbol{p} \tag{1.17}$$

with $E = \boldsymbol{p}^2/2m$. For $t = 0$ this is just the usual Fourier transform. For the exercises, remember that in one dimension the Fourier transform and its inverse are given by (Plancherel's theorem),

$$\boxed{f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} F(k) e^{ikx} \mathrm{d}k \quad \Longleftrightarrow \quad F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-ikx} \mathrm{d}x} \tag{1.18}$$

If we replace $\boldsymbol{p}$ with $-\boldsymbol{p}$ in the plane wave definition

$$\psi_{\text{out}} = N \, e^{i(-\boldsymbol{p}\cdot\boldsymbol{x} - Et)/\hbar} \tag{1.19}$$

we still have a solution to the Schrödinger equation, since the latter is quadratic in coordinate derivatives. These solutions are already included in the decomposition in Eq. (1.17).

By convention when describing scattering in terms of plane waves we identify those with $+\boldsymbol{p}\cdot\boldsymbol{x}$ in the exponent as *incoming* waves and those with $-\boldsymbol{p}\cdot\boldsymbol{x}$ as *outgoing* waves. In one dimension, incoming waves travel in the positive x direction and outgoing waves in the negative x direction.

Waves with $E \to -E$ are not solutions of the Schrödinger equation, but only to its complex conjugate. That is different for solutions to the Klein-Gordon equation, which we will describe next.

1.3 The Klein-Gordon equation

To find the wave equation for massive particles Schrödinger and others had originally started from the relativistic relation between energy and momentum, Eq. (1.6). Using again the operator substitution in Eq. (1.8) one obtains a wave equation

$$\boxed{-\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\phi = -\nabla^2\phi + \frac{m^2c^2}{\hbar^2} \, \phi} \tag{1.20}$$

This equation is called the *Klein-Gordon equation.* Having once seen it with the factors $\hbar$ and c included, we will from now on omit them. The Klein-Gordon equation can then be efficiently written in four-vector notation as

$$\left(\Box + m^2\right)\ \phi(x) = 0\ , \tag{1.21}$$

where

$$\Box\ \equiv\ \partial_\mu \partial^\mu\ \equiv\ \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 \tag{1.22}$$

is the so-called d'Alembert operator.

Unlike the Schrödinger equation, the KG equation does not contain factors i. Consequently, it can have both real and complex solutions. These have different applications. In chapter 3 we shall see an example of the KG equation for a real field. In this section we assume that the waves are complex.

Planes waves of the form

$$\phi(x)\ =\ N\, e^{i(\boldsymbol{px}-Et)} = e^{-ip_\mu x^\mu} \tag{1.23}$$

with $p^\mu = (E, \boldsymbol{p})$ are solutions of the KG equation provided that they satisfy the dispersion relation $E^2 = p^2 + m^2$. Nothing restricts solution to have positive energy: we discuss the interpretation of negative energy solutions later in this lecture.

Any solution to the KG equation can be written as a superposition of plane waves, like for the Schrödinger equation. However, in contrast to the classical case, the complex conjugate of the plane wave above

$$\phi^*(x)\ =\ N\, e^{i(-\boldsymbol{px}+Et)} = e^{ip_\mu x^\mu} \tag{1.24}$$

is also a solution to the KG equation and need to be accounted for in the decomposition. It is not independent though, since $\phi^*(p,E) = \phi(-p,-E)$. Consequently, we can write the generic decomposition restricting ourselves to positive energy solutions, if we write

$$\phi(x)\ =\ \int \mathrm{d}^3\boldsymbol{p}\ \left[A(p)\, e^{-ip_\mu x^\mu}\ +\ B(p)\, e^{ip_\mu x^\mu}\right] \tag{1.25}$$

with $E = +\sqrt{p^2+m^2}$. By popular convention, motivated later, we identify the first exponent as an incoming particle wave, or an outgoing anti-particle wave, and vice-versa for the second exponent.

In analogy to Eq. 1.13 we now derive a continuity equation. We multiply the Klein Gorden equation for ϕ from the left by $-i\phi^*$, then add to the complex conjugate equation:

$$\begin{aligned}
-i\phi^* \left(-\frac{\partial^2\phi}{\partial t^2}\right) &= -i\phi^*\ \left(-\nabla^2\phi + m^2\,\phi\right) \\
i\phi\ \left(-\frac{\partial^2\phi^*}{\partial t^2}\right) &= i\phi\ \left(-\nabla^2\phi^* + m^2\phi^*\right) \\
\hline
\frac{\partial}{\partial t}\underbrace{i\left(\phi^*\,\frac{\partial\phi}{\partial t} - \phi\frac{\partial\phi^*}{\partial t}\right)}_{\rho} &= \boldsymbol{\nabla}\cdot\underbrace{\left[i\ \left(\phi^*\,\boldsymbol{\nabla}\phi - \phi\,\boldsymbol{\nabla}\phi^*\right)\right]}_{-\boldsymbol{j}}
\end{aligned} \; + \tag{1.26}$$

where we can recognize again the continuity equation. In four-vector notation the *conserved current* becomes

$$j^\mu \;=\; (\rho, \boldsymbol{j}) = i\left[\phi^*\left(\partial^\mu\phi\right) - \left(\partial^\mu\phi^*\right)\phi\right] \tag{1.27}$$

while the continuity equation is simply

$$\partial_\mu j^\mu \;=\; 0 \tag{1.28}$$

You may wonder why we introduced the factor i in the current: this is in order to make the density real.

Substituting the plane wave solution gives

$$\begin{aligned} \rho &= 2\,|N|^2\,E \\ \boldsymbol{j} &= 2\,|N|^2\,\boldsymbol{p} \end{aligned} \tag{1.29}$$

or in four-vector notation

$$j^\mu \;=\; 2\,|N|^2\,p^\mu. \tag{1.30}$$

Like for the the classical Schrödinger equation, the ratio of the current to the density is still a velocity since $\boldsymbol{v} = \boldsymbol{p}/E$. However, in contrast to the non-relativistic case, the density of the Klein-Gordon wave is proportional to the energy. This is a direct consequence of the Klein-Gordon equation being second order in the time derivative.

We write the conserved current as a four-vector assuming that it transforms under Lorentz transformation the way four-vector are supposed to do. It is not so hard to show this by looking at how a volume and velocity change under Lorentz transformations (see e.g. the discussion in Feynman's Lectures, Vol. 2, sec. 13.7.) The short argument is that since ϕ is a Lorentz-scalar, and ∂^μ a Lorentz vector, their product must be a Lorentz vector.

You may remember that conservation rules in physics are related to symmetries. That makes you wonder which symmetry leads to the conserved currents for the Schrödinger and Klein-Gordon equations. In Lecture 8 we discuss Noether's theorem and show that it is the *phase* invariance of the Lagrangian, a so-called $U(1)$ symmetry. The phase of the wave functions is not a physical observable. For QM wave functions the conserved current implies that probability is conserved. For the QED Lagrangian it implies that charge is conserved.

1.4 Interpretation of negative energy solutions

The dispersion relation $E^2 = p^2 + m^2$ leaves the *sign* of the energy ambiguous. This leads to an interpretation problem: what is the meaning of the states with $E = -\sqrt{p^2+m^2}$ which have a negative density? We cannot just leave those states away since we need to work with a complete set of states.

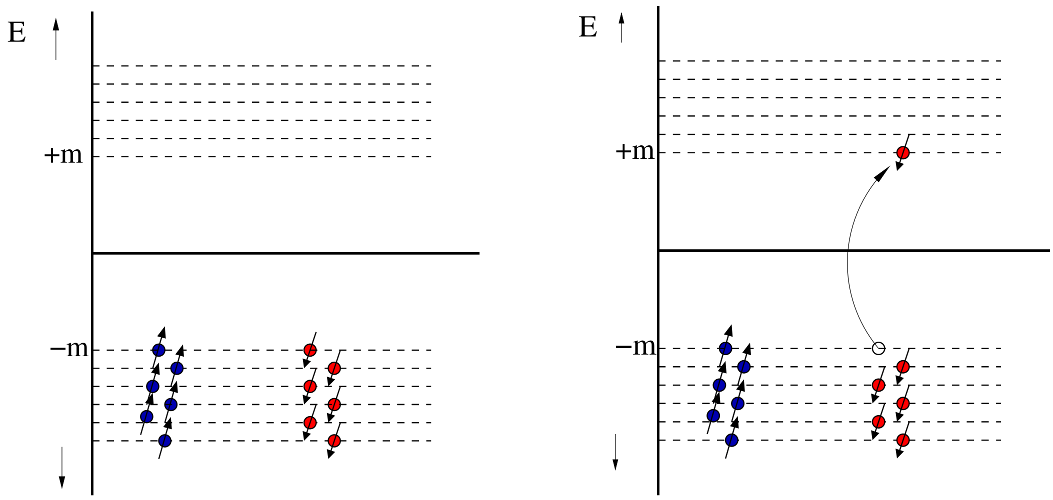


Figure 1.1: Dirac's interpretation of negative energy solutions: "holes"

1.4.1 Dirac's interpretation

In 1927 Dirac offered an interpretation of the negative energy states. To circumvent the problem of a negative density he developed a wave equation that was *linear* in time and space. The 'Dirac equation' turned out to describe particles with spin 1/2. (At this point in the course we consider spinless particles. The wave function ϕ is a scalar quantity as there is no individual spin "up" or spin "down" component. We shall discuss the Dirac equation later in this course.) Unfortunately, this did not solve the problem of negative energy states.

In a feat that is illustrative for his ingenuity Dirac turned to Pauli's exclusion principle. The exclusion principle states that identical fermions cannot occupy the same quantum state. Dirac's picture of the vacuum and of a particle are schematically represented in Fig. 1.1.

The plot shows all the available energy levels of an electron. Its lowest absolute energy level is given by $|E| = m$. Dirac imagined the vacuum to contain an infinite number of states with negative energy which are all occupied. Since an electron is a spin-1/2 particle each state can only contain one spin "up" electron and one spin-"down" electron. All the negative energy levels are filled. Such a vacuum ("sea") is not detectable since the electrons in it cannot interact, i.e. go to another state.

If energy is added to the system, an electron can be kicked out of the sea. It now gets a positive energy with $E > m$. This means this electron becomes visible as it can now interact. At the same time a "hole" in the sea has appeared. This hole can be interpreted as a positive charge at that position: an anti-electron! Dirac's original hope was that he could describe the proton in such a way, but it is essential that the anti-particle mass is identical to that of the electron. Thus, Dirac predicted the positron, a particle that can be created by 'pair production'. The positron was discovered in 1931 by Anderson.

There is one problem with the Dirac interpretation: it only works for fermions!

1.4.2 Pauli-Weisskopf interpretation

Pauli and Weiskopf proposed in 1934 that the density should be regarded as a *charge* density. For an electron the charge density is written as

$$j^{\mu} = -ie(\phi^{*}\partial^{\mu}\phi - \phi\partial^{\mu}\phi^{*}). \tag{1.31}$$

To describe electromagnetic interactions of charged particles we do not need to consider anything but the movement of 'charge'. This motivates the interpretation as a charge current. Clearly, in this interpretation solutions with a negative density pose no longer a concern. However, it does not yet solve the issue of negative energies.

1.4.3 Feynman-Stückelberg interpretation

Stückelberg and later Feynman took this approach one step further. Consider the current for a plane wave describing an electron with momentum $\boldsymbol{p}$ and energy E. Since the electron has charge $-e$, this current is

$$j^{\mu}(-e) = -2e\ |N|^{2}\ p^{\mu} = -2e\ |N|^{2}\ (E, \boldsymbol{p})\,. \tag{1.32}$$

Now consider the current for a positron with momentum $\boldsymbol{p}$. Its current is

$$j^{\mu}(+e) = +2e\ |N|^{2}\ p^{\mu} = -2e\ |N|^{2}\ (-E, -\boldsymbol{p})\,. \tag{1.33}$$

Consequently, the current for the positron is identical to the current for the electron but with negative energy and traveling in the opposite direction. Or, in terms of the plane waves, to go from the positron current to the electron current, we just need to change the sign in the exponent of $e^{ix_{\mu}p^{\mu}}$. By our earlier convention, this is equivalent to saying that the incoming plane wave of a positron is identical to the outgoing wave of an electron.

Now consider what happens to the electron wave function if we change the direction of time: We will have $ct \to -ct$ and $\boldsymbol{p} \to -\boldsymbol{p}$. You immediately notice that this has exactly the same effect on the plane wave exponent as the transformation $(E, p) \to (-E, -p)$. In other words, we can interprete the negative energy current of the electron as an electron moving *backward in time.* This current is identical to that of a positron moving forward in time.

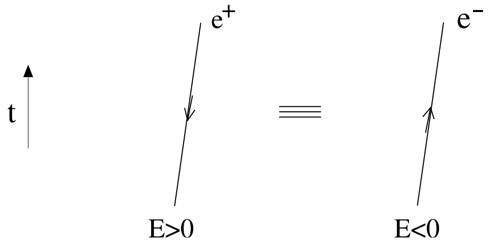


Figure 1.2: A positron travelling forward in time **is** an electron travelling backwards in time.

This interpretation, illustrated in Fig. 1.2, is very convenient when computing scattering amplitudes: in our calculations with Feynman diagrams we can now express everything in terms of particle waves, replacing every anti-particle with momentum p^μ by a particle with momentum $-p^\mu$, as if it were traveling backward in time. For example, the process of an absorption of a positron with energy E is the same as the emission of an electron with energy $-E$ (see Fig.1.3). Likewise, the process of an incoming positron scattering off a potential will be calculated as that of a scattering electron travelling back in time (see Fig. 1.4).

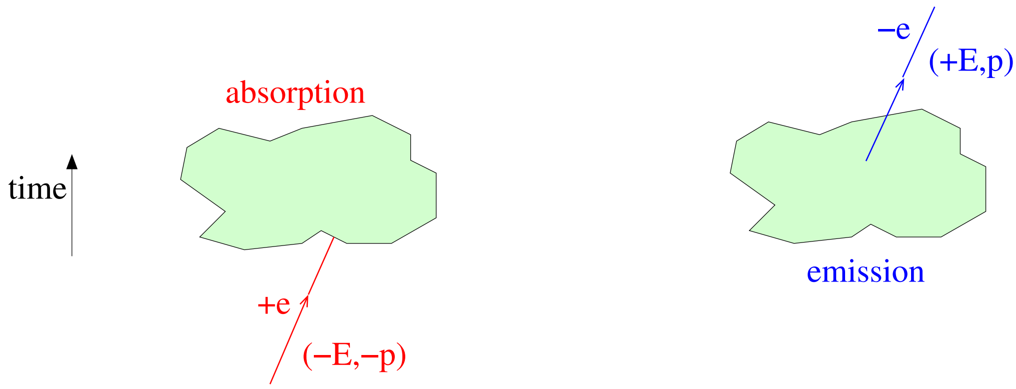


Figure 1.3: There is no difference between the process of an absorption of a positron with $p^\mu = (-E, -\boldsymbol{p})$ and the emission of an electron with $p^\mu = (e, \boldsymbol{p})$.

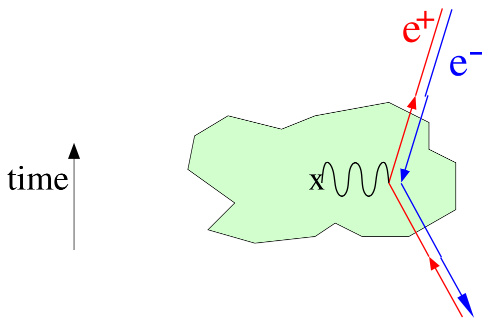


Figure 1.4: In terms of the charge current density $j^\mu_{+(E,\boldsymbol{p})}(+e) \equiv j^\mu_{-(E,\boldsymbol{p})}(-e)$

The advantage of this approach becomes more apparent when one considers higher order corrections to the amplitudes. Consider the scattering of an electron on a localized potential, illustrated in Fig. 1.5. To first order the interaction of the electron with the perturbation is described by the exchange of a single photon. When the calculation is extended to second order the electron interacts twice with the field. It is important to note that this second order contribution can occur in two time orderings as indicated in the figure. These two contributions are different and both of them must be included in a relativistically covariant computation.

The time-ordering on the right can be viewed in two ways:

- The electron scatters at time t_2 runs back in time and scatters at t_1.

- First at time t_1 "spontaneously" an e^-e^+ pair is created from the vacuum. Later-on, at time t_2, the produced positron annihilates with the incoming electron, while the produced electron emerges from the scattering process.

The second interpretation would allow the process to be computed in terms of particles and anti-particles that travel forward in time. However, the first interpretation is just more economic. We realize that the vacuum has become a complex environment since particle pairs can spontaneously emerge from it and dissolve into it!

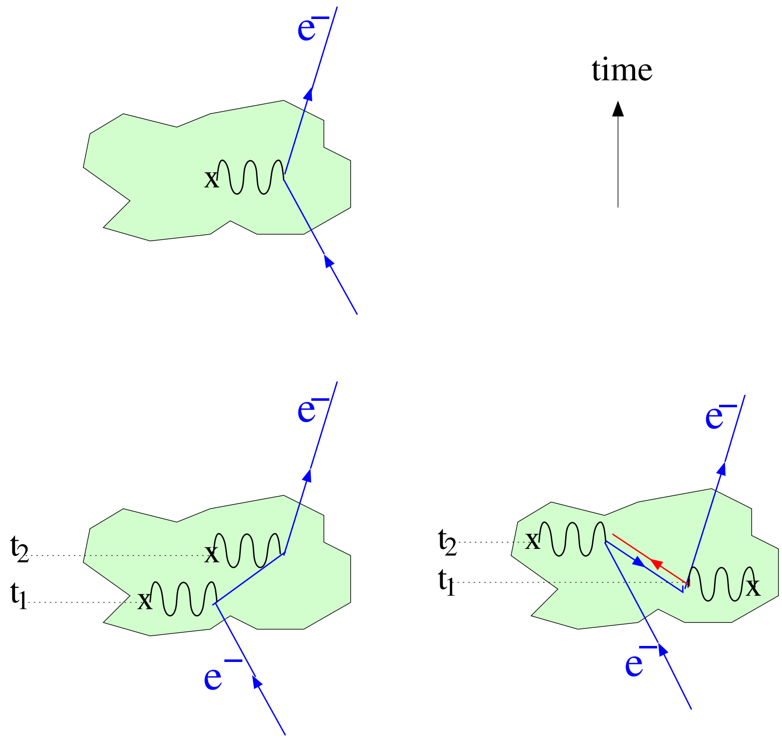


Figure 1.5: First and second order scattering.

Exercises

Exercise 1.1 (Conversion factors)
Derive the conversion factors for mass, length and time in table i.6.

Exercise 1.2 (Kinematics of Z production)
The Z-boson has a mass of 91.1 GeV. It can be produced by annihilation of an electron and a positron. The mass of an electron, as well as that of a positron, is 0.511 MeV.

(a) Draw the (dominant) Feynman diagram for this process.

(b) Assume that an electron and a positron are accelerated in opposite directions with equal beam energies and collide head-on. Calculate the beam energy required to produce a Z-boson.

(c) Assume that a beam of positron particles is shot on a target containing electrons. Calculate the beam energy required for the positron beam to produce Z-bosons.

(d) This experiment was carried out in the 1990's. Which method (b or c) do you think was used? Why?

Exercise 1.3 (The Yukawa potential)

(a) The wave equation for an electromagnetic wave in vacuum is given by:

$$\Box V = 0 \qquad ; \qquad \Box \equiv \partial_\mu \partial^\mu \equiv \frac{\partial^2}{\partial t^2} - \nabla^2$$

which in the static case can be written in the form of Laplace equation:

$$\nabla^2 V = 0$$

Now consider a point charge in vacuum. Exploiting spherical symmetry, show that this equation leads to a 'potential' $V(r) \propto 1/r$.
Hint: look up the expression for the Laplace operator in spherical coordinates.

(b) The wave equation for a massive field is the Klein Gordon equation:

$$\Box U + m^2 U = 0$$

which, again in the static case can be written in the form:

$$\nabla^2 U - m^2 U = 0$$

Show, again assuming spherical symmetry, that Yukawa's potential is a solution of the equation for a massive force carrier. What is the relation between the mass m of the force carrier and the range R of the force?

(c) Estimate the mass of the π-meson assuming that the range of the nucleon force is $1.5 \times 10^{-15}\,\mathrm{m} = 1.5\,\mathrm{fm}$.

Exercise 1.4 (Strangeness conservation (From A&H, chapter 1.))
Using the concept of strangeness conservation, explain why the threshold energy (for π^- incident on stationary protons) for

$$\pi^- + \mathrm{p} \to K^0 + \text{anything}$$

is less than for

$$\pi^- + \mathrm{p} \to \overline{K}^0 + \text{anything}$$

assuming that both processes proceed through the strong interaction.
Hint: Deduce what the minimal quark content of 'anything' is.

Exercise 1.5 (KG conserved current)
Verify that the conserved current Eq. (1.27) satisfies the continuity equation: compute $\partial_\mu j^\mu$ explicitly and make use of the KG equation to show that the result is zero.
Hint: note that that for any fourvectors A and B we have $A_\mu B^\mu = A^\mu B_\mu$.

Exercise 1.6 (From A&H, chapter 3. See also Griffiths, exercise 7.1)
In this exercise we derive expression Eq. (i.22) of the Introductory chapter.

(a) Start with the expressions for a Lorentz transformation along the x^1 axis in Eq. (i.17). Write down the *inverse* transformation (*i.e.* express (x^0, x^1) in $(x^{0\prime}, x^{1\prime})$)

(b) Use the chain rule to express the derivatives $\partial/\partial x^{0\prime}$ and $\partial/\partial x^{1\prime}$ in the derivatives $\partial/\partial x^0$ and $\partial/\partial x^1$.

(c) Use the result to show that $(\partial/\partial x^0, -\partial/\partial x^1)$ transforms in the same way as (x^0, x^1).

Exercise 1.7 (Wave packets (optional!))
In the coordinate representation a plane wave with momentum $p = \hbar k$ is infinitely dislocalised in space. This does not quite correspond to our picture of a particle, which is why we usually visualize particles as wave packets, superpositions of plane waves that have a finite spread both in momentum and coordinate space. As we shall see in the following, the irony is that such wave packets are *dispersive* in QM: their size increases as a function of time. So, even wave packets can hardly be thought of as representing particles.

(a) Consider a one-dimensional Gaussian wave packet that at time $t = 0$ is given by

$$\psi(x, 0) \;=\; A e^{-ax^2 + ik_0 x}$$

with a real and positive. Compute the normalization constant A such that

$$\int_{-\infty}^{+\infty} |\psi(x, 0)|^2 \mathrm{d}x = 1.$$

Hint:

$$\int_{\infty}^{\infty} e^{-y^2} \mathrm{d}y = \sqrt{\pi}$$

(b) Take the Fourier transform to derive the wave function in momentum space at $t = 0$,

$$\psi(k) = \left(\frac{1}{2a\pi}\right)^{1/4} e^{-(k-k_0)^2/4a}$$

Hint: You can write

$$\exp\left(-ax^2 + bx\right) = \exp\left[-a(x - b/2a)^2 + b^2/4a\right]$$

and then move the integration boundaries by $-b/2a$*. (Don't mind that* b *is complex.)*

(c) Use this result and Eq. (1.17) to show that the solution to the Schrödinger equation (with $E(p) = p^2/2m$ or $\omega(k) = \hbar k^2/2m$) is given by

$$\psi(x,t) = \left(\frac{2a}{\pi}\right)^{1/4} (1 + i\nu t)^{-1/2} \exp\left(\frac{-ax^2 + ik_0x - i\nu t k_0^2/4a}{1 + i\nu t}\right)$$

with $\nu \equiv 2\hbar a/m$.

(d) Compute $|\psi(x,t)|^2$. Qualitatively, what happens to ψ^2 as time goes on?

(e) Now compute the same for a solution to the massless Klein-Gordon equation ($\omega = ck$). Note that the wave packet maintains its size as a function of time.

Exercise 1.8 (The Quark Model (optional!))

(a) Quarks are fermions with spin 1/2. Show that the spin of a meson (2 quarks) can be either a triplet of spin 1 or a singlet of spin 0.
Hint: Remember the Clebsch Gordon coefficients in adding quantum numbers.
In group theory this is often represented as the product of two doublets leads to the sum of a triplet and a singlet: $\mathbf{2} \otimes \mathbf{2} = \mathbf{3} \oplus \mathbf{1}$ or, in terms of quantum numbers: $1/2 \otimes 1/2 = 1 \oplus 0$.

(b) Show that for baryon spin states we can write: $1/2 \otimes 1/2 \otimes 1/2 = 3/2 \oplus 1/2 \oplus 1/2$ or equivalently $\mathbf{2} \otimes \mathbf{2} \otimes \mathbf{2} = \mathbf{4} \oplus \mathbf{2} \oplus \mathbf{2}$

(c) Let us restrict ourselves to two quark flavours: u and d. We introduce a new quantum number, called *isospin* in complete analogy with spin, and we refer to the u quark as the isospin +1/2 component and the d quark to the isospin -1/2 component (or u= isospin "up" and d=isospin "down"). What are the possible isospin values for the resulting baryon?

(d) The Δ^{++} particle is in the lowest angular momentum state ($L = 0$) and has spin $J_3 = 3/2$ and isospin $I_3 = 3/2$. The overall wavefunction (L⇒space-part, S⇒spin-part, I⇒isospin-part) must be anti-symmetric under exchange of any of the quarks. The symmetry of the space, spin and isospin part has a consequence for the required symmetry of the Colour part of the wave function. Write down the colour part of the wave-function taking into account that the particle is colour neutral.

(e) In the case that we include the s quark the flavour part of the wave function becomes: $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}$. In the case that we include all 6 quarks it becomes: $\mathbf{6} \otimes \mathbf{6} \otimes \mathbf{6}$. However, this is not a good symmetry. Why not?

Lecture 2

Perturbation Theory and Fermi's Golden Rule

In this chapter we discuss Fermi's golden rule, which allows us to compute cross-sections and decay rates. A very readable account of this is given in Griffiths chapter 6 and Thomson section 2.3.6 and chapter 3.

2.1 Decay and scattering observables

Most species of particles do not live long. This holds for all baryons except the proton (even the neutron decays, when it is not inside a nucleus), but also for the muon and the tau. As particles do not age, the probability to decay is independent of time. Given a large number of particles N_0, the number of surviving particles is hence given by the exponential law

$$N(t) \;=\; N_0\, e^{-t/\tau} \;, \tag{2.1}$$

where τ is the mean lifetime. For particles that decay via the weak interaction, the mean lifetime is typically $10^{-12} - 10^{-9}$ seconds. A notable exception is the neutron which lives for about 15 minutes.

The mean lifetime is inversely proportional to the *decay width*

$$\Gamma \;=\; \frac{\hbar}{\tau} \;, \tag{2.2}$$

which has units of energy. If the particle can decay through different decay channels (*e.g.* a charged pion can decay to $\mu^-\bar{\nu}_\mu$ and to $e^-\bar{\nu}_\mu$), the decay width can be written as the sum of the decay widths to the individual channels

$$\Gamma \;=\; \sum_i \Gamma_i \;. \tag{2.3}$$

The ratio Γ_i/Γ is called the branching fraction. The *particle data book* is full of branching fractions of species in the particle zoo. The *decay rate* to a particular final state i, *i.e.* the average number of decays per unit of time into i, is just $\Gamma_i/\hbar$.

If the decay is to more than two particles, the distribution of angles and energies of particles in the final state becomes an observable as well. That is why we often consider *partial* or *differential decay widths*,

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}\boldsymbol{p}_1 \cdots \mathrm{d}\boldsymbol{p}_N} \,, \tag{2.4}$$

where $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_N$ are the momenta of the N particles in the final state.

Besides decay widths we also measure *scattering cross-sections*. (In fact, in our computations, decays and scattering are quite similar, so we deal with both at once.) In scattering experiments we collide beams of particles and study the collision rate. Consider an experiment in which we scatter a beam of particles A on a target of particles B. If n_A is the particle number density in the beam, and v_A is the particle velocity, the number of collisions per second per unit volume of B is

$$\frac{\mathrm{d}N}{\mathrm{d}t} = v_A \, n_A \, n_B \, \sigma_{\mathrm{tot}} \,. \tag{2.5}$$

The quantity σ_{tot} is called the total scattering cross-section. It has units of surface. In most cases we do not study the total collision rate, but rather the rate of particular final states. The total cross-section is a sum of cross-sections for all possible final states, such that

$$\sigma_{\mathrm{tot}} = \sum_i \sigma_i \,. \tag{2.6}$$

At lower collision energies the total proton-proton cross-section is about 40 *milli-barn*, or $4 \cdot 10^{-30}$ square meter . If you interprete the cross-section as the surface of the silhouet of the proton, then the proton radius is about one femto-meter.

Since the energy and direction of final state particles can be measured as well, we usually consider differential scattering cross-sections,

$$\frac{\mathrm{d}\sigma(A + B \to f_1 + \cdots + f_N)}{\mathrm{d}\boldsymbol{p}_1 \cdots \mathrm{d}\boldsymbol{p}_N} \,. \tag{2.7}$$

The expression for the calculation of a (differential) cross-section can be written schematically as

$$\boxed{\mathrm{d}\sigma = \frac{W_{\mathrm{fi}}}{\mathrm{flux}} \, \mathrm{d}\Phi} \tag{2.8}$$

The ingredients to this expression are:

1. the *transition rate* W_{fi}. You can think of this as the probability per unit time and unit volume to go from an initial state i to a final state f;
2. a *flux* factor that accounts for the 'density' of the incoming states;

3. the Lorentz invariant *phase space* factor $d\Phi$, sometimes referred to as 'dLIPS'. It accounts for the density of the outgoing states. (It takes care of the fact that experiments cannot observe individual states but integrate over a number of states with nearly equal momenta.)

The 'physics' (the dynamics of the interaction) is contained in the transition rate W_{fi}. The flux and the phase space factors are just 'bookkeeping', required to compare the result with the measurements.

The rigorous computation of the transition rate requires quantum field theory, which is outside the scope of this course. However, to illustrate the concepts we discuss non-relativistic scattering of a single particle in a time-dependent potential and formulate the result in a Lorentz covariant way. In the next chapter we will derive the lowest order amplitude for the scattering of $A + B \to A + B$, which can still be done without field theory. We can link that result to the 'Feynman rules' derived in field theory.

2.2 Non-relativistic scattering

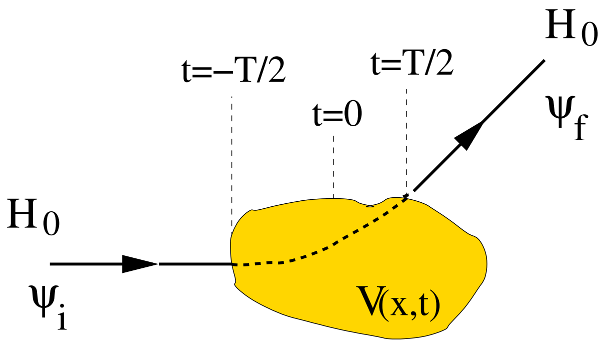


Figure 2.1: Scattering of a single particle in a potential.

Consider the scattering of a particle in a potential as depicted in Fig. 2.1 Assume that both long *before* and long *after* the interaction takes place, the system is described by the free Schrödinger equation,

$$i\hbar \frac{\partial \psi}{\partial t} = H_0 \, \psi \tag{2.9}$$

where H_0 is the unperturbed, *time-independent* Hamiltonian for a free particle. Let $\phi_m(\boldsymbol{x})$ be a normalized eigenstate of H_0 with eigenvalue E_m,

$$H_0 \phi_m(\boldsymbol{x}) = E_m \phi_m(\boldsymbol{x}). \tag{2.10}$$

The states ϕ_m form an orthonormal basis,

$$\int \phi_m^*(\boldsymbol{x}) \, \phi_n(\boldsymbol{x}) \, d^3 x = \delta_{mn}. \tag{2.11}$$

We use the Kronecker delta, as if the spectrum of eigenstates is discrete. In chapter 2 we considered a continuous spectrum of eigenstates for the free Hamiltonian, 'numbered' by the wave number k. Eventually, we could do that here, too, replacing the Kronecker delta by a Dirac delta-function. However, it is trivial to change between the two and the notation is a bit easier when we work with a discrete set of states.

The time-dependent wave function

$$\psi_m(\boldsymbol{x}, t) = \phi_m(\boldsymbol{x})\, e^{-iE_m t/\hbar}. \tag{2.12}$$

is a solution to the Schrödinger equation. Since these states form a complete set, any other wave function can be written as a superposition of the wave functions ψ_m.

Now consider a Hamiltonian that includes a time-dependent perturbation,

$$i\hbar \frac{\partial \psi}{\partial t} \;=\; \left(H_0 + V\left(\boldsymbol{x}, t\right)\right)\, \psi\ . \tag{2.13}$$

Any solution ψ can be written as

$$\psi = \sum_{n=0}^{\infty} a_n(t)\, \phi_n(\boldsymbol{x})\, e^{-iE_n t}. \tag{2.14}$$

where the $a_n(t)$ are time-dependent complex coefficients. We require ψ to be normalized, which implies that $\sum |a_n(t)|^2 = 1$. The probability to find ψ in state n at time t is just $|a_n(t)|^2$.

To determine the coefficients $a_n(t)$ we substitute (2.14) in (2.13) and find

$$i\hbar \sum_{n=0}^{\infty} \frac{\mathrm{d}a_n(t)}{\mathrm{d}t}\, \phi_n(\boldsymbol{x})\, e^{-iE_n t} \;=\; \sum_{n=0}^{\infty} V(\boldsymbol{x}, t)\, a_n(t)\, \phi_n(\boldsymbol{x})\, e^{-iE_n t}\ , \tag{2.15}$$

where we have used that the ψ_m are solutions of the free Schrödinger equation. Multiply the resulting equation from the left with $\psi_f^* = \phi_f^*(\boldsymbol{x})\, e^{iE_f t}$ and integrate over $\boldsymbol{x}$ to obtain

$$\begin{aligned} i\hbar \sum_{n=0}^{\infty} \frac{\mathrm{d}a_n(t)}{\mathrm{d}t} \underbrace{\int \mathrm{d}^3x\, \phi_f^*(\boldsymbol{x})\, \phi_n(\boldsymbol{x})}_{\delta_{fn}}\; e^{-i\left(E_n - E_f\right)t\hbar} \;=\; & \\ \sum_{n=0}^{\infty} a_n(t) \int \mathrm{d}^3x\, \phi_f^*(\boldsymbol{x})\, V(\boldsymbol{x}, t)\, \phi_n(\boldsymbol{x})\, e^{-i\left(E_n - E_f\right)t/\hbar} & \end{aligned} \tag{2.16}$$

Using the orthonormality relation for ϕ_m we then arrive at the following coupled linear differential equation for $a_k(t)$,

$$i\hbar \frac{\mathrm{d}a_k(t)}{\mathrm{d}t} = \sum_{n=0}^{\infty} a_n(t)\, V_{kn} e^{i\omega_{kn} t}\ , \tag{2.17}$$

where we have defined

$$\omega_{kn} \;=\; (E_k - E_n)/\hbar \tag{2.18}$$

and what is sometimes called the *transition matrix element*

$$V_{kn}(t) \;=\; \int \mathrm{d}^3x \; \phi_k^*(\boldsymbol{x}) \, V(\boldsymbol{x},t) \, \phi_n(\boldsymbol{x}) \; . \tag{2.19}$$

In some cases the set of equations (2.17) can be solved explicitly. A general solution is obtained in perturbation theory, by expanding in V_{kn}. The approximation of order $p+1$ can be obtained by inserting the p-th order result on the right hand side of Eq. (2.17),

$$i\hbar \frac{\mathrm{d}a_k^{(p+1)}(t)}{\mathrm{d}t} \;\approx \sum_n a_n^{(p)}(t) V_{kn}(t) e^{i\omega_{kn}t} \tag{2.20}$$

Without loss of generality we now assume that the incoming wave is prepared in eigenstate i of the free Hamiltonian, *i.e.* $a_k(-\infty) = \delta_{ki}$. The zeroeth order approximation then is $a_k^{(0)}(t) = \delta_{ki}$ (no interaction occurs) and the first order result becomes

$$i\hbar \frac{\mathrm{d}a_k^{(1)}(t)}{\mathrm{d}t} \;=\; V_{ki}(t) e^{i\omega_{ki}t} \tag{2.21}$$

Using that $a_f(-\infty) = 0$ and integrating this equation we obtain for the coefficient $a_k^{(1)}(t)$ at time t,

$$a_k^{(1)}(t) \;=\; \int_{-\infty}^{t} \frac{\mathrm{d}a_f(t')}{\mathrm{d}t} \, \mathrm{d}t' \;=\; \frac{1}{i\hbar} \int_{-\infty}^{t} V_{ki}(t') e^{i\omega_{ki}t'} \, \mathrm{d}t' \qquad \text{for } k \neq i \tag{2.22}$$

Higher order approximations can be obtained by inserting the lowest order solution in the right side of Eq. (2.20). (See textbooks.) A graphical illustration of the first and second order perturbation is given in Fig. 2.2. The lowest order approximation makes one 'quantum step' from the initial state i to the final state f, while the second order approximation includes all amplitudes $i \to n \to f$.

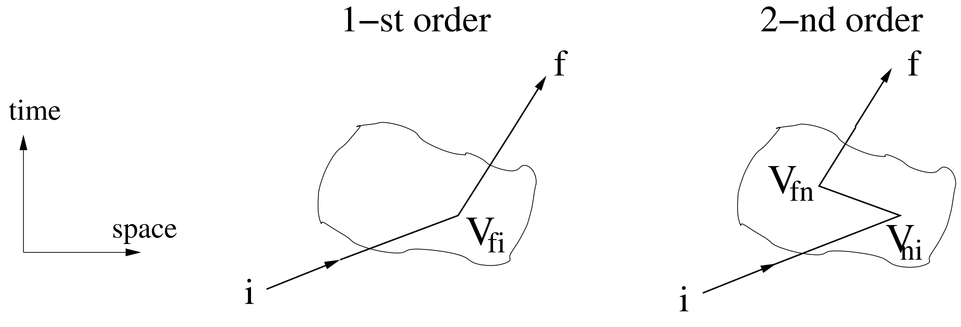


Figure 2.2: First and second order approximation in scattering.

In the following we only consider the first order approximation (Born approximation). We define the *transition amplitude* T_{fi} as the amplitude to go from a state i to a final state f at large times,

$$T_{\mathrm{fi}} \;\equiv\; a_{\mathrm{f}}(t \to \infty) = \frac{1}{i\hbar} \int_{-\infty}^{\infty} \mathrm{d}t \int \mathrm{d}^3x \; \psi_{\mathrm{f}}^*(\boldsymbol{x},t) \, V(\boldsymbol{x},t) \, \psi_{\mathrm{i}}(\boldsymbol{x},t) \tag{2.23}$$

where we substituted the definitions of V_{kn} and ω_{kn}. We can write the result more compactly as

$$\boxed{T_{\text{fi}} \;=\; \frac{1}{i\hbar}\int \mathrm{d}^4x\,\psi_{\text{f}}^*(x)\,V(x)\,\psi_{\text{i}}(x)} \tag{2.24}$$

Somewhat deceptively, the expression for T_{fi} seems to have a Lorentz covariant form. However, as we have seen in the previous lecture, the 'classical' free particle waves corresponds to a density that does not correctly transform under Lorentz transformations. Therefore, T_{fi} is actually not yet a proper Lorentz scalar.

We now make a simplification and consider a potential that is time-independent. The expression for the transition amplitude then becomes

$$T_{\text{fi}} \;=\; \frac{V_{\text{fi}}}{i\hbar}\int_{-\infty}^{\infty} e^{i\omega_{\text{fi}}t}\mathrm{d}t \;=\; -2\pi\, i\, V_{\text{fi}}\,\delta\left(E_f - E_i\right) \tag{2.25}$$

where we have used that the integral is an important representation of the Dirac δ function

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ikx}\,dk \tag{2.26}$$

and substituted our definition of ω_{fi}. The δ function expresses conservation of energy.

The transition amplitude T_{fi} is dimensionless. Can we interprete $|T_{\text{fi}}|^2$ as a probability? Well, there is one conceptual problem and one pragmatic problem. The conceptual problem is that if the potential is time-independent, then this probability will just grow with time. The pragmatic problem is that there is the δ function. These issues can be solved by considering a potential that is turned on for a 'finite time' T. We define the *mean transition rate* in the limit for large T as

$$W_{\text{fi}} \;\equiv\; \lim_{T\to\infty}\frac{|T_{\text{fi}}|^2}{T}\,. \tag{2.27}$$

For an interaction that is turned on at time $-T/2$ and turned off at time $T/2$, the equation above can be integrated to give for the transition amplitude at $T/2$,

$$a_f(T/2) \;=\; \frac{V_{\text{fi}}}{i\hbar}\int_{-T/2}^{T/2} e^{i\omega_{\text{fi}}t'}\,\mathrm{d}t' \;=\; \frac{2V_{\text{fi}}}{i\hbar}\,\frac{\sin(\omega_{\text{fi}}T/2)}{\omega_{\text{fi}}}\,. \tag{2.28}$$

Inserting this in the definition for the transition rate gives

$$W_{\text{fi}} \;\equiv\; \lim_{T\to\infty}\frac{4|V_{\text{fi}}|^2}{\hbar^2}\,\frac{\sin^2(\omega_{\text{fi}}T/2)}{\omega_{\text{fi}}^2 T}\,. \tag{2.29}$$

The function on the right is strongly peaked near $\omega_{\text{fi}} = (E_f - E_i)/\hbar = 0$, again enforcing energy conservation. In fact, for $T\to\infty$ it is yet another representation of the Dirac δ function,

$$\delta(x) = \lim_{\alpha\to\infty}\,\frac{1}{\pi}\frac{\sin^2\alpha x}{\alpha x^2}\,. \tag{2.30}$$

Substituting this in the equation, we obtain

$$\boxed{W_{\rm fi} \;=\; \frac{2\pi}{\hbar^2}\,|V_{\rm fi}|^2\,\delta(\omega_{\rm fi}) \;=\; \frac{2\pi}{\hbar}\,|V_{\rm fi}|^2\,\delta(E_f - E_i)\ .} \tag{2.31}$$

You can verify that $W_{\rm fi}$ is indeed a rate: $V_{\rm fi}$ is an energy, one of the factors of energy is canceled by the δ function and the other one is divided by $\hbar$ to turn it into reciprocal time.

As indicated before we can never actually probe final states with definite energy in a measurement with finite duration. In general, there will be a number of states with energy close to E_i that can be reached. Assuming that these states can be numbered by a continuos variable n, the total transition rate can be written as an integral over these final states

$$\begin{aligned} \overline{W}_{\rm fi} &\equiv \int W_{\rm fi}\,{\rm d}n \\ &= \frac{2\pi}{\hbar}\int |V_{\rm fi}|^2\,\delta(E_n - E_i)\,{\rm d}n\ . \end{aligned} \tag{2.32}$$

If $\rho(E_f)$ is the density of states per unit energy near E_f, the number of final states with energy between E_f and $E_f + {\rm d}E_f$ is given by

$${\rm d}n \;=\; \rho(E_f)\,{\rm d}E_f\ . \tag{2.33}$$

Inserting this in the expression above, we obtain Fermi's (Second) Golden Rule,

$$\boxed{\begin{aligned} \overline{W}_{\rm fi} &\equiv \int W_{\rm fi}\,\rho\left(E_f\right)\,{\rm d}E_f \\ &= \frac{2\pi}{\hbar}\,\left|V_{\rm fi}\right|^2\,\rho\left(E_i\right)\ . \end{aligned}} \tag{2.34}$$

In this expression $\rho(E_i)$ is really the density of *final* states at the energy E_i. Some textbooks therefore write this as $\rho(E_f)|_{E_f=E_i}$.

Above, we encountered a δ function in the transition amplitude. To deal with the square of that δ function we considered a finite time interval and went back to the expression for $a_k^{(0)}(T/2)$ for finite times T, taking the limit $T \to \infty$ only after taking the square. To make the final step you need to recognize the special representation of the δ function. For future applications it is useful to know that one can also solve this problem differently, namely by taking the limit $T \to \infty$ one integral at a time:

$$\begin{aligned} |W_{\rm fi}| &= \lim_{T\to\infty}\frac{1}{T}\left(\frac{V_{\rm fi}}{i\hbar}\int_{-T/2}^{T/2} e^{i\omega_{\rm fi}t}{\rm d}t\right)\left(\frac{V_{\rm fi}}{i\hbar}\int_{-T/2}^{T/2} e^{i\omega_{\rm fi}t'}{\rm d}t'\right)^* \\ &= \frac{|V_{\rm fi}|^2}{\hbar^2}2\pi\delta(\omega_{\rm fi})\,\lim_{T\to\infty}\frac{1}{T}\underbrace{\int_{-T/2}^{T/2} dt'}_{T} \end{aligned}$$

The final result is of course identical. We will encounter this trick at various places when going from a transition amplitude to a transition rate.

You may wonder why we need to consider a finite time interval T. The reason is that when we assume that the initial state is an eigenstate of the free Hamiltonian with fixed momentum (or energy), we have lost track of where a particle is in both space *and* time. A moving wave packet would see the static potential during a finite time, but the plane waves do not. Just like we will need to normalize the wave functions on a finite volume, we will need to normalize the potential to a finite time. A proper treatment is rather lengthy and relies on the use of wave packets. (See e.g. the book by K.Gottfried, "Quantum Mechanics" (1966), Volume 1, sections 12, 56.) In the end, we can write transition probabilities in terms of plane waves, provided that we normalize to T and V. We discuss the normalization in more detail below.

2.3 Relativistic scattering

Fermi's golden rule allows us to compute the scattering rate of non-relativistic particles on a static potential. In scattering experiments at high energies we need to deal with two scattering particles, rather than single particles scattering on a source. As an example, consider two spin-less electrons scatter in their mutual electromagnetic field, as depicted in Fig. 2.3.

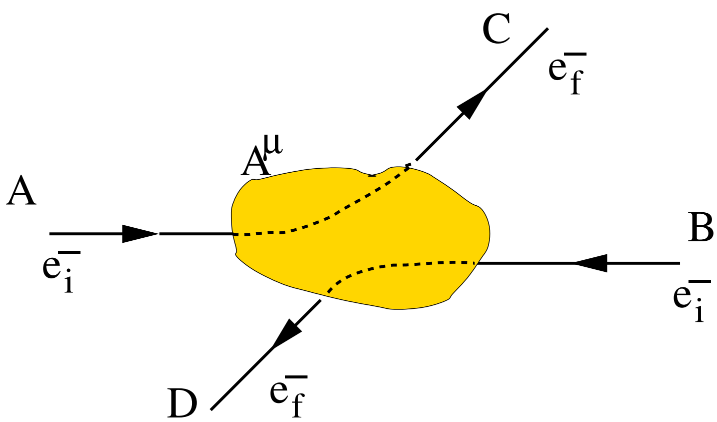


Figure 2.3: Scattering of two electrons in an electromagnetic potential.

Such scattering processes can be described by the exchange of virtual particles, Yukawa's force carriers. Even without understanding the details of the interaction, we can readily identify one place where it should differ from the discussion above: the result must somehow encode four-momentum conservation and not just energy conservation.

Our master formula for the differential cross-section, Eq. (2.8), is essentially a generalization to problems with more than one particle in the initial or final state. We cannot derive the expressions for a scattering cross-section at high energies without going through the machinery of quantum field theory. (This is not entirely true: see

Thomson, chapter 3 and section 5.1.) Instead, we will sketch the main results, then work through the electrodynamics of spin-less particles as an example in the next lectures.

In quantum electrodynamics with scalar particles the transition amplitude T_{fi} for the process $A+B \to C+D$ still takes the form in Eq. (2.24). Performing the integral using incoming and outgoing plane waves $\phi = Ne^{-ipx}$ the result can be written as

$$T_{\text{fi}} = -i\, N_A N_B N_C N_D\, (2\pi)^4\, \delta^4(p_A + p_B - p_C - p_D)\, \mathcal{M}\ . \tag{2.35}$$

where N_i are the plane wave normalization factors, which we will discuss shortly. The δ-function takes care of energy and momentum conservation in the process. (Note that the momentum vectors are four-vectors).

The quantity $\mathcal{M}$ is called the *(Lorentz) invariant amplitude.* It is computed using Feynman diagrams. For topologies with n particles (counting both incident and final state), the dimension of M is p^{4-n}. Using the convention for the wave function normalization described below, the invariant amplitude does not depend on arbitrary time intervals T or normalization volumes V.

To find the transition probability we square the expression for T_{fi},

$$\begin{aligned}
|T_{\text{fi}}|^2 &= |N_A N_B N_C N_D|^2\ |\mathcal{M}|^2 \int \mathrm{d}^4x\, e^{-i(p_A+p_B-p_C-p_D)x} \times \int \mathrm{d}^4x'\, e^{-i(p_A+p_B-p_C-p_D)x'} \\
&= |N_A N_B N_C N_D|^2\ |\mathcal{M}|^2\, (2\pi)^4\, \delta^4(p_A + p_B - p_C - p_D) \times \lim_{T,V\to\infty} \int_{TV} d^4x \\
&= |N_A N_B N_C N_D|^2\ |\mathcal{M}|^2\, (2\pi)^4\, \delta^4(p_A + p_B - p_C - p_D) \times \lim_{T,V\to\infty} TV
\end{aligned} \tag{2.36}$$

Since we now have a δ-function over 4 dimensions (the four-momentum rather than just the energy), the integral becomes proportional to both T and V. To get rid of them we consider a transition probability per unit time *and* per unit volume:

$$\begin{aligned}
W_{\text{fi}} &\equiv \lim_{T,V\to\infty} \frac{|T_{\text{fi}}|^2}{TV} \\
&= |N_A N_B N_C N_D|^2\ |\mathcal{M}|^2\ (2\pi)^4\, \delta\,(p_A + p_B - p_C - p_D)
\end{aligned} \tag{2.37}$$

To use this result in our master formula, we now need to discuss a few remaining ingredients, namely the normalization of the wave functions, the flux factor and the phase space factor.

2.3.1 Normalisation of the Wave Function

Above we defined the eigenstates of the free Hamiltonian to have unit normalization. As we have seen in lecture 2 the eigenstates for free particles (for both the Schrödinger equation and the Klein-Gordon equation) are plane waves

$$\psi(\boldsymbol{x}, t) \ = \ Ne^{-i(Et - \boldsymbol{x}\cdot\boldsymbol{p})}\ . \tag{2.38}$$

In contrast to wave packets the plane waves cannot be normalized over full space x (which further on leads to problems when computing the square of δ-functions as above). The solution is to apply so-called *box normalization*: we choose a finite volume V and normalize all wave functions such that

$$\int_V \psi^*(\boldsymbol{x},t)\psi(\boldsymbol{x},t)\mathrm{d}^3\boldsymbol{x} = 1 \ . \tag{2.39}$$

For the plane waves this gives $N = 1/\sqrt{V}$. Like the time interval T, the volume V is arbitrary and must drop out once we compute an observable cross-section or decay rate.

For the classical wave function the density $\rho = |\psi|^2$ so that the normalization gives one particle per volume V. This normalization is not Lorentz invariant: under a Lorentz transformation the volume element $\mathrm{d}^3\boldsymbol{x}$ shrinks with a factor $\gamma = E/m$.

For the plane wave solutions of the Klein-Gordon equation, we had $\rho = 2|N|^2E$. It is now customary to still use $N = 1/\sqrt{V}$, but define the box normalization such that there are $2E$ particles in the volume V. In other words,

$$\rho = 2E/V. \tag{2.40}$$

We refer to this as $2E$ *particles per unit volume*. If a Lorentz transformation is applied, the change in the energy exactly cancels the contraction of the volume, such that the number of particles in a given volume is Lorentz invariant.

Above we have introduced the Lorentz invariant amplitude without an explicit definition, which is how we have found it in text books that do not derive the formalism with field theory. Thomson takes an alternative approach: the plane wave functions in the classical and relativistic case only differ by the normalization constant $\sqrt{2E}$. If we label classical by ψ and relativistic by ψ', then we have $\psi' = \sqrt{2E}\psi$. For a process $A + B + \cdots \to 1 + 2 + \cdots$, we now *define* the Lorentz-invariant matrix element in terms of the wave functions with relativistic normalization,

$$\mathcal{M}_{\mathrm{fi}} \ = \ \langle \psi'_1\psi'_2 \cdots | V | \psi'_A\psi'_B \cdots \rangle \tag{2.41}$$

where V is the perturbation to the free Hamiltonian (and not the volume!). As the name suggests, with this construction $\mathcal{M}$ is Lorentz invariant. The non-relativistic transition element that appears in Fermi's golden rule is then related to $\mathcal{M}$ by

$$\mathcal{M}_{\mathrm{fi}} \ = \ \sqrt{2E_1 \cdot 2E_2 \cdots 2E_A \cdot 2E_B}\, V_{\mathrm{fi}} \ . \tag{2.42}$$

2.3.2 Density of states and phase space factor

In the final step to Fermi's golden rule we introduced the density of final states $\rho(E)$. In the more general expression for the cross-section, it is the phase space factor that accounts for the density of final states. It depends on the volume V and on the momentum p of each final state particle.

Consider a cross-section measurement in which we measure the 3 components (p_x, p_y, p_z) of the momenta of all final state particles. As stated above it is customary to express the cross-section as a differential cross-section to the final state momenta,

$$\mathrm{d}\sigma \;=\; \ldots \prod_f \mathrm{d}^3\boldsymbol{p}_f \tag{2.43}$$

where the product runs over all final state particles. To compute an actual number for our experiment, we now convolute with experimental resolutions and integrate over eventual particles or momentum components that we do not measure. (For example, we often just measure the number of particles in a solid angle element $\mathrm{d}\Omega$.) For the differential cross-section the question of the number of accessible states should then be rephrased as "how many states fit in the momentum-space volume $V\mathrm{d}^3\boldsymbol{p}$".

Assume that our volume V is rectangular with sides L_x, L_y, L_z. Using periodic boundary conditions to ensure no net particle flow out of the volume we need to require that $L_x p_x = 2\pi\hbar n_x$ with n_x integer. Hence, the total number of states in the range p_x to $p_x + \mathrm{d}p_x$ is $\mathrm{d}n_x = L_x \mathrm{d}p_x/2\pi\hbar$. Since the total number of available states is $n = n_x n_y n_z$, we find that the number of states with momentum between $\boldsymbol{p}$ and $\boldsymbol{p} + \mathrm{d}\boldsymbol{p}$ (i.e. between (p_x, p_y, p_z) and $(p_x + \mathrm{d}p_x, p_y + \mathrm{d}p_y, p_z + \mathrm{d}p_z)$) is:

$$\mathrm{d}n = \frac{V\mathrm{d}^3\boldsymbol{p}}{(2\pi\hbar)^3} \;. \tag{2.44}$$

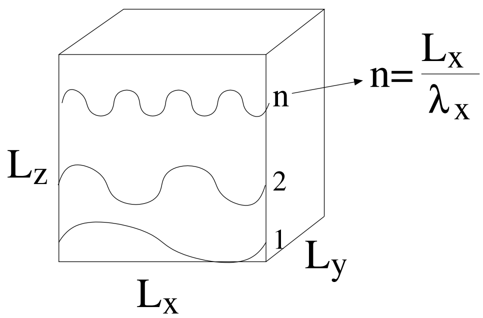


Figure 2.4: Schematic calculation of the number of states in a box of volume V.

As explained above, in the relativistic case the wave functions are normalized such that the volume V contains $2E$ particles. Therefore, the number of states per particle is:

$$\#\ \text{states/particle} = \frac{V}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\boldsymbol{p}}{2E} \tag{2.45}$$

If there are more particles in the final state, the density of states in Fermi's rule must account for each of those. Consequently, the phase space factor for a process with N final state particles becomes

$$\mathrm{d}\Phi = \mathrm{dLIPS} = \prod_{f=1}^{N} \frac{V}{(2\pi\hbar)^3}\,\frac{\mathrm{d}^3\boldsymbol{p}_f}{2E_f} \;. \tag{2.46}$$

In exercise 2.3 you will show that (ignoring V) the phase space factor is indeed Lorentz invariant. We will omit the factors $\hbar$ in what follows.

2.3.3 The Flux Factor

The flux factor corresponds to the number of particles that pass each other per unit area and per unit time. It can be most easily computed in a frame in which one of the particles is at rest. Consider the case that a beam of particles (A) is shot on a target (B), see Fig. 2.5.

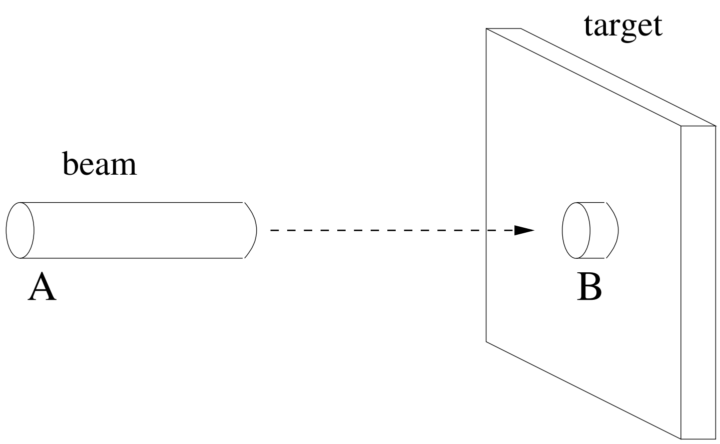


Figure 2.5: A beam incident on a target.

The number of beam particles that pass through unit area per unit time is given by $|\boldsymbol{v}_A|\, n_A$. The number of target particles per unit volume is n_B. For relativistic plane waves the density of particles n is $\rho = \frac{2E}{V}$ such that

$$\text{flux} \;=\; |\boldsymbol{v}_A|\, n_a\, n_b \;=\; \frac{2|\boldsymbol{p}_A|}{V}\,\frac{2m_B}{V} \tag{2.47}$$

(Remember that in relativity $\boldsymbol{v} = \boldsymbol{p}/E$, modulo a factor c. For the KG waves we had indeed that the current density was $\boldsymbol{j} = \rho\,\boldsymbol{p}/E$.) In exercise 2.2 you will show that the kinematic factor $|\boldsymbol{p}_A| m_B$ is Lorentz invariant and that this expression can be rewritten as

$$\boxed{\text{flux} = 4\,\sqrt{(p_{A,\mu}p_B^{\mu})^2 - m_A^2 m_B^2}\;/\;V^2} \tag{2.48}$$

The volume factor is not Lorentz invariant, but it will drop out later, as explained above.

The incident flux as defined here is not actually a certain number of particles per unit surface per unit time per unit volume: we need to account for the fact that it is proportional to the square of an energy. The factors of energy will be accounted for by the other ingredients to the cross-section formula.

2.3.4 Golden rules for cross-section and decay

Putting this all together, we arrive at the formula to calculate a cross-section for the process $A_i + B_i \to C_f + D_f + ...$:

$$\boxed{\begin{aligned} \mathrm{d}\sigma_{\mathrm{fi}} &= \frac{1}{\mathrm{flux}} W_{\mathrm{fi}}\,\mathrm{d}\Phi \\ T_{\mathrm{fi}} &= \frac{1}{i\hbar}\int \mathrm{d}^4x\, \psi_f^*(x)\, V(x)\, \psi_i(x) \\ W_{\mathrm{fi}} &= \lim_{V,T\to\infty} \frac{|T_{\mathrm{fi}}|^2}{TV} \\ \mathrm{d}\Phi &= \prod_{f=1}^{N} \frac{V}{(2\pi)^3}\frac{\mathrm{d}^3\boldsymbol{p}_f}{2E_f} \\ \mathrm{flux} &= 4\sqrt{(p_A\cdot p_B)^2 - m_A^2 m_B^2}\;/\;V^2 \end{aligned}} \tag{2.49}$$

In exercise 2.5 you will show that the cross-section is indeed independent on the volume V.

Inserting the expression for the transition rate per unit time and volume, Eq. (2.37), we find for the differential cross-section of the process $A + B \to C + D$

$$\boxed{\mathrm{d}\sigma = \frac{(2\pi)^4\ \delta^4\left(p_A + p_B - p_C - p_D\right)}{4\sqrt{(p_A\cdot p_B)^2 - m_A^2 m_B^2}} \cdot |\mathcal{M}|^2 \cdot \frac{\mathrm{d}^3\boldsymbol{p}_C}{(2\pi)^3\ 2E_C}\,\frac{\mathrm{d}^3\boldsymbol{p}_D}{(2\pi)^3\ 2E_D}} \tag{2.50}$$

The integrals of the flux factors are only over the spatial part of the outgoing four-momentum vectors. The energy component has been integrated out, using the fact that the outgoing particles are on the mass shell. Therefore, E_f is not an independent variable, but equal to $\sqrt{|\boldsymbol{p}_f|^2 + m_f^2}$. This is important when performing integrals over phase space.

In exercise 2.4 we calculate the integrals and flux factors in the centre-of-momentum system, where $\boldsymbol{p_A} + \boldsymbol{p_B} = \boldsymbol{p_C} + \boldsymbol{p_D} = 0$. The result is

$$\boxed{\left.\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right|_{cm} = \frac{1}{64\pi^2 s}\,\frac{|\boldsymbol{p}_f|}{|\boldsymbol{p}_i|}\;|\mathcal{M}|^2} \tag{2.51}$$

where we defined $\boldsymbol{p}_i \equiv \boldsymbol{p}_A = -\boldsymbol{p}_B$, $\boldsymbol{p}_f \equiv \boldsymbol{p}_C = -\boldsymbol{p}_D$ and $s = (E_A + E_B)^2$.

The computation of a decay rate for the process $A \to C + D$ follows a similar strategy. The result for the partial decay rate is

$$\boxed{\mathrm{d}\Gamma = \frac{(2\pi)^4\ \delta^4\left(p_A - p_C - p_D\right)}{2E_A} \cdot |\mathcal{M}|^2 \cdot \frac{\mathrm{d}^3\boldsymbol{p}_C}{(2\pi)^3\ 2E_C}\,\frac{\mathrm{d}^3\boldsymbol{p}_D}{(2\pi)^3\ 2E_D}} \tag{2.52}$$

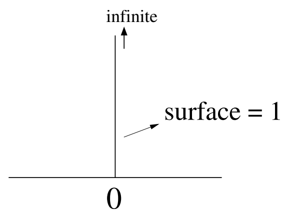
which after integration of one of the momenta gives ($4p_i\sqrt{s} \to 2E_A = 2m_A$)

$$\boxed{\left.\frac{\mathrm{d}\Gamma}{\mathrm{d}\Omega}\right|_{cm} = \frac{1}{32\pi^2 m_A^2}\,|\boldsymbol{p}_f|\ |\mathcal{M}|^2} \tag{2.53}$$

Exercises

Exercise 2.1 (The Dirac δ-Function)

Consider a function defined by the following prescription



$$\delta(x) \;=\; \lim_{\Delta\to 0}\begin{cases} 1/\Delta & \text{for } |x| < \Delta/2 \\ 0 & \text{otherwise}\end{cases}$$

The integral of this function is normalized

$$\int_{-\infty}^{\infty}\delta(x)\,\mathrm{d}x = 1 \tag{2.54}$$

and for any (reasonable) function $f(x)$ we have

$$\int_{-\infty}^{\infty} f(x)\,\delta(x)\,\mathrm{d}x = f(0). \tag{2.55}$$

These last two properties define the *Dirac δ-function.* The prescription above gives an approximation of the δ-function. We shall encounter more of those prescriptions which all have in common that they are the limit of a sequence of functions whose properties converge to those given here.

(a) Starting from the defining properties of the δ-function, prove that

$$\delta(kx) = \frac{1}{|k|}\delta(x)\ . \tag{2.56}$$

(b) Prove that

$$\delta\left(g\left(x\right)\right) = \sum_{i=1}^{n}\frac{1}{|g'\left(x_i\right)|}\ \delta\left(x - x_i\right)\ , \tag{2.57}$$

where the sum i runs over the 0-points of $g(x)$, i.e.:$g(x_i) = 0$.
Hint: make a Taylor expansion of g around the 0-points.

Exercise 2.2 (Lorentz invariance of the flux)

Prove that (ignoring transformations of the volume V) the flux factor derived in the lab frame in Eq. (2.47) is indeed Lorentz-invariant by proving the identity

$$\sqrt{\left(p_A\cdot p_B\right)^2 - m_A^2 m_B^2} \;=\; |\boldsymbol{p}_A| m_B \tag{2.58}$$

Hint: Since the left hand side is Lorentz invariant, you can compute it in any frame. Note that $p_A \cdot p_B$ *is an inner product of the four-vectors, not the three-vectors.*

Exercise 2.3 (Lorentz invariance of the phase space factor)
Show that for any Lorentz invariant function $M(p)$ of the Lorentz vector p, we have the identity

$$\int M(p)\, \mathrm{d}^4p\, \delta(p^2 - m^2)\, \theta(p^0) = \int M(E, \boldsymbol{p}) \frac{\mathrm{d}^3\boldsymbol{p}}{2E}. \tag{2.59}$$

where $\theta(p^0)$ is the Heavyside function and on the right hand-side $E = \sqrt{m^2 + |\boldsymbol{p}|^2}$ is a function of $\boldsymbol{p}$. (On the left hand side, p^0 it is still an independent integration variable.) Argue that this result implies that

$$\frac{\mathrm{d}^3\boldsymbol{p}}{2E} \tag{2.60}$$

is Lorentz invariant.

Exercise 2.4 ($AB \to CD$ cross-section in the c.m.s. See also H&M, Ex. 4.2)
In this exercise we derive a simplified expression for the $A + B \to C + D$ cross-section in the center-of-momentum frame.

(a) Start with the expression:

$$\mathrm{d}\Phi = \int (2\pi)^4\ \delta^4\left(p_A + p_B - p_C - p_D\right)\ \frac{\mathrm{d}^3\boldsymbol{p_C}}{(2\pi)^3\ 2E_C}\, \frac{\mathrm{d}^3\boldsymbol{p_D}}{(2\pi)^3\ 2E_D} \tag{2.61}$$

Do the integral over d^3p_D using the δ function and show that we can write:

$$\mathrm{d}\Phi = \int \frac{1}{(2\pi)^2}\, \frac{p_f^2\, \mathrm{d}p_f\, \mathrm{d}\Omega}{4E_C E_D}\, \delta\left(E_A + E_B - E_C - E_D\right) \tag{2.62}$$

where we have made use of spherical coordinates (*i.e.* $\mathrm{d}^3\boldsymbol{p}_C = |\boldsymbol{p}_C|^2 \mathrm{d}|\boldsymbol{p}_C|\, \mathrm{d}\Omega$) and defined $p_f \equiv |\boldsymbol{p}_C|$.

(b) In the C.M. frame we have $|\boldsymbol{p}_A| = |\boldsymbol{p}_B| = p_i$ and $|\boldsymbol{p}_C| = |\boldsymbol{p}_D| = p_f$. Furthermore, in this frame $\sqrt{s} \equiv |p_A + p_B| = E_A + E_B \equiv W$. Show that the expression becomes (hint: calculate dW/dp_f):

$$\mathrm{d}\Phi = \int \frac{1}{(2\pi)^2}\, \frac{p_f}{4} \left(\frac{1}{E_C + E_D}\right) \mathrm{d}W\, \mathrm{d}\Omega\, \delta\left(W - E_C - E_D\right) \tag{2.63}$$

So that we finally get:

$$\mathrm{d}\Phi = \frac{1}{4\pi^2}\, \frac{p_f}{4\sqrt{s}}\, \mathrm{d}\Omega \tag{2.64}$$

(c) Show that the flux factor in the C.M. frame is:

$$\text{flux} = 4p_i\sqrt{s} \tag{2.65}$$

and hence that the differential cross-section for a $2 \to 2$ process in the center-of-momentum frame is given by

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right|_{cm} = \frac{1}{64\pi^2 s}\,\frac{p_f}{p_i}\,|\mathcal{M}|^2 \tag{2.66}$$

Exercise 2.5 (Box volume is arbitrary (optional!))
Show that the cross-section does not depend on the arbitrary size of the volume V: identify all places where factors V enters in the summary in Eq. (2.49) and show that they cancel.

Exercise 2.6 (Important representations of the δ-function (optional!))

(a) The delta-function can have many forms. One of them is:

$$\delta(x) = \lim_{\alpha\to\infty} \frac{1}{\pi}\frac{\sin^2 \alpha x}{\alpha x^2} \tag{2.67}$$

Make this plausible by sketching the function $\sin^2(\alpha x)/(\pi\alpha x^2)$ for two relevant values of α.

(b) Remember the Fourier transform,

$$\begin{aligned} f(x) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} g(k)\, e^{ikx}\,\mathrm{d}k \\ g(k) &= \int_{-\infty}^{+\infty} f(x)\, e^{-ikx}\,\mathrm{d}x \end{aligned} \tag{2.68}$$

Use this to show that another (important!) representation of the Dirac delta function is given by

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ikx}\,\mathrm{d}k \tag{2.69}$$

Lecture 3

The Electromagnetic Field

3.1 The Maxwell Equations

In classical electrodynamics the movement of a point particle with charge q in an electric field $\boldsymbol{E}$ and magnetic field $\boldsymbol{B}$ follows from the equation of motion

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = q\,(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}). \tag{3.1}$$

The Maxwell equations tell us how electric and magnetic field are induced by static charges and currents. In vacuum they can be written as:

$$\text{Gauss' law} \qquad \boldsymbol{\nabla} \cdot \boldsymbol{E} = \frac{\rho}{\epsilon_0} \tag{3.2}$$

$$\text{No magnetic charges} \qquad \boldsymbol{\nabla} \cdot \boldsymbol{B} = 0 \tag{3.3}$$

$$\text{Faraday's law of induction} \qquad \boldsymbol{\nabla} \times \boldsymbol{E} + \frac{\partial \boldsymbol{B}}{\partial t} = 0 \tag{3.4}$$

$$\text{Modified Ampére's law} \qquad c^2\, \boldsymbol{\nabla} \times \boldsymbol{B} - \frac{\partial \boldsymbol{E}}{\partial t} = \frac{\boldsymbol{j}}{\epsilon_0} \tag{3.5}$$

where ϵ_0 is the vacuum permittivity. From the first and the fourth equation we can 'derive' the continuity equation for electric charges, $\boldsymbol{\nabla} \cdot \boldsymbol{j} = -\frac{\partial \rho}{\partial t}$. It was the continuity equation that lead Maxwell to add the time dependent term to Ampère's law.

The constant c in the Maxwell equations is, of course, the velocity of light. When Maxwell formulated his laws, he did not anticipate this. He did realize that c is the velocity of a propagating electromagnetic wave. The value of c^2 can be computed from measurements of ϵ_0 (e.g. with the force between static charges) and measurements of ϵ_0/c^2 (e.g. from measurements of the force between static currents). From the fact that the result was close to the known speed of light Maxwell concluded that electromagnetic waves and light were closely related. He had, in fact, made one of the great unifications of physics! For a very readable account, including an explanation of how electromagnetic waves travel, see the Feynman lectures, Vol.2, section 18. From now on we choose units of charge such that we can set $\epsilon_0 = 1$ and velocities such that $c = 1$. (We use so-called 'Heaviside-Lorentz rationalised units': see section i.5.)

For what follows it is convenient to write the Maxwell equations in a covariant way (i.e. in a manifestly Lorentz invariant way). As shown below we can formulate them in terms of a single 4 component vector field, which we denote by $A^\mu = (V/c, \boldsymbol{A})$. As suggested by the notation, the components of this field transform as a Lorentz vector.

You may prove for yourself that for any vector field $\boldsymbol{A}$ and scalar field V

$$\text{divergence of rotation is 0:} \qquad \boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \boldsymbol{A}) \;=\; 0 \tag{3.6}$$

$$\text{rotation of gradient is 0:} \qquad \boldsymbol{\nabla} \times (\boldsymbol{\nabla} V) \;=\; 0\,. \tag{3.7}$$

From your electrostatics course you may remember that, because the rotation of $\boldsymbol{E}$ is zero (which is the same as saying that $\boldsymbol{E}$ is a conservative vector field), all physics can be derived by considering a scalar potential field V. The electric field becomes the gradient of the potential, $\boldsymbol{E} = -\boldsymbol{\nabla} V$. The potential V is not unique: we can add an arbitrary constant and the physics will not change. Likewise, because the divergence of the $\boldsymbol{B}$ field is zero, we can always find a vector field $\boldsymbol{A}$ such that $\boldsymbol{B}$ is the rotation of $\boldsymbol{A}$.

So, let's choose a vector field $\boldsymbol{A}$ such that

$$\boldsymbol{B} \;=\; \boldsymbol{\nabla} \times \boldsymbol{A} \tag{3.8}$$

and a scalar field V such that

$$\boldsymbol{E} = -\frac{\partial \boldsymbol{A}}{\partial t} - \boldsymbol{\nabla} V \tag{3.9}$$

Then, by virtue of the vector identities above, the Maxwell equations 3.3 and 3.4 are automatically satisfied.

What remains is to write the other two equations, those that involve the charge density and the charge current density, in components of $\boldsymbol{A}$ and $A^0 = V/c$ as well. You will show in exercise 3.1 that these can be written very efficiently as

$$\partial_\mu \partial^\mu A^\nu - \partial^\nu \partial_\mu A^\mu = j^\nu. \tag{3.10}$$

The current for electric charge j^μ is a conserved current and transforms as a Lorentz vector. (See also Feynman, Vol.2, section 13.6.) The derivative ∂^μ also transforms as a Lorentz vector. Therefore, if the equation above is Lorentz covariant, then A^ν must transform as a Lorentz vector as well. Showing that the electromagnetic field indeed transform this way is outside the scope of these lecture notes, but you may know that the transformation properties of the electric and magnetic fields were an important clue when Einstein formulated his theory of special relativity.

The expressions can be made even more compact by introducing the tensor

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{3.11}$$

such that

$$\partial_\mu F^{\mu\nu} = j^\nu. \tag{3.12}$$

Just as the potential V in electrostatics was not unique, neither is the field A^μ. Imposing additional constrains on A^μ is called *choosing a gauge*. In the next section we shall discuss this in more detail. Written out in terms of the components $\boldsymbol{E}$ and $\boldsymbol{B}$ the (4×4) matrix for the electromagnetic field tensor $F^{\mu\nu}$ is given by

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{3.13}$$

The field tensor is uniquely specified in terms of E and B. In other words, it does not depend on the choice of the gauge.

3.2 Gauge transformations

For a given $\boldsymbol{E}$ and $\boldsymbol{B}$ field, the field A^μ is not unique. Transformations of the field A^μ that leave the electric and magnetic fields invariant are called *gauge transformations*. In exercise 3.2 you will show that for any scalar field $\lambda(t, \boldsymbol{x})$, the transformations

$$\begin{aligned} V' &= V + \frac{\partial \lambda}{\partial t} \\ \boldsymbol{A}' &= \boldsymbol{A} - \boldsymbol{\nabla} \lambda. \end{aligned} \tag{3.14}$$

or in terms of four-vectors

$$A^\mu \to A'^\mu \;=\; A^\mu \;+\; \partial^\mu \lambda \tag{3.15}$$

do not change $\boldsymbol{E}$ and $\boldsymbol{B}$.

If the laws of electrodynamics only involve the electric and magnetic fields, then, when expressed in terms of the field A, the laws must be gauge 'invariant': physical observables should not depend on λ. Sometimes we choose a particular gauge in order to make the expressions in calculations simpler. In other cases, we exploit gauge invariance to impose constraints on a solution, as we will do with the photon below.

A common gauge choice is the so-called *Lorentz gauge*[1]. In exercise 3.3 you will show that it is always possible to choose the gauge field λ such that A^μ satisfies the condition

$$\text{Lorentz condition:} \qquad \partial_\mu A^\mu = 0. \tag{3.16}$$

[1]It is actually called the *Lorenz* condition, named after Ludvig Lorenz (without the letter 't'). It is a Lorentz invariant condition, and is frequently called the "Lorentz condition" because of confusion with Hendrik Lorentz, after whom Lorentz covariance is named. Since almost every reference has this wrong, we will use 'Lorentz' as well.

With this choice A^μ becomes a conserved current. In the Lorentz gauge the Maxwell equations simplify further:

$$\text{Maxwell equations in the Lorentz gauge:} \qquad \partial_\mu \partial^\mu A^\nu \;=\; j^\nu \tag{3.17}$$

However, as you will see in the exercise, A^μ still has some freedom since the Lorentz condition fixes only $\partial_\mu \left(\partial^\mu \lambda\right)$ and not $\partial^\mu \lambda$ itself. In other words a gauge transformation of the form

$$A^\mu \to A'^\mu = A^\mu + \partial^\mu \lambda \qquad \text{with} \qquad \Box\lambda = \partial_\mu \partial^\mu \lambda = 0 \tag{3.18}$$

is still allowed within the Lorentz gauge $\partial_\mu A^\mu = 0$. Consequently, we can in addition impose the *Coulomb condition*:

$$\text{Coulomb condition:} \qquad A^0 = 0 \tag{3.19}$$

(In combination with the Lorentz condition, also $\boldsymbol{\nabla} \cdot \boldsymbol{A} = 0$ with this choice of gauge.) This choice of gauge is not Lorentz invariant. This is allowed since the choice of the gauge is irrelevant for the physics observables, but it is sometimes considered less elegant.

3.3 The photon

Let us now turn to electromagnetic waves and consider Maxwell's equations in vacuum in the Lorentz gauge,

$$\text{vacuum:} \quad j^\mu = 0 \qquad \Longrightarrow \qquad \Box A^\mu = 0. \tag{3.20}$$

Each of the four components of A^μ satisfies the Klein-Gordon equation of a particle with zero mass. (See Eq. (1.20).) This particle is the photon. It represents an electromagnetic wave, a bundle of electric and magnetic field that travels freely through space, no longer connected to the source. Using results below you can show that the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields of such a wave are perpendicular to the wave front and perpendicular to each other. Furthermore, the magnitudes are related by the speed of light, $|E| = c|B|$.

We have seen before that the following complex plane waves are solutions of the Klein-Gorden equation,

$$\phi(x) \sim e^{-ip_\mu x^\mu} \qquad \text{and} \qquad \phi(x) \sim e^{ip_\mu x^\mu} \tag{3.21}$$

For a given momentum vector p any solution in the complex plane is a linear combination of these two plane waves. However, you may have noticed that, in contrast to the Schrödinger equation, the Klein-Gorden equation is actually real. Since the E and B fields are real, we restrict ourselves to solutions with a *real* field A^μ.

We could write down the solution to $\Box A^\mu = 0$ considering only the real axis, but it is customary (and usually more efficient) to form the real solutions by combining the two complex solutions,

$$\begin{aligned} A^\mu(x) \;&=\; a^\mu(p)\, e^{-ipx} \;+\; a^\mu(p)^*\, e^{ipx} \\ &=\; 2\, \Re\big(a^\mu(p) e^{-ipx}\big) \end{aligned} \tag{3.22}$$

(Note that the second term is the complex conjugate of the first.) The four-vector $a^\mu(p)$ depends only on the momentum vector. It has four components but due to the gauge transformation not all of those are physically meaningful. The Lorentz condition gives

$$0 = \partial_\mu A^\mu = -ip_\mu a^\mu e^{-ipx} + ip_\mu a^{\mu *} e^{ipx}, \tag{3.23}$$

which leads to

$$p_\mu a^\mu = 0. \tag{3.24}$$

The Lorentz condition therefore reduces the number of independent complex components to three. However, as explained above, we have not yet exhausted all the gauge freedom: we are still free to make an additional shift $A^\mu \to A^\mu + \partial^\mu \lambda$, provided that λ itself satisfies the Klein-Gordon equation. If we choose it to be

$$\lambda = i\alpha e^{-ipx} - i\alpha^* e^{+ipx} \tag{3.25}$$

with α a complex constant, then its derivative is

$$\partial^\mu \lambda = \alpha p^\mu e^{-ipx} + \alpha^* p^\mu e^{ipx}. \tag{3.26}$$

With a bit of algebra we see that the result of the gauge transformation corresponds to

$$a^{\mu\prime} = a^\mu - \alpha p^\mu \tag{3.27}$$

Note that $a^{\mu\prime}$ still satisfies the Lorentz condition *only because* $p^2 = 0$ *for a massless photon.*

As we have already seen, this additional freedom allows us to apply the Coulomb condition and choose $A^0 = 0$, or equivalently $a^0(p) = 0$. In combination with the Lorentz condition this leads to

$$\boldsymbol{a} \cdot \boldsymbol{p} = 0 \tag{3.28}$$

or $\boldsymbol{p} \cdot \boldsymbol{A} = 0$.

At this point it is customary to write $\boldsymbol{a}(p)$ as a product of two terms

$$\boldsymbol{a}(\boldsymbol{p}) \;\equiv\; N(p)\, \boldsymbol{\epsilon}(\boldsymbol{p}) \tag{3.29}$$

where $\boldsymbol{\epsilon}$ is a vector of unit length and $N(p)$ is real. The normalization $N(p)$ depends only on the magnitude of the momentum and corresponds to the energy density of the wave. The vector $\boldsymbol{\epsilon}$ depends only on the direction of $\boldsymbol{p}$ and is called the *polarization vector*. Choosing the z axis along the direction of the momentum vector and imposing the gauge conditions, the latter can be parameterized as

$$\boldsymbol{\epsilon} = \left(c_1 e^{i\phi_1}, c_2 e^{i\phi_2}, 0\right) . \tag{3.30}$$

where c_i and ϕ_i are all real and $c_1^2 + c_2^2 = 1$. We can remove one phase by moving the origin. (Just look at how a shift of the origin affects the factors $e^{\pm ipx}$.) Therefore, only two parameters of the polarization vector are physically meaningful: these are the two polarization degrees of freedom of the photon.

Any polarization vector can be written as a (complex) linear combination of the two *transverse polarization vectors*

$$\boldsymbol{\epsilon}_1 = (1,0,0) \qquad \boldsymbol{\epsilon}_2 = (0,1,0)\,. \tag{3.31}$$

If the phases of the two components are identical, the light is said to be linearly polarized. If the two components have equal size ($c_1 = c_2 = \sqrt{2}$) but a phase difference of $\pm\pi/2$, the light wave is circularly polarized. The corresponding *circular polarization vectors* are

$$\boldsymbol{\epsilon}_+ = \frac{-\boldsymbol{\epsilon}_1 - i\boldsymbol{\epsilon}_2}{\sqrt{2}} \qquad \boldsymbol{\epsilon}_- = \frac{+\boldsymbol{\epsilon}_1 - i\boldsymbol{\epsilon}_2}{\sqrt{2}} \tag{3.32}$$

You will show in exercise 3.4 that the circular polarization vectors ϵ_+ and ϵ_- transform under a rotation with angle θ around the z-axis (the momentum direction) as

$$\begin{aligned}\boldsymbol{\epsilon}_+ \to \boldsymbol{\epsilon}'_+ &= e^{-i\theta}\boldsymbol{\epsilon}_+ \\ \boldsymbol{\epsilon}_- \to \boldsymbol{\epsilon}'_- &= e^{i\theta}\boldsymbol{\epsilon}_-\end{aligned} \tag{3.33}$$

We now show that this means that these polarization states correspond to the two helicity eigenstates of the photon.

You may remember from your QM course that the z component of the angular momentum operator J_z is the *generator* of rotations around the z-axis. That means that for a wavefunction $\psi(x)$ the effect of an infinitesimal rotation around the z axis is given by

$$\psi(x) \;\to\; U(\varepsilon)\psi(x) \equiv (1 - i\varepsilon J_3)\psi(x). \tag{3.34}$$

An arbitrarily large rotation θ may be built up from infinitesimal rotations by dividing it in small steps

$$\begin{aligned}U(\theta) &= \lim_{n\to\infty} \left(U(\theta/n)\right)^n \\ &= \lim_{n\to\infty} \left(1 - i\left(\frac{\theta}{n}\right) J_z)\right)^n = e^{-i\theta J_z}\end{aligned} \tag{3.35}$$

Consequently, if ψ is an eigen vector of J_z with eigenvalue m, then for any rotation around the z-axis we have

$$\psi(x) \to U(\theta)\psi(x) = e^{-im\theta}\psi(x) \tag{3.36}$$

Comparing this to the effect of rotations on the polarization states above we now identify $\boldsymbol{\epsilon}_+$ with an $m = +1$ state and $\boldsymbol{\epsilon}_-$ with an $m = -1$ state.

Apparently, the polarization states belong to a representation of the rotation group: they are spin states. Since we find ± 1 for the J_z quantum number the photon must be a spin-1 representation: it could not be spin zero, because than you would have only have a state with $m = 0$. And it could not have higher spin state, because there are no degrees of freedom in the photon field that could be identified with larger values of m.

Since the photon is spin-1, one could have expected to find 3 spin states, namely for $m_z = -1, 0, +1$. You may wonder what happened to the $m_z = 0$ component. This component was removed when we applied the Coulomb gauge condition, exploiting $p^2 = 0$, leading to $\boldsymbol{A} \cdot p = 0$. For massive vector fields (or virtual photons!), there is no corresponding gauge freedom and a component parallel to the momentum (a longitudinal polarization) remains. Massive vector fields have one spin degree of freedom more.

Another way to look at this is to say that to define spin properly one needs to boost to the rest frame of the particle. For the massless photon this is not possible. Therefore, we can talk only about helicity (spin projection on the momentum) and not about spin. The equivalent of the $m_z = 0$ state does not exist for the photon.

Finally, we compute the electric and magnetic fields. Substituting the generic expression for A^μ in the definitions of $\boldsymbol{E}$ and $\boldsymbol{B}$ and exploiting the coulomb condition $A^0 \equiv V = 0$, we find

$$\begin{aligned} \boldsymbol{E} &= \; i\, \boldsymbol{a}\, p^0\, e^{-ipx} \;+\; \text{c.c.} \\ \boldsymbol{B} &= \; -i\, (\boldsymbol{p} \times \boldsymbol{a})\, e^{-ipx} \;+\; \text{c.c.} \end{aligned} \tag{3.37}$$

Indeed, for the electromagnetic waves, the E and B fields are perpendicular to each other and to the momentum, while the ratio of their amplitudes is 1 (or rather, c).

3.4 Electrodynamics in quantum mechanics

In classical mechanics an elegant way to introduce electrodynamics is via a method called *minimal substitution.* The method states that the equation of motion of a charged particle under the influence of a vector field A^μ can be obtained by making the substitution

$$p^\mu \to p^\mu - qA^\mu. \tag{3.38}$$

in the equations of motion of the *free* particle. Written out in terms of the potential V and vector potential $\boldsymbol{A}$, the free Hamiltonian is then replaced by

$$H = \frac{1}{2m} \left(\boldsymbol{p} - q\boldsymbol{A} \right)^2 + qV \tag{3.39}$$

It can be shown (see e.g. Jackson §12.1, page 575) that this indeeds leads to the Lorentz force law, Eq. (3.1).

Performing the operator substitution, the Schrödinger equation for the Hamiltonian above becomes

$$\left(\frac{1}{2m} \left(-i\boldsymbol{\nabla} - q\boldsymbol{A} \right)^2 + qV \right) \psi(\boldsymbol{x}, t) \;=\; i\frac{\partial}{\partial t}\psi(\boldsymbol{x}, t) \tag{3.40}$$

Comparing this to the Schrödinger equation for the free particle, we note that we have essentially made the substitution

$$\begin{aligned} -\boldsymbol{\nabla} &\to -\boldsymbol{\nabla} + iq\boldsymbol{A} \\ \partial/\partial t &\to \partial/\partial t + iqV \end{aligned} \tag{3.41}$$

In four-vector notation this can be written as

$$\boxed{\partial^\mu \to D^\mu \equiv \partial^\mu + iqA^\mu} \tag{3.42}$$

The 'derivative' D is called the *covariant derivative*. We may now wonder what the effect of the gauge transformation in Eq. (3.15) is on the wave function ϕ. We have just established that the classical $\boldsymbol{E}$ and $\boldsymbol{B}$ fields do not depend on gauge transformations. However, that does not mean that the wave function is invariant as well. In fact, one can show (see e.g. Aichison and Hey, section 2.4) that the combined transformation, required to make the Schrödinger equation invariant, is given by

$$\begin{aligned} A^\mu \to A'^\mu &= \; A^\mu \; + \; \partial^\mu \lambda \\ \psi \to \psi' &= \; \exp\left(-i\lambda q/\hbar\right) \; \psi \end{aligned} \tag{3.43}$$

The gauge transformation leads to a change of the phase of the wave function. If λ is not constant, then the change in phase is different at different points in space-time. That is why we also call the gauge transformation a *local phase transformation*.

This result is at the heart of the application of *gauge symmetries* in quantum field theory. Because, as we will see in more detail in Lecture 8, one can turn this argument around: Since the phase of the wave function is not an observable, the equations that describe the dynamics (a Schrödinger equation, or a Lagrangian) must be invariant to such arbitrary phase transformations. If we impose this requirement, then we are forced to introduce an A^μ field in the Hamiltonian via the substitution above and with transformation properties defined above. In other words, the requirement of local phase invariance imposes the form of the interaction!

3.5 The Aharanov-Bohm Effect

In the classical equations of motion only the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields appear. Therefore, you may wonder, is the A^μ field 'real', or merely a mathematical construct that simplifies expressions? Or phrased differently, does it contribute anything to our description of moving charges that the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields do not? The answer to that question is beautifully illustrated by what is called the Aharanov-Bohm effect.

In quantum mechanics we do not have forces: it is the amplitude of a wave function that tells us where we are likely to find the particle in space and time. In Feynman's path integral picture, the particle follows all possible trajectories to get from point x_1 to point x_2, accumulating a phase $e^{iS/\hbar}$, where S is the action along the path. Different paths have different phases. It is only around the classical trajectory (obtained by requiring the action to be minimal) that these phases interfere constructively. The size of deviations along the classical trajectory is determined by $\hbar$.

As we have seen above the vector potential appears in the Schrödinger equation and affects the wave function. In the presence of a magnetic field, the phase of the wave function is changed along a trajectory according to

$$\Delta\alpha(\boldsymbol{A}) \;=\; \frac{q}{\hbar}\int_{x_1}^{x_2} \boldsymbol{A}(r,t)\cdot \mathrm{d}\boldsymbol{r} \tag{3.44}$$

where the integral runs along the trajectory. (We do not prove this here. See also Feynman Vol 2, section 15-5.) Although we do not need it here, for completeness we also mention that the change in phase due an electric field is given by the integral of the potential over the time:

$$\Delta\alpha(V) \;=\; -\frac{q}{\hbar}\int_{t_1}^{t_2} V(r,t)\mathrm{d}t \tag{3.45}$$

This last equation can easily be derived from the SE for a constant electric field. When combined these two equations lead to a Lorentz covariant formulation if the integral is performed over space and time.

Let us now consider Feynman's famous two-slit experiment demonstrating the interference between two electron trajectories. In the absence of external fields, the intensity at a detection plate positioned behind the two slits shows an interference pattern. This is most easily understood by considering the two 'classical' trajectories, depicted by ψ_1 and ψ_2 in Fig. 3.1. The relative length of these trajectories differs as a function of the position along the detection screen. The resulting phase difference leads to the interference pattern. For a great description see chapter 1 of the "Feynman Lectures on Physics" volume 3 ("2-slit experiment") and pages 15-8 to 15-14 in volume 2 ("Bohm-Aharanov").

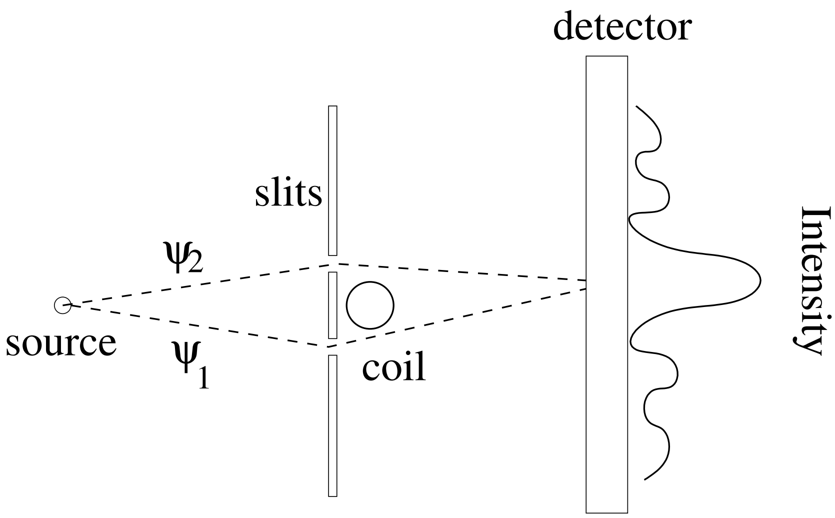


Figure 3.1: The schematical setup of an experiment that investigates the effect of the presence of an A field on the phase factor of the electron wave functions.

Now consider the presence of a magnetic field in the form of vector field $\boldsymbol{A}$. (We choose the electric field zero, so $A^0 = 0$.) Due to the $\boldsymbol{A}$ field, the phases of the two contributions to the wave functions change,

$$\psi = \psi_1\, e^{i\alpha_1(\boldsymbol{r},t)} + \psi_2\, e^{i\alpha_2(\boldsymbol{r},t)} = \left(\psi_1\, e^{i(\alpha_1-\alpha_2)} + \psi_2\right)\, e^{i\alpha_2}\ . \tag{3.46}$$

The extra contribution to the relative phase is given by

$$\begin{aligned} \alpha_1 - \alpha_2 &= \frac{q}{\hbar}\left(\int_{r_1} d\boldsymbol{r}'_1 A_1 - \int_{r_2} d\boldsymbol{r}'_2 A_2\right) = \frac{q}{\hbar}\oint d\boldsymbol{r}' \cdot \boldsymbol{A}(\boldsymbol{r}', t) \\ &= \frac{q}{\hbar}\int_S \boldsymbol{\nabla} \times \boldsymbol{A}(\boldsymbol{r}', t) \cdot d\boldsymbol{S} = \frac{q}{\hbar}\int_S \boldsymbol{B} \cdot d\boldsymbol{S} = \frac{q}{\hbar}\Phi \end{aligned} \tag{3.47}$$

where we have used Stokes' theorem to relate the integral around a closed loop to the magnetic flux Φ through the surface. The magnetic field shifts the interference pattern on the screen. In exercise 3.5 you will show that for a homogenous magnetic field this leads to the same deflection as the classical force law.

Let us now consider the case that a very long and thin solenoid is positioned in the setup of the two-slit experiment. Inside the solenoid the $\boldsymbol{B}$-field is homogeneous and outside it is zero (or sufficiently small). However, the $\boldsymbol{A}$ field is **not** zero outside the coil, as illustrated in Fig. 3.2. The classical trajectories do not pass through the B field, but they do pass through the A field, leading to a shift in the relative phase. Experimentally it has been verified (in a technically difficult experiment) that the interference pattern indeed shifts.

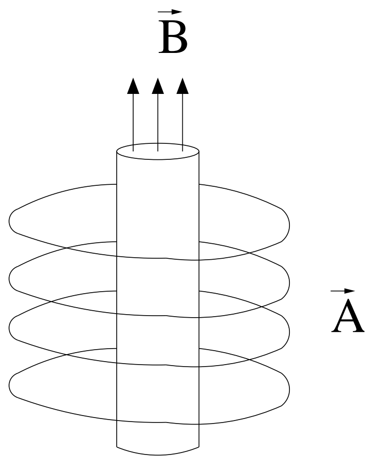


Figure 3.2: Magnetic field and vector potential of a long solenoid.

We introduced the vector potential as a mathematical tool to write Maxwell's equations in a Lorentz covariant form. However, the vector field A^μ is not just an alternative formulation. It is the only correct way to implement the Maxwell equations in quantum mechanics. The gauge freedom may seem an undesirable feature now, but will turn out to be a fundamental concept in our description of interactions.

Exercises

Exercise 3.1 (Maxwell equations)
Using the vector identity

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \boldsymbol{A}) = -\nabla^2 \boldsymbol{A} + \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \boldsymbol{A}) \tag{3.48}$$

(which one can prove using $\varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$) show that (with $c = 1$ and $\epsilon_0 = 1$) Maxwell's equations can be written as:

$$\partial_\mu \partial^\mu A^\nu - \partial^\nu \partial_\mu A^\mu = j^\nu \tag{3.49}$$

Hint: Derive the expressions for ρ and $\boldsymbol{j}$ explicitly.

Exercise 3.2 (Gauge transformation)
Verify that the transformation in Eq. (3.14) does not change the $\boldsymbol{E}$ and $\boldsymbol{B}$ fields.

Exercise 3.3 (Lorentz gauge)
In this exercise you will show that it is always possible to choose a gauge such that the field A^μ satisfies the Lorentz condition, Eq. (3.16).

(a) Suppose that for a given A^μ field one has $\partial_\mu A^\mu = g(x)$, with $g(x) = e^{ikx}$. Find the function $\lambda(x)$ such that $A^\mu + \partial^\mu \lambda$ satisfies the Lorentz condition.

(b) Now consider an arbitrary function $g(x)$. How does your result in (a) tell you what $\lambda(x)$ is?

Exercise 3.4 (Helicity)
Show that the circular polarization vectors ϵ_+ and ϵ_- transform under a rotation of angle θ around the z-axis as

$$\boldsymbol{\epsilon}_\pm \;\rightarrow\; \boldsymbol{\epsilon}'_\pm \;=\; e^{\mp i\theta}\,\boldsymbol{\epsilon}_\pm \tag{3.50}$$

Hint: First consider how the rotations transforms $\boldsymbol{\epsilon}_{1,2}$.

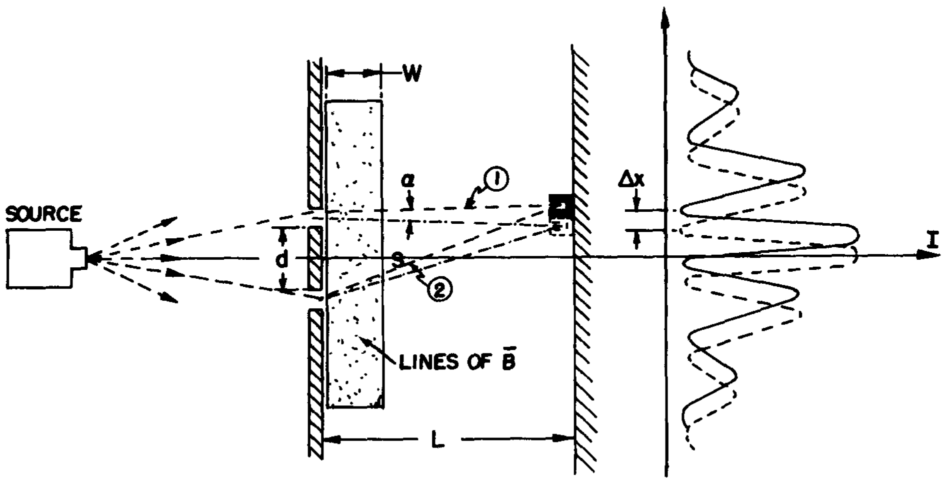


Fig. 15–8. The shift of the interference pattern due to a strip of magnetic field.

Figure 3.3: From "The Feynman Lectures in Physics", volume II.

Exercise 3.5 (Deflection in magnetic field. Feynman, Vol II, sec. 15-5.)
We have stated above that with the minimal substitution recipe the Schrödinger equation leads to the Lorentz force law. We have also stated (not proven) how the change of the phase of a wavefunction due to a vector field $\boldsymbol{A}$ can be obtained by integrating the vector field along the trajectory. Let's take these things for given and see if we can reproduce the deflection of a particle in a magnetic field. Feynman beautifully illustrates that by looking at the famous two-split experiment.

Consider the setup in Fig. 3.3. Particles with charge q, mass m and momentum p travel from a source, via two slits, to a photographic plate. The interference of the two paths leads to a diffraction pattern. The distance between the slits and the plate is L. Directly behind the slits is a thin strip of magnetic field. The thickness of the strip is w and $w \ll L$. The B field is homogenous, coming out of the plane of the figure. We label the coordinate along the photographic plate by x.

(a) For very small deflections, compute the deflection of the particles using the Lorentz force law. Translate this into the displacement Δx at the photographic plate. *Hint: Assume that the plate is thin enough that direction of the force is along the x-axis. The force lasts for a time w/v.*

(b) Consider two classical (shortest distance) trajectories through the two slits (indicated by ① and ②). For small deflections, compute the phase shift between the two trajectories as a function of x, in the absence of a magnetic field. Compute the distance between two maxima in the diffraction pattern. *Hint: The reduced wavelength of the particles is $\lambda/2\pi = \hbar/p$.*

(c) Assuming again small deflections use equation (3.47) to compute the increase in phase shift between the two trajectories as a result of the B field. Translate the phase shift in a shift Δx of the diffraction pattern.

Lecture 4

Electromagnetic Scattering of Spinless Particles

In this lecture we discuss electromagnetic scattering of spin-0 particles. First we compute the scattering of a charged particle on a static point charge. We show that in the non-relativistic limit the result is in agreement with the well known formula for Rutherford scattering. Subsequently, we derive the cross-section for two particles that scatter in each others field. We end the lecture with a prescription for treating anti-particles.

4.1 Electromagnetic current

As we discussed in section 3.4 the laws of electrodynamics can be introduced in the equations of motions of free particles by the method of minimal substitution,

$$p^\mu \to p^\mu - qA^\mu, \tag{4.1}$$

which in terms of operators in coordinate space takes the form

$$\boxed{\partial^\mu \to \partial^\mu + iqA^\mu} \; . \tag{4.2}$$

Now consider a spinless particle with mass m and charge $-e$ scattering in a vector field A^μ, as in figure 2.1. (It is conventional to consider a charge $-e$ as for a hypothetical spin-0 electron.) The wave equation for the free particle is the Klein-Gordon equation,

$$\left(\partial_\mu \partial^\mu + m^2\right) \phi \;=\; 0 \tag{4.3}$$

Substituting $\partial^\mu \to \partial^\mu + iqA^\mu$ with $q = -e$, we obtain

$$\left(\partial_\mu - ieA_\mu\right)\left(\partial^\mu - ieA^\mu\right) \phi + m^2\phi \;=\; 0 \; . \tag{4.4}$$

She operators ∂^μ act on all fields on their right, so both on ϕ and A^μ. This equation can be rewritten as

$$\left(\partial_\mu \partial^\mu + m^2 + V(x)\right) \phi = 0 \tag{4.5}$$

with a perturbation potential $V(x)$ given by

$$V(x) \equiv -ie\,(\partial_\mu A^\mu + A_\mu \partial^\mu) - e^2 A^2 \ . \tag{4.6}$$

The sign of V is chosen such that compared to the kinetic energy it gets the same sign as in the Schrödinger equation, Eq. (3.39). Since e^2 is small ($\alpha = e^2/4\pi = 1/137$) and we only consider the Born level cross-section, we neglect the second order term, $e^2 A^2 \approx 0$.

From Lecture 2 we take the general expression for the transition amplitude in the Born approximation and insert the expression for $V(x)$,

$$\begin{aligned} T_{fi} &\equiv -i \int \mathrm{d}^4x\, \phi_f^*(x)\, V(x)\, \phi_i(x) \\ &= -i \int \mathrm{d}^4x\, \phi_f^*(x)\, (-ie)\, \left(A_\mu \partial^\mu + \partial_\mu A^\mu\right)\, \phi_i(x). \end{aligned} \tag{4.7}$$

The second ∂_μ operator on the right hand side acts on both A^μ and ϕ. However, we can use integration by parts to write

$$\int \mathrm{d}^4x\, \phi_f^*\, \partial_\mu \left(A^\mu\, \phi_i\right) = \left[\phi_f^*\, A^\mu\, \phi_i\right]_{-\infty}^{\infty} - \int \partial_\mu \left(\phi_f^*\right)\, A^\mu\, \phi_i\, \mathrm{d}^4x \tag{4.8}$$

Requiring the A field to be zero at $t = \pm\infty$, the first term on the left vanishes, such that the transition amplitude becomes

$$T_{fi} = -i \int (-ie)\, \left[\phi_f^*(x)\, (\partial_\mu \phi_i(x)) - \left(\partial_\mu \phi_f^*(x)\right) \phi_i(x)\right]\, A^\mu\, \mathrm{d}^4x \ . \tag{4.9}$$

In this expression the derivatives no longer act on the field A^μ. Remember the definition of the charge current density for the Klein-Gordon field of the electron, Eq. (1.31),

$$j_\mu = (-ie)\, \left[\phi^*\, (\partial_\mu \phi) - (\partial_\mu \phi^*)\, \phi\right] \ .$$

In analogy we define the *electromagnetic transition current*' to go from initial state i to final state f as

$$j_\mu^{fi} \equiv (-ie)\, \left[\phi_f^*\, (\partial_\mu \phi_i) - \left(\partial_\mu \phi_f^*\right) \phi_i\right] \ . \tag{4.10}$$

You may verify that if ϕ_f and ϕ_i are both solutions to the Klein-Gordon equation with the same mass m, this current satisfies the continuity equation $\partial^\mu j_\mu^{fi} = 0$ as well.

The transition amplitude can now be written as

$$\boxed{T_{fi} = -i \int j_\mu^{fi}\, A^\mu\, \mathrm{d}^4x} \tag{4.11}$$

This is the expression for the transition amplitude for going from free particle solution i to free particle solution f in the presence of a perturbation caused by an electromagnetic field. Restricting ourselves to plane wave solutions of the unperturbed Klein-Gordon equation,

$$\phi_i = N_i\, e^{-ip_i x} \qquad \text{and} \qquad \phi_f^* = N_f^* e^{ip_f x} \ , \tag{4.12}$$

we find for the transition current of spinless particles

$$j_\mu^{fi} = (-e)\, N_i N_f^* \,\left(p_\mu^i + p_\mu^f\right)\, e^{i\left(p_f - p_i\right)x} \,. \tag{4.13}$$

Inserting this in the transition amplitudes gives

$$T_{fi} \;=\; -i \int (-e)\, N_i N_f^* \,\left(p_i^\mu + p_f^\mu\right)\, A_\mu \, e^{i\left(p_f - p_i\right)x}\, \mathrm{d}^4 x \tag{4.14}$$

4.2 Coulomb scattering

Consider the case that the external field is a static field of a point charge Ze located at the origin,

$$A_\mu = (V, \boldsymbol{A}) = (V, \boldsymbol{0}) \qquad \text{with} \qquad V(x) = \frac{Ze}{4\pi|\boldsymbol{x}|}. \tag{4.15}$$

With a vector field of this form, we have $p_k^\mu \, A_\mu = E_k \, V(\boldsymbol{x})$. Consequently, we find for the transition amplitude

$$T_{fi} = i \int N_i N_f^* \,\left(E_i + E_f\right)\, e^{i\left(p_f - p_i\right)x} \,\frac{Ze^2}{4\pi|\boldsymbol{x}|}\, \mathrm{d}^4 x \tag{4.16}$$

Since $V(x)$ is time independent, we split the integral over space and time. As we have seen before, the integral over time turns into a δ function, expressing energy conservation,

$$\int e^{i\left(E_f - E_i\right)t} dt = 2\pi\, \delta\left(E_f - E_i\right) \tag{4.17}$$

For the integral over $\boldsymbol{x}$ we use an important Fourier transform

$$\frac{1}{|\boldsymbol{q}|^2} = \int \mathrm{d}^3\boldsymbol{x}\, e^{i\boldsymbol{q}\boldsymbol{x}}\, \frac{1}{4\pi|\boldsymbol{x}|} \tag{4.18}$$

Using this with $\boldsymbol{q} \equiv (\boldsymbol{p}_f - \boldsymbol{p}_i)$ we obtain

$$T_{fi} \;=\; i\, N_i N_f^* \,\left(E_i + E_f\right)\, 2\pi\, \delta\left(E_f - E_i\right)\, \frac{Ze^2}{\left|\boldsymbol{p}_f - \boldsymbol{p}_i\right|^2} \,. \tag{4.19}$$

As we have seen before for a time-independent potential, we consider a time-averaged transition rate,

$$W_{fi} \;=\; \lim_{T\to\infty} \frac{\left|T_{fi}\right|^2}{T} \tag{4.20}$$

where the time-averaging effectively takes care of one of the δ functions when taking the square of the amplitude. The result is

$$W_{fi} = \left|N_i N_f\right|^2\, 2\pi\, \delta\left(E_f - E_i\right) \left(\frac{Ze^2\left(E_i + E_f\right)}{\left|\boldsymbol{p}_f - \boldsymbol{p}_i\right|^2}\right)^2 \tag{4.21}$$

Working with normalization of the plane waves over a box with volume V, we have $|N_i N_f|^2 = 1/V^2$. The flux factor for a single particle is given by

$$\text{flux} \;=\; |\boldsymbol{v_i}|\,\frac{2E_i}{V} \;=\; \frac{2|\boldsymbol{p}_i|}{V}, \tag{4.22}$$

while the phase space factor is

$$\text{dLips} \;=\; \frac{V}{(2\pi)^3}\frac{\mathrm{d}^3\boldsymbol{p}_f}{2E_f}. \tag{4.23}$$

Inserting these expressions in our master formula for the cross-section, Eq. (2.8), we find

$$\mathrm{d}\sigma \;=\; \frac{2\pi\,\delta\left(E_f - E_i\right)}{2|\boldsymbol{p}_i|}\left(\frac{Ze^2\left(E_i + E_f\right)}{|\boldsymbol{p}_f - \boldsymbol{p}_i|^2}\right)^2 \frac{\mathrm{d}^3\boldsymbol{p}_f}{(2\pi)^3\,2E_f} \tag{4.24}$$

where we have canceled all factors V.

We can still simplify this by integrating over the outgoing momentum. Choose the z-axis along $\boldsymbol{p}_i$ and switch to polar coordinates for $\boldsymbol{p}_j$ such that

$$\mathrm{d}^3\boldsymbol{p}_f \;=\; p_f^2\,\mathrm{d}p_f\,\mathrm{d}\!\cos\theta\,\mathrm{d}\phi \;=\; p_f^2\,\mathrm{d}p_f\,\mathrm{d}\Omega \tag{4.25}$$

where $p_{f,i}$ now refers to the size of the three-momentum. Since $E_f^2 = m^2 + \boldsymbol{p}_f^2$, we have $p_f\,\mathrm{d}p_f = E_f\,\mathrm{d}E_f$, and therefore,

$$\delta\left(E_f - E_i\right)\;\mathrm{d}p_f = \frac{E_f}{p_f}\delta\left(E_f - E_i\right)\;\mathrm{d}E_f \tag{4.26}$$

Energy conservation will imply that $p_f = p_i$, such that

$$|\boldsymbol{p}_f - \boldsymbol{p}_i|^2 \;=\; 2\,p_i^2\,(1-\cos\theta) \;=\; 4\,p_i^2\,\sin^2\!\theta/2\;. \tag{4.27}$$

The differential cross-section then becomes

$$\boxed{\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{Z^2 E_i^2 e^4}{64\pi^2|p_i|^4\sin^4\theta/2} = \frac{Z^2|E_i|^2\alpha^2}{4\,|p_i|^4\,\sin^4\theta/2}} \tag{4.28}$$

where we defined $\alpha \equiv e^2/4\pi$.

In the non-relativistic limit we have $E \to m$ and $p^2 = 2mE_{\text{kin}}$, giving

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{Z^2\alpha^2}{16E_{\text{kin}}^2\sin^4\theta/2}\;, \tag{4.29}$$

which is the well-known Rutherford scattering formula.

Above we have not explicitly shown the solution of the wave function itself. Without deriving it here, we just state that for a spherically symmetric potential, $V(\boldsymbol{r}) = V(r)$, and an incident wave with momentum along the z-axis, the wave function for large r takes the form

$$\phi(\boldsymbol{r}) \xrightarrow{r\to\infty} e^{ikz} + f(\theta,\phi)\frac{e^{ikr}}{r} \tag{4.30}$$

where the first term is the incident 'unscattered' wave and the second term the 'scattered' wave. Expressed in terms of the $f(\theta,\phi)$, the differential cross-section can be written as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f(\theta,\phi)|^2 \tag{4.31}$$

The interference between the scattered and the unscattered wave leads to a 'shadow' behind the scattering potential. The flux that is missing in the shadow is exactly the total scattered flux. This is expressed in the *optical theorem*, which states that the total cross-section is proportional to value of f in the forward direction,

$$\mathrm{Im} f(0) \;=\; \frac{k}{4\pi}\sigma\,. \tag{4.32}$$

See also appendix H of Aichison and Hey, and references therein.

4.3 Spinless $\pi - K$ Scattering

We now proceed with the electromagnetic scattering of two particles, $A + B \to C + D$. As an example we consider the scattering of a π^- particle and a K^- particle. We ignore the fact that pions and kaons also are subject to the strong interaction, which is fine as long as the recoil momentum is small compared to the binding energy. We could equally well consider a process like $e^-\mu^-$ scattering, provided that we ignore the lepton spin. For the computation presented here, the essential restrictions are that the incident particles carry no spin and that they are of different type.

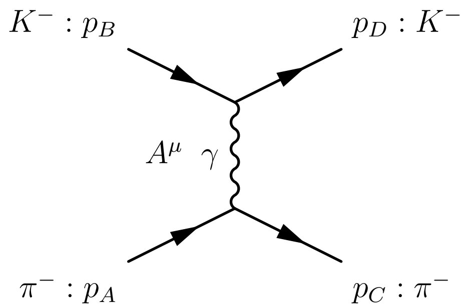


Figure 4.1: Schematic presentation of the leading order electromagnetic scattering of a pion and a kaon.

We have seen above how a particle scatters in an external field. In this case the field is not external as the particles scatter in each others field. How do we deal with this?

First consider a pion scattering in the vector field A^μ generated by the current of the kaon. The transition current of the kaon is given by (see Eq. (4.13))

$$j^\mu_{BD} = -eN_B N_D^* \; (p_B^\mu + p_D^\mu) \; e^{i(p_D - p_B)x} \tag{4.33}$$

We now assume that the field generated by the kaon can be computed by inserting this current in the Maxwell equations for the vector potential, *i.e.*

$$\partial_\nu \partial^\nu A^\mu = j^\mu_{BD} \tag{4.34}$$

where we have adopted the Lorentz gauge. (A proof that this indeed works requires the full theory.) Since $\partial_\nu \partial^\nu e^{iqx} = -q^2 \, e^{iqx}$, we can easily verify that the solution is given by

$$A^\mu = -\frac{1}{q^2} \, j^\mu_{BD} \quad , \tag{4.35}$$

where we defined $q = p_D - p_B$. The latter corresponds to the *four-momentum* transfered by the photon from the kaon to the pion. The transition probability becomes

$$T_{fi} = -i \int j^\mu_{AC} \, A_\mu \, \mathrm{d}^4 x = -i \int j^\mu_{AC} \, \frac{-1}{q^2} j^{BD}_\mu \, \mathrm{d}^4 x = -i \int j^\mu_{AC} \, \frac{-g_{\mu\nu}}{q^2} \, j^\nu_{BD} \, \mathrm{d}^4 x \, . \tag{4.36}$$

Four-momentum conservation (which appears as a result of the integral when we substitute plane waves in the currents) makes that the momentum transfer is also equal to $q = -(p_C - p_A)$. Therefore, T_{fi} is indeed symmetric in the two currents. It does not matter whether we scatter the pion in the field of the kaon or the kaon in the field of the pion.

The expression has a pole for $q^2 = 0$, the mass of a 'real' photon: zero momentum transfer (non-scattered waves) has 'infinite' probability. The only contribution to scattering under non-zero angles comes from photons that are "off the mass-shell". We call these *virtual* photons.

Inserting the plane wave solutions

$$T_{fi} = -ie^2 \int (N_A N_C^*) \, (p_A^\mu + p_C^\mu) \, e^{i(p_C - p_A)x} \cdot \frac{-1}{q^2} \cdot (N_B N_D^*) \, (p_B^\mu + p_D^\mu) \, e^{i(p_D - p_B)x} \mathrm{d}^4 x \tag{4.37}$$

and performing the integral over x we obtain

$$T_{fi} = -ie^2 \, (N_A N_C^*) \; (p_A^\mu + p_C^\mu) \; \frac{-1}{q^2} \, (N_B N_D^*) \; \left(p^B_\mu + p^D_\mu\right) \; (2\pi)^4 \, \delta^4 \left(p_A + p_B - p_C - p_D\right) \tag{4.38}$$

where the δ-function that takes care of four-momentum conservation appears. Usually this is written in terms of the invariant amplitude $\mathcal{M}$ (sometimes called 'matrix element') as

$$T_{fi} \; = \; -i \; N_A N_B N_C^* N_D^* \; (2\pi)^4 \, \delta^4 \left(p_A + p_B - p_C - p_D\right) \cdot \mathcal{M} \tag{4.39}$$

with the invariant amplitude given by

$$-i\mathcal{M} \; = \; \underbrace{ie \, (p_A + p_C)^\mu}_{\text{vertex factor}} \cdot \; \underbrace{\frac{-ig_{\mu\nu}}{q^2}}_{\text{propagator}} \; \cdot \underbrace{ie \, (p_B + p_D)^\nu}_{\text{vertex factor}} \; . \tag{4.40}$$

The signs and factors i are assigned such that the expressions for vertex factors and propagator are also appropriate for higher orders. These are, in fact, our first set of Feynman rules!

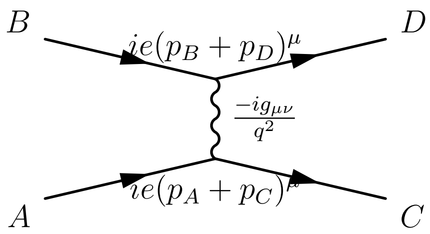


Figure 4.2: Feynman rules for the t-Channel contribution to electromagnetic scattering of spinless particles.

Feynman rules allow us to specify the amplitude corresponding to a particular Feynman diagram without going through the explicit computation of the amplitude in quantum field theory. In figure 4.2 we illustrate the rules for the diagram that we are considering. The invariant amplitude contains:

a vertex factor: for each vertex we introduce the factor iep^μ, where

- ▷ e is the intrinsic coupling strength of the particle to the e.m. field
- ▷ p^μ is the sum of the 4-momenta before and after the scattering (remember the particle/anti-particle convention).

a propagator: for each internal line (photon) we introduce a factor $\frac{-ig_{\mu\nu}}{q^2}$, where q is the 4-momentum of the virtual photon.

The ingoing and outgoing four-momenta, and the four-momenta of internal particles are free, but we add a δ function at each vertex to ensure energy-momentum conservation. In the end, all internal momentum vectors are integrated over, and what remains is a single δ function over ingoing and outgoing momenta. By convention, the latter does not belong to $\mathcal{M}$. The full set of rules also specify how this works for higher order diagrams.

The expression of the amplitude in terms of propagators and vertex factors is part of what we refer to as the Feynman calculus. To prove that this works requires field theory. However, the attractiveness of this approach is that once you have established the recipe, you can derive the Feynman rules (the expressions for the propagators and vertices) directly from the Lagrangian that specifies the dynamics of your favourite theory: if you insert a new type of particle or interaction in your Lagrangian, you do not really need field theory anymore to compute cross-sections. We discuss the role of the Lagrangian in more detail in Lecture 8. In appendix B we sketch how the propagators and vertex factors are obtained from the Lagrangian.

We can now insert the invariant amplitude into the expression for the $A + B \to C + D$

cross-section that we derived in lecture 2,

$$\mathrm{d}\sigma = \frac{(2\pi)^4\,\delta^4\,(p_A + p_B - p_C - p_D)}{4\sqrt{(p_A\cdot p_B)^2 - m_A^2 m_B^2}}\;|\mathcal{M}|^2\;\frac{\mathrm{d}^3 p_C}{(2\pi)^3\,2E_C}\;\frac{\mathrm{d}^3 p_D}{(2\pi)^3\,2E_D}\;. \tag{4.41}$$

In the centre-of-momentum frame ($\boldsymbol{p}_A = -\boldsymbol{p}_B$) this expression became (see Eq. 2.51)

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{64\pi^2}\,\frac{1}{s}\,\left|\frac{\boldsymbol{p}_f}{\boldsymbol{p}_i}\right|\;|\mathcal{M}|^2\;. \tag{4.42}$$

where $s = (p_A + p_B)^2$. As we consider elastic scattering, $|\boldsymbol{p}_f|/|\boldsymbol{p}_i| = 1$.

Finally, consider the limit of massless particles. Define $\boldsymbol{p} \equiv \boldsymbol{p}_A$ and $\boldsymbol{p\prime} \equiv \boldsymbol{p}_C$. In the centre-of-momentum frame the four-vectors are given by

$$\begin{aligned}
p_A^\mu &= (|\boldsymbol{p}|, \boldsymbol{p})\\
p_B^\mu &= (|\boldsymbol{p}|, -\boldsymbol{p})\\
p_C^\mu &= (|\boldsymbol{p\prime}|, \boldsymbol{p'})\\
p_D^\mu &= (|\boldsymbol{p\prime}|, -\boldsymbol{p'})
\end{aligned}$$

Define $p \equiv |\boldsymbol{p}|$ which, by four-vector conservation is also equal to $|\boldsymbol{p\prime}|$. Define θ as the angle between $\boldsymbol{p}_A$ and $\boldsymbol{p}_C$, which means that $\cos\theta = \boldsymbol{p_A}\cdot\boldsymbol{p}_C/|p_A||p_C| = \boldsymbol{p_A}\cdot\boldsymbol{p}_C/p^2$. We then have

$$\begin{aligned}
(p_A + p_C)^\mu\;g_{\mu\nu}\;(p_B + p_D)^\nu &= (p_A)_\mu(p_B)^\mu + (p_A)_\mu(p_D)^\mu + (p_C)_\mu(p_B)^\mu + (p_C)_\mu(p_D)^\mu\\
&= 2p^2 + p^2(1+\cos\theta) + p^2(1+\cos\theta) + 2p^2\\
&= p^2(6 + 2\cos\theta)
\end{aligned}$$

Likewise, we get for q^2,

$$\begin{aligned}
q^2 &= (p_A - p_C)^2\\
&= p_A^2 + p_C^2 - 2(p_A)_\mu(p_C)^\mu\\
&= -2p^2(1-\cos\theta)
\end{aligned}$$

Consequently, we obtain for the invariant amplitude defined above

$$\mathcal{M} \;=\; e^2\,\frac{p^2\,(6+2\cos\theta)}{2p^2\,(1-\cos\theta)} = e^2\left(\frac{3+\cos\theta}{1-\cos\theta}\right)\;. \tag{4.43}$$

Inserting this in Eq. (4.42) gives (with $\alpha = e^2/4\pi$),

$$\boxed{\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{64\pi^2}\,\frac{1}{s}\,\left(e^2\right)^2\left(\frac{3+\cos\theta}{1-\cos\theta}\right)^2 = \frac{\alpha^2}{4s}\,\left(\frac{3+\cos\theta}{1-\cos\theta}\right)^2} \tag{4.44}$$

This is the leading order QED cross-section for the scattering of massless spin-0 particles in the centre-of-momentum frame. In exercise 4.1 you will derive the formula for particles with non-zero mass.

4.4 Form factors

In the previous sections we studied 'point charges', objects with their charge located in an infinitely small region. If the charge distribution has a finite size, the differential cross-section is different from that of a point source. Consequently, the measured differential cross-section can tell us important information over the substructure of particles. For example, most information about the structure of the proton has been obtained in electron-proton scattering experiments, most notably at the Hera collider in Hamburg.

Consider a static source with a charge distribution $Ze\rho(\boldsymbol{x})$, normalized so that

$$\int \rho(\boldsymbol{x})\, \mathrm{d}^3x \;=\; 1 \tag{4.45}$$

By following the same procedure as above for the static source, one can show that the differential cross-section can be written as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \;=\; \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}_{\mathrm{point}} |F(\boldsymbol{p}_i - \boldsymbol{p}_f)|^2 \tag{4.46}$$

where $F(\boldsymbol{q})$ is called the *form factor*. It is given by the Fourier transform of the charge distribution

$$F(\boldsymbol{q}) \;=\; \int \rho(\boldsymbol{x}) e^{i\boldsymbol{q}\cdot\boldsymbol{x}} \mathrm{d}^3x \tag{4.47}$$

In real electron-proton scattering we also need to account for the spin and the magnetic moment of the proton. The form factor will then become more complicated. You will learn more about this in the Particle Physics II course.

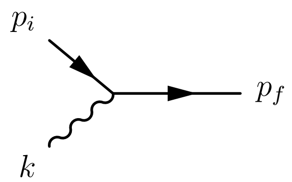
4.5 Particles and Anti-Particles

We have seen that the negative energy state of a particle can be interpreted as the positive energy state of its anti-particle. How does this effect energy conservation that we encounter in the δ-functions? We have seen that the invariant amplitude takes the form

$$\mathcal{M} \propto \int \phi_f^*(x)\, V(x)\, \phi_i(x)\, dx \tag{4.48}$$

Let us examine four cases:

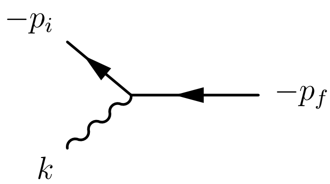
1. Scattering of an electron and a photon:



$$
\begin{aligned}
\mathcal{M} &\propto \int \left(e^{-ip_f x}\right)^* \, e^{-ikx} \, e^{-ip_i x} \, dx \\
&= \int e^{-i\left(p_i + k - p_f\right)x} \, dx \\
&= (2\pi)^4 \; \delta\left(E_i + \omega - E_f\right) \; \delta^3\left(\boldsymbol{p}_i + \boldsymbol{k} - \boldsymbol{p}_f\right)
\end{aligned}
$$

$\Rightarrow$ Energy and momentum conservation are enforced by the δ-function.

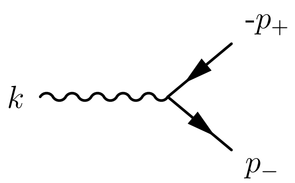
2. Scattering of a positron and a photon:



Replace the anti-particles always by particles by reversing $(E, \boldsymbol{p} \to -E, -\boldsymbol{p})$ such that now: incoming state $= -p_f$, outgoing state $= -p_i$:

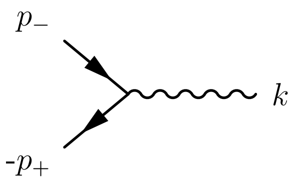
$$
\begin{aligned}
\mathcal{M} &\propto \int \left(e^{-i(-p_i)x}\right)^* \, e^{-ikx} \, e^{-i(-p_f)x} \, dx \\
&= \int e^{-i\left(p_i - p_f + k\right)x} \, dx \\
&= (2\pi)^4 \; \delta\left(E_i + \omega - E_f\right) \; \delta^3\left(\boldsymbol{p_i} + \boldsymbol{k} - \boldsymbol{p_f}\right)
\end{aligned}
$$

3. Electron positron pair production:



$$
\begin{aligned}
\mathcal{M} &\propto \int \left(e^{-ip_- x}\right)^* \, e^{-i(-p_+ + k)x} \, dx \\
&= \int e^{-i(k - p_+ - p_-)x} \, dx \\
&= (2\pi)^4 \; \delta\left(k - p_- - p_+\right)
\end{aligned}
$$

4. Electron positron annihilation:



$$
\begin{aligned}
\mathcal{M} &\propto \int \left(e^{-i(k - p_+)x}\right)^* \, e^{-i(p_-)x} \, dx \\
&= \int e^{-i(p_- + p_+ - k)x} \\
&= (2\pi)^4 \; \delta\left(p_- + p_+ - k\right)
\end{aligned}
$$

Exercises

Exercise 4.1 (Scattering off a static source as a limit)
Above we derived the expression for $A+B \to C+D$ scattering in the special case that A, B, C and D are massless. We will now show that in the limit in which $m_B \gg m_A$ and $m_B \gg p$ the cross-section for the process $A+B \to A+B$ reduces to the formula that we derived for scattering of a static source.

Call m_A the mass of A and m_B the mass of B. We consider scattering in the CMS. Choose a coordinate system such that the initial momenta are along the z-axis, with A going in the positive z direction and B in the negative z direction. Label the outgoing particles with a prime. Call θ the polar angle of $\boldsymbol{p}'_A$. Momentum conservation makes that all momenta have the same size, which we label by p.

(a) Express $(p_A + p'_A)_\mu (p_B + p'_B)^\mu$ in m_A, m_B, p and $\cos\theta$. (You may of course also use $E_A = \sqrt{m_A^2 + p^2}$ and $E_B = \sqrt{m_B^2 + p^2}$.)
Hint: one method is to first just write down all fourvectors in these symbols. Take the x,y-coordinates together, because we do not care about the azimuthal angle (ϕ).Now you can either first add and them take the inner product, or vice versa.

(b) Do the same for $q^2 = (p_A - p'_A)^2$ and for $s = (p_A + p_B)^2$.

(c) Write down the differential cross-section $\mathrm{d}\sigma/\mathrm{d}\Omega$ using Eq. (4.42). Note that this result is more general than our 'massless particle' result in Eq. (4.44).

(d) Take the limit $m_B \gg p$ and $m_B \gg m_A$. Compare to the result for scattering of a static source, Eq. (4.28).

Exercise 4.2 (Equal particles)
When computing the scattering of two particles A and B above, we explicitly restricted the computation to the case where the particles were different.

(a) What changes if B is an anti-A?

(b) What changes if B is an A?

Hint: Look at figure 4.2. Can you imagine additional 'leading order' diagrams? If so, draw them!

Exercise 4.3 (Decay rate of $\pi^0 \to \gamma\gamma$ (see also Griffiths, ex.6.6))

(a) Write down the expression for the total decay width Γ for the decay: $A \to C + D$

(b) Assume that particle A is a π^0 particle with a mass of 140 MeV and that particles C and D are photons. Draw the Feynman diagram for this decay

(i) assuming the pion is a $u\bar{u}$ state.

(ii) assuming the pion is a $d\bar{d}$ state.

(c) We do not know the matrix element $\mathcal{M}$. However, we know that it is proportional to $e^2 \sim \sqrt{\alpha}^2$. Why?

(d) We also know something about the dimension of $\mathcal{M}$: for a two-body decay the dimension is mc (or p or E/c). The constant with dimension $[mc]$ in the amplitude is called the *decay constant* and denoted by f_π, such that

$$\mathcal{M} \propto f_\pi \alpha.$$

where α is the dimensionless coupling constant, and additional factors are dimensionless as well. If you do not know anything else about the π^0 decay constant but its dimension, what value would you use?

(e) Assuming that the π^0 is a $u\bar{u} + d\bar{d}$ state

(i) give the expression for the decay width (by adding up the amplitudes);

(ii) calculate the decay width expressed in GeV;

(iii) convert the rate into a mean lifetime in seconds.

(iv) How does the value compare to the Particle Data Group (PDG) value?

Remark: Do not be disappointed if your prediction is completely wrong! It turns out that the π^0 lifetime is quite hard to compute.

Lecture 5

The Dirac Equation

In Lecture 1 we have seen how the Klein-Gordon equation leads to solutions with negative energy and negative 'probability density'. This is a consequence of the fact that the wave equation is quadratic in $\partial/\partial t$. In 1928 in an attempt to avoid this problem Dirac developed a relativistic wave equation that is linear in $\partial/\partial t$. Lorentz invariance requires that such a wave equation is also linear in $\partial/\partial \boldsymbol{x}$. What Dirac found was an equation that describes particles with spin-$\frac{1}{2}$, just what was needed for electrons. We now think that all fundamental fermions are described by this wave equation. Dirac also predicted the existence of anti-particles, an idea that was not taken seriously until 1932, when Anderson discovered the positron.

5.1 Spin, spinors and the gyromagnetic ratio

Before we proceed with the derivation of the Dirac equation, we briefly discuss spin and the Pauli-Schrödinger equation, since this allows us to introduce some important concepts.

A lump of charge rotating around an axis through its centre-of-gravity is a magnetic dipole. It has a magnetic moment

$$\boldsymbol{\mu} = \gamma \boldsymbol{S} \tag{5.1}$$

where S is the classical spin, the integral of $\boldsymbol{r} \times \boldsymbol{p}$ over the mass density distribution. The factor γ is called the gyromagnetic ratio. If the charge distribution follows the mass distribution, this ratio is given by

$$\gamma_{\text{classic}} = \frac{q}{2m} \,. \tag{5.2}$$

When placed in a magnetic field $\boldsymbol{B}$, the particle experiences a torque, $\boldsymbol{\mu} \times \boldsymbol{B}$. The potential energy associated with the torque is

$$H \;=\; -\,\boldsymbol{\mu} \cdot \boldsymbol{B} \;=\; -\,\gamma\, \boldsymbol{S} \cdot \boldsymbol{B} \tag{5.3}$$

So far, this is just classical electrodynamics. The classical spin S is nothing but the total angular momentum of all bits and pieces that the particle is made up from. However, as you remember from your quantum mechanics course, elementary particles also carry *intrinsic* spin. Though we sometimes imagine it as a result of a charged particle spinning around an axis, this interpretation actually falls short. In particular, the prediction of the gyromagnetic ratio that would come out of this picture is wrong.

On the other hand, elementary particles do feel a torque in a magnetic field, as demonstrated in the Stern-Gerlach experiment in 1922. So, in 1927 Pauli tried to address the question of how to describe their magnetic moment in quantum mechanics.

Pauli considered a spin-$\frac{1}{2}$ system. As you know, such a system has two values for the eigenvalue of spin, namely $\pm\frac{1}{2}\hbar$. An arbitrary spin wave function is a superposition of the two eigenstates. Pauli represented it as a two-component vector, called a *spinor*,

$$\chi \;=\; \begin{pmatrix} a \\ b \end{pmatrix} \;=\; a\chi^{(1)} \;+\; b\chi^{(2)} \tag{5.4}$$

where the basis vectors

$$\chi^{(1)} \;\equiv\; \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \chi^{(2)} \;\equiv\; \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{5.5}$$

represent the spin-up and spin-down state respectively. The hermitian operator that measures spin, the spin operator $\boldsymbol{S}$, satisfies the same algebra as for orbital angular momentum in quantum mechanics, namely $[S_i, S_j] = i\hbar\epsilon_{ijk}S_k$. In the basis above, $\boldsymbol{S}$ is represented by 2x2 matrices. Choosing the z-axis as the quantization axis, $\boldsymbol{S}$ is given by

$$\boldsymbol{S} \;=\; \frac{\hbar}{2}\boldsymbol{\sigma} \tag{5.6}$$

where the Pauli spin matrices are

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{5.7}$$

The matrices σ_i all have zero trace, are hermitian, and satisfy

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k \tag{5.8}$$

which implies as well that they anti-commute ($\{\sigma_i, \sigma_j\} = 2\delta_{ij}$).

Pauli then proposed to replace the momentum operator in the Schrödinger equation by a new operator in spinor space,

$$\boldsymbol{p} \longrightarrow \boldsymbol{\sigma} \cdot \boldsymbol{p} \tag{5.9}$$

Take a careful look at what is written on the right-hand-side: it is an inner product of a vector of 2x2 matrices with the momentum operator. The result is again a 2x2 matrix, which is more apparent when written out,

$$\boldsymbol{\sigma} \cdot \boldsymbol{p} = \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix} . \tag{5.10}$$

You will show in an exercise that $(\boldsymbol{\sigma}\cdot\boldsymbol{p})^2 = |\boldsymbol{p}|^2\mathbb{1}_2$, where $\mathbb{1}_2$ is the 2x2 identity matrix. Therefore, the Schrödinger equation for free spinors is just the ordinary Schrödinger equation,

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} a \\ b \end{pmatrix} \;=\; \frac{1}{2m}\,(\boldsymbol{\sigma}\cdot\boldsymbol{p})^2\begin{pmatrix} a \\ b \end{pmatrix} \;=\; \frac{\boldsymbol{p}^2}{2m}\begin{pmatrix} a \\ b \end{pmatrix} \tag{5.11}$$

and the two spin states are degenerate in energy.

This is no longer the case if we introduce the vector field. Using again minimal substitution, the Hamiltonian (a matrix in spinor space) for a particle in a vector field $(V, \boldsymbol{A})$ becomes

$$H = \frac{1}{2m}\left[\boldsymbol{\sigma}\cdot(\boldsymbol{p}-q\boldsymbol{A})\right]^2 + qV \tag{5.12}$$

It is a not entirely trivial exercise to show[1] that this equation can be rewritten as

$$H = \frac{1}{2m}\,(\boldsymbol{p}-q\boldsymbol{A})^2 \;+\; qV \;-\; \frac{\hbar q}{2m}\boldsymbol{\sigma}\cdot\boldsymbol{B} \tag{5.13}$$

The Schrödinger equation with this Hamiltonian is called the Pauli-Schrödinger equation, or simply Pauli equation. Comparing this Hamiltonian to the Hamiltonian of the classical spin, we find that the gyromagnetic ratio for a spin-$\frac{1}{2}$ particle is

$$\gamma_{\text{spin-1/2}} = \frac{q}{m} \tag{5.14}$$

exactly a factor 2 larger than for the classical picture of a spinning charge distribution.

The ratio of the magnetic moment relative to that of the classical case is called the g-factor. For spin-$\frac{1}{2}$ particles the Schrödinger-Pauli equation predicts $g = 2$. In QED the magnetic moment is modified by higher order corrections. The predictions and measurements of the magnetic moment of the electron and muon are so precise that they make QED the most precisely tested theory in physics.

Pauli introduced the spin matrices in the Hamiltonian on purely phenomenological grounds. As we shall see in the rest of this Lecture and the next, Dirac found a theoretical motivation: His construction of a wave equation that is linear in space and time derivatives, leads (in its simplest form) to the description of spin-$\frac{1}{2}$ particles and anti-particles. As you will prove in exercise 5.8, the Pauli-Schrödinger equation can be obtained as the non-relativistic limit of the equation of motion of Dirac particles in a vector field A^μ.

5.2 The Dirac equation

In order to appreciate what Dirac discovered we follow (a modern interpretation of) his approach leading to a linear wave equation. (For a different approach, which may be

[1] Take into account that $\boldsymbol{p} = -i\hbar\boldsymbol{\nabla}$ and $\boldsymbol{A}$ do not commute, then use the chain rule to see that in coordinate space only a term $\boldsymbol{B} = \boldsymbol{\nabla}\times\boldsymbol{A}$ survives.

closer to what Dirac actually did, see Griffiths, §7.1.) Consider the usual form of the Schrödinger equation,

$$i\frac{\partial}{\partial t}\psi \;=\; H\psi \;. \tag{5.15}$$

The classical Hamiltonian is quadratic in the momentum. Dirac searched for a Hamiltonian that is linear in the momentum. We start from the following ansatz[2]:

$$H \;=\; (\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta m) \tag{5.16}$$

with coefficients $\alpha_1, \alpha_2, \alpha_3, \beta$. In order to satisfy the relativistic relation between energy and momentum, we must have for any eigenvector $\psi(p)$ of H that

$$H^2\psi = \left(\boldsymbol{p}^2 + m^2\right)\,\psi \tag{5.17}$$

where $\boldsymbol{p}^2 + m^2$ is the eigenvalue. What should H look like such that these eigenvectors exist? Squaring Dirac's ansatz for the Hamiltonian gives

$$\begin{aligned} H^2 &= \left(\sum_i \alpha_i p_i + \beta m\right)\left(\sum_j \alpha_j p_j + \beta m\right) \\ &= \left(\sum_{i,j} \alpha_i\alpha_j p_i p_j + \sum_i \alpha_i\beta p_i m + \sum_i \beta\alpha_i p_i m + \beta^2 m^2\right) \\ &= \left(\sum_i \alpha_i^2 p_i^2 + \sum_{i>j}(\alpha_i\alpha_j + \alpha_j\alpha_i)p_i p_j + \sum_i(\alpha_i\beta + \beta\alpha_i)p_i m + \beta^2 m^2\right) \end{aligned} \tag{5.18}$$

where we on purpose did not impose that the coefficients (α_i, β) commute. In fact, comparing to equation (5.17) we find that the coefficients must satisfy the following requirements:

- $\alpha_1^2 = \alpha_2^2 = \alpha_3^2 = \beta^2 = 1$
- $\alpha_1, \alpha_2, \alpha_3, \beta$ anti-commute with each other.

With the following notation of the anti-commutator

$$\{A, B\} \;=\; AB + BA. \tag{5.19}$$

we can also write these requirements as

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij} \qquad \{\alpha_i, \beta\} = 0 \qquad \beta^2 = 1\;. \tag{5.20}$$

Clearly, the α_i and β cannot be ordinary numbers. At this point Dirac had a brilliant idea, possibly motivated by Pauli's picture of fermion wave functions as spinors: what if the α_i and β are matrices that act on a wave function that is a column vector? As

[2] Note that $\boldsymbol{\alpha}\cdot\boldsymbol{p} = \alpha_x p_x + \alpha_y p_y + \alpha_z p_z$.

we require the Hamiltonian to be hermitian (such that its eigenvalues are real), the matrices α_i and β must be hermitian as well,

$$\alpha_i^\dagger = \alpha_i \qquad \text{and} \qquad \beta^\dagger = \beta \; . \tag{5.21}$$

Furthermore, we can show using just the anti-commutation relations and normalization above that they all have eigenvalues ± 1 and zero trace. It then follows that they must have *even* dimension.

It can be shown that the lowest dimensional matrices that have the desired behaviour are 4x4 matrices. (See exercise 5.6 and also Aitchison (1972) §8.1). The choice of the matrices α_i and β is not unique. Here we choose the Dirac-Pauli representation,

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix} \qquad \text{and} \qquad \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \; . \tag{5.22}$$

Of course, we may expect that the final expressions for the amplitudes are independent of the representation: all the physics is in the anti-commutation relations themselves. Another frequently used choice is the Weyl representation,

$$\boldsymbol{\alpha} = \begin{pmatrix} -\boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \qquad \text{and} \qquad \beta = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \; . \tag{5.23}$$

5.3 Covariant form of the Dirac equation

With Dirac's Hamiltonian and the substitution $\boldsymbol{p} = -i\hbar\boldsymbol{\nabla}$ we arrive at the following relativistic Schrödinger-like wave equation,

$$i\frac{\partial}{\partial t}\psi = \left(-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m\right)\,\psi \; . \tag{5.24}$$

Multiplying on the left by β and using $\beta^2 = 1$ we can write this equation as

$$\left(i\beta\frac{\partial}{\partial t}\psi + i\beta\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} - m\right)\psi = 0 \; . \tag{5.25}$$

We now define the four Dirac γ-matrices by

$$\gamma^\mu \equiv (\beta, \beta\boldsymbol{\alpha}) \tag{5.26}$$

The wave equation then takes the simple form

$$\boxed{\left(i\gamma^\mu\partial_\mu - m\right)\,\psi \;=\; 0} \tag{5.27}$$

This equation is called the Dirac equation. Note that ψ is a four-element vector. We call it a *bi-spinor* or Dirac spinor. We shall see later that the solutions of the Dirac equation have four degrees of freedom, corresponding to spin-up and spin-down for a particle and its anti-particle.

The Dirac equation is actually a set of 4 coupled differential equations,

$$\text{for each j=1,2,3,4} \quad : \quad \sum_{k=1}^{4}\left[\sum_{\mu=0}^{3} i\,(\gamma^\mu)_{jk}\;\partial_\mu - m\delta_{jk}\right](\psi_k) = 0$$

$$\text{or} \quad : \quad \left[i \underbrace{\begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}}_{\gamma^\mu} \cdot \partial_\mu - \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot m \right] \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

or even more explicit, in the Dirac-Pauli representation,

$$\left[\begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \frac{i\partial}{\partial t} + \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix} \frac{i\partial}{\partial x} + \begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix} \frac{i\partial}{\partial y} + \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix} \frac{i\partial}{\partial z} - \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} m \right] \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Take note of the use of the Dirac (or spinor) indices ($j, k = 1, 2, 3, 4$) simultaneously with the Lorentz indices ($\mu = 0, 1, 2, 3$). As far as it concerns us, it is a coincidence that both types of indices assume four different values.

To simplify notation even further we define the 'slash' operator of a four-vector a^μ as

$$\not{a} \;=\; \gamma_\mu a^\mu \;. \tag{5.28}$$

The wave equation for spin-$\frac{1}{2}$ particles can then be written very concisely as

$$(i\not{\partial} - m)\,\psi = 0\;. \tag{5.29}$$

Although we write γ^μ like a contra-variant four-vector, it actually is not a four-vector. It is a set of four constant matrices that are identical in each Lorentz frame. For a four-vector a^μ, $\not{a}$ is a (4×4) matrix, but it is *not* a Lorentz invariant and still depends on the frame. The behaviour of Dirac spinors under Lorentz transformations is not entirely trivial. (See also Griffiths §7.3, Halzen and Martin §5.6 or Thomson appendix B.)

5.4 Dirac algebra

From the definitions of $\boldsymbol{\alpha}$ and β we can derive the following relation for the anti-commutator of two γ-matrices

$$\boxed{\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu \;=\; 2\,g^{\mu\nu}\,\mathbb{1}_4} \tag{5.30}$$

where the identity matrix on the right-hand side is the 4×4 identity in bi-spinor space. Text books usually leave such identity matrices away. However, it is important to realize that the equation above is a matrix equation for every value of μ and ν. In particular, $g^{\mu\nu}$ is not a matrix in spinor space. (In the equation, it is just a number!)

Using this result we find

$$\left(\gamma^0\right)^2 = \mathbb{1}_4 \qquad \left(\gamma^1\right)^2 = \left(\gamma^2\right)^2 = \left(\gamma^3\right)^2 = -\mathbb{1}_4 \,. \tag{5.31}$$

The hermitian conjugates are

$$\gamma^{0\dagger} = \gamma^0 \tag{5.32}$$
$$\gamma^{i\dagger} = \left(\beta\alpha^i\right)^\dagger = \alpha^{i\dagger}\beta^\dagger = \alpha^i\beta = -\gamma^i \,. \tag{5.33}$$

Using the anti-commutation relation once more then gives

$$\boxed{\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0} \tag{5.34}$$

In words this means that we can undo a hermitian conjugate $\gamma^{\mu\dagger}\gamma^0$ by moving a γ^0 "through it", $\gamma^{\mu\dagger}\gamma^0 = \gamma^0\gamma^\mu$. Finally, we define

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{5.35}$$

which has the characteristics

$$\gamma^{5\dagger} = \gamma^5 \qquad \left(\gamma^5\right)^2 = \mathbb{1}_4 \qquad \left\{\gamma^5, \gamma^\mu\right\} = 0 \tag{5.36}$$

In the Dirac-Pauli representation γ^5 is

$$\gamma^5 = \begin{pmatrix} 0 & \mathbb{1}_2 \\ \mathbb{1}_2 & 0 \end{pmatrix} \tag{5.37}$$

5.5 Adjoint spinors and current density

Similarly to the case of the Schrödinger and the Klein-Gordon equations we can define a current density j^μ that satisfies a continuity equation. First, we write the Dirac equation as

$$i\gamma^0\frac{\partial\psi}{\partial t} + i\sum_{k=1,2,3}\gamma^k\frac{\partial\psi}{\partial x^k} - m\psi = 0 \tag{5.38}$$

As we now work with matrices, we use hermitian conjugates rather than complex conjugates and find for the conjugate equation

$$-i\frac{\partial\psi^\dagger}{\partial t}\gamma^0 - i\sum_{k=1,2,3}\frac{\partial\psi^\dagger}{\partial x^k}\left(-\gamma^k\right) - m\psi^\dagger = 0 \tag{5.39}$$

However, we now see a potential problem: the additional minus sign in $\left(-\gamma^k\right)$ disturbs the Lorentz invariant form of the equation. We can restore Lorentz covariance by multiplying the equation from the right by γ^0. Therefore, we define the *adjoint Dirac spinor*

$$\boxed{\overline{\psi} = \psi^\dagger \gamma^0} \,. \tag{5.40}$$

Note that the adjoint spinor is a row-vector:

$$\text{Dirac spinor :} \quad \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \qquad \text{Adjoint Dirac spinor:} \quad \left(\overline{\psi}_1, \overline{\psi}_2, \overline{\psi}_3, \overline{\psi}_4\right)$$

The adjoint Dirac spinor satisfies the equation

$$-i\frac{\partial\overline{\psi}}{\partial t}\gamma^0 - i\sum_{k=1,2,3}\frac{\partial\overline{\psi}}{\partial x^k}\gamma^k - m\overline{\psi} = 0 \tag{5.41}$$

which can be written in short-hand as

$$\boxed{i\partial_\mu\overline{\psi}\gamma^\mu + m\overline{\psi} \;=\; 0.} \tag{5.42}$$

Now we multiply the Dirac equation from the left by $\overline{\psi}$ and we multiply the adjoint Dirac equation from the right by ψ:

$$\begin{array}{rcl} \left(i\partial_\mu\overline{\psi}\gamma^\mu + m\overline{\psi}\right)\,\psi &=& 0 \\ \overline{\psi}\,\left(i\partial_\mu\gamma^\mu\psi - m\psi\right) &=& 0 \\ \hline & + & \\ \overline{\psi}\left(\partial_\mu\gamma^\mu\psi\right) + \left(\partial_\mu\overline{\psi}\gamma^\mu\right)\psi &=& 0 \end{array}$$

Consequently, we realize that if we define a current as

$$\boxed{j^\mu \;=\; \overline{\psi}\gamma^\mu\psi} \tag{5.43}$$

then this current satisfies a continuity equation, $\partial_\mu j^\mu = 0$. The first component of this current is simply

$$j^0 = \overline{\psi}\gamma^0\psi = \psi^\dagger\psi = \sum_{i=1}^{4}|\psi_i|^2 \tag{5.44}$$

which is always positive. This property was the original motivation of Dirac's work.

The form Eq. (5.43) suggests that the Dirac probability current density transforms as a contravariant four-vector. In contrast to the Klein-Gordon case, this is not so easy to show since the Dirac spinors transform non-trivially. We will leave the details to the textbooks.

5.6 Bilinear covariants

The Dirac probability current in Eq. (5.43) is an example of a so-called *bilinear covariant*: a quantity that is a product of components of $\overline{\psi}$ and ψ and obeys the standard transformation properties of Lorentz scalars, vectors or tensors. The bilinear covariants represent the most general form of currents consistent with Lorentz covariance.

Given that $\overline{\psi}$ and ψ each have four components, we have 16 independent combinations. Requiring the currents to be covariant, then leads to the following types of currents:

$$\begin{array}{lll} & & \text{\# of components} \\ \text{scalar} & \overline{\psi}\,\psi & 1 \\ \text{vector} & \overline{\psi}\,\gamma^\mu\,\psi & 4 \\ \text{tensor} & \overline{\psi}\,\sigma^{\mu\nu}\,\psi & 6 \\ \text{axial vector} & \overline{\psi}\,\gamma^5\,\gamma^\mu\,\psi & 4 \\ \text{pseudo scalar} & \overline{\psi}\,\gamma^5\,\psi & 1 \end{array} \tag{5.45}$$

where the (anti-symmetric) tensor is defined as

$$\sigma^{\mu\nu} \;\equiv\; \frac{i}{2}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right) \tag{5.46}$$

The names 'axial' and 'pseudo' refer to the behaviour of these objects under the parity transformation, $\boldsymbol{x} \to -\boldsymbol{x}$. The scalar is invariant under parity, while the pseudo scalar changes sign. The space components of the vector change sign under parity, while those of the axial vector do not. We shall discuss the bilinear covariants and their transformation properties in more detail in Lecture 7.

5.7 Solutions to the Dirac Equation

5.7.1 Plane waves solutions with $\mathbf{p} = 0$

We now consider explicit expressions for the solutions of the Dirac equation, Eq. 5.27. In exercise 5.1 you will show that each of the components of the Dirac wave satisfies the Klein-Gordon equation. Therefore, we try to construct the solutions as plane wave solutions

$$\psi(x) \;=\; u(p)\,e^{-ipx} \tag{5.47}$$

where $u(p)$ is a 4-component column-vector that does not depend on x. After substitution in the Dirac equation we find the Dirac equation in the momentum representation,

$$\boxed{\left(\gamma^\mu p_\mu - m\right)\; u(p) \;=\; 0} \tag{5.48}$$

which, using the 'slash notation' can also be written as

$$\left(\not{p} - m\right)\; u(p) \;=\; 0\;. \tag{5.49}$$

In momentum-space the coupled differential equations reduce to a set of coupled linear equations. In the Pauli-Dirac representation we have

$$\left[\begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} E - \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} p_i - \begin{pmatrix} \mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} m \right] \begin{pmatrix} u_A \\ u_B \end{pmatrix} = 0 \tag{5.50}$$

where u_A and u_B are two-component spinors. We can rewrite this as a set of two equations

$$\begin{aligned} (\boldsymbol{\sigma} \cdot \boldsymbol{p}) \; u_B &= (E - m)\, u_A \\ (\boldsymbol{\sigma} \cdot \boldsymbol{p}) \; u_A &= (E + m)\, u_B \;, \end{aligned} \tag{5.51}$$

where $\boldsymbol{\sigma} \equiv (\sigma_1, \sigma_2, \sigma_3)$.

Now consider a particle with non-zero mass in its restframe, $p = 0$. In this case, the two equations decouple,

$$\begin{aligned} E\, u_A &= m\, u_A \\ E\, u_B &= -m\, u_B \;. \end{aligned} \tag{5.52}$$

There are four independent solutions, which we write as

$$u^{(1)} = N \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} , \quad u^{(2)} = N \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} , \quad u^{(3)} = N \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} , \quad u^{(4)} = N \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{5.53}$$

The first two have a positive energy eigenvalue $E = m$ and the second two a negative energy $E = -m$. We discuss the normalization constant N later.

5.7.2 Plane wave solutions for $\mathbf{p} \neq 0$

In order to extend the solution to particles with non-zero momentum, consider two Dirac spinors for which the two upper coordinates $u_A(p)$ of $u(p)$ are given by

$$u_A^{(1)} \;=\; \chi^{(1)} \qquad \text{and} \qquad u_A^{(2)} \;=\; \chi^{(2)} \;. \tag{5.54}$$

with the basis spinors $\chi^{1,2}$ defined in Eq. 5.5. Substituting this into the second equation of (5.51) gives for the lower two components

$$u_B^{(1,2)} = \frac{\boldsymbol{\sigma} \cdot \boldsymbol{p}}{E + m}\, u_A^{(1,2)} \;=\; \frac{\boldsymbol{\sigma} \cdot \boldsymbol{p}}{E + m}\, \chi^{(1,2)} \;. \tag{5.55}$$

To prove that these are indeed solutions of the equations, one can use the identity

$$(\boldsymbol{\sigma} \cdot \boldsymbol{p})\, (\boldsymbol{\sigma} \cdot \boldsymbol{p}) \;=\; |\boldsymbol{p}|^2\, \mathbb{1}_2 \tag{5.56}$$

such that $u_A^{(1,2)}$ and $u_B^{(1,2)}$ also satisfy the first equation in (5.51). (See also exercise 5.2.)

In the Pauli-Dirac representation the Hamiltonian is given by

$$H = \boldsymbol{\alpha} \cdot \boldsymbol{p} + \beta m = \begin{pmatrix} 0 & \boldsymbol{\sigma} \cdot \boldsymbol{p} \\ \boldsymbol{\sigma} \cdot \boldsymbol{p} & 0 \end{pmatrix} + \begin{pmatrix} m\mathbb{1}_2 & 0 \\ 0 & -m\mathbb{1}_2 \end{pmatrix} \tag{5.57}$$

With a bit of algebra we obtain for our solution

$$Hu^{(1)} = \begin{pmatrix} \left[m + \frac{\boldsymbol{p}^2}{E+m}\right] u_A^{(1)} \\ E\, u_B^{(1)} \end{pmatrix} , \tag{5.58}$$

which illustrates two things: In order that $u^{(1)}$ be a solution we need indeed that $E^2 = \boldsymbol{p}^2 + m^2$. Furthermore, in the limit that $p \to 0$, the energy eigenvalue is $+m$, such that this is a positive energy solution. The calculation for $u^{(2)}$ is identical. Hence, two orthogonal positive-energy solutions are

$$\boxed{u^{(1)} = N \begin{pmatrix} \chi^{(1)} \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi^{(1)} \end{pmatrix} \qquad \text{and} \qquad u^{(2)} = N \begin{pmatrix} \chi^{(2)} \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi^{(2)} \end{pmatrix}} \tag{5.59}$$

where N is again a normalization constant.

In an exactly analogous manner, we can start for our $E < 0$ solutions with the *lower* component given by $\chi^{(s)}$,

$$u_B^{(3)} = \chi^{(1)} \qquad , \qquad u_B^{(4)} = \chi^{(2)} . \tag{5.60}$$

Using the first of the equations in Eq. (5.51) gives for the upper coordinates

$$u_A^{(3,4)} = \frac{\boldsymbol{\sigma} \cdot \boldsymbol{p}}{E - m}\, u_B^{(3,4)} = -\frac{\boldsymbol{\sigma} \cdot \boldsymbol{p}}{(-E) + m}\, \chi^{(1,2)} \tag{5.61}$$

Note the difference in the enumerator: it has become $(E - m)$ rather than $(E + m)$. Evaluating the energy eigenvalue, we now find *e.g.*

$$Hu^{(3)} = \begin{pmatrix} E\, u_A^{(3)} \\ \left[-m + \frac{\boldsymbol{p}^2}{E-m}\right] u_B^{(3)} \end{pmatrix} , \tag{5.62}$$

which again requires $E^2 = \boldsymbol{p}^2 + m^2$ and in the limit $p \to 0$ gives $E = -m$, a negative energy solution. Consequently, two negative-energy orthogonal solutions are given by

$$\boxed{u^{(3)} = N \begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E-m}\chi^{(1)} \\ \chi^{(1)} \end{pmatrix} \qquad \text{and} \qquad u^{(4)} = N \begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E-m}\chi^{(2)} \\ \chi^{(2)} \end{pmatrix}} \tag{5.63}$$

To gain slightly more insight, let's write them out in momentum components. Using the definition of the Pauli matrices we have

$$\boldsymbol{\sigma} \cdot \boldsymbol{p} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} p_x + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} p_y + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} p_z \tag{5.64}$$

to find

$$(\boldsymbol{\sigma}\cdot\boldsymbol{p})\ u_A^{(1)} = \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} p_z \\ p_x + ip_y \end{pmatrix} \tag{5.65}$$

and similar for $u_A^{(2)}$, $u_B^{(3)}$, $u_B^{(4)}$. The solutions can then be written as

$$E>0 \text{ spinors} \qquad u^{(1)}(p) = N \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ \frac{p_x+ip_y}{E+m} \end{pmatrix} \quad , \quad u^{(2)}(p) = N \begin{pmatrix} 0 \\ 1 \\ \frac{p_x-ip_y}{E+m} \\ \frac{-p_z}{E+m} \end{pmatrix}$$

$$E<0 \text{ spinors} \qquad u^{(3)}(p) = N \begin{pmatrix} \frac{-p_z}{-E+m} \\ \frac{-p_x-ip_y}{-E+m} \\ 1 \\ 0 \end{pmatrix} \quad , \quad u^{(4)}(p) = N \begin{pmatrix} \frac{-p_x+ip_y}{-E+m} \\ \frac{p_z}{-E+m} \\ 0 \\ 1 \end{pmatrix}$$

You can verify that these solutions are indeed orthogonal, *i.e.* that

$$u^{(i)\dagger}u^{(j)} = 0 \qquad \text{for } i \neq j. \tag{5.66}$$

5.8 Anti-particle spinors

As for the solutions of the K.-G. equation, we interprete $u^{(1)}$ and $u^{(2)}$ as the positive energy solutions of a particle (electron, charge e^-) and $u^{(3)}$, $u^{(4)}$ as the positive energy solutions of the corresponding anti-particle (the positron). We define the anti-particle components of the wave function as

$$\begin{aligned} v^{(1)}(p) &\equiv u^{(4)}(-p) \\ v^{(2)}(p) &\equiv u^{(3)}(-p)\ . \end{aligned} \tag{5.67}$$

Using this definition we can replace the two negative energy solutions by the following anti-particle spinors with positive energy, $E = +\sqrt{\boldsymbol{p}^2 + m^2}$,

$$v^{(1)}(p) = N \begin{pmatrix} \frac{p_x-ip_y}{E+m} \\ \frac{-p_z}{E+m} \\ 0 \\ 1 \end{pmatrix} \quad , \quad v^{(2)}(p) = N \begin{pmatrix} \frac{p_z}{E+m} \\ \frac{(p_x+ip_y)}{E+m} \\ 1 \\ 0 \end{pmatrix} .$$

The spinors $u(p)$ of matter waves are solutions of the Dirac equation in momentum space, Eq. (5.48). Replacing p with $-p$ in the Dirac equation we find that our positive energy anti-particle spinors satisfy another Dirac equation,

$$\boxed{(\not{p} + m)\ v(p) = 0} \tag{5.68}$$

5.9 Normalization of the wave function

As for the Klein-Gordon case we choose a normalization such that there are $2E$ particles per unit volume. The density is given by the time-component of the bi-spinor current (see Eq. 5.44)

$$\rho(x) \;=\; \psi^\dagger(x)\,\psi(x)\,. \tag{5.69}$$

Substituting the plane wave solution $\psi = u(p)\,e^{-ipx}$, and integrating over a volume V we find

$$\int_V \rho\,\mathrm{d}^3x \;=\; \int_V u^\dagger(p)\,e^{ipx}\,u(p)\,e^{-ipx}\,\mathrm{d}^3x \;=\; u^\dagger(p)\,u(p)\,V \tag{5.70}$$

Consequently, to find $2E$ particles per unit volume we must normalize such that

$$u^\dagger(p)\,u(p) \;=\; 2E \tag{5.71}$$

Explicit calculation for the positive energy solutions ($s \in \{1, 2\}$) gives

$$\begin{aligned} u^{(s)\dagger}\,u^{(s)} &= N^2 \left(\chi^{(s)T} \chi^{(s)} + \chi^{(s)T} \frac{(\boldsymbol{\sigma}\cdot\boldsymbol{p})^\dagger\,(\boldsymbol{\sigma}\cdot\boldsymbol{p})}{(E+m)^2} \chi^{(s)} \right) \\ &= N^2 \left(1 + \frac{\boldsymbol{p}^2}{(E+m)^2} \right) \;=\; N^2 \frac{2E}{E+m} \end{aligned}$$

Consequently, in order to have $2E$ particles per unit volume we choose

$$N \;=\; \sqrt{E+m}\,. \tag{5.72}$$

The computation for the positive energy anti-particle waves $v(p)$ leads to the same normalization. We can now write the orthogonality relations as (with $r, s \in \{1, 2\}$)

$$\begin{aligned} u^{(r)\dagger}\,u^{(s)} &= 2E\,\delta_{rs} \\ v^{(r)\dagger}\,v^{(s)} &= 2E\,\delta_{rs} \end{aligned} \tag{5.73}$$

5.10 The completeness relation

We now consider the Dirac equation for the adjoint spinor $\overline{u}$, $\overline{v}$. Taking the hermitian conjugate of Eq. (5.48) and multiplying on the right by γ^0 we have

$$u^\dagger \gamma^{\mu\dagger} \gamma^0 p_\mu - u^\dagger \gamma^0 m = 0 \tag{5.74}$$

Using that $\gamma^{\mu\dagger}\gamma^0 = \gamma^0\gamma^\mu$ we then find for the Dirac equation of the adjoint spinor $\overline{u} = u^\dagger\gamma^0$,

$$\boxed{\overline{u}\;(\not{p} - m) = 0} \tag{5.75}$$

In the same manner we find for the adjoint anti-particle spinors

$$\boxed{\overline{v}\;(\not{p} + m) = 0} \tag{5.76}$$

Using these results you will derive in exercise 5.4 the so-called *completeness relations*

$$\boxed{\begin{aligned}\sum_{s=1,2} u^{(s)}(p)\,\overline{u}^{(s)}(p) &= (\not{p}+m)\\ \sum_{s=1,2} v^{(s)}(p)\,\overline{v}^{(s)}(p) &= (\not{p}-m)\end{aligned}} \tag{5.77}$$

These relations will be used later on in the calculation of amplitudes with Feynman diagrams. Note that the left-hand side is *not* an inner product. Rather, on both sides we have a (4×4) matrix, or schematically

$$\begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \cdot (....) = \begin{pmatrix} & & \\ & \gamma^\mu & \\ & & \end{pmatrix} \cdot p_\mu + \begin{pmatrix} & & \\ & \mathbb{1} & \\ & & \end{pmatrix} \cdot m$$

For the negative-energy solutions we find $\sum_{s=3,4} u^{(s)}(p)\,\overline{u}^{(s)}(p) = \sum_{s=1,2} v^{(s)}(-p)\,\overline{v}^{(s)}(-p) = -(\not{p}+m)$.

5.11 Helicity

The Dirac spinors for a given momentum p have a two-fold degeneracy. This implies that there must be an additional observable with an operator that commutes with H and p, and the eigenvalues of which distinguish between the degenerate states. It is tempting to identify this extra quantum number with the intrinsic spin, but the relation is subtle.

The spin operator in bi-spinor space is defined as $S = \frac{1}{2}\hbar\Sigma$, with

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} . \tag{5.78}$$

In exercise 5.3 you will show that Σ does not commute with the Hamiltonian in Eq. (5.57). We can also realize this by looking directly at our Dirac spinor solutions: If spin is a good quantum number then those solutions should be eigenstates of the spin operator,

$$\boldsymbol{\Sigma}\, u^{(i)} = s\, u^{(i)} \ \ ?$$

where s is the spin eigenvalue. Now insert one of the solutions, for example $u^{(1)}$,

$$\begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} p_z/\,(E+m) \\ (p_x+ip_y)\,/\,(E+m) \end{pmatrix} \end{pmatrix} \stackrel{?}{=} s \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} p_z/\,(E+m) \\ (p_x+ip_y)\,/\,(E+m) \end{pmatrix} \end{pmatrix}$$

and you realize that this could never be true for arbitrary p_x, p_y, p_z.

The orbital angular momentum operator is defined as usual as

$$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p} \tag{5.79}$$

You will also show in exercise 5.3 that the total angular momentum

$$\boldsymbol{J} \;=\; \boldsymbol{L} + \tfrac{1}{2}\boldsymbol{\Sigma} \tag{5.80}$$

does commute with the Hamiltonian. Now, as we can choose an arbitrary axis to get the spin quantum numbers, we can choose an axis such that the orbital angular momentum vanishes, namely along the direction of the momentum. Consequently, we define the *helicity* operator λ as

$$\lambda = \frac{1}{2}\boldsymbol{\Sigma}\cdot\hat{\boldsymbol{p}} \equiv \frac{1}{2}\begin{pmatrix} \boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}} \end{pmatrix} \tag{5.81}$$

where $\hat{\boldsymbol{p}} \equiv \boldsymbol{p}/|\boldsymbol{p}|$. We can interpret the helicity as the "spin component in the direction of movement". One can verify that indeed λ commutes with the Hamiltonian in (5.57).

As λ and H commute, they have a common set of eigenvectors. However, that does not necessarily mean that our solutions $u^{(i)}$ are indeed also eigenvectors of λ. In fact, with our choice above, they are only eigenvectors of λ if we choose the momentum along the z-axis. The reason is that the two-component spinors $\chi^{(s)}$ are eigenvectors of σ_3 only. For other directions of the momentum, we would need to choose a different linear combination of the $u^{(i)}$ to form a set of states that are eigenvectors for both H and λ.

Now, consider a momentum vector $\boldsymbol{p} = (0, 0, p)$. Applying the helicity operator on $u^{(i)}$ gives

$$\begin{aligned} \frac{1}{2}\left(\boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}}\right)\, u_A^{\pm} &= \frac{1}{2}\sigma_3 u_A^{\pm} = \pm\frac{1}{2}u_A^{\pm} \\ \frac{1}{2}\left(\boldsymbol{\sigma}\cdot\hat{\boldsymbol{p}}\right)\, u_B^{\pm} &= \frac{1}{2}\sigma_3 u_B^{\pm} = \pm\frac{1}{2}u_B^{\pm} \end{aligned}$$

where the plus sign holds for $u^{(1,3)}$ and the minus sign for $u^{(2,4)}$. So you see that indeed u is an eigenvector of λ with eigenvalues $\pm 1/2$. Positive helicity states have spin and momentum parallel, while negative helicity states have them anti-parallel.

It is not so difficult to derive the spinors that are eigenvectors of both λ and the Dirac Hamilitonian for arbitrary momentum $\boldsymbol{p}$. (See for instance §4.8.1 in Thomson.) We save you the algebra and just give the result. To simplify notation we switch to polar coordinates,

$$\boldsymbol{p} \;=\; (p\sin\theta\cos\phi, p\sin\theta\sin\phi, p\cos\theta)\ . \tag{5.82}$$

The particle spinors for helicity $+1/2$ and helicity $-1/2$ become, respectively,

$$u_{\uparrow} \;=\; N\begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ e^{i\phi}\,\sin\left(\frac{\theta}{2}\right) \\ \frac{p}{E+m}\,\cos\left(\frac{\theta}{2}\right) \\ \frac{p}{E+m}\,e^{i\phi}\,\sin\left(\frac{\theta}{2}\right) \end{pmatrix} \qquad u_{\downarrow} \;=\; N\begin{pmatrix} -\sin\left(\frac{\theta}{2}\right) \\ e^{i\phi}\,\cos\left(\frac{\theta}{2}\right) \\ \frac{p}{E+m}\,\sin\left(\frac{\theta}{2}\right) \\ -\frac{p}{E+m}\,e^{i\phi}\,\cos\left(\frac{\theta}{2}\right) \end{pmatrix}\ , \tag{5.83}$$

while those for the anti-particles are

$$v_\uparrow = N \begin{pmatrix} \frac{p}{E+m} \sin\left(\frac{\theta}{2}\right) \\ -\frac{p}{E+m} e^{i\phi} \cos\left(\frac{\theta}{2}\right) \\ -\sin\left(\frac{\theta}{2}\right) \\ e^{i\phi} \cos\left(\frac{\theta}{2}\right) \end{pmatrix} \qquad v_\downarrow = N \begin{pmatrix} \frac{p}{E+m} \cos\left(\frac{\theta}{2}\right) \\ \frac{p}{E+m} e^{i\phi} \sin\left(\frac{\theta}{2}\right) \\ \cos\left(\frac{\theta}{2}\right) \\ e^{i\phi} \sin\left(\frac{\theta}{2}\right) \end{pmatrix} . \tag{5.84}$$

5.12 Charge current and anti-particles

Once we consider interactions of fermions in QED, we are interested in charge density rather than probability density. Following the Pauli-Weiskopf prescription for the complex scalar field, we multiply the current of (negatively charged) particles by $-e$,

$$j^\mu_{\text{em}} = -e\,\overline{\psi}\,\gamma^\mu\,\psi\,. \tag{5.85}$$

The interaction current density 4-vector takes the form

$$j^\mu_{fi} = -e \left(\quad \overline{u}_f \quad \right) \begin{pmatrix} & & \\ & \gamma^\mu & \\ & & \end{pmatrix} \begin{pmatrix} \\ u_i \\ \\ \end{pmatrix} e^{i(p_f - p_i)x} \tag{5.86}$$

We have seen above that although the probability density of the Dirac fields is positive, the negative energy solutions just remain. Following the Feynman-Stückelberg interpretation the solution with negative energy is again seen as the anti-particle solution with positive energy. However, when it comes to the Feynman rules, there is an additional subtlety for fermions.

In the case of Klein-Gordon waves the current of an anti-particle ($j^\mu = 2|N|^2 p^\mu$) gets a minus sign with respect to the current of the particle, due to reversal of 4-momentum. This cancels the change in the sign of the charge and that is how we came to the nice property of 'crossing': simply replace any anti-particle by a particle with opposite momentum. For fermions this miracle does not happen: the current does not automatically change sign when we go to anti-particles. As a result, if we want to keep the convention that allows us to replace anti-particles by particles, we need an additional 'ad-hoc' minus sign in the Feynman rule for the current of the spin-$\frac{1}{2}$ anti-particle.

This additional minus sign between particles and anti-particles is only required for fermionic currents and not for bosonic currents. It is related to the spin-statistics connection: bosonic wavefunctions are symmetric, and fermionic wavefunctions are antisymmetric. In field theory[3] the extra minus sign is a result of the fact that bosonic field operators follow *commutation* relations, while fermionic field operators follow *anti-commutation* relations. This was realized first by Pauli in 1940.

[3]See Aitchison & Hey, 3rd edition §7.2

5.13 The charge conjugation operation

The Dirac equation for a particle in an electromagnetic field is obtained by substituting $\partial_\mu \to \partial_\mu + iqA_\mu$ in the free Dirac equation. For an electron ($q = -e$) this leads to:

$$[\gamma^\mu (i\partial_\mu + eA_\mu) - m]\, \psi = 0 \ . \tag{5.87}$$

Similarly, there must be a Dirac equation describing the positron ($q = +e$):

$$[\gamma^\mu (i\partial_\mu - eA_\mu) - m]\, \psi_C = 0 \ , \tag{5.88}$$

where the positron wave function ψ_C is obtained by a one-to-one correspondence with the electron wave function ψ.

To find the relation between ϕ_C and ψ, let's take the complex conjugate of the electron equation,

$$[-\gamma^{\mu *} (i\partial_\mu - eA_\mu) - m]\, \psi^* = 0 \ . \tag{5.89}$$

Now suppose that there is a matrix M such that

$$\gamma^{\mu *} \ = \ M^{-1} \gamma^\mu M \ . \tag{5.90}$$

We can then rewrite the equation above as

$$M^{-1} [\gamma^\mu (i\partial_\mu - eA_\mu) - m]\, M\psi^* = 0 \ . \tag{5.91}$$

and we obtain the relation

$$\psi_C = M\psi^* = M\gamma^0 \overline{\psi}^T \equiv C\overline{\psi}^T \tag{5.92}$$

where we have used the definition of the adjoint spinor (see Lecture 6) and defined the charge conjugation matrix $C = M\gamma^0$. It can be shown (see Halzen and Martin exercise 5.6) that in the Pauli-Dirac representation a possible choice of M is

$$M = C\gamma^0 = i\gamma^2 = \begin{pmatrix} & & & 1 \\ & & -1 & \\ & -1 & & \\ 1 & & & \end{pmatrix} \ . \tag{5.93}$$

Interpreting the probability current as a charge current, we define the electron current as

$$j^\mu_{-e} \ = \ -e\overline{\psi}\gamma^\mu\psi \tag{5.94}$$

The current of the charge conjugate wave function is then

$$j^{\mu\ C}_{-e} \ = \ -e\overline{\psi_C}\gamma^\mu\psi_C \ = \ \ldots \ = \ e\overline{\psi}\gamma^\mu\psi \tag{5.95}$$

Exercises

Exercise 5.1 (From Dirac to Klein-Gordon)
Each of the four components of the Dirac equation satisfies the Klein Gordon equation, $(\partial_\mu \partial^\mu + m^2)\, \psi_i = 0$. Show this explicitly by operating on the Dirac equation from the left with $(i\gamma^\nu \partial_\nu + m)$.

Hint: For any a^μ and b^ν we can write $\gamma^\nu \gamma^\mu a_\nu b_\mu = \frac{1}{2}(\gamma^\nu \gamma^\mu a_\nu b_\mu + \gamma^\mu \gamma^\nu a_\mu b_\nu)$ by just 'renaming' indices'. Now take the special case that $b = a$ and the a^μ commute. Then we can write $\gamma^\nu \gamma^\mu a_\nu a_\mu = \frac{1}{2}(\gamma^\nu \gamma^\mu + \gamma^\mu \gamma^\nu) a_\mu a_\nu$. Now use the anti-commutation relation of the γ-matrices.

Exercise 5.2 (Energy eigenvalue of solutions to Dirac equation)
Starting from the Dirac equation in momentum space, Eq. (5.51), show ...

(a) ... by eliminating u_B that a solution to the Dirac equation satisfies the relativistic relation between energy and momentum.

(b) ... for a non-relativistic particle with velocity β, u_B is a factor $\frac{1}{2}\beta$ smaller than u_A. (In a non-relativistic description u_A and u_B are often called respectively the "large" and "small" components of the wave function.)

Hint: Use the result from exercise 5.7b.

Exercise 5.3 (See also exercise 5.4 of H& M and exercise 7.8 of Griffiths)
The purpose of this problem is to demonstrate that particles described by the Dirac equation carry "intrinsic" angular momentum ($\boldsymbol{S}$) in addition to their orbital angular momentum ($\boldsymbol{L}$). We will see that $\boldsymbol{L}$ and $\boldsymbol{S} = \boldsymbol{\Sigma}/2$ are not conserved individually but that their *sum* is.

(a) Consider the Hamiltonian that leads to the Dirac equation,

$$H \;=\; \boldsymbol{\alpha} \cdot \boldsymbol{p} + \beta m$$

Use the fundamental commutator $[x_i, p_j] = i\hbar \delta_{ij}$ (with $\hbar = 1$) to show that

$$[H, \boldsymbol{L}] = -i\; \boldsymbol{\alpha} \times \boldsymbol{p} \tag{5.96}$$

where $\boldsymbol{L} = \boldsymbol{x} \times \boldsymbol{p}$.
Hint: To do this efficiently use the Levi-Civita tensor to write out the cross product as $L_i \;=\; \sum_{j,k} \epsilon_{ijk} x_j p_k$. Now evaluate the commutator $[H, L_i]$.

(b) Show that

$$[\alpha_k, \Sigma_l] = 2i \sum_m \epsilon_{klm} \alpha_m$$

where the operator $\boldsymbol{\Sigma}$ (see also Eq. (5.78)) and $\boldsymbol{\alpha}$ in the Pauli-Dirac representation were

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}$$

Hint: Use the commutation relation for the Pauli spin matrices $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$ *(which follows from* $\sigma_i\sigma_j = i\sum_k \epsilon_{ijk}\sigma_k$*).*

(c) Use the result in *(b)* to show that

$$[H, \boldsymbol{\Sigma}] = 2\, i\, \boldsymbol{\alpha} \times \boldsymbol{p} \tag{5.97}$$

We see from (a) and (c) that the Hamiltonian commutes with $J = L + \frac{1}{2}\Sigma$.

Exercise 5.4 (See also H&M p.110-111 and Griffiths p. 242)
The spinors u, v, $\bar{u}$ and $\bar{v}$ are solutions of respectively:

$$\begin{aligned} (\not{p} - m)\, u &= 0 \\ (\not{p} + m)\, v &= 0 \\ \bar{u}\, (\not{p} - m) &= 0 \\ \bar{v}\, (\not{p} + m) &= 0 \end{aligned}$$

(a) Use the orthogonality relations:

$$\begin{aligned} u^{(r)\dagger}\, u^{(s)} &= 2E\, \delta_{rs} \\ v^{(r)\dagger}\, v^{(s)} &= 2E\, \delta_{rs} \end{aligned}$$

to show that:

$$\begin{aligned} \bar{u}^{(s)}\, u^{(s)} &= 2m \\ \bar{v}^{(s)}\, v^{(s)} &= -2m \end{aligned}$$

Hint: evaluate the sum of $\overline{u}\gamma^0(\not{p} - m)u$ *and* $\overline{u}(\not{p} - m)\gamma^0 u$ *and use* $\gamma^0\gamma^k = -\gamma^k\gamma^0$ *(*$k = 1, 2, 3$*).*

(b) Derive the completeness relations:

$$\begin{aligned} \sum_{s=1,2} u^{(s)}(p)\, \bar{u}^{(s)}(p) &= \not{p} + m \\ \sum_{s=1,2} v^{(s)}(p)\, \bar{v}^{(s)}(p) &= \not{p} - m \end{aligned}$$

Hint: For $s = 1, 2$ *take the solution* $u^{(s)}$ *from Eq.* (5.59) *and write out the row-vector for* $\overline{u}^s$ *using the explicit form of* γ^0 *in the Dirac-Pauli representation. Then write out the matrix* $u^{(s)}\overline{u}^{(s)}$ *and use that* $\sum_{s=1,2}\chi^{(s)}\chi^{(s)\dagger} = \mathbb{1}_2$*. Finally, note that*

$$\not{p} = \begin{pmatrix} E\, \mathbb{1}_2 & -\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -E\, \mathbb{1}_2 \end{pmatrix} \tag{5.98}$$

Exercise 5.5 (Dirac algebra: traces and products of γ matrices)
Use the anti-commutator relation for Dirac γ-matrices in Eq. (5.30) (or anything that follows from that) to show that:

(a) $\not{a}\ \not{b} + \not{b}\ \not{a} = 2\,(a \cdot b)\,\mathbb{1}_4$

(b) i) $\gamma_\mu \gamma^\mu = 4\,\mathbb{1}_4$

ii) $\gamma_\mu \not{a}\, \gamma^\mu = -2\not{a}$

iii) $\gamma_\mu \not{a}\ \not{b}\, \gamma^\mu = 4\,(a \cdot b)\,\mathbb{1}_4$

iv) $\gamma_\mu \not{a}\ \not{b}\ \not{c}\, \gamma^\mu = -2\not{c}\ \not{b}\ \not{a}$

(c) i) Tr (odd number of γ-matrices) $= 0$

ii) $\mathrm{Tr}(\not{a}\ \not{b}) = 4\,(a \cdot b)$

iii) $\mathrm{Tr}(\not{a}\ \not{b}\ \not{c}\ \not{d}) = 4\ [\,(a \cdot b)(c \cdot d) - (a \cdot c)(b \cdot d) + (a \cdot d)(b \cdot c)\,]$

(d) i) $\mathrm{Tr}\,\gamma^5 = \mathrm{Tr}\, i\, \gamma^0\gamma^1\gamma^2\gamma^3 = 0$

ii) $\mathrm{Tr}\,\gamma^5 \not{a}\ \not{b} = 0$

iii) **Optional!**, for the die-hards:

$$\mathrm{Tr}\,\gamma^5 \not{a}\ \not{b}\ \not{c}\ \not{d} \ = -4\, i\, \varepsilon_{\alpha\beta\gamma\delta}\, a^\alpha b^\beta c^\gamma d^\delta$$

where $\varepsilon_{\alpha\beta\gamma\delta} = +1\,(-1)$ for an even (odd) permutation of 0,1,2,3; and 0 if two indices are the same.

Exercise 5.6 (Representations of Dirac matrices)

(a) Write a general hermitian 2×2 matrix in the form $\begin{pmatrix} a & b \\ b^* & c \end{pmatrix}$ where a and c are real. Write then $b = s + it$ and show that the matrix can be written as: $\{(a+c)/2\}\, I + s\sigma_1 - t\sigma_2 + \{(a-c)/2\}\,\sigma_3$
How can we conclude that $\boldsymbol{\alpha}$ and β cannot be 2×2 matrices?

(b) Show that the $\boldsymbol{\alpha}$ and β matrices in both the Dirac-Pauli as well as in the Weyl representation have the required anti-commutation behaviour.

Exercise 5.7 (Pauli vector identity)

(a) Prove the Pauli vector identity

$$(\boldsymbol{\sigma} \cdot \boldsymbol{a})(\boldsymbol{\sigma} \cdot \boldsymbol{b}) \ = \ \boldsymbol{a} \cdot \boldsymbol{b}\, \mathbb{1}_2 \ + \ i\boldsymbol{\sigma} \cdot (\boldsymbol{a} \times \boldsymbol{b}) \tag{5.99}$$

Hint: Write the inner products as sums over i and j. Use that $\sigma_i\sigma_j = \delta_{ij} + i\sum_k \epsilon_{i,jk}\sigma_k$. Use that $\sum_{i,j} \epsilon_{ijk} a_i b_j = (\boldsymbol{a} \times \boldsymbol{b})_k$.

(b) Prove the identity in Eq. (5.56), *i.e.*

$$(\boldsymbol{\sigma} \cdot \boldsymbol{p})\,(\boldsymbol{\sigma} \cdot \boldsymbol{p}) \ = \ |\boldsymbol{p}|^2\, \mathbb{1}_2$$

Exercise 5.8 (Pauli equation as non-relativistic limit)
In this exercise you will show that in a non-relativistic approximation, the Dirac equation combined with minimal substitution leads to the Schrödinger-Pauli equation, and hence to the prediction of the gyromagnetic ratio of the electron.

Consider again minimal substitution for a charge q and a field $A^\mu = (V, \boldsymbol{A})$:

$$p^\mu \to p^\mu - qA^\mu \qquad \Longrightarrow \qquad \begin{cases} E & \to E - qV \\ \boldsymbol{p} & \to \boldsymbol{p} - q\boldsymbol{A} \end{cases}$$

In the following we concentrate on the u_A component because in exercise 5.2 you have shown that in the non-relativistic limit the other component is small.

(a) Starting from the Dirac equation in momentum space Eq. (5.51),write down the equations for u_A and u_B after minimal substitution.

(b) In coordinate space $\boldsymbol{p}$ and E are operators. Therefore, they do not commute with $\boldsymbol{A}$ and V. That means that unlike in the case for free particles (exercise 5.2) you cannot just eliminate u_B! However, in the non-relativistic limit and assuming $|qV| \ll m$ we can use the approximation $E - qV + m \approx 2m$. Use this to eliminate u_B and obtain the Dirac equation for u_A

$$[\boldsymbol{\sigma} \cdot (\boldsymbol{p} - q\boldsymbol{A})]\,[\boldsymbol{\sigma} \cdot (\boldsymbol{p} - q\boldsymbol{A})]\, u_A = 2m(E_{\text{kin}} - qV)u_A \tag{5.100}$$

where $E_{\text{kin}} = E - m$.

(c) Use the Pauli vector identity, Eq. (5.99), to show that

$$[\boldsymbol{\sigma} \cdot (\boldsymbol{p} - q\boldsymbol{A})]\,[\boldsymbol{\sigma} \cdot (\boldsymbol{p} - q\boldsymbol{A})] \;=\; (\boldsymbol{p} - q\boldsymbol{A})^2 - iq\boldsymbol{\sigma} \cdot (\boldsymbol{p} \times \boldsymbol{A} + \boldsymbol{A} \times \boldsymbol{p}) \tag{5.101}$$

(d) Show that in coordinate space

$$\boldsymbol{\sigma} \cdot (\boldsymbol{p} \times \boldsymbol{A} + \boldsymbol{A} \times \boldsymbol{p})\, u_A = -i\hbar\boldsymbol{\sigma} \cdot (\boldsymbol{\nabla} \times \boldsymbol{A})\, u_A \tag{5.102}$$

where on the right hand side the derivative $\boldsymbol{\nabla}$ works only on $\boldsymbol{A}$ and not on u_A. Therefore, we can replace it with $\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}$.

(e) Using these results, show that the Dirac equation for an electron with charge $q = -e$ in the non-relativistic limit in an electromagnetic field $A^\mu = (A^0, \mathbf{A})$ reduces to the Schrödinger-Pauli equation

$$i\frac{\mathrm{d}}{\mathrm{d}t}\psi_A = \left(\frac{1}{2m}\,(\boldsymbol{p} + e\boldsymbol{A})^2 + \frac{e}{2m}\boldsymbol{\sigma} \cdot \boldsymbol{B} - eA^0\right)\psi_A\;, \tag{5.103}$$

where we have identified E_{kin} as the classical operator $i\mathrm{d}/\mathrm{d}t$.

The term with eA^0 in (5.103) *is a constant potential energy that is of no further importance. The term with $\boldsymbol{B}$ arises due to the fact that $\boldsymbol{p}$ and $\boldsymbol{A}$ do not commute. In this term we recognise the magnetic field:*

$$-\,\boldsymbol{\mu} \cdot \boldsymbol{B} = -g\frac{e}{2m}\boldsymbol{S} \cdot \boldsymbol{B}\;. \tag{5.104}$$

Here g is the gyromagnetic ratio*, i.e. the ratio between the magnetic moment of a particle and its spin. Classically we have $g = 1$, but according to the Dirac equation ($\boldsymbol{S} = \frac{1}{2}\boldsymbol{\sigma}$) one finds $g = 2$. The current value of $(g-2)/2$ is according to the Particle Data Book*

$$(g-2)/2 = 0.001159652193 \pm 0.000000000010 \tag{5.105}$$

This number, and its precision, make QED the most accurate theory in physics. The deviation from $g = 2$ is caused by high order corrections in perturbation theory.

Lecture 6

Spin-1/2 Electrodynamics

6.1 Feynman rules for fermion scattering

With the spinor solutions of the Dirac equation we finally have the tools to calculate cross-sections for fermions, spin-$\frac{1}{2}$ particles. In analogy to the case for spin-0 particles we determine the solutions of the Dirac equations in the presence of an electromagnetic field A^μ by starting from the free equation of motion and applying 'minimal substitution', $p^\mu \to p^\mu - qA^\mu$. For a particle with mass m and charge $q = -e$, the perturbed Dirac equation then becomes

$$(\gamma_\mu p^\mu - m)\,\psi + e\gamma_\mu A^\mu \psi = 0\ . \tag{6.1}$$

To isolate the perturbation term we write this again in terms of a Hamiltonian,

$$(H_0 \,+\, V)\,\psi \;=\; E\,\psi\ . \tag{6.2}$$

One can either start from the Dirac equation in terms of $\boldsymbol{\alpha}$ and β, or, work towards that form by multiplying the Dirac equation on the left by γ^0. The result is

$$E\psi = \underbrace{\left(\gamma^0\gamma^k p^k + \gamma^0 m\right)}_{H_0 = \boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta m} \psi - \underbrace{e\gamma^0\gamma_\mu A^\mu}_{V} \psi\ . \tag{6.3}$$

Consequently, the perturbation potential is

$$\boxed{\;V(x) \;=\; -e\gamma^0\gamma_\mu A^\mu\;}\ . \tag{6.4}$$

We define the transition amplitude by

$$T_{fi} = -i \int \psi_f^\dagger(x)\, V(x)\, \psi_i(x)\, \mathrm{d}^4 x \tag{6.5}$$

Compare this to the transition amplitude for spin-0 scattering, Eq. 4.7: The wave function now has four components and the perturbation potential $V(x)$ is a (4×4) matrix. We take a hermitian conjugate of the wave ψ, rather than its complex conjugate. The transition amplitude is still just a scalar.

Substituting the expression for $V(x)$ we obtain

$$\begin{aligned} T_{fi} &= -i \int \psi_f^\dagger(x) \, \left(-e\gamma^0\gamma_\mu A^\mu(x)\right) \, \psi_i(x) \mathrm{d}^4 x \\ &= -i \int \overline{\psi}_f(x) \, (-e) \, \gamma_\mu \psi_i(x) A^\mu(x) \, \mathrm{d}^4 x \end{aligned} \tag{6.6}$$

In Lecture 5 we defined the charge current density of the Dirac wave as

$$j^\mu(x) \;=\; -e \, \overline{\psi}(x) \, \gamma^\mu \, \psi(x)$$

In analogy to the spinless particle case we define the electromagnetic *transition current* between states i and $f$as

$$j^\mu_{fi}(x) \;=\; -e \, \overline{\psi}_f(x) \, \gamma^\mu \, \psi_i(x) \; , \tag{6.7}$$

such that the transition amplitude can be written as

$$T_{fi} \;=\; -i \int j^{fi}_\mu \, A^\mu \, \mathrm{d}^4 x \; . \tag{6.8}$$

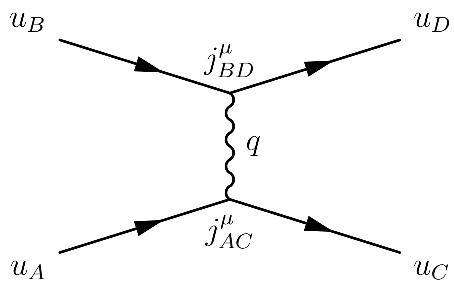
After inserting the plane wave decomposition $\psi(x) \;=\; u(p) e^{-ipx}$, the transition current becomes

$$j^\mu_{fi} \;=\; -e \overline{u}_f \, \gamma^\mu \, u_i \, e^{i\left(p_f - p_i\right)x} \; . \tag{6.9}$$

Note that the current is a 'scalar' in Dirac spinor space, or schematically,

$$j^\mu_{fi} \;=\; \begin{pmatrix} \quad \overline{u}_f \quad \end{pmatrix} \left(\quad \gamma^\mu \quad \right) \left(\; u_i \; \right) \tag{6.10}$$

Now consider again the two-body scattering $A + B \to C + D$:



Just as we did for the scattering of spinless particles, we obtain the vector potential A^μ by using the Maxwell equation with the transition current of one of the two particles (say 'particle AC') as a source. That is, we take

$$\Box A^\mu \;=\; j^\mu_{AC} \; .$$

and obtain for the potential

$$A^\mu = -\frac{1}{q^2} j^\mu_{AC} \; .$$

where $q \equiv p_i - p_f$ is the momentum transfer. The transition amplitude becomes

$$T_{fi} \;=\; -i \int j_\mu^{BD} \,\frac{-1}{q^2}\, j^\mu_{AC} \,\mathrm{d}^4x \;=\; -i \int j^\mu_{BD} \,\frac{-g_{\mu\nu}}{q^2}\, j^\nu_{AC} \,\mathrm{d}^4x \tag{6.11}$$

which is symmetric in terms of particle BD and AC. Inserting the expressions of the plane wave currents using Eq. (6.9) we obtain

$$T_{fi} = -i \int -e\, \overline{u}_C \gamma^\mu u_A \, e^{i(p_C - p_A)x} \;\cdot\; \frac{-g_{\mu\nu}}{q^2} \;\cdot\; -e\, \overline{u}_D \gamma^\nu u_B \, e^{i(p_D - p_B)x} \,\mathrm{d}^4x \tag{6.12}$$

Performing the integral (and realizing that nothing depends on x except the exponentials) leads us to the expression

$$T_{fi} \;=\; -i\,(2\pi)^4 \;\delta^4\left(p_D + p_C - p_B - p_A\right)\; \mathcal{M} \tag{6.13}$$

with the matrix element given by

$$-\,i\mathcal{M} \;=\; \underbrace{ie\,(\overline{u}_C\gamma^\mu u_A)}_{\text{vertex}} \;\; \underbrace{\frac{-ig_{\mu\nu}}{q^2}}_{\text{propagator}} \;\; \underbrace{ie\,(\overline{u}_D\gamma^\nu u_B)}_{\text{vertex}} \tag{6.14}$$

From the matrix element we can now read off the Feynman rules. Again, as for the spinless case, the various factors are defined such that the rules can also be applied to higher order diagrams.

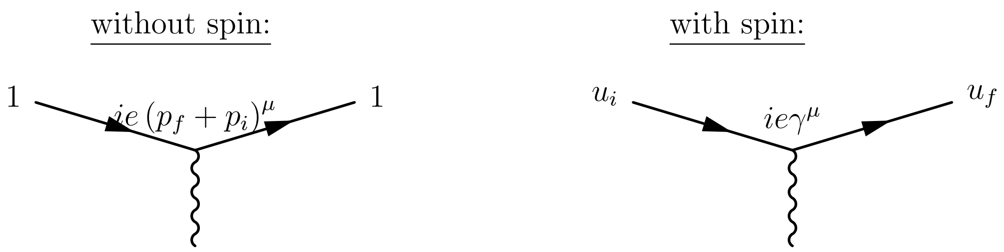


Figure 6.1: Diagrams for a spin-0 (left) and spin-$\frac{1}{2}$ (right) particle with charge e interacting with the electromagnetic field.

The rules for the vertex factors for spin-0 and spin-$\frac{1}{2}$ particles are shown side-by-side in Fig. 6.1. A spinless electron can interact with A^μ only via its charge. The coupling is proportional to $(p_f + p_i)^\mu$. However, an electron with spin can also interact with the magnetic field, via its magnetic moment. As you will prove in exercise 6.1, we can rewrite the Dirac current as

$$\overline{u}_f \gamma^\mu u_i = \frac{1}{2m} \overline{u}_f \left[(p_f + p_i)^\mu + i\sigma^{\mu\nu} \left(p_f - p_i\right)_\nu \right] u_i \tag{6.15}$$

where the tensor $\sigma^{\mu\nu}$ was defined in Eq. 5.46. This formulation of the current is called the 'Gordon decomposition'. We observe that in addition to the contribution that appears for the spinless wave, there is a new contribution that involves the factor $i\sigma^{\mu\nu}\,(p_f - p_i)$. In the non-relativistic limit this leads indeed to a term proportional to the magnetic field component of A^μ, just as you would expect from a magnetic moment.

6.2 Electron-muon scattering

We will now use the Feynman rules to calculate the cross-section of the process $e^-\mu^- \to e^-\mu^-$. The Feynman diagram is drawn in Fig. 6.2.

$-i\mathcal{M} =$

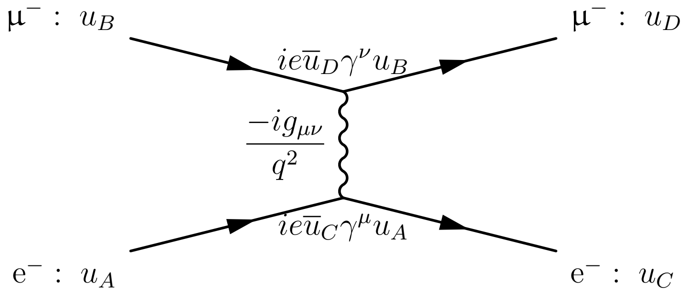


Figure 6.2: Lowest order Feynman diagram for $e^-\mu^-$ scattering.

Applying the Feynman rules we find for the lowest-order amplitude

$$- i\mathcal{M} \;=\; -e^2\,\overline{u}_C\gamma^\mu u_A\,\frac{-i}{q^2}\,\overline{u}_D\gamma_\mu u_B \tag{6.16}$$

and for its square

$$|\mathcal{M}|^2 \;=\; e^4\left[(\overline{u}_C\gamma^\mu u_A)\;\frac{1}{q^2}\;(\overline{u}_D\gamma_\mu u_B)\right]\left[(\overline{u}_C\gamma^\nu u_A)\;\frac{1}{q^2}\;(\overline{u}_D\gamma_\nu u_B)\right]^* \tag{6.17}$$

For a given value of μ and ν the currents are just complex numbers. (The γ-matrices are sandwiched between the bi-spinors.) Therefore, we can reorder them and write the amplitude as

$$|\mathcal{M}|^2 \;=\; \frac{e^4}{q^4}\sum_{\mu\nu}\left[(\overline{u}_C\gamma^\mu u_A)\;(\overline{u}_C\gamma^\nu u_A)^*\right]\left[(\overline{u}_D\gamma_\mu u_B)\;(\overline{u}_D\gamma_\nu u_B)^*\right] \tag{6.18}$$

We have factorized the right hand side into two tensors, each of which only depends on one of the leptons. We call these the *polarized lepton tensors*.

Up to now we have ignored the fact that the particle spinors come in two flavours, namely one for positive and one for negative helicity. Assuming that we do not measure the helicity (or spin) of the incoming and outgoing particles, the cross-section that we need to compute is a so-called 'unpolarized cross-section':

- If the incoming beams are unpolarized, we have no knowledge of initial spins. Therefore, we *average* over all spin configurations of the initial state;
- If the spin states of the outgoing particles are not measured, we should *sum* over spin configurations of the final state.

Performing the summation and averaging leads to the following 'unpolarized' matrix element

$$|\mathcal{M}|^2 \to \overline{|\mathcal{M}|^2} = \frac{1}{(2s_A+1)\,(2s_B+1)}\sum_{\text{spin}}|\mathcal{M}|^2 \tag{6.19}$$

where $2s_A + 1$ is the number of spin states of particle A and $2s_B + 1$ for particle B. So the product $(2s_A + 1)(2s_B + 1)$ is the number of spin states in the initial state.

Some of you may wonder why in the spin summation we add up the squares of the amplitudes, rather than square the total amplitude. The reason is that it does not make a difference since the final states over which we average are orthogonal: there is no interference between states with different helicity. It turns out that the math is easier when we sum over amplitudes squared.

Both the electron and the muon have $s = \frac{1}{2}$. Inserting the expression for the amplitude above, we can write the spin averaged amplitude as

$$\overline{|\mathcal{M}|^2} \;=\; \frac{1}{4}\,\frac{e^4}{q^4}\,L^{\mu\nu}_{\text{electron}}\,L^{\text{muon}}_{\mu\nu} \tag{6.20}$$

where the *unpolarized lepton tensors* are defined as

$$\begin{aligned} L^{\mu\nu}_{\text{electron}} &= \sum_{e-\text{spin}} \left[\overline{u}_C\gamma^\mu u_A\right]\,\left[\overline{u}_C\gamma^\nu u_A\right]^* \\ L^{\mu\nu}_{\text{muon}} &= \sum_{\mu-\text{spin}} \left[\overline{u}_D\gamma^\mu u_B\right]\,\left[\overline{u}_D\gamma^\nu u_B\right]^* \;. \end{aligned} \tag{6.21}$$

The spin summation is unfortunately rather tedious. The rest of the lecture is basically just the calculation to do this!

First, take a look at the complex conjugate of the transition current that appears in the tensor. Since it is just a (four-vector of) numbers, complex conjugation is the same as hermitian conjugation. Consequently, we have

$$\begin{aligned} \left[\overline{u}_C\gamma^\nu u_A\right]^* &= \left[\overline{u}_C\gamma^\nu u_A\right]^\dagger \\ &= \left[u_C^\dagger\gamma^0\gamma^\nu u_A\right]^\dagger = \left[u_A^\dagger\gamma^{\nu\dagger}\gamma^0 u_C\right] \\ &= \left[\overline{u}_A\gamma^0\gamma^{\nu\dagger}\gamma^0 u_C\right] = \left[\overline{u}_A\gamma^\nu u_C\right] \end{aligned} \tag{6.22}$$

In other words, by reversing the order of the spinors, we can get rid of the complex conjugation and find

$$L^{\mu\nu}_{\text{e}} = \sum_{e\text{ spin}} \left(\overline{u}_C\gamma^\mu u_A\right)\,\left(\overline{u}_A\gamma^\nu u_C\right) \tag{6.23}$$

Next, we apply what is called *Casimir's trick*. Write out the matrix multiplications in the tensors explicitly in terms of the components of the matrices and the incoming spins s and outgoing spins s',

$$L^{\mu\nu}_{\text{e}} = \sum_{s'}\sum_{s}\sum_{klmn} \overline{u}^{(s')}_{C,k}\,\gamma^\mu_{kl}\,u^{(s)}_{A,l}\,\overline{u}^{(s)}_{A,m}\,\gamma^\nu_{mn}\,u^{(s')}_{C,n} \tag{6.24}$$

All of the factors on the right are just complex numbers, so we can manipulate their order and write this as

$$L^{\mu\nu}_{\text{e}} = \sum_{klmn}\sum_{s'} u^{(s')}_{C,n}\,\overline{u}^{(s')}_{C,k}\,\gamma^\mu_{kl}\,\sum_{s} u^{(s)}_{A,l}\,\overline{u}^{(s)}_{A,m}\,\gamma^\nu_{mn} \tag{6.25}$$

Now remember the completeness relation, Eq. (5.77), that we derived in the previous lecture[1],

$$\sum_s u^{(s)}\,\overline{u}^{(s)} = \not{p} + m \tag{6.26}$$

Substituting this expression for the spin sums gives

$$L_{\rm e}^{\mu\nu} = \sum_{klmn} (\not{p}_C + m_{\rm e})_{nk}\ \gamma^{\mu}_{kl}\ (\not{p}_A + m_{\rm e})_{lm}\ \gamma^{\nu}_{mn} \tag{6.27}$$

where $m_{\rm e}$ is the electron mass.

Let's look more carefully at this expression: the right hand side contains products of components of (4×4) matrices. Call the product of these matrices 'A'. We could obtain the components of A by summing over the indices k, l and m. The final expression for the tensor would then be $L = \sum_n A_{nn}$, which is nothing else but the *trace* of A. Consequently, we can write the expression for the lepton tensor also as

$$L_{\rm e}^{\mu\nu} = {\rm Tr}\left[(\not{p}_C + m)\ \gamma^{\mu}\ (\not{p}_A + m)\ \gamma^{\nu}\right] \tag{6.28}$$

You now realize why we made you compute the traces of products of γ-matrices in exercise 5.5. We briefly repeat here the properties that we need:

- In general, for matrices A, B and C and any complex number z
  - ${\rm Tr}(zA) \;=\; z\,{\rm Tr}(A)$
  - ${\rm Tr}\,(A+B) \;=\; {\rm Tr}(A) + {\rm Tr}(B)$
  - ${\rm Tr}\,(ABC) \;=\; {\rm Tr}\,(CAB) \;=\; {\rm Tr}\,(BCA)$
- For γ-matrices (from the anti-commutator $\gamma^{\mu}\gamma^{\nu} + \gamma^{\nu}\gamma^{\mu} = 2g^{\mu\nu}$):
  - ${\rm Tr}(\text{odd number of }\gamma\text{-matrices} = 0)$
  - ${\rm Tr}\,(\gamma^{\mu}\gamma^{\nu}) = 4\,g^{\mu\nu}$
  - ${\rm Tr}(\gamma^{\alpha}\gamma^{\beta}\gamma^{\mu}\gamma^{\nu}) \;=\; 4\left(g^{\alpha\beta}g^{\mu\nu} - g^{\alpha\mu}g^{\beta\nu} + g^{\alpha\nu}g^{\beta\mu}\right)$

Using the first rule we can write out the tensor as a sum of traces,

$$\begin{aligned} L_{\rm e}^{\mu\nu} =& \ {\rm Tr}\left[(\not{p}_C + m)\ \gamma^{\mu}\ (\not{p}_A + m)\ \gamma^{\nu}\right] \\ =& \ \underbrace{{\rm Tr}\left[\not{p}_C\,\gamma^{\mu}\ \not{p}_A\,\gamma^{\nu}\right]}_{\text{case 1}} + \underbrace{{\rm Tr}\left[m\gamma^{\mu}m\gamma^{\nu}\right]}_{\text{case 2}} + \underbrace{{\rm Tr}\left[\not{p}_C\,\gamma^{\mu}m\gamma^{\nu}\right]}_{3\gamma' s\Rightarrow 0} + \underbrace{{\rm Tr}\left[m\gamma^{\mu}\ \not{p}_A\,\gamma^{\nu}\right]}_{3\gamma' s\Rightarrow 0} \end{aligned} \tag{6.29}$$

The last two terms vanish because they contain an odd number of γ-matrices. For the second term ('case 2') we find

$${\rm Tr}\left[m\gamma^{\mu}m\gamma^{\nu}\right] = m^2\,{\rm Tr}\left[\gamma^{\mu}\gamma^{\nu}\right] \;=\; 4\,m^2\,g^{\mu\nu}\ . \tag{6.30}$$

[1] for anti-fermions this gives an overall "$-$" sign in the tensor: $L_{\rm e}^{\mu\nu} \to -L_{\bar{\rm e}}^{\mu\nu}$ for each particle $\to$ anti-particle.

Finally, for the first term ('case 1') we have

$$
\begin{aligned}
\mathrm{Tr}\left[\not{p}_C\,\gamma^\mu\;\not{p}_A\,\gamma^\nu\right] \equiv\; & \mathrm{Tr}\left[\gamma^\alpha p_{C,\alpha}\gamma^\mu\gamma^\beta p_{A,\beta}\gamma^\nu\right] \\
=\; & Tr\left[\gamma^\alpha\gamma^\mu\gamma^\beta\gamma^\nu\right]\; p_{C,\alpha}\, p_{A,\beta} \\
=\; & 4\,\left(g^{\alpha\mu}g^{\beta\nu} - g^{\alpha\beta}g^{\mu\nu} + g^{\alpha\nu}g^{\beta\mu}\right)\, p_{C,\alpha}\, p_{A,\beta} \\
=\; & 4\,(p_C^\mu p_A^\nu + p_C^\nu p_A^\mu - g^{\mu\nu}(p_A\cdot p_C))\,,
\end{aligned}
\tag{6.31}
$$

where we used the trace formula for four γ-matrices in the third step. Adding the two contributions gives for the lepton tensor

$$
L_{\mathrm{e}}^{\mu\nu} \;=\; 4\,\left[p_C^\mu p_A^\nu \;+\; p_C^\nu p_A^\mu \;+\; \left(m_{\mathrm{e}}^2 - p_C\cdot p_A\right) g^{\mu\nu}\right] \tag{6.32}
$$

The expression for the muon tensor is obtained with the substitution $(p_A, p_C, m_{\mathrm{e}}) \to (p_B, p_D, m_{\mu})$,

$$
L_{\mu}^{\mu\nu} \;=\; 4\,\left[p_D^\mu p_B^\nu \;+\; p_D^\nu p_B^\mu \;+\; \left(m_{\mu}^2 - p_D\cdot p_B\right) g^{\mu\nu}\right] \tag{6.33}
$$

To compute the contraction of the two tensors, which appears in the amplitude, we just write everything out

$$
\begin{aligned}
L_{\mathrm{e}}^{\mu\nu}\, L^{\mu}_{\mu\nu} \;=\;& 4\left[p_C^\mu p_A^\nu + p_C^\nu p_A^\mu + \left(m_{\mathrm{e}}^2 - p_C\cdot p_A\right)\; g^{\mu\nu}\right]\cdot 4\;\left[p_{D\mu}p_{B\nu} + p_{D\nu}p_{B\mu} + \left(m_{\mu}^2 - p_D\cdot p_B\right)\; g_{\mu\nu}\right] \\
=\;& 16\left[(p_C\cdot p_D)\,(p_A\cdot p_B) + (p_C\cdot p_B)\,(p_A\cdot p_D) - (p_C\cdot p_A)\,(p_D\cdot p_B) + (p_C\cdot p_A)\, m_{\mu}^2\right. \\
& + (p_C\cdot p_B)\,(p_A\cdot p_D) + (p_C\cdot p_D)\,(p_A\cdot p_B) - (p_C\cdot p_A)\,(p_D\cdot p_B) + (p_C\cdot p_A)\, m_{\mu}^2 \\
& - (p_C\cdot p_A)\,(p_D\cdot p_B) - (p_C\cdot p_A)\,(p_D\cdot p_B) + 4\,(p_C\cdot p_A)\,(p_D\cdot p_B) - 4\,(p_C\cdot p_A)\, m_{\mu}^2 \\
& \left.+ m_e^2\;(p_D\cdot p_B) + m_e^2\;(p_D\cdot p_B) - 4m_e^2\;(p_D\cdot p_B) + 4m_e^2 m_{\mu}^2\right] \\
=\;& 32\left[(p_A\cdot p_B)\,(p_C\cdot p_D) + (p_A\cdot p_D)\,(p_C\cdot p_B) - m_e^2\,(p_D\cdot p_B) - m_{\mu}^2\,(p_A\cdot p_C) + 2m_e^2 m_{\mu}^2\right]
\end{aligned}
$$

Combining everything we obtain for the square of the unpolarized amplitude for electron-muon scattering

$$
\begin{aligned}
\overline{|\mathcal{M}|^2} \;=\; 8\,\frac{e^4}{q^4}\;\Big[\,&(p_C\cdot p_D)\,(p_A\cdot p_B) + \\
&(p_C\cdot p_B)\,(p_A\cdot p_D) - m_e^2\,(p_D\cdot p_B) - m_{\mu}^2\,(p_A\cdot p_C) + 2m_e^2 m_{\mu}^2\,\Big]
\end{aligned}
\tag{6.34}
$$

We now consider the ultra-relativistic limit and ignore the rest masses of the particles. The amplitude squared then becomes

$$
\overline{|\mathcal{M}|^2} \;\simeq\; 8\,\frac{e^4}{q^4}\;\Big[\,(p_C\cdot p_D)\,(p_A\cdot p_B) + (p_C\cdot p_B)\,(p_A\cdot p_D)\,\Big] \tag{6.35}
$$

Furthermore, we define the Mandelstam variables

$$
\begin{aligned}
s \;&\equiv\; (p_A + p_B)^2 = p_A^2 + p_B^2 + 2\,(p_A\cdot p_B) && \simeq 2\,(p_A\cdot p_B) \\
t \;&\equiv\; (p_D - p_B)^2 \equiv q^2 && \simeq -2\,(p_D\cdot p_B) \\
u \;&\equiv\; (p_A - p_D)^2 && \simeq -2\,(p_A\cdot p_D)
\end{aligned}
\tag{6.36}
$$

where the approximation on the right follows in the ultra-relativistic limit ($m \approx 0$). From energy-momentum conservation ($p_A^\mu + p_B^\mu = p_C^\mu + p_D^\mu$) we have

$$\begin{array}{rcl} (p_A + p_B)^2 &=& (p_C + p_D)^2 \\ (p_D - p_B)^2 &=& (p_C - p_A)^2 \\ (p_A - p_D)^2 &=& (p_B - p_C)^2 \end{array} \quad \Longrightarrow \quad \begin{array}{rcl} p_A \cdot p_B &=& p_C \cdot p_D \\ p_D \cdot p_B &=& p_C \cdot p_A \\ p_A \cdot p_D &=& p_B \cdot p_C \end{array} \tag{6.37}$$

which gives

$$(p_A \cdot p_B)\ (p_C \cdot p_D) = \frac{1}{2}s\,\frac{1}{2}s = \frac{1}{4}s^2 \tag{6.38}$$

$$(p_A \cdot p_D)\ (p_C \cdot p_B) = \left(-\frac{1}{2}u\right)\left(-\frac{1}{2}u\right) = \frac{1}{4}u^2 \tag{6.39}$$

$$\tag{6.40}$$

Inserting this in the amplitude, we find

$$\boxed{\overline{|\mathcal{M}|^2} \simeq 2\,e^4 \left(\frac{s^2+u^2}{t^2}\right)} \tag{6.41}$$

Finally, as we did for the spinless scattering in Lecture 4, consider again the scattering process in the centre-of-momentum system. The four-vectors can then be written as

$$p_A^\mu = (|\boldsymbol{p}_A|, \boldsymbol{p}_A) \qquad p_B^\mu = (|\boldsymbol{p}_A|, -\boldsymbol{p}_A)$$
$$p_C^\mu = (|\boldsymbol{p}_C|, \boldsymbol{p}_C) \qquad p_D^\mu = (|\boldsymbol{p}_C|, -\boldsymbol{p}_C)$$

Define $p \equiv |\boldsymbol{p}_A|$ which, by four-vector conservation is also equal to $|\boldsymbol{p}_{B,C,D}|$. Define θ as the angle between $\boldsymbol{p}_A$ and $\boldsymbol{p}_C$ (see Fig. 6.3), such that

$$\boldsymbol{p}_A \cdot \boldsymbol{p}_C = \boldsymbol{p}_B \cdot \boldsymbol{p}_D = p^2 \cos\theta$$

We then find for the Mandelstam variables

$$\begin{array}{rl} s = & 4\,p^2 \\ t = & -2\,p^2\ (1-\cos\theta) \\ u = & -2\,p^2\ (1+\cos\theta) \end{array} \tag{6.42}$$

which gives for the amplitude squared

$$\overline{|\mathcal{M}|^2} \simeq 8\,e^4\, \frac{4 + (1+\cos\theta)^2}{(1-\cos\theta)^2}. \tag{6.43}$$

Inserting this in the expression for the differential cross-section (which we obtained after integrating over the final state momenta in exercise 2.4) we find

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right|_{\mathrm{c.m.}} = \frac{1}{64\pi^2}\,\frac{1}{s}\,\overline{|\mathcal{M}|^2} \simeq \frac{\alpha^2}{2s}\,\frac{4+(1+\cos\theta)^2}{(1-\cos\theta)^2} \tag{6.44}$$

with $\alpha \equiv e^2/4\pi$.

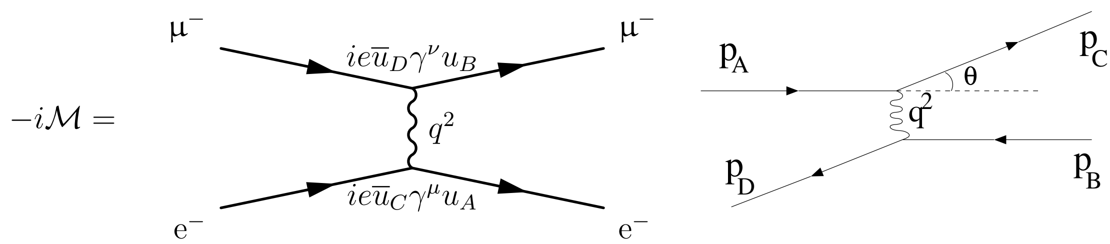


Figure 6.3: $e^-\mu^- \to e^-\mu^-$ scattering. Left: the Feynman diagram. Right: definition of scattering angle in C.M. frame.

6.3 Crossing: the process $e^-e^+ \to \mu^-\mu^+$

We now introduce the method of "crossing" by computing the amplitude for $e^-e^+ \to \mu^-\mu^+$ scattering from the amplitude of $e^-\mu^- \to e^-\mu^-$ scattering. The rules are the following:

1. assuming that you had computed the original amplitude in terms of particles, replace $p \to -p$ for every anti-particle in the diagram

2. now relabel the momenta such that the ingoing and outgoing lines correspond to those in the original diagram

3. for every crossed fermion line, *i.e.* for every outgoing fermion that became incoming or vice-versa, multiply the amplitude squared by a factor (-1). (This has to do with the sign of the current which we discussed in section 5.12.)

The procedure is illustrated in Fig. 6.4. Labeling the momenta of the 'target' process with primes, we use crossing to derive for the amplitude

$$\mathcal{M}\left[e^-(p'_A)e^+(p'_B) \to \mu^+(p'_C)\mu^-(p'_D)\right] \;=\; \mathcal{M}\left[e^-(p'_A)\mu^-(-p'_C) \to e^-(-p'_B)\mu^-(p'_D)\right] \tag{6.45}$$

In other words, we can use the original computation of the amplitude provide that we relabel the momenta as follows:

$$p_A = p'_A \qquad p_B = -p'_C \qquad p_C = -p'_B \qquad p_D = p'_D$$

Consequently, the Mandelstam variables of the 'original' particle diagram are

$$\begin{array}{rclcccl} s & \equiv & (p_A + p_B)^2 & = & (p'_A - p'_C)^2 & \equiv & t' \\ t & \equiv & (p_D - p_B)^2 & = & (p'_C + p'_D)^2 & = & s' \\ u & \equiv & (p_A - p_D)^2 & = & (p'_A - p'_D)^2 & = & u' \end{array} \tag{6.46}$$

Using the result in Eq. (6.41) the amplitude squared for the two processes are then

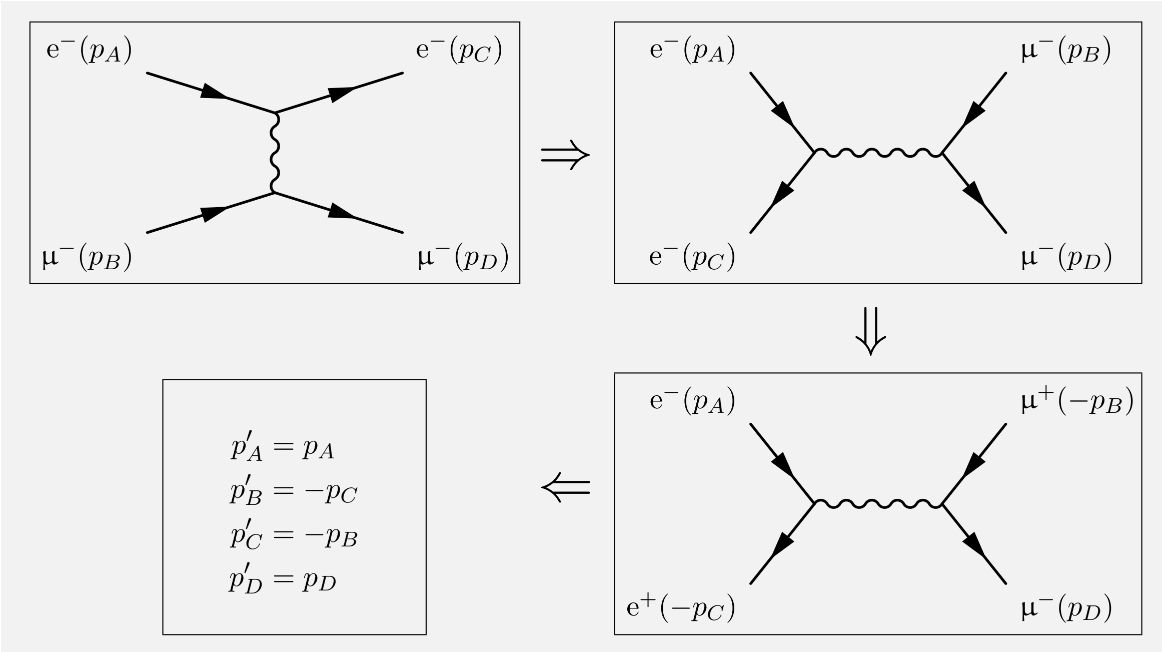
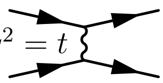


Figure 6.4: Illustration of crossing. Use the anti-particle interpretation of a particle with the 4-momentum reversed in order to related the Matrix element of the “crossed” reaction to the original one.

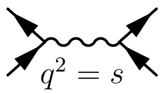
$$\overline{|\mathcal{M}|}^2_{\mathrm{e}^-\mu^-\to\mathrm{e}^-\mu^-} = 2\,e^4\,\frac{s^2+u^2}{t^2}$$

”t-channel”: $q^2 = t$



$$\overline{|\mathcal{M}|}^2_{\mathrm{e}^-\mathrm{e}^+\to\mu^-\mu^+} = 2\,e^4\,\frac{u'^2+t'^2}{s'^2}$$

”s-channel”:



It is customary to label these as the *t-channel* and the *s-channel* process, because we have $q^2 = t$ and $q^2 = s$, respectively.

We can express the momenta in the centre-of-momentum frame in terms of an initial momentum p and a scattering angle θ, where θ is now the angle between the incoming e^- (p'_A) and the outgoing μ^- (p'_C). The expressions for u', s' and t' are identical to those in (6.42).

We immediately get for the matrix element:

$$\overline{|\mathcal{M}|}^2_{\mathrm{c.m.}} = 2\,e^4\,\frac{t'^2+u'^2}{s'^2} = e^4\,\left(1+\cos^2\theta\right) \tag{6.47}$$

The differential cross-section becomes

$$\boxed{\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\alpha^2}{4s}\,\left(1+\cos^2\theta\right)} \tag{6.48}$$

Finally, to calculate the total cross-section for the process we integrate over the azimuthal angle ϕ and the polar angle θ:

$$\boxed{\sigma\left(\mathrm{e}^+\mathrm{e}^- \rightarrow \mu^+\mu^-\right) = \frac{4\pi}{3}\,\frac{\alpha^2}{s}} \tag{6.49}$$

Note that the 'shape' of the angular distribution does not depend on the available energy, but that the total cross-section scales as $1/s$: the higher the cms energy, the smaller the cross-section. If you look back to our original formulation of the golden rule, you'll find that the $1/s$ dependence comes from the density of the incoming waves. The faster the relative velocity, the shorter the particles are in each others vicinity!

Figure 6.5 shows a comparison of the kinematic factors in the differential cross-section of the t-channel process $\mathrm{e}^-\mu^- \rightarrow \mathrm{e}^-\mu^-$ and the s-channel process $\mathrm{e}^-\mathrm{e}^+ \rightarrow \mu^-\mu^+$ for spin-0 and spin-$\frac{1}{2}$ leptons. For the t-channel process the difference is only visible in the very backward region, while in the s-channel process there is a constant offset.

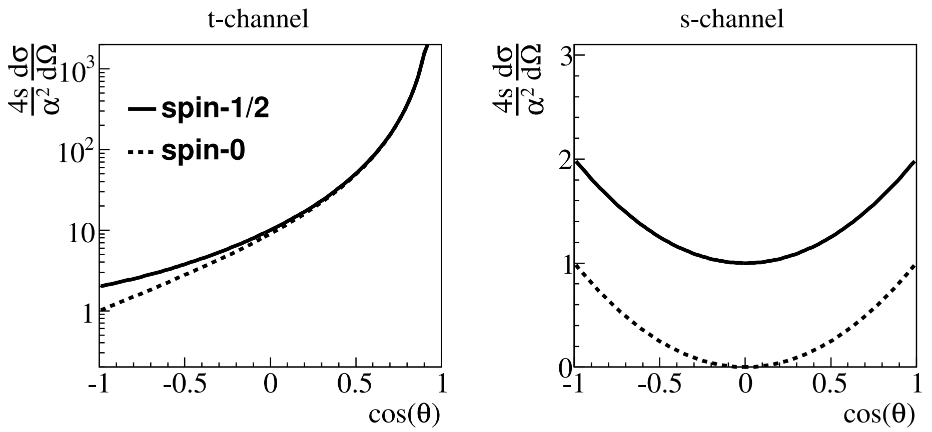


Figure 6.5: Leading order QED differential cross-section $\mathrm{d}\sigma/\mathrm{d}\Omega$ divided by $\alpha^2/4s$ as function of $\cos\theta$ for the t-channel process $\mathrm{e}^-\mu^- \rightarrow \mathrm{e}^-\mu^-$ (left) and the s-channel process $\mathrm{e}^-\mathrm{e}^+ \rightarrow \mu^-\mu^+$ (right) in the ultra-relativistic limit ($m_e = m_\mu = 0$).

Figure 6.6 shows a table copied from Halzen and Martin with the kinematic factors for important leading order QED processes. These processes are related by crossing. The interference terms follow via crossing procedure as well, provided that you add up amplitudes (not amplitudes squared).

6.4 Summary of QED Feynman rules

In the computation of the $\mathrm{e}^-\mu^-$ above we have seen only a subset of the Feynman rules for QED. As an example of things we missed, consider the annihilation process $\mathrm{e}^+\mathrm{e}^- \rightarrow \gamma\gamma$. (Draw it!) To compute the cross-section for this process we need more Feynman rules, namely those for the electron propagator and those for external photon lines. We now briefly summarize the rules for QED. You can find these in more detail in

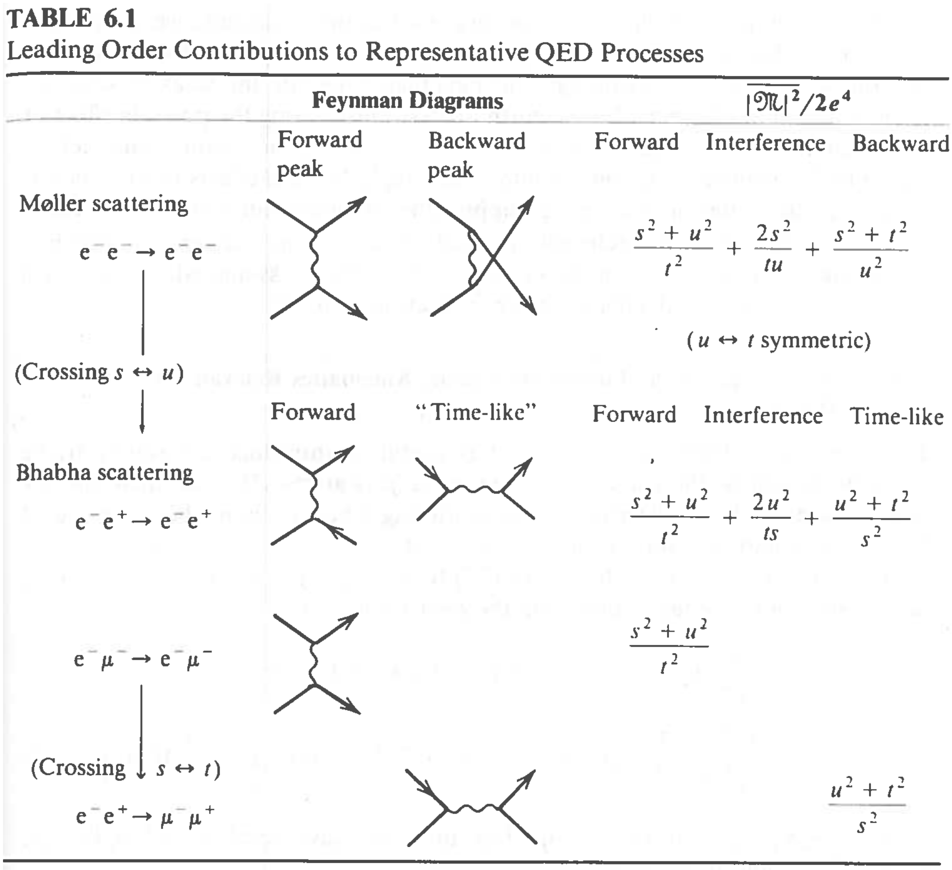


TABLE 6.1
Leading Order Contributions to Representative QED Processes

	Feynman Diagrams		$\overline{\lvert\mathcal{M}\rvert^2}/2e^4$		
	Forward peak	Backward peak	Forward	Interference	Backward
Møller scattering $e^-e^- \to e^-e^-$			$\frac{s^2+u^2}{t^2}$ +	$\frac{2s^2}{tu}$ +	$\frac{s^2+t^2}{u^2}$
				($u \leftrightarrow t$ symmetric)	
(Crossing $s \leftrightarrow u$)	Forward	"Time-like"	Forward	Interference	Time-like
Bhabha scattering $e^-e^+ \to e^-e^+$			$\frac{s^2+u^2}{t^2}$ +	$\frac{2u^2}{ts}$ +	$\frac{u^2+t^2}{s^2}$
$e^-\mu^- \to e^-\mu^-$			$\frac{s^2+u^2}{t^2}$		
(Crossing $s \leftrightarrow t$) $e^-e^+ \to \mu^-\mu^+$					$\frac{u^2+t^2}{s^2}$

Figure 6.6: Leading order QED processes and their relations via crossing. From Halzen and Martin, "Quarks and Leptons".

the textbooks, *e.g.* in appendix D of Griffiths and on the inside of the cover of Halzen and Martin, or Thomson.

For the external lines, we have in the matrix element:

$$
\begin{array}{ll}
\text{spin-0:} & \text{nothing} \\
\text{spin-}\frac{1}{2}\text{:} & \left\{ \begin{array}{ll} \text{incoming particle:} & u \\ \text{outgoing particle:} & \overline{u} \\ \text{incoming anti-particle:} & \overline{v} \\ \text{outgoing anti-particle:} & v \end{array} \right. \\
\text{spin-1:} & \left\{ \begin{array}{ll} \text{incoming:} & \epsilon_\mu \\ \text{outgoing:} & \epsilon^*_\mu \end{array} \right.
\end{array}
\tag{6.50}
$$

We have seen the photon polarization vectors in Lecture 3. Both the spin-$\frac{1}{2}$ and spin-1 external lines carry also an index for the helicity. In calculations for cross-sections or decays in which we measure the spin, we need explicit forms of the Dirac spinors and the photon polarization vectors. However, often we sum over all incoming and outgoing spins ('spin averaging') and we can use the completeness relations.

For the internal lines (the propagators) we have

$$
\begin{aligned}
&\text{spin-0:} \quad \frac{i}{q^2 - m^2} \\
&\text{spin-}\tfrac{1}{2}\text{:} \quad \frac{i(\not{q} + m)}{q^2 - m^2} \\
&\text{spin-1:} \quad \begin{cases} \text{massless:} & \dfrac{-ig^{\mu\nu}}{q^2} \\ \text{masssive:} & \dfrac{i\left[-g^{\mu\nu} + q^\mu q^\nu / m^2\right]}{q^2 - m^2} \end{cases}
\end{aligned}
\tag{6.51}
$$

Finally, the QED vertex factors are

$$
\begin{aligned}
&\text{spin-0:} \quad ig_e(p^\mu_{\text{in}} + p^\mu_{\text{out}}) \\
&\text{spin-}\tfrac{1}{2}\text{:} \quad ig_e\gamma^\mu
\end{aligned}
\tag{6.52}
$$

with g_e the charge of the particle in the vertex. Section 7.6 of Griffiths contains worked out examples of several key QED processes, both with and without spin averaging.

Exercises

Exercise 6.1 (The Gordon decomposition)
A spinless electron can interact with A^μ only via its charge; the coupling is proportional to $(p_f + p_i)^\mu$. An electron with spin, on the other hand, can also interact with the magnetic field via its magnetic moment. This coupling involves the factor $i\sigma^{\mu\nu}\,(p_f - p_i)$. The relation between the Dirac current and the Klein-Gordon current can be studied as follows:

(a) Show that the Dirac current can be written as

$$\overline{u}_f \gamma^\mu u_i = \frac{1}{2m}\overline{u}_f \left[(p_f + p_i)^\mu + i\sigma^{\mu\nu}\,(p_f - p_i)_\nu\right] u_i$$

where the tensor is defined as

$$\sigma^{\mu\nu} = \frac{i}{2}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right)$$

Hint: Start with the term proportional to $\sigma^{\mu\nu}$ and use: $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$ and use the Dirac equations: $\gamma^\nu p_{i\nu} u_i = m u_i$ and $\overline{u}_f \gamma^\nu p_{f\nu} = m\overline{u}_f$.

(b) (**optional!**) Make exercise 6.2 on page 119 of H&M which shows that the Gordon decomposition in the non-relativistic limit leads to an electric and a magnetic interaction. (Compare also exercise 5.8.)

Exercise 6.2
Can you easily obtain the cross-section of the process $e^+e^- \to e^+e^-$ from the result of $e^+e^- \to \mu^+\mu^-$? If **yes**: give the result, if **no**: why not?

Exercise 6.3 (The process $e^+e^- \to \pi^+\pi^-$)
We consider scattering of spin 1/2 electrons with spin-0 pions. We assume point-particles; i.e. we forget that the pions have a substructure consisting of quarks. Also we only consider electromagnetic interaction and we assume that the particle masses can be neglected.

(a) Consider the process of electron - pion scattering: $e^-\pi^- \to e^-\pi^-$. Draw the Feynman diagram and write down the expression for the $-i\mathcal{M}$ using the Feynman rules.

(b) Perform the spin averaging of the electron and compute $\overline{|\mathcal{M}|}^2$.

(c) Use the principle of crossing to find $\overline{|\mathcal{M}|}^2$ for $e^+e^- \to \pi^+\pi^-$
(Note the extra minus sign that appears from the 3rd crossing rule.)

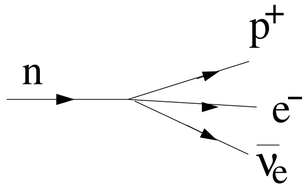
(d) Determine the differential cross-section $d\sigma/d\Omega$ for $e^+e^- \to \pi^+\pi^-$ in the centre-of-momentum of the e^+e^--system.

Lecture 7

The Weak Interaction

In 1896 Henri Becquerel studied the effect of fluorescence, which he thought was related to X-rays that had been discovered by Wilhelm Röntgen. To test his hypothesis he wrapped a photographic plate in black paper and placed various phosphorescent salts on it. All results were negative until he used uranium salts. These affected photographic plates even when put in the dark, such that the effects clearly had nothing to do with fluorescence. Henri Becquerel had discovered natural radioactivity, and thereby the *weak interaction.*

We know now that the most nuclear decays are the result of the transition of a neutron to an electron, a proton and an anti-neutrino,



or in a formula,

$$n \to p + e^- + \overline{\nu}_e \ . \tag{7.1}$$

A 'free' neutron has a lifetime of about 15 minutes, but the lifetime of various weakly decaying isotopes spans a very wide range.

7.1 Lifetimes and couplings

Compare the lifetime of the following particles:

particle	lifetime [sec]	dominant decay mode
ρ^0	$4.4 \cdot 10^{-23}$	$\rho^0 \to \pi^+\pi^-$
π^0	$8.4 \cdot 10^{-17}$	$\pi^0 \to \gamma\gamma$
π^-	$2.6 \cdot 10^{-8}$	$\pi^- \to \mu^-\overline{\nu}_\mu$
μ^-	$2.2 \cdot 10^{-6}$	$\mu^- \to e^-\overline{\nu}_e\nu_\mu$

The listed decay modes are the *dominant* decay modes. Other decay modes exist, but they contribute marginally to the total decay width. As we have seen before, the lifetime of a particle is inversely proportional to the total decay width,

$$\tau = \frac{\hbar}{\Gamma} . \tag{7.2}$$

We have also seen that the decay width to a particular final state is proportional to the matrix element squared. For example, for the two-body decay $A \to B + C$ we had (in particle A's rest frame)

$$\Gamma(A \to B + C) = \int \frac{|\mathcal{M}|^2}{2E_A} \mathrm{d}\Phi = \frac{p_B}{8\pi m_A^2} |\mathcal{M}|^2 . \tag{7.3}$$

In Yukawa's picture of scattering by particle exchange, the leading order contribution to the matrix element is proportional to the square of the coupling constant. Consequently, the lifetime of particles tells us something about the strength of the interaction that is responsible for the decay.

All fundamental fermions in the standard model 'feel' the weak interaction. However, in processes that can also occur via the strong or electromagnetic interaction, those interactions will dominate. The reason that we still see the effects of the weak interaction is because the strong and electromagnetic interaction do not change quark and lepton flavour. Consequently, if a particle cannot decay to a lighter state obeying the 'flavour-conservation' rule, then it can only decay through the weak interaction. In contrast to quarks and charged leptons, neutrinos feel only the weak interaction. That is the reason why they are so hard to detect!

We can now understand the hierarchy of the lifetimes above as follows:

- The ρ^0 particle (which is an excited meson consisting of u and d quarks and their anti-quarks) decays via the strong interaction to two pions.
- The π^0 is the lightest neutral hadron such that it cannot decay to hadrons. It decays via the electromagnetic interaction to two photons, as we have seen in exercise 4.3.
- The π^+ is the lightest charged hadron. Because it is charged, it cannot decay two photons. Instead, it decays via the weak interaction to a μ^+ and a neutrino. (It could also decay to an e^+ and a neutrino, but for reasons explained later that mode is kinematically suppressed, despite the larger 'phase space'.)
- The μ^+ is a lepton and therefore does not couple to the strong interaction. It cannot decay to an electron and photon, as the electromagnetic interaction conserves lepton flavour. Its dominant decay is via the weak interaction to an electron and neutrinos.

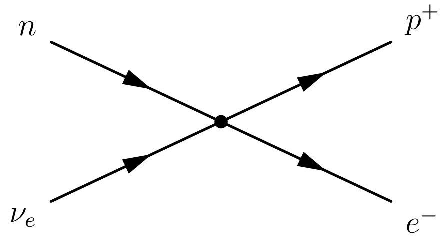
Considerations like these explain the gross features in the hierarchy of lifetimes. However, as you can also judge from the wide range in lifetimes of particles that decay weakly, kinematic effects must be important as well.

Besides the fact proper that the weak interaction unlike the electromagnetic and strong interaction does not 'honour' the quantum numbers for quark and lepton flavour, the weak interaction is special in at least two more ways:

- it violates parity symmetry P. Until 1956, when the parity violating aspects of the weak interaction were demonstrated, physicists were convinced that at least at the level of fundamental interactions our world was left-right symmetric;
- in the quark sector, it even violates CP symmetry. That means, because of CPT invariance, that it also violates T (time-reversal) symmetry. As we shall see, the existence of a third quark family was predicted from the observation that neutral Kaon decays exhibit CP violation.

7.2 The 4-point interaction

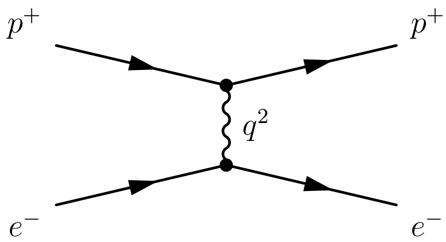
In 1932 Fermi tried to formulate a theory to describe nuclear decays quantitatively. He proposed a so-called *4-point interaction*, introducing the Fermi constant as the strength of the interaction: $G_F \approx 1.166 \cdot 10^{-5}\text{GeV}^{-2}$.



The "Feynman diagram" of the 4-point interaction "neutrino scattering on a neutron" has the following matrix element:

$$\boxed{\mathcal{M} = G_F \; (\overline{u}_p \gamma^\mu u_n) \; (\overline{u}_e \gamma_\mu u_\nu)} \tag{7.4}$$

This is to be compared to the electromagnetic diagram for electron proton scattering:



Here the matrix element was:

$$\mathcal{M} = \frac{4\pi\alpha}{q^2} \; (\overline{u}_p \gamma^\mu u_p) \; (\overline{u}_e \gamma_\mu u_e) \tag{7.5}$$

1. $e^2 = 4\pi\alpha$ is replaced by G_F
2. $1/q^2$ is removed

We take note of the following properties of the weak interaction diagram:

1. The matrix element involves a hadronic current and a leptonic current. In contrast to electromagnetic scattering, these currents change the charge of the particles involved. In this particular process we have $\Delta Q = 1$ for the hadronic current and $\Delta Q = -1$ for the leptonic curent. Since there is a net charge from the hadron to the lepton current we refer to this process as a *charged current* interaction. We will see later that there also exists a neutral current weak interaction.
2. There is no propagator; ie. a "4-point interaction".

3. There is a coupling constant G_F, which plays a similar role as α in QED. Since there is no propagator, the coupling constant is not dimensionless.
4. The currents have what is called a "vector character" similar as in QED. This means that the currents are of the form $\overline{\psi}\gamma^\mu\psi$.

The vector character of the interaction was just a guess that turned out successful to describe many aspects of β-decay. There was no reason for this choice apart from similarity with quantum electrodynamics. In QED the reason that the interaction has a vector behaviour is because the force mediator, the photon, is a spin-1 (or 'vector') particle.

In the most general case the matrix element of the 4-point interaction can be written as

$$\mathcal{M} = G_F \,\left(\overline{\psi}_p \;(4\times 4)\; \psi_n\right)\, \left(\overline{\psi}_e \;(4\times 4)\; \psi_\nu\right)\ , \tag{7.6}$$

where the (4×4) is a matrix. Lorentz invariance puts restrictions on the form of these matrices. We have seen these already in lecture 5: Any such matrix needs to be a so-called bilinear covariant. The bilinear covariants all involve 4×4 matrices that are products of γ matrices:

	current	# components	# γ-matrices	spin
Scalar	$\overline{\psi}\psi$	1	0	0
Vector	$\overline{\psi}\gamma^\mu\psi$	4	1	1
Tensor	$\overline{\psi}\sigma^{\mu\nu}\psi$	6	2	2
Axial vector	$\overline{\psi}\gamma^\mu\gamma^5\psi$	4	3	1
Pseudo scalar	$\overline{\psi}\gamma^5\psi$	1	4	0

The last column in the table shows the spin of the mediator in the interaction, if the 4-point interaction is interpreted as the result of a force-carrier exchange.

The most general 4-point interaction now takes the following form:

$$\mathcal{M} = G_F \sum_{i,j}^{S,P,V,A,T} C_{ij}\left(\overline{u}_p\, O_i\, u_n\right)\ \left(\overline{u}_e\, O_j\, u_\nu\right) \tag{7.7}$$

where O_i, O_j are operators of the form S, V, T, A, P. The kinematics of a decay depend on the type of operator involved. For example, it can be shown (see eg. Perkins: "Introduction to High Energy Physics", 3rd edition, appendix D) that for the decay $n \to p e^- \overline{\nu}_e$

- S, P and T interactions imply that the helicity of the e^- and the $\overline{\nu}_e$ have the same sign;
- V and A interactions imply that they have opposite sign.

Fermi had assumed that the weak interaction was of the V type. In a number of experiments performed in the late fifties it was established that the weak interaction was a combination of V and A. Before we look at that in more detail, we need to discuss the concept of *parity*.

7.3 Parity

Parity, or (space) inversion, is the operation that multiplies all spatial coordinates by -1, so $\boldsymbol{x} \to -\boldsymbol{x}$. It is closely related to reflection in a mirror: the parity operation is identical to a reflection in a plane through the origin, followed by a rotation under 180 degrees around an axis through the origin perpendicular to the mirror. Therefore, for systems that are rotation and translation invariant, the two are equivalent. When illustrating parity violation in pictures, we usually use an image with a reflection in a mirror. Yet, when formulating the effect of parity in a physics theory, we work with space inversion.

Now, consider a process $\phi_i \to \phi_f$ for some initial state i and final state f. The relation between i and f given by an operator that describes the time evolution,

$$\phi_f \;=\; \hat{U}_{fi}\,\phi_i \tag{7.8}$$

(We can look at the process at any time scale. So $\hat{U}$ can just be a continuous function of time.) Denoting the parity operation by $\hat{P}$ we can also consider the mirror process, characterised by $\phi'_i = \hat{P}\phi_i$ and $\phi'_f = \hat{P}\phi_f$. We define the process to be 'symmetric under parity' when it does not make any difference whether we first transform ϕ_i to its mirror image and then look at its time evolution, $\phi_i \to \phi'_i \to \phi'_f$, or first wait for the system to evolve and then reflect it, $\phi_i \to \phi_f \to \phi'_f$. Or, in terms more common in quantum mechanics, the process is symmetric under parity when $\hat{P}$ and $\hat{U}$ commute,

$$[\hat{U}, \hat{P}] \;=\; 0 \tag{7.9}$$

Because for small times t we have $\hat{U}(t) = e^{-iHt/\hbar} \approx 1 - iHt/\hbar$, it follows that such $\hat{P}$ also commutes with the Hamiltonian. This definition of a symmetry is not limited to mirror symmetry, but holds for any operator: if an operator $\hat{Q}$ commutes with H then it is called a symmetry operator.

If $\hat{P}$ and H commute, then they have a common set of eigenvectors. If we consider eigenvectors with energy E that are *not degenerate* (that is, there is no other state with equal energy) then this immediately implies that these states have definite parity: they are eigenstates of the parity operator and there is an observable property (a quantum number) associated with the parity operation.

If we apply the parity operator twice, then we put the system back in its original state. Consequently, if p is the eigenvalue of our state under the parity operator, then $p^2 = 1$. (Strictly speaking, the system would be in the same state even if we had changed the wave function by an arbitrary phase. However, for simplicity we will not deal with the minor complications that this introduces.) Therefore, the eigenvalue is either $+1$ or -1. We call such states states of *even* and *odd* parity respectively.

Until 1956 all the known laws of physics were invariant under inversion symmetry. At the scale of elementary particles our world was perfectly left-right symmetric. This symmetry was well tested for the electromagnetic and strong interaction and it was generally assumed that it held for the weak interaction as well.

Since all *all* our leptons, mesons, and baryons are characterised by different masses (*e.g.* by different eigenvalues of the total Hamiltonian) they all have definite parity: they are either odd or even under the parity operation. (You will find their quantum numbers for parity in the PDG.) These facts, definite parity for all (stable and meta-stable) particles, and parity conservation in known interactions, is exactly what lead in the early fifties to what was called the 'theta-tau puzzle'.

The θ and the τ were charged particles with strangeness one that decayed through the weak interaction to two and three pions respectively,

$$\begin{aligned} \theta^+ &\to \pi^+ + \pi^0 \\ \tau^+ &\to \pi^+ + 2\pi^0 \quad \text{or} \quad 2\pi^+ + \pi^- \end{aligned} \tag{7.10}$$

The pions were all known to have parity -1. Then, assuming parity to be conserved in these processes the theta had even parity and the tau odd parity. However, what was truly strange is that the theta and tau were otherwise seemingly identical particles: they had the same mass and same lifetime.

After verifying that there had never been any experimental tests of parity conservation in the weak interaction, Lee and Yang hypothesized in 1956, that the tau and theta were actually the same particle, and that the weak interaction was responsible for the apparent violation of parity. They also proposed a number of experiments that could establish parity violation in weak decays directly. Within half a year two of these experiments were performed (Wu *et al.* (1957), Garmin, Lederman and Weinrich (1957)) and the parity violating character of the weak interaction was firmly established.

7.4 Covariance of the wave equations under parity

Let us now consider the parity transformation of the solutions to the wave equations in more detail. In our notation the parity operation transforms

$$\begin{aligned} \boldsymbol{x} &\to \boldsymbol{x}' = -\boldsymbol{x} \\ t &\to t' = t \\ \phi(\boldsymbol{x}) &\to \phi'(\boldsymbol{x}') \end{aligned} \tag{7.11}$$

We would now like to establish the relation between $\phi'(\boldsymbol{x}')$ and $\phi(\boldsymbol{x})$. For the wave equation to be 'covariant' (that is, the same form in every frame) this relation is to be chosen such that if $\phi(\boldsymbol{x})$ satisfies the wave equation, then $\phi'(\boldsymbol{x}')$ satisfies the *same* wave equation.

Now remember the Klein-Gordon equation for a free particle:

$$(\partial_\mu \partial^\mu + m^2)\, \phi(x) \;=\; 0 \tag{7.12}$$

Because this equation is quadratic in ∂_μ, the equation itself does not change under parity. Hence, one possible solution is simply

$$\phi'_{\mathrm{KG}}(\boldsymbol{x}') = \phi_{\mathrm{KG}}(\boldsymbol{x}) = \phi_{\mathrm{KG}}(-\boldsymbol{x}') \tag{7.13}$$

Any global change in the phase for ϕ' is allowed as well: such a phase change can be considered part of our definition of parity and we can choose it to be zero. If we take ϕ to be a positive energy plane-wave solution with momentum $\boldsymbol{p}$, then ϕ' is also a positive energy solution but with momentum $-\boldsymbol{p}$. (Try!)

Now consider the more complicated case of the Dirac equation. Split in its time and space part, the Dirac equation in coordinate space can be written as

$$i\frac{\partial\psi(\boldsymbol{x},t)}{\partial t} = (-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m)\,\psi(\boldsymbol{x},t) \tag{7.14}$$

In order for this equation to be covariant under parity transformations we must find the field $\psi'(\boldsymbol{x'},t)$ such that it satisfies the transformed Dirac equation,

$$i\frac{\partial\psi'(\boldsymbol{x}',t)}{\partial t} = (-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}' + \beta m)\,\psi'(\boldsymbol{x}',t) \tag{7.15}$$

Applying our definition of space inversion, which implies $\boldsymbol{\nabla}' = -\boldsymbol{\nabla}$, we find

$$i\frac{\partial\psi'(-\boldsymbol{x},t)}{\partial t} = (i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m)\,\psi'(-\boldsymbol{x},t) \tag{7.16}$$

Note that $\psi'(-\boldsymbol{x},t)$ does not satisfy the Dirac equation due to the additional minus sign in front of the spatial derivative. However, we now multiply the equation on the left by β and use that fact that $\boldsymbol{\alpha}$ and β ant-commute. The result is

$$i\frac{\partial\beta\psi'(-\boldsymbol{x},t)}{\partial t} = (-i\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta m)\,\beta\psi'(-\boldsymbol{x},t) \tag{7.17}$$

Hence $\beta\psi'(-\boldsymbol{x},t)$ does satisfy the Dirac equation. Consequently, one choice for the parity operation is

$$\psi'(\boldsymbol{x}',t) = \beta\psi(-\boldsymbol{x}',t) \tag{7.18}$$

Again, we could insert any constant phase factor in the transformation. By convention, we choose that factor to be one.

We now look at the solutions to the Dirac equation in the Pauli-Dirac representation. In this representation, we have for the $\beta = \gamma^0$ matrix:

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \tag{7.19}$$

Consequently, the parity operator has opposite sign for the positive and negative energy solutions. In other words, fermions and anti-fermions have opposite parity. With our choice of the phase of the parity transformation, fermions have positive parity and anti-fermions have negative parity.

What does this mean for the currents in the interactions? Under the parity operation we find

$$
\begin{array}{lccccl}
S: & \overline{\psi}\psi & \rightarrow & \overline{\psi}\gamma^0\gamma^0\psi & = & \overline{\psi}\psi & \text{Scalar} \\
P: & \overline{\psi}\gamma^5\psi & \rightarrow & \overline{\psi}\gamma^0\gamma^5\gamma^0\psi & = & -\overline{\psi}\gamma^5\psi & \text{Pseudo Scalar} \\
V: & \overline{\psi}\gamma^\mu\psi & \rightarrow & \overline{\psi}\gamma^0\gamma^\mu\gamma^0\psi & = & \left\{ \begin{array}{c} \overline{\psi}\gamma^0\psi \\ -\overline{\psi}\gamma^k\psi \end{array} \right. & \text{Vector} \\
A: & \overline{\psi}\gamma^\mu\gamma^5\psi & \rightarrow & \overline{\psi}\gamma^0\gamma^\mu\gamma^5\gamma^0\psi & = & \left\{ \begin{array}{c} -\overline{\psi}\gamma^0\psi \\ \overline{\psi}\gamma^k\psi \end{array} \right. & \text{Axial Vector.}
\end{array}
$$

Experiments in the fifties had shown that the weak interaction was of the type vector or axial vector. However, if only a single bi-linear covariant contributes to the interaction, a parity transformation does not affect the cross-section or decay width as these are always proportional to the amplitude squared. Consequently, the experiments by Wu and others implied that the weak interaction received contributions from both the vector and the axial vector covariants,

$$\mathcal{M} = G_F \sum_{i,j}^{V,A} C_{ij} \ (\overline{u}_p O_i u_p) \ (\overline{u}_e O_j u_\nu) \tag{7.20}$$

Which combination of V and A appears in the weak interaction was established with a famous experiment by Goldhaber.

7.5 The $V - A$ interaction

Goldhaber and collaborators studied in 1958 the decay of *Europium*-152 to *Samorium* via a so-called electron capture reaction, ${}^{152}\mathrm{Eu} + e^- \rightarrow^{152} \mathrm{Sm}^*(\mathrm{J}=1) + \nu$:

$$
\begin{array}{lccccccc}
 & {}^{152}\mathrm{Eu} & + & e^- & \longrightarrow & {}^{152}\mathrm{Sm}^* & + & \nu_e \\
\text{direction of travel:} & & & & & \longleftarrow & & \longrightarrow \\
\text{spin configuration A:} & \circ & & \overset{-1/2}{\Leftarrow} & & \overset{-1}{\Longleftarrow} & & \overset{+1/2}{\Rightarrow} \\
\text{spin configuration B:} & \circ & & \overset{+1/2}{\Rightarrow} & & \overset{+1}{\Longrightarrow} & & \overset{-1/2}{\Leftarrow}
\end{array}
$$

The excited Samorium nucleus is in a $J = 1$ state. To conserve angular momentum, $\boldsymbol{J}$ must be parallel to the spin of the electron, but opposite to that of the electron-neutrino. The neutrino in this decay cannot be observed. However, the spin of the Samorium nucleus can be probed with the photon that is emitted in its decay to the ground state,

$${}^{152}\mathrm{Sm}^* \rightarrow^{152} \mathrm{Sm} + \gamma \ .$$

If we imagine the photon traveling left and project spins onto the photon direction, then situations A and B in the picture correspond to two different helicity configurations for the photon and the neutrino. Therefore, a measurement of the photon helicity is also a measurement of the neutrino helicity. Measuring the photon spin is a work-of-art by itself (for a good description of the experiment, see Perkins ed 3, §7.5.), but assuming it can be done, this allows to distinguish topologies A and B. The measurement by Goldhaber's group showed that only case 'B' actually occurs: the neutrinos in this decay always have helicity $-\frac{1}{2}$. It is therefore said that neutrinos are *left-handed*.

The cumulative evidence from Goldhaber's and other experiments involving weak interactions led to the conclusion that the weak interaction violates parity *maximally*. Rather than the vector form assumed by Fermi, the charge-lowering lepton current is actually

$$J^{\mu} \; = \; \overline{u}_e \, \gamma^{\mu} \, \tfrac{1}{2} \left(1 - \gamma^5\right) \, u_{\nu} \; . \tag{7.21}$$

The current for quarks looks identical, for example for a $u \rightarrow d$ transition

$$J^{\mu} \; = \; \overline{u}_u \, \gamma^{\mu} \, \tfrac{1}{2} \left(1 - \gamma^5\right) \, u_{d} \; . \tag{7.22}$$

We call this the "V-A" form.

For massless particles (or in the ultra-relativistic limit), the projection operator

$$P_L \; \equiv \; \tfrac{1}{2} \left(1 - \gamma^5\right) \tag{7.23}$$

selects the 'left-handed' helicity state of a particle spinor and the right-handed helicity state of an anti-particle spinor. As a result, only left handed neutrinos (ν_L) and right-handed anti-neutrinos ($\overline{\nu}_R$) are involved in weak interactions.

For decays of nuclei the structure is more complicated since the constituents are not free elementary particles. The matrix element for neutron decays can be written as

$$\mathcal{M} = \frac{G_F}{\sqrt{2}} \; \left(\overline{u}_p \gamma^{\mu} \left(C_V - C_A \gamma^5\right) u_n\right) \; \left(\overline{u}_e \gamma_{\mu} \left(1 - \gamma^5\right) u_{\nu}\right) \; . \tag{7.24}$$

For neutron decay, the measured vector and axial vector couplings are $C_V = 1.000 \pm 0.003$, $C_A = 1.260 \pm 0.002$

7.6 The propagator of the weak interaction

The Fermi theory has a 4-point interaction, unlike the Yukawa theory: there is no propagator to 'transmit' the interaction from the lepton current to the hadron current. However, we know now that forces are carried by bosons:

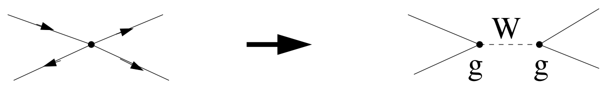
- the electromagnetic interaction is carried by the massless photon. It has a propagator

$$\frac{-ig^{\mu\nu}}{q^2}$$

- the weak interaction is carried by the massive W, Z bosons. For vector boson with non-zero mass M the propagator is

$$\frac{-i\left(g^{\mu\nu} - q^\mu q^\nu / M^2\right)}{M^2 - q^2} \,.$$

At low energies, *i.e.* when $q^2 \ll M_{Z,W}^2$, the q^2 dependence of the propagator vanishes, and the interaction looks like a four-point interaction. Therefore, Fermi's coupling constant is related to the actual coupling constant 'g' of the weak interaction:



strength: $\sim \frac{G_F}{\sqrt{2}}$ $\qquad$ $\sim \frac{g^2}{8M_W^2}$

It is an experimental fact that the strength of the coupling of the weak interaction, the coupling constant "g", is *identical* for quarks and leptons of all flavours. For leptons this is sometimes called 'lepton-universality'.

How "weak" is the weak interaction? For the electromagnetic coupling we have $\alpha = \frac{e^2}{4\pi} \approx 1/137$. It turns out that the weak coupling is equal to $\alpha_w = \frac{g^2}{4\pi} \approx 1/29$. We see that at low energies, the weak interaction is 'weak' compared to the electromagnetic interaction not because the coupling is small, but because the propagator mass is large! At high energies $q^2 \gtrsim M_W^2$ the weak interaction is comparable in strength to the electromagnetic interaction.

7.7 Muon decay

Similar to the process $e^+e^- \to \mu^+\mu^-$ in QED, the muon decay process $\mu^- \to e^-\overline{\nu}_e\nu_\mu$ is the standard example of a weak interaction process. The Feynman diagram is shown in Fig. 7.1.

Figure 7.1: Muon decay: *left:* Labelling of the momenta, *right:* Feynman diagram. For the spinor of the outgoing anti-particle we use $u_{\nu_e}(-k') = v_{\nu_e}(k')$.

Using the Feynman rules we can write for the matrix element:

$$\mathcal{M} = \frac{g}{\sqrt{2}} \left(\underbrace{\overline{u}(k)}_{\text{outgoing } \nu_\mu} \;\; \gamma^\mu \frac{1}{2}\left(1-\gamma^5\right) \;\; \underbrace{u(p)}_{\text{incoming } \mu} \right) \;\; \underbrace{\frac{1}{M_W^2}}_{\text{propagator}} \;\; \frac{g}{\sqrt{2}} \left(\underbrace{\overline{u}(p')}_{\text{outgoing } e} \;\; \gamma_\mu \frac{1}{2}\left(1-\gamma^5\right) \;\; \underbrace{v(k')}_{\text{outgoing } \overline{\nu}_e} \right) \tag{7.25}$$

Next we square the matrix element and sum over the spin states, just like we did for $e^+e^- \to \mu^+\mu^-$. Then we use again Casimir's tric, as well as the completeness relations, to convert the sum over spins into a trace. The result is:

$$\overline{|\mathcal{M}|}^2 = \frac{1}{2} \sum_{\text{Spin}} |\mathcal{M}|^2 = \frac{1}{2} \left(\frac{g^2}{8M_W^2} \right)^2 \quad \cdot \quad \mathrm{Tr} \left\{ \gamma^\mu \left(1-\gamma^5\right) \left(\not{p}\,' + m_e\right) \gamma^\nu \left(1-\gamma^5\right) \not{k}\,' \right\}$$
$$\cdot \quad \mathrm{Tr} \left\{ \gamma_\mu \left(1-\gamma^5\right) \not{k}\, \gamma_\nu \left(1-\gamma^5\right) \left(\not{p} + m_\mu\right) \right\}$$

Now we use some more trace theorems (see below) and also $\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2}$ to find the result:

$$\boxed{\overline{|\mathcal{M}|}^2 = 64\, G_F^2 \; (k \cdot p') \; (k' \cdot p)} \tag{7.26}$$

Intermezzo: Trace theorems used (see also Halzen & Martin p 261):

$$\mathrm{Tr}\left(\gamma^\mu \not{a}\, \gamma^\nu \not{b}\,\right) \cdot \mathrm{Tr}\left(\gamma_\mu \not{c}\, \gamma_\nu \not{d}\,\right) = 32 \; \left[(a \cdot c) \; (b \cdot d) + (a \cdot d) \; (b \cdot c)\right]$$
$$\mathrm{Tr}\left(\gamma^\mu \not{a}\, \gamma^\nu \gamma^5 \not{b}\,\right) \cdot \mathrm{Tr}\left(\gamma_\mu \not{c}\, \gamma_\nu \gamma^5 \not{d}\,\right) = 32 \; \left[(a \cdot c) \; (b \cdot d) - (a \cdot d) \; (b \cdot c)\right]$$
$$\mathrm{Tr}\left(\gamma^\mu \left(1-\gamma^5\right) \not{a}\, \gamma^\nu \left(1-\gamma^5\right) \not{b}\,\right) \cdot \mathrm{Tr}\left(\gamma_\mu \left(1-\gamma^5\right) \not{c}\, \gamma_\nu \left(1-\gamma^5\right) \not{d}\,\right) = 256 \; (a \cdot c) \; (b \cdot d)$$

The decay width is computed with Fermi's golden rule:

$$\mathrm{d}\Gamma \;=\; \frac{1}{2E} \, \overline{|\mathcal{M}|}^2 \, \mathrm{d}Q \tag{7.27}$$

where

$$\mathrm{d}Q \;=\; \frac{\mathrm{d}^3 p'}{(2\pi)^3 \; 2E'} \cdot \frac{\mathrm{d}^3 k}{(2\pi)^3 \; 2\omega} \cdot \frac{\mathrm{d}^3 k'}{(2\pi)^3 \; 2\omega'} \cdot (2\pi)^4 \, \delta^4 \left(p - p' - k' - k\right) \tag{7.28}$$

with

$$\begin{aligned} E &= \text{muon energy} \\ E' &= \text{electron energy} \\ \omega' &= \text{electron neutrino energy} \\ \omega &= \text{muon neutrino energy} \end{aligned}$$

First we evaluate the expression for the matrix element. Working in the rest frame of the muon and ignoring the mass of the electron and the neutrinos, we find (Eq. 7.48 in exercise 7.3),

$$\overline{|\mathcal{M}|}^2 = 64\, G_F^2\ (k \cdot p')\ (k' \cdot p) = 32\, G_F^2\ \left(m^2 - 2m\omega'\right)\ m\omega' \tag{7.29}$$

where m is the muon mass. Inserting this in the expression for the differential decay width, we obtain

$$\mathrm{d}\Gamma = \frac{1}{2E}\, \overline{|\mathcal{M}|}^2\, \mathrm{d}Q = \frac{16 G_F^2}{m}\ \left(\left(m^2 - 2m\omega'\right) m\omega'\, \mathrm{d}Q \tag{7.30}$$

where we used that $E = m$ in the muon rest frame. To obtain the total decay width we must integrate over the phase space,

$$\Gamma = \int \frac{1}{2E}\, \overline{|\mathcal{M}|}^2\, \mathrm{d}Q = \frac{16 G_F^2}{m} \int \left(\left(m^2 - 2m\omega'\right) m\omega'\, \mathrm{d}Q \tag{7.31}$$

The integrand only depends on the neutrino energy ω'. So, let us first perform the integral in $\mathrm{d}Q$ over the <u>other</u> energies and momenta:

$$\begin{aligned}
\int_{\text{other}} \mathrm{d}Q &= \frac{1}{8\,(2\pi)^5} \int \delta\left(m - E' - \omega' - \omega\right)\ \delta^3\left(\boldsymbol{p}' + \boldsymbol{k}' + \boldsymbol{k}\right)\ \frac{\mathrm{d}^3\boldsymbol{p}'}{E'}\, \frac{\mathrm{d}^3\boldsymbol{k}'}{\omega'}\, \frac{\mathrm{d}^3\boldsymbol{k}}{\omega} \\
&= \frac{1}{8\,(2\pi)^5} \int \delta\left(m - E' - \omega' - \omega\right)\ \frac{\mathrm{d}^3\boldsymbol{p}'\, \mathrm{d}^3\boldsymbol{k}'}{E'\omega'\omega}
\end{aligned}$$

since the δ-function gives 1 for the integral over $\boldsymbol{k}$.

We also have the relation:

$$\omega = |k| = |\boldsymbol{p}' + \boldsymbol{k}'| = \sqrt{E'^2 + \omega'^2 + 2E'\omega' \cos\theta} \tag{7.32}$$

where θ is the angle between the electron and the electron neutrino. We choose the z-axis along $\boldsymbol{k}'$, the direction of the electron neutrino. From the equation for ω we derive:

$$\mathrm{d}\omega = \frac{-2E'\omega' \sin\theta}{2\underbrace{\sqrt{E'^2 + \omega'^2 + 2E'\omega' \cos\theta}}_{\omega}}\, \mathrm{d}\theta \qquad \Leftrightarrow \qquad \mathrm{d}\theta = \frac{-\omega\, \mathrm{d}\omega}{E'\omega' \sin\theta} \tag{7.33}$$

Next we integrate over $\mathrm{d}^3\boldsymbol{p}' = E'^2 \sin\theta\, \mathrm{d}E'\, \mathrm{d}\theta\, \mathrm{d}\phi$ with $\mathrm{d}\theta$ as above:

$$\begin{aligned}
\mathrm{d}Q &= \frac{1}{8\,(2\pi)^5} \int \delta\left(m - E' - \omega' - \omega\right)\ \frac{E'^2 \sin\theta}{E'}\, \mathrm{d}E'\, \mathrm{d}\theta\, \mathrm{d}\phi\, \frac{\mathrm{d}^3\boldsymbol{k}'}{\omega'}\, \frac{1}{\omega} \\
&= \frac{1}{8\,(2\pi)^5}\, 2\pi \int \delta\left(m - E' - \omega' - \omega\right)\ \mathrm{d}E'\, \mathrm{d}\omega\, \frac{\mathrm{d}^3\boldsymbol{k}'}{\omega'^2}
\end{aligned}$$

(using the relation: $E' \sin\theta\, \mathrm{d}\theta = -\frac{\omega}{\omega'}\, \mathrm{d}\omega$).

Since we integrate over ω, the δ-function will cancel:

$$\mathrm{d}Q = \frac{1}{8\,(2\pi)^4} \int \mathrm{d}E' \, \frac{\mathrm{d}^3\boldsymbol{k}'}{\omega'^2} \tag{7.34}$$

such that the full expression for Γ becomes:

$$\Gamma = \frac{2G_F^2}{(2\pi)^4} \int \left(m^2 - 2m\omega'\right) \omega' \, \mathrm{d}E' \, \frac{\mathrm{d}^3\boldsymbol{k}'}{\omega'^2} \tag{7.35}$$

Next we do the integral over k' as far as possible with:

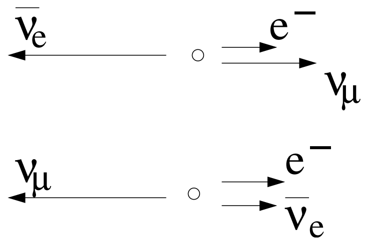
$$\int \mathrm{d}^3\boldsymbol{k}' = \int \omega'^2 \, \sin\theta' \, \mathrm{d}\omega' \, \mathrm{d}\theta' \, \mathrm{d}\phi' = 4\pi \int \omega'^2 \, \mathrm{d}\omega \tag{7.36}$$

so that we get:

$$\Gamma = \frac{G_F^2 m}{(2\pi)^3} \int (m - 2\omega') \; \omega' \, \mathrm{d}\omega' \, \mathrm{d}E' \tag{7.37}$$

Before we do the integral over ω' we have to determine the limits:

- maximum electron neutrino energy: $\omega' = \frac{1}{2}m$
- minimum electron neutrino energy: $\omega' = \frac{1}{2}m - E'$



Therefore, we obtain for the distribution of the electron energy in the muon rest frame

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}E'} = \frac{G_F^2 m}{(2\pi)^3} \int_{\frac{1}{2}m-E'}^{\frac{1}{2}m} (m - 2\omega') \; \omega' \, \mathrm{d}\omega' = \frac{G_F^2 m^2}{12\pi^3} \, E'^2 \left(3 - 4\frac{E'}{m}\right) \, . \tag{7.38}$$

Figure 7.2 shows a comparison between this prediction and an actual measurement. Finally, integrating the expression over the electron energy we find for the total decay width of the muon

$$\boxed{\Gamma \equiv \frac{1}{\tau} = \frac{G_F^2 \, m^5}{192 \, \pi^3}} \tag{7.39}$$

The measurement of the muon lifetime is the standard method to determine the coupling constant of the weak interaction. The muon lifetime has been measured to be $\tau = 2.19703 \pm 0.00004\ \mu s$. From this we derive for the Fermi coupling constant $G_F = (1.16639 \pm 0.00002) \cdot 10^{-5}\mathrm{GeV}^{-2}$.

7.8 Quark mixing

The strong and electromagnetic interaction do not couple to currents that connect leptons or quarks of different flavour. These interactions *conserve* the type of lepton or quark at the interaction vertex.

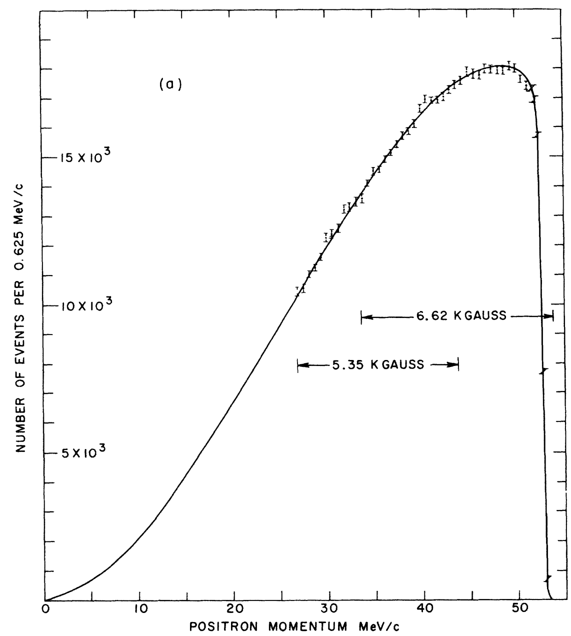
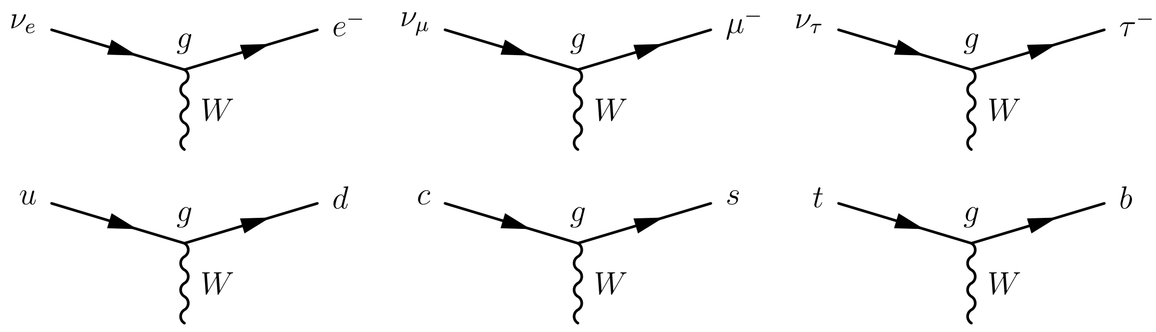


Figure 7.2: Experimental measurement of the energy spectrum of a positron in the decay of μ^+. The superimposed curve is the prediction. From Bardon *et al.*, PRL 14, 449 (1965).

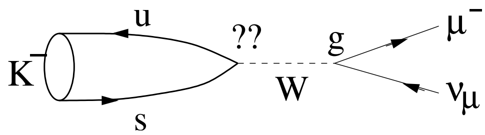
This is different for the weak interaction. As the W is charged, it necessarily couples to a current that contains two particles that differ by one unit in charge. For the quarks and leptons in the standard model, the Feynman diagrams for the interactions are



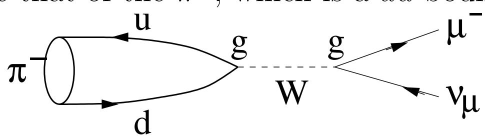
Leptons and quarks are usually ordered in three 'generations' to show how the weak interaction couples:

$$\underline{\text{Leptons:}} \quad \begin{pmatrix} \nu_e \\ e \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} \qquad \underline{\text{Quarks:}} \quad \begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \tag{7.40}$$

If the only couplings of the W are those shown in the Feynman diagrams, then the lightest hadrons with a strange quark (such as the K^- which is a $s\bar{u}$ bound state) would be stable. However, K^- mesons do decay, for instance to a muon and a muon neutrino:



This decay looks a lot like that of the π^-, which is a $d\bar{u}$ bound state:



Experimentally the K^- decay is found to have a much smaller decay width than the pion decay.

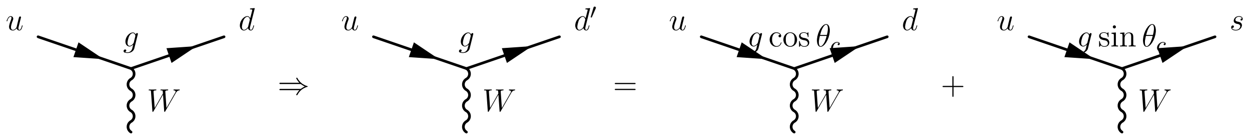
In 1963 Nicola Cabibbo provided a solution that explained most available data on strange hadron decay by presenting the d quark in the current that couples to the W as a linear combination of a d quark and an s quark:

$$\begin{aligned} d \rightarrow \; d' &= d\,\cos\theta_c + s\,\sin\theta_c \\ s \rightarrow \; s' &= -d\,\sin\theta_c + s\,\cos\theta_c \end{aligned} \tag{7.41}$$

where θ_c is a 'mixing' angle now known as the Cabibbo angle. In matrix representation the mixing can be written as

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \tag{7.42}$$

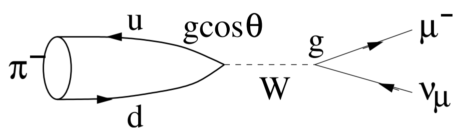
while in terms of the diagrams the replacement looks like:



Due to the mixing the amplitudes for pion and kaon decay contain factors $\cos\theta_c$ and $\sin\theta_c$:

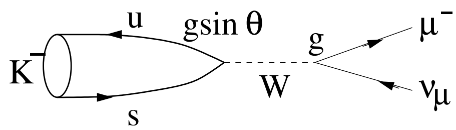
1. Pion decay

   $\pi^- \rightarrow \mu^- \overline{\nu}_\mu$
   $\Gamma_{\pi^-} \propto G_F^2 \cos^2\theta_c$



2. Kaon decay

   $K^- \rightarrow \mu^- \overline{\nu}_\mu$
   $\Gamma_{K^-} \propto G_F^2 \sin^2\theta_c$



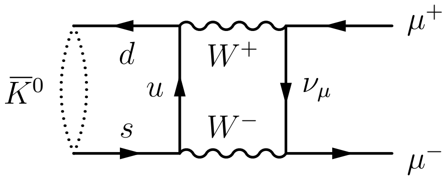
A proper calculation gives for the ratio of the decay rates

$$\frac{\Gamma\left(K^-\right)}{\Gamma\left(\pi^-\right)} \approx \tan^2\theta_c \cdot \left(\frac{m_\pi}{m_K}\right)^3 \left(\frac{m_K^2 - m_\mu^2}{m_\pi^2 - m_\mu^2}\right)^2 \tag{7.43}$$

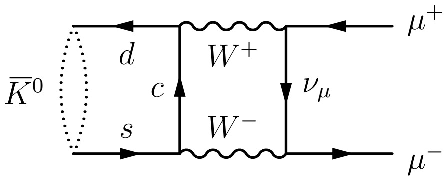
From the experimental result on the lifetime ratios, the Cabibbo angle is then found to be

$$\boxed{\theta_C \approx 13.0^\circ} \tag{7.44}$$

Even though Cabibbo's theory explained strange decays, it did not quite get everything right. The proposed quark mixing would allow neutral kaons ($s\bar{d}$ combinations) to decay to muons, via the amplitude represented by this Feynman diagram:

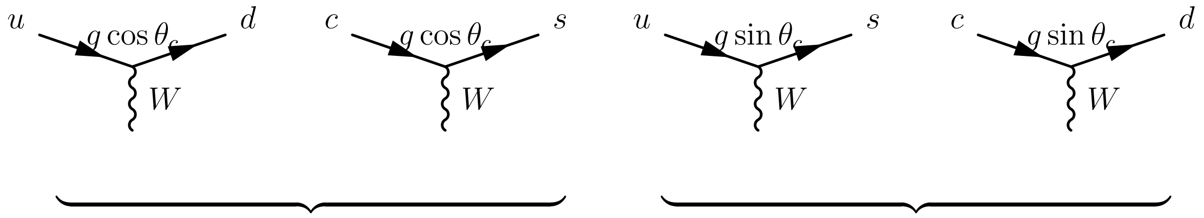


According to Cabibbo's calculation this decay should have an appreciable rate, but it was never found! An explanation was provided by Glashow, Iliopoulis and Maiani in 1970: They hypothesised the existence of the charm (c) quark, contributing with a diagram



The up and charm quark amplitudes have opposite sign, which leads to a nearly vanishing decay rate. This mechanism, which is now known as the GIM mechanism, was the first well-motivated prediction for a fourth quark. The charm quark was discovered 3 years later.

Including charm quarks, the couplings for the first two generations are:



Cabibbo "favoured" decay

Cabibbo "suppressed" decay

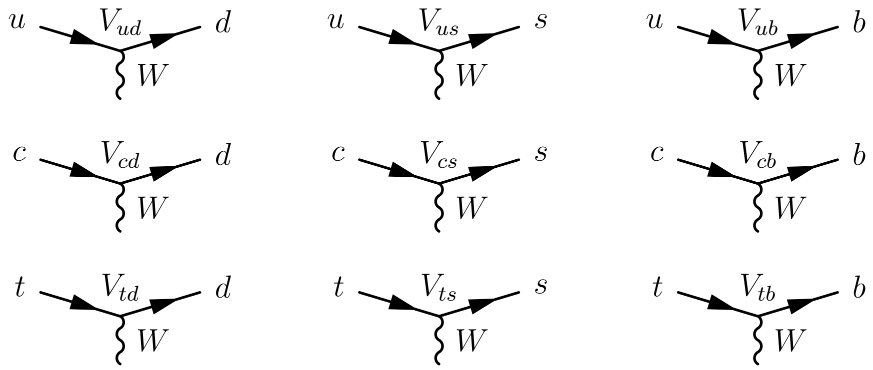
The flavour eigenstates u, d, s, c are the mass eigenstates of the total Hamiltonian describing quarks. The states $\begin{pmatrix} u \\ d' \end{pmatrix}$, $\begin{pmatrix} c \\ s' \end{pmatrix}$ are the eigenstates of the weak interaction Hamiltonian, which affects the decay of the particles. By convention mixing is presented for 'down' quarks, but in fact that choice is arbitrary: We could also consider the mixing matrix to mix the u and c quarks.

Of course, the story of quarks did not stop with the discovery of the charm quark. In 1964 Cronin and Fitch had shown in experiments that CP symmetry is violated in neutral kaon decays. Kobayashi and Maskawa found a solution in 1973: They extended

Cabibbo's picture of quark mixing with a third family of quarks,

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{\text{CKM matrix}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{7.45}$$

The three-generation mixing matrix is nowadays called the "Cabibbo-Kobayashi-Maskawa" matrix, or simply the CKM matrix. The couplings in the Feynman diagram for charged current interactions all get modified by elements of the mixing matrix:



The mixing matrix V_{CKM} is a 3×3 *unitary* matrix. This matrix is not uniquely defined since the phases of the quark field can be chosen arbitrarily. If the phases are "'absorbed' in the quark fields, the matrix can be parametrized by four real parameters, which are usually chosen to be three mixing angles between the quark generations θ_{12}, θ_{13}, θ_{23}, and one complex phase δ,

$$V_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}s_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{7.46}$$

where $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$.

Kobayashi and Maskawa realized that the fact elements of V_{CKM} can have a non-trivial complex phase — i.e. a phase that can not be removed by redefining the phase of the quark fields — leads to CP violation in charged current decays. For CP violation to occur this way, at least three generations of quarks are required. The bottom and top quark were eventually discovered in 1977 and 1994, respectively.

In case neutrino particles have a non-zero mass, mixing occurs in the lepton sector as well. Just like the down-type quarks were chosen to describe mixing in the quark sector, the neutrinos are chosen for the lepton sector:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \underbrace{\begin{pmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \\ U_{31} & U_{32} & U_{33} \end{pmatrix}}_{\text{PMNS-matrix}} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{7.47}$$

The lepton mixing matrix is called the Pontecorvo-Maki-Nakagawa-Sakata matrix. This matrix relates the mass eigenstates of the neutrinos (ν_1, ν_2, ν_3) to the weak interaction eigenstates (ν_e, ν_μ, ν_τ). There is an interesting open question whether neutrino's are their own anti-particles ("Majorana" neutrino's) or not ("Dirac" neutrino's). In case neutrinos are of the Dirac type, the U_{PMNS} matrix has one complex phase, similar to the quark mixing matrix. If neutrinos are Majorana particles, the U_{PMNS} matrix includes three complex phases.

Violation of CP in fundamental interactions is needed to understand why we live in a 'matter-dominated' universe — rather than universe with equal amounts of matter and anti-matter. It is still unknown if the matter-dominance is a result of CP violation in the quark sector ("baryogenesis") or in the lepton sector ("leptogenesis").

Exercises

Exercise 7.1 (Helicity versus chirality. See also H&M exercise 5.15)

(a) Write out the **chirality operator** γ^5 in the Dirac-Pauli representation.

(b) The helicity operator is defined as $\lambda = \frac{1}{2}\boldsymbol{\Sigma} \cdot \hat{\boldsymbol{p}}$, where $\hat{\boldsymbol{p}}$ is a unit vector along the momentum and Σ is

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} .$$

Show that in the ultra-relativistic limit ($E \gg m$) the helicity operator and the chirality operator have the same effect on a spinor solution, *i.e.*

$$\gamma^5\,\psi = \gamma^5 \begin{pmatrix} \chi^{(s)} \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi^{(s)} \end{pmatrix} \qquad \approx \qquad 2\lambda \begin{pmatrix} \chi^{(s)} \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi^{(s)} \end{pmatrix} = 2\lambda\,\psi$$

(c) Explain why the weak interaction is called **left-handed**.

Exercise 7.2 (Pion Decay)
Usually at this point the student is asked to calculate pion decay, which requires again quite some calculations. The ambitious student is encourage to try and do it (using some help from the literature). However, the exercise below requires little or no calculation but instead insight in the formalism.

(a) Draw the Feynman diagram for the decay of a pion to a muon and an anti-neutrino: $\pi^- \to \mu^- \overline{\nu}_\mu$.

(b) Compute the momentum of the muon in the rest frame of the pion. You may ignore the neutrino mass, but not the muon mass.

Due to the fact that the quarks in the pion are not free particles we cannot just apply the Dirac formalism for free particle waves. However, we know that the interaction is transmitted by a W^- and therefore the coupling must be of the type: V or A. (Also, the matrix element must be a Lorentz scalar.) It turns out the decay amplitude has the form:

$$\mathcal{M} = \frac{G_F}{\sqrt{2}} \left(q^\mu f_\pi\right) \left(\overline{u}(p)\gamma_\mu \left(1-\gamma^5\right) v(k)\right)$$

where p^μ and k^μ are the 4-momenta of the muon and the neutrino respectively, and q is the 4-momentum carried by the W boson. f_π is called the decay constant.

(c) Can the pion also decay to an electron and an electron-neutrino? Write down the Matrix element for this decay.
Would you expect the decay width of the decay to electrons to be larger, smaller, or similar to the decay width to the muon and muon-neutrino?
Base your argument on the available phase space in each of the two cases.

The decay width to a muon and muon-neutrino is found to be:

$$\Gamma = \frac{G_F^2}{8\pi}\, f_\pi^2\, m_\pi\, m_\mu^2 \left(\frac{m_\pi^2 - m_\mu^2}{m_\pi^2}\right)^2$$

The measured lifetime of the pion is $\tau_\pi = 2.6 \cdot 10^{-8} s$ which means that $f_\pi \approx m_\pi$. An interesting observation is to compare the decay width to the muon and to the electron:

$$\frac{\Gamma(\pi^- \to e^-\overline{\nu}_e)}{\Gamma(\pi^- \to \mu^-\overline{\nu}_\mu)} = \left(\frac{m_e}{m_\mu}\right)^2 \left(\frac{m_\pi^2 - m_e^2}{m_\pi^2 - m_\mu^2}\right)^2 \approx 1.2 \cdot 10^{-4} \quad !!$$

(d) Can you give a reason why the decay rate into an electron and an electron-neutrino is strongly suppressed in comparison to the decay to a muon and a muon-neutrino. Consider the spin of the pion, the handedness of the W coupling and the helicity of the leptons involved.

Exercise 7.3 (Kinematics of muon decay)
We consider the decay in Fig. 7.1.

(a) Starting from four-momentum conservation ($p = p' + k' + k$) and by ignoring the electron and neutrino masses, show that

$$2k \cdot p' \;=\; m^2 \;-\; 2p \cdot k'$$

where m is the muon mass.
Hint: write the equation as $p - k' = p' + k$, then square both sides.

(b) Use this result to show that *in the muon rest frame*

$$2(k \cdot p')(k' \cdot p) \;=\; (m^2 - 2m\omega')\, m\omega' \tag{7.48}$$

Lecture 8

Local Gauge Invariance

In the next two lectures we discuss the theory of the electroweak interaction, the so-called "Glashow-Salam-Weinberg model". This theory can be formulated starting from the principle of local gauge invariance.

8.1 Symmetries

Symmetries play a fundamental role in particle physics. A symmetry is always related to a quantity that is fundamentally unobservable. In general one can distinguish[1] four types of symmetries:

- permutation symmetries: These lead to Bose-Einstein statistics for particles with integer spin (bosons) and to Fermi-Dirac statistics for particles with half integer spin (fermions). The unobservable is the absolute identity of a particle;
- continuous space-time symmetries: translation, rotation, acceleration, etc. The related unobservables are respectively: absolute position in space, absolute direction and the equivalence between gravity and acceleration;
- discrete symmetries: space inversion, time inversion, charge inversion. The unobservables are absolute left/right handedness, the direction of time and an absolute definition of the sign of charge;
- unitary symmetries or internal symmetries, also called 'gauge invariance': These are the symmetries discussed in this lecture. As an example of an unobservable quantity think of the complex phase of a wave function in quantum mechanics.

The relation between symmetries and conservation laws is expressed in a fundamental theorem by Emmy Noether: each continuous symmetry transformation under which the Lagrangian is invariant in form leads to a conservation law. Invariances under *external* operations as time and space translation lead to conservation of energy and momentum,

[1] T.D. Lee: "Particle Physics and Introduction to Field Theory"

and invariance under rotation to conservation of angular momentum. Invariances under *internal* operations, like the shift of the complex phase of wave functions, lead to conserved currents, or more specific, conservation of charge. We discuss the application of Noether's theorem to phase transformations in section 8.4.

In the standard model the elementary interactions of the quarks and leptons (electromagnetic, weak and strong) are all the result of gauge symmetries. Starting from a Lagrangian that describes free quarks and leptons, the interactions can be constructed by requiring the Lagrangian to be symmetric under particular transformations. The idea of local gauge invariance will be discussed in this lecture and will be applied in the unified electroweak theory in the next lecture.

8.2 The principle of least action

In classical mechanics the equations of motion can be derived using the variational principle of Hamilton. This principle states that the action integral S should be stationary under arbitrary variations of the so-called generalized coordinates q_i. For a pedagogical discussion of the principle of least action read the Feynman lectures, Vol.2, chapter 19.

Generalized coordinates are coordinates that correspond to the actual degrees of freedom of a system. As an example, consider a swinging pendulum in two dimensions. We could describe the movement of the weight of the pendulum in terms of both its horizontal coordinate x and its vertical coordinate y. However, only one of those is independent since the length of the pendulum is fixed. Therefore, we say that the movement of the pendulum can be described by one 'generalized' coordinate. We could choose x or y, but also the angle of the pendulum with the vertical axis (usually called the amplitude). We denote generalized coordinates with the symbol q and call the evolution of q with time a *trajectory* or *path*.

The Lagrangian of the system can be defined as the kinetic energy minus the potential energy,

$$L(q, \dot{q}, t) \;=\; T(\dot{q}) \;-\; V(q) \;, \tag{8.1}$$

where the potential energy only depends on q (and eventually t) and the kinetic energy only on the generalized velocity $\dot{q} = \mathrm{d}q/\mathrm{d}t$. We denote the action (or 'action integral') of a path that starts at t_1 and ends at t_2 with

$$S(q) \;=\; \int_{t_0}^{t_1} L(q, \dot{q}, t)\, \mathrm{d}t \;. \tag{8.2}$$

Hamilton's principle states that the actual trajectory $q(t)$ followed by the system is the trajectory $q(t)$ that *minimizes* the action. (It is said that the action is 'stationary' around this trajectory.) This is equivalent to requiring that for a given point $q, \dot{q}$ on this trajectory, the change in the action following from a small deviation $\delta q, \delta\dot{q}$ is zero:

$$\delta S(q, \delta q, \dot{q}, \delta\dot{q}) \;=\; 0 \;. \tag{8.3}$$

You will show in exercise 8.1 that for each of the coordinates q_i, this leads to the so-called Euler-Lagrange equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i}. \tag{8.4}$$

This may be written more symmetrically as

$$\dot{p}_i = \frac{\partial L}{\partial q_i} \qquad \text{with} \qquad p_i = \frac{\partial L}{\partial \dot{q}_i} , \tag{8.5}$$

where p_i is called the generalized momentum, or the *momentum canonical to* q_i. In terms of these coordinates, the Hamiltonian takes the form

$$H(p, q, t) = \sum_i p_i \dot{q}_i - L. \tag{8.6}$$

and the equation of motion can also be written in the form of Hamilton's equations,

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \qquad \text{and} \qquad \dot{q}_i = \frac{\partial H}{\partial p_i} . \tag{8.7}$$

Finally, the classical system can be quantized by imposing the fundamental postulate of quantum mechanics,

$$[q_i, p_j] = i\hbar\delta_{ij}. \tag{8.8}$$

8.3 Lagrangian density for fields

The classical theory does not treat space and time symmetrically as the Lagrangian might depend on the *parameter* t. This causes a problem if we want to make a Lorentz covariant theory. The solution is to go to *field theory*: Rather than a finite set of degrees of freedom we consider an infinite set of degrees of freedom, represented by the values of a field ϕ, that is a function of the space-time coordinates x^μ. The Lagrangian is replaced by a Lagrangian density $\mathcal{L}$ (usually just called Lagrangian),

$$L(q, \dot{q}, t) \quad \longrightarrow \quad \mathcal{L}(\phi, \partial^\mu\phi, x^\mu) \tag{8.9}$$

such that the action becomes

$$S = \int_{x_1}^{x_2} \mathrm{d}^4x \, \mathcal{L}\big(\phi(x), \partial^\mu\phi(x), x^\mu\big) \tag{8.10}$$

Following the principle of least action we obtain the Euler-Lagrange equation for the fields:

$$\frac{\partial \mathcal{L}}{\partial \phi(x)} = \partial_\mu \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi(x)\right)} \tag{8.11}$$

(If at this point you are confused about the position of Lorentz indices on the right-hand-side, then remember that what is meant is

$$\frac{\partial \mathcal{L}}{\partial \phi\,(x)} = \frac{\partial}{\partial x_0}\frac{\partial L}{\partial(\partial\phi/\partial x_0)} + \frac{\partial}{\partial x_1}\frac{\partial L}{\partial(\partial\phi/\partial x_1)} + \frac{\partial}{\partial x_2}\frac{\partial L}{\partial(\partial\phi/\partial x_2)} + \frac{\partial}{\partial x_3}\frac{\partial L}{\partial(\partial\phi/\partial x_3)} \quad (8.12)$$

You could also use upper indices as long as you are consistent.)

To create a Lorentz covariant theory, we require the Lagrangian to be a Lorentz scalar. (This also means that in the expression for $\mathcal{L}$ above the 'loose' Lorentz indices must somehow be contracted with others.) This requirement imposes certain conditions on the Lorentz transformation properties of the fields. (We have not discussed these in detail. See textbooks.) Furthermore, although we consider complex fields, we always require the Lagrangian to be real.

In *quantum* field theory, the coordinates ϕ become operators that obey the standard quantum mechanical commutation relation with their associated generalized momenta. The commutation relations lead to *creation* and *annihilation* operators. The *wave functions* that we have considered before can be viewed as single particle excitations that occur when the creation operators of the field act on the vacuum. For the discussions here we do not need field theory. What is important to know is that field theory tells us that, given a Lagrangian, we can find a set of Feynman rules that can be used to draw diagrams and compute amplitudes.

Now consider the following Lagrangian for a complex scalar field:

$$\mathcal{L} \;=\; (\partial_\mu \phi^*)(\partial^\mu \phi) - m^2 \phi^* \phi \quad (8.13)$$

You will show in an exercise that the equation of motion corresponding to this Lagrangian is the Klein-Gordon equation. Because the field is complex, it has two separate components. We could choose these to be the real and imaginary part of the field, such that $\phi = \phi_1 + i\phi_2$ with $\phi_{1,2}$ real. It is easy to see what the Lagrangian looks like and what the equations of motion become. However, rather than choosing ϕ_1 and ϕ_2 we can also choose ϕ and ϕ^* to represent the 'independent' components of the field.

A similar argument can be made for the bi-spinor and the adjoint bi-spinor in the Lagrangian of the Dirac field. The latter is given by

$$\mathcal{L} \;=\; \overline{\psi}\,(i\gamma^\mu \partial_\mu - m)\,\psi \quad (8.14)$$

and its equation of motion (treating $\overline{\psi}$ and ψ as independent components of the field) is the Dirac equation.

8.4 Global phase invariance and Noether's theorem

Lagrangians for complex fields are constructed such that the Lagrangian is real (and the Hamiltonian hermitian). As a result, the Lagrangian is not sensitive to a *global* shift in the complex phase of the field. Such a global phase change is called a $U(1)$ symmetry. (The group $U(1)$ is the group of unitary matrices of dimension 1, *e.g.* complex numbers with unit modulus.) Noether's theorem tells us that there must be a conserved quantity associated with such a phase invariance. For 'internal' symmetries of the Lagrangian this works as follows.

Consider a transformation of the field components with a small variation ϵ,

$$\phi_i \to \phi_i + \epsilon_i(x) \tag{8.15}$$

The resulting change in the Lagrangian is

$$\delta\mathcal{L} \;=\; \sum_i \left(\frac{\partial\mathcal{L}}{\partial\phi_i}\epsilon_i + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)}\partial_\mu\epsilon_i \right) \;=\; \partial_\mu \left(\sum_i \epsilon_i \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)} \right) \tag{8.16}$$

where in the second step we have used the Euler-Lagrange equation to remove $\partial\mathcal{L}/\partial\phi$. Consequently, if the Lagrangian is insensitive to the transformation, then the quantity

$$j^\mu \;=\; \sum_i \epsilon_i \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_i)} \tag{8.17}$$

is a conserved current.

Let's now apply this to the complex scalar field for a (small) $U(1)$ phase translation. The two independent field components are ϕ and ϕ^*. Under the phase translation these change as

$$\begin{aligned} \phi \to \phi e^{i\alpha} \;&\approx\; \phi(1 + i\alpha) \\ \phi^* \to \phi^* e^{-i\alpha} \;&\approx\; \phi^*(1 - i\alpha) \end{aligned} \tag{8.18}$$

Consequently, we have $\epsilon_\phi = i\alpha\phi$ and $\epsilon_{\phi^*} = -i\alpha\phi^*$. Inserting these into the expression for the Noether current, Eq. (8.17), we find

$$j^\mu \;=\; \alpha\, i \left(\phi(\partial^\mu\phi^*) - \phi^*(\partial^\mu\phi) \right) \tag{8.19}$$

Since α is an arbitrary constant, we omit it from the current. We have obtained exactly the current that we constructed for the Klein-Gordon wave in Lecture 1.

8.5 Local phase invariance

We can also look at the $U(1)$ symmetry from a slightly more general perspective. The expectation value of a quantum mechanical observable (such as the Hamiltonian) is typically of the form:

$$\langle O \rangle = \int \psi^* O \psi \tag{8.20}$$

If we make the replacement $\psi(x) \to e^{i\alpha}\psi(x)$ the expectation value of the observable remains the same. We say that we cannot measure the absolute phase of the wave function. (We can only measure *relative* phases between wave functions in interference experiments.)

However, this only holds for a phase that is constant in space and time. Are we allowed to choose a different phase convention on, say, the moon and on earth, for a wave function $\psi(x)$? In other words, can we choose a phase that depends on space-time,

$$\psi(x) \to \psi'(x) = e^{i\alpha(x)}\psi(x) \quad ? \tag{8.21}$$

In general, we cannot do this without breaking the symmetry. The problem is that the Lagrangian density $\mathcal{L}\left(\psi(x), \partial_\mu\psi(x)\right)$ depends on both the fields $\psi(x)$ and the derivatives $\partial_\mu\psi(x)$. The derivative term yields:

$$\partial_\mu\psi(x) \to \partial_\mu\psi'(x) = e^{i\alpha(x)}\,\big(\,\partial_\mu\psi(x) \;+\; i\left(\partial_\mu\alpha(x)\right)\psi(x)\,\big) \tag{8.22}$$

and therefore the Lagrangian is not invariant.

However, suppose that we now introduce 'local $U(1)$ symmetry' as a requirement. Is it possible to modify the Lagrangian such that it obeys this symmetry? The answer is 'yes', provided that we introduce a new field, the so-called *gauge field*. The recipe consists of two steps.

First, we replace the derivative ∂_μ by the so-called *gauge-covariant derivative*:

$$\partial_\mu \to D_\mu \equiv \partial_\mu + iqA_\mu(x)\;, \tag{8.23}$$

where A_μ is a new field and q is (for now) an arbitrary constant. Second, we require that the field A_μ transforms as

$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) - \frac{1}{q}\partial_\mu\alpha(x)\;. \tag{8.24}$$

By inserting the expression for A in the covariant derivative, we find that it just transforms with the *local* phase $\alpha(x)$:

$$\begin{aligned} D_\mu\psi(x) \;\to\; D'_\mu\psi'(x) \;&=\; e^{i\alpha(x)}\left(\partial_\mu\psi(x) + i\partial_\mu\alpha(x)\psi(x) + iqA_\mu(x)\psi(x) - iq\frac{1}{q}\partial_\mu\alpha(x)\psi(x)\right) \\ &=\; e^{i\alpha(x)}D_\mu\psi(x) \end{aligned} \tag{8.25}$$

As a consequence, terms in the derivative that look like $\psi^* D_\mu\psi$ are phase invariant. With the substitution $\partial_\mu \to D_\mu$ the Klein-Gordon and Dirac Lagrangians (and any other real Lagrangian that we can construct with 2nd order terms from a complex field and its derivatives) satisfy the local phase symmetry.

8.6 Application to the Dirac Lagrangian

Now consider the effect of the substitution of the derivative with the covariant derivative in Eq. (8.23) in the Lagrangian for the Dirac field (Eq. (8.14)),

$$\begin{aligned} \mathcal{L} &= \overline{\psi}\ (i\gamma^\mu D_\mu - m)\ \psi \\ &= \overline{\psi}\ (i\gamma^\mu \partial_\mu - m)\ \psi\ -\ q\, A_\mu\, \overline{\psi}\, \gamma^\mu\, \psi \\ &\equiv\ \mathcal{L}_{\text{free}}\ +\ \mathcal{L}_{\text{int}}\ , \end{aligned} \tag{8.26}$$

where we defined the interaction term

$$\mathcal{L}_{\text{int}} = -J^\mu A_\mu \tag{8.27}$$

with the familiar Dirac current

$$J^\mu = q\overline{\psi}\gamma^\mu\psi\ . \tag{8.28}$$

This is exactly the form of the electromagnetic interaction that we discussed in the previous lectures. Furthermore, since we require the Lagrangian to be real and since the conserved current is real, the field A^μ must be a real as well. We can now identify q with the charge and the gauge field A^μ with the electromagnetic vector potential, *i.e.* the photon field. The transformation of the field in Eq. (8.24) is just the gauge freedom that we identified in the electromagnetic field in Lecture 3 (with $\lambda = q\alpha$).

The picture is not entirely complete yet, though. We know that the photon field satisfies its own 'free' Lagrangian. This is the Lagrangian that leads to the Maxwell equations in vacuum. It is given by

$$\mathcal{L}_A^{\text{free}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{8.29}$$

with $F^{\mu\nu} = \partial^\nu A^\mu - \partial^\mu A^\nu$. We call this the 'kinetic' term of the gauge field Lagrangian and we simply add it to the total Lagrangian. The full Lagrangian for a theory that has one Dirac field and obeys local $U(1)$ symmetry is then given by

$$\mathcal{L}_{\text{QED}} = \overline{\psi}\,(i\gamma^\mu\partial_\mu - m)\,\psi - qA_\mu\overline{\psi}\gamma^\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{8.30}$$

This is called the *QED Lagrangian.*

At this point you may wonder if we could also add a mass term for the photon field. If the photon would have a mass, the corresponding term in the Lagrangian would be

$$\mathcal{L}_\gamma\ =\ \frac{1}{2}m^2 A^\mu A_\mu\ . \tag{8.31}$$

However, this term violates local gauge invariance, since:

$$A^\mu A_\mu \to \left(A^\mu - \tfrac{1}{q}\partial^\mu\alpha\right)\left(A_\mu - \tfrac{1}{q}\partial_\mu\alpha\right) \neq A^\mu A_\mu \tag{8.32}$$

Therefore, the requirement of local $U(1)$ invariance automatically implies that the photon is massless. This actually holds for other gauge symmetries as well. In chapters 11 to 14 we discuss how masses of vector bosons can be generated in the Higgs mechanism by 'breaking' the symmetry.

You may wonder why we put so much emphasis on the principle of local gauge invariance. After all, it looks like all we have done is find a different way of arriving at the equations of motions of electrodynamics: is it really so attractive to formulate QED as a symmetry?

The reason that local gauge symmetries are so important is because of what is called 'renormalizability'. By way of the Feynman rules, the Lagrangian encodes the information to compute scattering and decay processes to arbitrary order. However, if you compute anything beyond leading order you will quickly find that the result is not finite. This can be solved by a number of different techniques, called collectively 'renormalization'. It was shown by 't Hooft and Veldman in the early seventies that only Lagrangians with interaction terms generated by local gauge symmetries are renormalizable. In other words, if we want to have a theory in which we can compute something, then we cannot have any other interactions than those derived from internal symmetries.

8.7 Yang-Mills theory

The $U(1)$ symmetry discussed above is the simplest local gauge symmetry. To extend the gauge principle to the weak and strong interaction, we need to consider more complicated symmetries. We introduce these so-called 'non-abelian' symmetries in a somewhat historical fashion as this helps to understand the origin of the term 'weak isospin' and the relation to (strong-) isospin.

In the 1950s Yang and Mills tried to describe the strong interaction in the proton-neutron system in terms of a gauge symmetry. Ignoring the electric charge, the free Lagrangian for the nucleons can be written as

$$\mathcal{L} = \overline{p}\ (i\gamma^\mu \partial_\mu - m)\ p + \overline{n}\ (i\gamma^\mu \partial_\mu - m)\ n\ . \tag{8.33}$$

In terms of the bi-spinor *doublet*

$$\psi = \begin{pmatrix} p \\ n \end{pmatrix}, \tag{8.34}$$

the Lagrangian becomes

$$\mathcal{L} \ = \ \overline{\psi} \begin{pmatrix} i\gamma^\mu \partial_\mu -\ m & 0 \\ 0 & i\gamma^\mu \partial_\mu -\ m \end{pmatrix} \psi \tag{8.35}$$

Have a careful look at what is written here: The doublet ψ is a 2-component column vector with a Dirac spinor for each component. Each of the entries in the matrix in the Lagrangian is again a 4x4 matrix.

Note that we have taken the two components to have identical mass m. Because they have identical mass and no charge the nucleons are indistinguishable. Therefore, we consider a global transformation of the field ψ with a complex unitary (2×2) matrix U that 'rotates' the proton-neutron system,

$$\psi \to U\psi \qquad \text{and} \qquad \overline{\psi} \to \overline{\psi} U^\dagger$$

Since $U^\dagger U = 1$, our Lagrangian is invariant to this transformation.

Any complex unitary (2×2) matrix can be written in the form

$$U \;=\; e^{i\theta} \exp\left(\frac{i}{2}\boldsymbol{\alpha} \cdot \boldsymbol{\tau}\right) \;, \tag{8.36}$$

where $\boldsymbol{\alpha}$ and θ are *real* and $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3)$ are the Pauli spin matrices [2]. We have already considered the effects of a phase transformation $e^{i\theta}$, which was the $U(1)$ symmetry. Therefore, we concentrate on the case where $\theta = 0$. Since the matrices τ all have zero trace, the matrices U with this property all have determinant 1. They form the group $SU(2)$ and the matrices τ are the generators of this group. (In group theory language we say that $SU(2)$ is an irriducible subgroup of $U(2)$ and $U(2) = U(1) \otimes SU(2)$.)

Note that members of $SU(2)$ do in general not commute. This holds in particular for the generators. We call such groups "non-abelian". In contrast, the $U(1)$ group is abelian since complex numbers just commute.

Using the same prescription as for the $U(1)$ symmetry, we can derive a conserved current. Consider a small $SU(2)$ transformation in doublet space

$$\psi \to \psi' \;=\; \left(\mathbb{1} + \frac{i}{2}\boldsymbol{\alpha} \cdot \boldsymbol{\tau}\right)\psi \tag{8.37}$$

and similar for $\overline{\psi}$. The Lagrangian transforms as

$$\begin{aligned} \delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\psi}\delta\psi + \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi\right)}\delta\left(\partial_\mu\psi\right) + \frac{\partial\mathcal{L}}{\partial\overline{\psi}}\delta\overline{\psi} + \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\overline{\psi}\right)}\delta\left(\partial_\mu\overline{\psi}\right) \\ &= \partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\delta\psi\right] + \partial_\mu\left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\overline{\psi})}\delta\overline{\psi}\right] \end{aligned} \tag{8.38}$$

where in the second line we have used the Euler-Lagrange equation to eliminate $\partial\mathcal{L}/\partial\psi$, just as we did in Eq. (8.16). Computing the derivatives of the Lagrangian, we find that the right term vanishes, while the left term gives

$$\delta\mathcal{L} \;=\; \partial_\mu\left[-\frac{1}{2}\,\overline{\psi}\,\gamma^\mu\,\boldsymbol{\alpha}\cdot\boldsymbol{\tau}\,\psi\right] \tag{8.39}$$

Since α is a constant and since the requirement of phase invariance must hold for any α, we can drop α and obtain three continuity equations $\partial_\mu \boldsymbol{J}^\mu = 0$ for the three conserved currents

$$\boldsymbol{J}^\mu \;=\; \overline{\psi}\,\gamma^\mu\,\frac{\boldsymbol{\tau}}{2}\,\psi \;. \tag{8.40}$$

---

[2]We had labeled these by $\boldsymbol{\sigma}$, but for some obscure reason the textbooks also switch from $\boldsymbol{\sigma}$ to $\boldsymbol{\tau}$ at this point. Our default representation is: $\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$,

As for the $U(1)$ symmetry, we now try to promote the global symmetry to a local symmetry. The strategy is similar to that for $U(1)$, but because the group is non-abelian, the implementation is more complicated. The first step is to make the parameters α depend on space time. To simplify the notation we define the gauge transformation as follows,

$$\begin{aligned} \psi(x) \rightarrow \psi'(x) = & \; G(x)\psi(x) \\ \text{with} \quad G(x) = & \; \exp\left(\frac{i}{2}\boldsymbol{\tau}\cdot\boldsymbol{\alpha}(x)\right) \end{aligned} \tag{8.41}$$

We have again, as in the case of QED, that the derivative transforms non-trivially

$$\partial_\mu\psi(x) \;\rightarrow\; G\left(\partial_\mu\psi\right) \;+\; \left(\partial_\mu G\right)\,\psi \tag{8.42}$$

such that the Lagrangian is not phase invariant. To restore phase invariance, we introduce the 2×2 covariant derivative

$$D_\mu \;=\; \mathbb{1}\partial_\mu + igB_\mu \tag{8.43}$$

where g is a (so far arbitrary) coupling constant and B_μ a gauge field. In spinor space the latter is a 2×2 unitary matrix with determinant 1. It is customary to parametrize it in terms of three new *real* vector fields b_1 , b_2 and b_3,

$$B_\mu \;=\; \frac{1}{2}\,\boldsymbol{\tau}\cdot\boldsymbol{b}_\mu \;=\; \frac{1}{2}\sum_k \tau^k b^k_\mu \;=\; \frac{1}{2}\begin{pmatrix} b_3 & b_1 - ib_2 \\ b_1 + ib_2 & -b_3 \end{pmatrix}. \tag{8.44}$$

We call the fields b_i the gauge fields of the $SU(2)$ symmetry. We need three fields rather than one, because $SU(2)$ has three generators.

In terms of the covariant derivative the Lagrangian is

$$\mathcal{L} = \overline{\psi}\;\left(i\gamma^\mu\, D_\mu - \mathbb{1}m\right)\;\psi \tag{8.45}$$

In order to obtain a Lagrangian that is invariant, we need the gauge transformation to take the form

$$D_\mu\psi \rightarrow D'_\mu\psi' = G\left(D_\mu\psi\right) \tag{8.46}$$

Inserting the expression for the covariant derivative, we find for the term on the left side of the equality

$$\begin{aligned} D'_\mu\psi' \;&=\; \left(\partial_\mu + igB'_\mu\right)\;\psi' \\ &=\; G\left(\partial_\mu\psi\right) + \left(\partial_\mu G\right)\psi + igB'_\mu G\psi \end{aligned} \tag{8.47}$$

while the right side is

$$G\left(D_\mu\psi\right) \;=\; G\partial_\mu\psi + igGB_\mu\psi \tag{8.48}$$

Comparing these two expressions we find that invariance implies that

$$igB'_\mu\left(G\psi\right) = igG\left(B_\mu\psi\right) - \left(\partial_\mu G\right)\psi \tag{8.49}$$

Since this expression must hold for all values of the field ψ, we can omit the field ψ from this expression. If we subsequently multiply both sides of the equation on the right by G^{-1} we find for the transformation of the gauge field

$$B'_\mu \;=\; G B_\mu G^{-1} \;+\; \frac{i}{g}\left(\partial_\mu G\right) G^{-1}\;. \tag{8.50}$$

Although this looks rather complicated we can again try to interpret this by comparing to the case of electromagnetism. For $G_{em} = e^{i\alpha(x)}$ we have

$$\begin{aligned} A'_\mu &= G_{em} A_\mu G_{em}^{-1} + \frac{i}{q}\left(\partial_\mu G_{em}\right) G_{em}^{-1} \\ &= A_\mu - \frac{1}{q}\partial_\mu \alpha \end{aligned} \tag{8.51}$$

which is exactly the transformation rule that we had before.

We see that for an $SU(2)$ symmetry the transformation of the gauge field B_μ involves both a rotation and a gradient. The gradient term was already present in QED. The rotation term is new. It arises due to the non-commutativity of the elements of $SU(2)$. If we write out the gauge field transformation formula in the components of the real vector fields

$$b_\mu^{k\,\prime} = b_\mu^k - \epsilon_{klm}\,\alpha^l\, b^m - \frac{1}{g}\partial_\mu \alpha^k \tag{8.52}$$

we can see that there is a coupling *between* the different components of the field. We call this the self-coupling. (To derive this start from Eq. 8.50, consider infinitesimal small $\boldsymbol{\alpha}(x)$ and use the commutation relation of the SU(2) generators, $[\tau_i, \tau_j] = 2\epsilon_{ijk}\tau_k$.)

The effect of the self-coupling becomes clear if one considers the kinetic term of the $SU(2)$ gauge field. Analogous to the QED case, the three new fields require their own free Lagrangian, which we write as

$$\mathcal{L}_b^{\text{free}} = -\frac{1}{4}\sum_l F_l^{\mu\nu} F_{\mu\nu,l} \;=\; -\frac{1}{4}\boldsymbol{F}^{\mu\nu}\cdot\boldsymbol{F}_{\mu\nu}\;. \tag{8.53}$$

Mass terms like $m^2 b^\nu b_\nu$ are again excluded by gauge invariance: as for the $U(1)$ symmetry, the gauge fields must be massless. However, while for the photon the field tensor in the kinetic term was given by $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$, this form does not work here because it would break the symmetry. Rather, the individual components of the field tensor are given by

$$F_l^{\mu\nu} = \partial^\nu b_l^\mu - \partial^\mu b_l^\nu + g\,\epsilon_{jkl}\, b_j^\mu\, b_k^\nu \tag{8.54}$$

or in vector notation

$$\boldsymbol{F}^{\mu\nu} \;=\; \partial^\mu \boldsymbol{b}^\nu - \partial^\nu \boldsymbol{b}^\mu \;-\; g\,\boldsymbol{b}^\mu \times \boldsymbol{b}^\nu \tag{8.55}$$

As a consequence of the last term the Lagrangian contains contributions with 2, 3 and 4 factors of the b-field. These couplings are respectively referred to as bilinear, trilinear and quadrilinear couplings. In QED there is only the bilinear photon propagator term. In the $SU(2)$ theory there are self interactions by a 3-gauge boson vertex and a 4 gauge boson vertex.

Summarizing, we started from the free Lagrangian for a doublet $\psi = \begin{pmatrix} p \\ n \end{pmatrix}$ of two fields with equal mass,

$$\mathcal{L}_\psi^{free} = \overline{\psi}\ (i\gamma^\mu \partial_\mu - m)\ \psi$$

This Lagrangian has a global $SU(2)$ symmetry. We then hypothesized a local $SU(2)$ phase invariance which we could implement by making the replacement $\partial_\mu \to D_\mu = \partial_\mu + igB_\mu$ with $B_\mu = \frac{1}{2}\boldsymbol{\tau} \cdot \boldsymbol{b}_\mu$. The full Lagrangian of the theory (which is called the Yang-Mills theory) is then given by

$$\begin{aligned} \mathcal{L}_{SU(2)} &= \ \overline{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}\boldsymbol{F}^{\mu\nu} \cdot \boldsymbol{F}_{\mu\nu} \\ &= \overline{\psi}(i\gamma^\mu \partial_\mu - m)\psi - g\boldsymbol{J}^\mu \boldsymbol{b}_\mu - \frac{1}{4}\boldsymbol{F}^{\mu\nu} \cdot \boldsymbol{F}_{\mu\nu} \\ &\equiv\ \mathcal{L}_\psi^{\text{free}}\ +\ \mathcal{L}^{\text{interaction}}\ +\ \mathcal{L}_b^{\text{free}} \end{aligned} \tag{8.56}$$

where we now absorbed the coupling constant g in the definition of the conserved current,

$$\boldsymbol{J}^\mu = \frac{g}{2}\overline{\psi}\gamma^\mu \boldsymbol{\tau}\psi \tag{8.57}$$

Comparing this to the QED Lagrangian

$$\mathcal{L}_{U(1)} = \mathcal{L}_\psi^{free} - A_\mu \cdot J^\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{8.58}$$

(with the electromagnetic current $J^\mu = q\overline{\psi}\gamma^\mu\psi$), we see that instead of one field, we now have three new fields. Furthermore, the kinetic term is more complicated and gives rise to self-coupling vertices with three and four b-field lines.

8.8 Isospin, QCD and weak isospin

Yang and Mills tried to extend the isospin symmetry in the proton-neutron system to a local isospin symmetry in the hope of formulating a viable theory of strong interactions. In order to do so, they needed three new gauge fields, and obviously they wondered what those were. They had to be massless vector bosons that couple to the proton and neutron. Clearly, it could not be the three pions, since those are pseudo-scalar particles rather then vector bosons.

As we know now, the $SU(2)$ theory cannot describe the strong interaction. Rather the strong interactions follow from an $SU(3)$ symmetry. The implementation is a carbon copy of the Yang-Mills theory for $SU(2)$ symmetry. The mediators of the force are the eight massless gluons, corresponding to the 8 generators of the fundamental representation of $SU(3)$, namely

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$
$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

In this case, too, the Lagrangian contains self-coupling terms for the gauge fields. The strong interaction is discusses in the Particle Physics II course.

The isospin symmetry in the proton-neutron is a *flavour symmetry*. Extended to the system of all other hadrons it is essentially just the symmetry between u and d quarks. We know that such a symmetry only exists if we ignore electromagnetic interactions, and the small difference in mass between the u and the d quark (or the proton and the neutron). Since the symmetry is not exact, we call it an *approximate* symmetry.

Although the Yang-Mills isospin theory is of no real use to the proton-neutron system, it turns out to be exactly what is needed to describe the weak interactions. For historical reasons the local $SU(2)$ symmetry applied to the Lagrangian of Dirac fermion doublets, discussed in the next lecture, is sometimes called 'weak isospin'. It should certainly not be confused with the $u - d$ flavour symmetry. In contrast with the flavour symmetry, gauge symmetries are *exact* symmetries of the Lagrangian.

8.9 The origin of the name "gauge theory"

The idea of gauge invariance as a dynamical principle is due to Hermann Weyl. He called it "*eichinvarianz*" ("gauge" = "calibration"). Weyl tried to find a geometrical basis for both gravitation and electromagnetism.[3] Although his effort was unsuccessful the terminology survived. His idea is summarized here.

Consider a change in a function $f(x)$ between point x_μ and point $x_\mu + \delta x_\mu$. If the space has a uniform scale we expect simply:

$$f(x + \delta x) = f(x) + \partial^\mu f(x)\, \delta x_\mu \tag{8.59}$$

But if in addition the scale, or the unit of measure, for f changes by a factor $(1 + S^\mu \delta x_\mu)$ between x and $x + \delta x$, then the value of f becomes:

$$\begin{aligned} f(x + \delta x) &= \left(f(x) + \partial^\mu f(x) \delta x_\mu\right) \, (1 + S^\nu \delta x_\nu) \\ &= f(x) + \left(\partial^\mu f(x) + f(x) S^\mu\right) \delta x_\mu + O(\delta x)^2 \end{aligned} \tag{8.60}$$

[3] H. Weyl, *Z. Phys.* **56**, 330 (1929)

So, to first order, the increment is:

$$\delta f = (\partial^\mu + S^\mu)\, f\, \delta x_\mu \tag{8.61}$$

In other words Weyl introduced a modified differential operator by the replacement: $\partial^\mu \to \partial^\mu + S^\mu$.

One can see this in analogy in electrodynamics in the replacement of the momentum by the canonical momentum parameter: $p^\mu \to p^\mu - qA^\mu$ in the Lagrangian, or in Quantum Mechanics: $\partial^\mu \to \partial^\mu + iqA^\mu$, as was discussed in the earlier lectures. In this case the "scale" is $S^\mu = iqA^\mu$. If we now require that the laws of physics are invariant under a change:

$$(1\, +\, S^\mu\, \delta x_\mu)\ \to\ (1\, +\, iq\, A^\mu\, \delta x_\mu)\ \approx\ \exp\left(iq\, A^\mu\, \delta x_\mu\right) \tag{8.62}$$

then we see that the change of scale gets the form of a change of a phase. When Weyl later studied the invariance under phase transformations, he kept using the term "gauge invariance".

Exercises

Exercise 8.1 (Classical Euler-Lagrange equations)
Use Hamilton's principle in Eq. (8.3) to derive the Euler-Lagrange equation of motion (Eq. (8.4)).
Hint: Write down the change in the action δS for an arbitrary small change in the trajectory $(\delta q, \delta\dot{q})$. Now use integration by parts to replace $\delta\dot{q}$ by δq. The boundary condition (the fact that we know where the trajectory starts and ends) makes that $\delta q(t_0) = \delta q(t_1) = 0$. The Euler-Lagrange equations then follow from the fact that for the trajectory that minimizes the action, the change in the action should be zero, independent of δq.

Exercise 8.2 (Lagrangians versus equations of motion)

(a) Show that the Euler-Lagrange equations of the Lagrangian

$$\mathcal{L}_{\mathrm{KG}}^{\mathrm{free}} = \frac{1}{2}\left(\partial_\mu\phi\right)\left(\partial^\mu\phi\right) - \frac{1}{2}m^2\phi^2 \tag{8.63}$$

of a *real* scalar field ϕ leads to the Klein-Gordon equation.

(b) Show the same for a *complex* scaler field starting from the Lagrangian

$$\mathcal{L}_{\mathrm{KG}}^{\mathrm{free}}\ =\ \left(\partial_\mu\phi\right)^*\left(\partial^\mu\phi\right) - m^2\phi^*\phi \tag{8.64}$$

taking ϕ and ϕ^* as the (two) independent fields. (Alternatively, you can take the real and imaginary part of ϕ. Note that you obtain two equations of motion, one for ϕ and one for ϕ^*.)

(c) Show that the Euler-Lagrange equations for the Lagrangian

$$\mathcal{L}_{\text{Dirac}}^{\text{free}} = i\overline{\psi}\gamma_\mu\partial^\mu\psi - m\overline{\psi}\psi \tag{8.65}$$

leads to the Dirac equations for ψ *and* for $\overline{\psi}$. Note again that you need to consider ψ and $\overline{\psi}$ as independent fields.

(d) Show that the Lagrangian

$$\mathcal{L} = \mathcal{L}_{EM} = -\frac{1}{4}\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right)\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) - j^\mu A_\mu = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - j^\mu A_\mu \tag{8.66}$$

leads to the Maxwell equations:

$$\partial_\mu\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) = j^\nu \tag{8.67}$$

Hence the current is conserved ($\partial_\nu j^\nu = 0$), since $F^{\mu\nu}$ is antisymmetric.

Exercise 8.3 (Global phase invariance)

(a) Show that the Dirac Lagrangian remains invariant under the *global* phase transformation

$$\psi(x) \to e^{iq\alpha}\psi(x) \quad ; \quad \overline{\psi}(x) \to e^{-iq\alpha}\overline{\psi}(x) \tag{8.68}$$

(b) Noether's Theorem: consider an infinitesimal transformation: $\psi \to \psi' = e^{i\alpha}\psi \approx (1 + i\alpha)\psi$. Show that the requirement of invariance of the Dirac Lagrangian $(\delta\mathcal{L}(\psi, \partial_\mu\psi, \overline{\psi}, \partial_\mu\overline{\psi}) = 0)$ leads to the conserved current

$$j^\mu = i\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi\right)}\psi - \overline{\psi}\frac{\mathcal{L}}{\partial\left(\partial_\mu\overline{\psi}\right)}\right) = -\overline{\psi}\gamma^\mu\psi \tag{8.69}$$

Consequently, the charge carried by the current is conserved.

Exercise 8.4 (Local phase invariance)

(a) (i) Start with the Lagrange density for a complex Klein-Gordon field

$$\mathcal{L} = \left(\partial_\mu\phi\right)^*\left(\partial^\mu\phi\right) - m^2\phi^*\phi \tag{8.70}$$

and show that a **local** field transformation:

$$\phi(x) \to e^{iq\alpha(x)}\phi(x) \quad ; \quad \phi^*(x) \to e^{-iq\alpha(x)}\phi^*(x) \tag{8.71}$$

does **not** leave the Lagrangian invariant.

(ii) Substitute the covariant derivative $\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu$ in the Lagragian. Show that the Lagrangian now remains invariant under the gauge transformation, provided that the additional field transforms as

$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) - \partial_\mu \alpha(x) \ . \tag{8.72}$$

(b) (i) Start with the Lagrange density for a Dirac field

$$\mathcal{L} = i\overline{\psi}\gamma^\mu \partial_\mu \psi - m\overline{\psi}\psi \tag{8.73}$$

and show that a **local** field transformation:

$$\psi(x) \to e^{iq\alpha(x)}\psi(x) \qquad ; \qquad \overline{\psi}(x) \to e^{-iq\alpha(x)}\overline{\psi}(x) \tag{8.74}$$

also does **not** leave the Lagrangian invariant.

(ii) Again make the replacement $\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu$ and show that the Lagrangian remains invariant provided that A_μ transforms are above.

Exercise 8.5 (Extra exercise, not obligatory!)
Consider an infinitesimal gauge transformation:

$$G = 1 + \frac{i}{2}\boldsymbol{\tau} \cdot \boldsymbol{\alpha} \qquad |\boldsymbol{\alpha}| \ll 1 \tag{8.75}$$

Use the general transformation rule for B'_μ and use $B_\mu = \frac{1}{2}\boldsymbol{\tau} \cdot \boldsymbol{b}_\mu$ to demonstrate that the fields transform as:

$$\boldsymbol{b}'_\mu = \boldsymbol{b}_\mu - \boldsymbol{\alpha} \times \boldsymbol{b}_\mu - \frac{1}{g}\partial_\mu \boldsymbol{\alpha} \tag{8.76}$$

(Hint: use the Pauli vector identity $(\boldsymbol{\tau} \cdot \boldsymbol{a})(\boldsymbol{\tau} \cdot \boldsymbol{b}) = \boldsymbol{a} \cdot \boldsymbol{b} + i\boldsymbol{\tau} \cdot (\boldsymbol{a} \times \boldsymbol{b})$.*)*

Lecture 9

Electroweak Theory

In the previous lecture we have seen how imposing a local gauge symmetry requires a modification of the free Lagrangian in such a way that a theory with interactions is obtained. We studied two symmetries, namely

- local $U(1)$ gauge invariance:

$$\overline{\psi}\;(i\gamma^\mu D_\mu - m)\;\psi \;=\; \overline{\psi}\;(i\gamma^\mu\partial_\mu - m)\;\psi \;-\; q\,\underbrace{\overline{\psi}\,\gamma^\mu\,\psi}_{J^\mu}\,A_\mu \tag{9.1}$$

- local $SU(2)$ gauge invariance for a Dirac spinor douplet Ψ:

$$\overline{\Psi}\;(i\gamma^\mu D_\mu - m)\;\Psi \;=\; \overline{\Psi}\;(i\gamma^\mu\partial_\mu - m)\;\Psi \;-\; \frac{g}{2}\,\underbrace{\overline{\Psi}\,\gamma^\mu\,\boldsymbol{\tau}\,\Psi}_{\boldsymbol{J}^\mu}\,\boldsymbol{b}_\mu\,. \tag{9.2}$$

For the $U(1)$ symmetry we can identify the A_μ field as the photon. The Feynman rules for QED, as we discussed them in previous lectures, follow automatically.

Yang and Mills implemented the $SU(2)$ local gauge symmetry, hoping that they could derive the strong interaction from proton-neutron isospin symmetry. Although that did not work, we now show that the $SU(2)$ gauge symmetry is still useful, but then to explain the *weak* interaction. (The strong interaction follows from $SU(3)$ gauge invariance.)

Before we continue with $SU(2)$, we make a small modification to the interaction terms above. First consider the $U(1)$ symmetry. Every fermion field has its own charge. Within the Standard Model we cannot explain why the charge of an up quark is two-thirds of the charge of an electron. This is why the symbol q appears in the interaction term above: it is a dimensionless parameter that signifies the strength of the interaction and it can be different for different fields.

At this point it is customary to introduce a *charge operator* Q which acts as the *generator* of the $U(1)$ symmetry group for electromagnetic interactions. It appears in the field transformation rule as

$$\psi' \;=\; e^{i\alpha(x)Q}\,\psi\,. \tag{9.3}$$

We define the conserved current as

$$J^{\mu}_{\rm EM} \;=\; \overline{\psi}\,\gamma^{\mu}\,Q\,\psi \tag{9.4}$$

and write the interaction term in the Hamiltonian as

$$\mathcal{L}_{\rm int} \;=\; -g_{\rm EM}\,J^{\mu}_{\rm EM}\,A_{\mu}\;, \tag{9.5}$$

where we now use the *same* coupling $g_{\rm EM}$ for all fields. The fields ψ are eigenstates of the charge operator Q with an eigenvalue equal to the charge in units of the positron charge. (Why we do this will become clear later.)

A similar strategy is taken for the isospin symmetry. Rather than $\boldsymbol{\tau}$ as the generator we consider an operator $\boldsymbol{T}$ which for unit isospin charge is given by $\boldsymbol{T} = \boldsymbol{\tau}/2$. It enters into the douplet transformation rule as

$$\Psi' \;=\; e^{i\boldsymbol{\alpha}(x)\boldsymbol{T}}\;\Psi \tag{9.6}$$

and finds its way into the conserved current as

$$\boldsymbol{J}^{\mu}_{T} \;=\; \overline{\Psi}\,\gamma^{\mu}\,\boldsymbol{T}\,\Psi\;, \tag{9.7}$$

while the interaction terms is given by

$$\mathcal{L}_{\rm int} \;=\; -g\;\boldsymbol{J}^{\mu}_{\rm T}\,\boldsymbol{b}_{\mu}\;. \tag{9.8}$$

When, in the following sections, we consider $SU(2)$ symmetry to generate the weak interaction, the coupling constant g is taken to be the same for all douplets, but the physical fields (which are eigenstates of T_3) each have their own value of 'weak-isospin charge', the eigenvalue for T_3. In the Standard Model, this eigenvalue is always $\pm 1/2$.

The coupling constants $g_{\rm EM}$ and g in the interaction terms are dimensionless. For the electromagnetic interaction the coupling is related to the unit charge and the fine-structure constant as

$$g_{\rm EM} = \frac{e}{\sqrt{\epsilon_0 \hbar c}} = \sqrt{4\pi\alpha}\;. \tag{9.9}$$

In our system of units $\epsilon_0\hbar c = 1$. Therefore, in the following we substitute $g_{\rm EM} = e$, as you will find in most textbooks.

9.1 SU(2) symmetry for left-handed douplets

We define for any Dirac field ψ the left- and right-handed *chiral* projections,

$$\psi_L \equiv \frac{1}{2}\left(1-\gamma^5\right)\,\psi \qquad \text{and} \qquad \psi_R \equiv \frac{1}{2}\left(1+\gamma^5\right)\,\psi\;. \tag{9.10}$$

As we have seen in Lecture 7, for particles with $E \gg m$ these correspond to the negative and positive helicity states, respectively. Using the fact that (see exercise 9.1)

$$\overline{\psi}\,\gamma^{\mu}\,\psi \;=\; \overline{\psi}_L\,\gamma^{\mu}\,\psi_L \;+\; \overline{\psi}_R\,\gamma^{\mu}\,\psi_R \tag{9.11}$$

where the chiral projections of the adjoint spinors are given by

$$\overline{\psi}_L \;\equiv\; \overline{(\psi_L)} \;=\; \frac{1}{2}\,\overline{\psi}\,\left(1+\gamma^5\right) \qquad \text{and} \qquad \overline{\psi}_R \;\equiv\; \overline{(\psi_R)} \;=\; \frac{1}{2}\,\overline{\psi}\,\left(1-\gamma^5\right) \tag{9.12}$$

we can rewrite the Dirac Lagrangian for ψ as

$$\mathcal{L} \;=\; \overline{\psi}_R\,(i\gamma^\mu\partial_\mu)\,\psi_R \;+\; \overline{\psi}_L\,(i\gamma^\mu\partial_\mu)\,\psi_L \;-\; \overline{\psi}_R\,m\,\psi_L \;-\; \overline{\psi}_L\,m\,\psi_R \tag{9.13}$$

The mass terms 'mix' the left- and right-handed components. That is incovenient for what we are going to do next. Therefore, in the following we consider only massless fields and deal with non-zero mass later.

Let us now introduce the following doublets for the left-handed chirality states of the leptons and quarks in the first family:

$$\Psi_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \qquad \text{and} \qquad \Psi_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \tag{9.14}$$

We call these "weak isospin" doublets. Again, Ψ is not a Dirac spinor, but a doublet of Dirac spinors. Consider the Lagrangian for the electron and neutrino and verify that it can be written as (*c.f.* Eq. 8.35)

$$\begin{aligned} \mathcal{L} = \;& \overline{e}_R\, i\gamma^\mu\partial_\mu\, e_R \;+\; \overline{\nu}_R\, i\gamma^\mu\partial_\mu\, \nu_R \;+ \\ & +\, (\overline{\nu}_L\; \overline{e}_L) \begin{pmatrix} i\gamma^\mu\partial_\mu & 0 \\ 0 & i\gamma^\mu\partial_\mu \end{pmatrix} \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \end{aligned} \tag{9.15}$$

Now it comes: We impose the $SU(2)$ gauge symmetry on the *left-handed* doublets only. That is, we require that the Lagrangian be invariant for local rotations of the doublet. To do this we need to ignore that the two components of a doublet have different charge, a problems that we will clearly need to deal with later. As in the Yang-Mills theory, we also need to ignore that they have different mass, which is another motivation for only considering massless fields. The right-handed fields are not combined into doublets. They are *singlets* under the $SU(2)$ transformation.

The fact that we only impose the gauge symmetry on left-handed states leads to a weak interaction that is completely left-right asymmetric. This is why it is referred to as *maximal violation of parity.*

To construct the weak $SU(2)_L$ theory[1] we start again with the free Dirac Lagrangian and we impose $SU(2)$ symmetry on the weak isospin doublets:

$$\mathcal{L}_{free} = \overline{\Psi}_L\, i\gamma^\mu\partial_\mu\, \Psi_L \tag{9.16}$$

After introducing the covariant derivative

$$\partial_\mu \to D_\mu \;=\; \partial_\mu + igB_\mu \qquad \text{with} \qquad B_\mu \;=\; \boldsymbol{T}\cdot\boldsymbol{b}_\mu \tag{9.17}$$

[1] The subscript L is used to indicate that we only consider $SU(2)$ transformations of the left-handed doublet.

the Dirac equation obtains an interaction term,

$$\mathcal{L}_{\text{free}} \;\rightarrow\; \mathcal{L}_{\text{free}} \;-\; g\,\boldsymbol{b}_\mu \cdot \boldsymbol{J}^\mu_{\text{weak}} \tag{9.18}$$

where the weak current is

$$J^\mu_{\text{weak}} \;=\; \overline{\Psi}_L\,\gamma^\mu\,\boldsymbol{T}\,\Psi_L \tag{9.19}$$

This is just a carbon copy of the Yang-Mills theory for "strong isospin" in the previous lecture.[2]

The gauge fields (b^1, b^2, b^3) couple to the left-handed doublets defined above. However, the particles in our real world do not appear as doublets: we scatter electrons, not electron-neutrino doublets. We now show, as Glashow first did in 1961, how these gauge fields can be recast into the 'physical' fields of the 3 vector bosons W^+, W^-, Z^0 in order to have them interact with currents of the physical electrons and neutrinos.

9.2 The Charged Current

We choose for the representation of the $SU(2)$ generators the Pauli spin matrices,

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,. \tag{9.20}$$

The generators τ_1 and τ_2 mix the components of a doublet, while τ_3 does not. We define the fields $W^\pm$ as

$$W^\pm_\mu \;\equiv\; \frac{1}{\sqrt{2}}\,\left(b^1_\mu \mp i\,b^2_\mu\right) \tag{9.21}$$

The $\pm$ index on the W refers to the *electric* charge. However, at this point we have not yet shown that these fields are indeed electrically charged: That would require us to look at the coupling of the W fields to the photon, which we will not do as part of these lectures. As an alternative, we now show that these W fields couple to charge-lowering and charge-raising currents. Charge conservation at each Feynman diagram vertex then implies the charge of the gauge boson.

We define the charged current term of the interaction Lagrangian as

$$\mathcal{L}_{CC} \;\equiv\; -\,g\,b^1_\mu J^{1\mu} \;-\; g\,b^2_\mu J^{2\mu} \tag{9.22}$$

with

$$J^{1\mu} \;=\; \overline{\Psi}_L\,\gamma^\mu\,\frac{\tau_1}{2}\,\Psi_L \qquad J^{2\mu} \;=\; \overline{\Psi}_L\,\gamma^\mu\,\frac{\tau_2}{2}\,\Psi_L \tag{9.23}$$

As you will show in exercise 9.2 we can rewrite the charged current Lagrangian as

$$\mathcal{L}_{CC} \;=\; -g\,W^+_\mu\,J^{+\mu} - g\,W^-_\mu\,J^{-\mu} \tag{9.24}$$

[2] Note that in terms of physics strong and weak isospin have nothing to do with one another. It is just that we use the same math!

with

$$J^{\mu,\pm} \;=\; \frac{1}{\sqrt{2}}\,\overline{\Psi}_L\,\gamma^\mu\,\tau^\pm\,\Psi_L \tag{9.25}$$

and $\tau^\pm \;=\; \frac{1}{2}(\tau_1 \pm i\tau_2)$, or in our representation

$$\tau^+ \;=\; \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \tau^- \;=\; \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\,. \tag{9.26}$$

The leptonic currents can then be written as

$$J^{+\mu} = \frac{1}{\sqrt{2}}\,\overline{\nu}_L\,\gamma^\mu\,e_L \qquad \text{and} \qquad J^{-\mu} = \frac{1}{\sqrt{2}}\,\overline{e}_L\,\gamma^\mu\,\nu_L \tag{9.27}$$

or written out with the left-handed projection operators:

$$J^{+\mu} = \frac{1}{\sqrt{2}}\,\overline{\nu}\,\frac{1}{2}\left(1+\gamma^5\right)\,\gamma^\mu\,\frac{1}{2}\left(1-\gamma^5\right)\,e \tag{9.28}$$

and similar for $J^{-\mu}$. Verify for yourself that

$$\left(1+\gamma^5\right)\gamma^\mu\left(1-\gamma^5\right) \;=\; 2\gamma^\mu\left(1-\gamma^5\right) \tag{9.29}$$

such that we can rewrite the leptonic charge raising current as

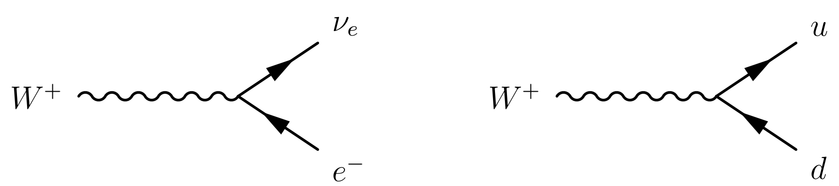
$$\boxed{J^{+\mu} = \frac{1}{2\sqrt{2}}\,\overline{\nu}\,\gamma^\mu\left(1-\gamma^5\right)e} \tag{9.30}$$

and the leptonic charge lowering current as

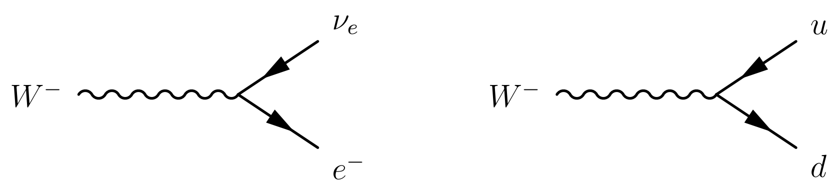
$$\boxed{J^{-\mu} = \frac{1}{2\sqrt{2}}\,\overline{e}\,\gamma^\mu\left(1-\gamma^5\right)\nu}\,. \tag{9.31}$$

Remembering that a vector interaction has an operator γ^μ in the current and an axial vector interaction a term $\gamma^\mu\gamma^5$, we recognize in the charged weak interaction the famous "V-A" interaction. The story for the quark doublet is identical. Drawn as diagrams, the charged currents then look as follows:

Charge raising:



Charge lowering:



9.3 The Neutral Current

The third component of the weak isospin gauge field leads to a neutral current interaction,

$$\mathcal{L}_{\text{int}} = -g\, b^3_\mu J^\mu_3 \tag{9.32}$$

with b^3_μ the third gauge boson (another real vector field) and the conserved current given by

$$J^\mu_3 = \overline{\Psi}_L \, \gamma^\mu \, \frac{\tau_3}{2} \, \Psi_L \, . \tag{9.33}$$

It is now tempting to identify this third component as the Z^0 boson and simply add the electromagnetic interaction term that we had previously constructed with a $U(1)$ symmetry with the electromagnetic charge operator Q as generator.

However, this is not a valid way to extend the symmetry of the Lagrangian: the left-handed doublets that we have constructed are not eigenstates of Q since they mix fields with different charge. Therefore, our $SU(2)_L$ invariant Lagrangian cannot be symmetric under a transformation with Q as generator.

The solution is to start from another $U(1)$ gauge symmetry, called 'weak hypercharge'. We denote its generator with the symbol Y and require that it commutes with the $SU_L(2)$ generators. The different members of the isospin multiplet then by construction obtain the same value of hypercharge.

We denote the combined symmetry by $SU(2)_L \otimes U(1)_Y$. Under this symmetry a left-handed doublet transform as

$$\Psi_L \;\to\; \Psi'_L \;=\; \exp\left[\, i\, \boldsymbol{\alpha}(x)\, \boldsymbol{T} \;+\; i\, \beta(x)\, Y \,\right] \Psi_L \, , \tag{9.34}$$

where $\boldsymbol{T} = \boldsymbol{\tau}/2$ are the $SU(2)$ generators and Y is the generator for $U(1)_Y$. At the same time, the right-handed components of the fields transform only under hypercharge,

$$\Psi_R \;\to\; \Psi'_R \;=\; e^{i\beta(x)Y}\, \Psi_R \, . \tag{9.35}$$

The conserved current corresponding to the $U(1)_Y$ symmetry is

$$J^\mu_Y = \overline{\Psi}\, \gamma^\mu \, Y \, \Psi \, . \tag{9.36}$$

The Lagrangian following from local $SU(2)_L \otimes U(1)_Y$ symmetry takes the form (see e.g. Halzen and Martin, Chapter 13)

$$\boxed{\mathcal{L}_{EW} \;=\; \mathcal{L}_{free} \;-\; g\, \boldsymbol{J}^\mu_T \cdot \boldsymbol{b}_\mu \;-\; \frac{g'}{2} J^\mu_Y \, a_\mu} \, , \tag{9.37}$$

where a_μ is the gauge field corresponding to $U(1)_Y$ and $g'/2$ is its coupling strength. The factor 2 appears just because of a convention.

The transformations corresponding to T_3 and Y both lead to neutral current interactions. As a result the gauge boson fields can actually 'mix'. Neither of them couples specifically to the electromagnetic charge. Therefore, an important question is whether we can recast these fields such that one becomes a physical 'photon' field A^μ that couples to the fermion fields via the charge operator Q and the other one becomes the Z^0 boson.

Weinburg and Salam showed indepently that the answer to this has everything to do with the Higgs mechanism. The gauge fields $\boldsymbol{b}^\mu$ and a^μ in the formalism above are all massless. An explicit mass term ($\mathcal{L}_M = K b_\mu b^\mu$) would break the gauge invariance of the theory. Their masses and the masses of all fermions can be generated in a mechanism that is called *spontaneous symmetry breaking* and involves a new scalar field, the Higgs field. The main idea of the Higgs mechanism is that the Lagrangian retains the full gauge symmetry, but that the ground state (or 'vacuum'), i.e. the state from which we start perturbation theory, is not symmetric. We discuss this in detail in Lectures 11 and Lectures 12. For now, you will have to do with some conjectures.

The Higgs mechanism leads to a solution in which *three out of four* vector fields ($\boldsymbol{b}_\mu$ and a_μ) acquire a mass. Two of the massive fields can be written as charge raising and charge lowering fields $W^\pm$ as we did before,

$$W^\pm_\mu \;=\; \frac{1}{\sqrt{2}}\left(b^1_\mu \mp i\, b^2_\mu\right) \;. \tag{9.38}$$

The physical neutral fields are linear combinations of the T_3 and Y gauge fields, written as

$$\begin{aligned} A_\mu &= \;\; a_\mu \cos\theta_W + b^3_\mu \sin\theta_W \\ Z_\mu &= -a_\mu \sin\theta_W + b^3_\mu \cos\theta_W \end{aligned} \tag{9.39}$$

where θ_W is called the weak mixing angle.

The Higgs mechanism predicts that in order that the massless field becomes the photon, the quantum number for the charge operator is related to the $SU(2)_L$ and $U(1)_Y$ quantum numbers by

$$Q \;=\; T_3 \;+\; \frac{Y}{2} \tag{9.40}$$

This relation is also called the Gell-Mann-Nishima relation. Interpreted in terms of quantum numbers, Q is the electromagnetic charge, Y is the hypercharge and T_3 is the charge associated to the third generator of $SU(2)$ (*i.e.* $\tau_3/2$). The quantum number for the hypercharge of the various fermion fields is *not* predicted by the theory, just as the electromagnetic charges are not predicted.

As an example, consider the neutrino-electron doublet. A left-handed neutrino state has $T_3 = 1/2$ while the left-handed electron has $T_3 = -1/2$. (If you don't understand this, consider the eigenvalues of the eigenvectors

$$\Psi_{L,\nu} = \begin{pmatrix} \nu \\ 0 \end{pmatrix} \qquad \text{and} \qquad \Psi_{L,e} = \begin{pmatrix} 0 \\ e \end{pmatrix}$$

for the $\tau_3/2$ generator.) The right-handed electron is a singlet under $SU(2)_L$ and has $T_3 = 0$. Given a coupling constant e, the observed electromagnetic charge of the electron is -1. Therefore, the hypercharge of the right-handed electron is -2 while the hypercharge of the left-handed electron and neutrino are both -1. (The latter two must be equal, since the $SU(2)_L$ doublet is a singlet under $U(1)_Y$.)

Expressing the interactions terms of the b_3^μ and a^μ fields in the Lagrangian above in terms of the physical fields, we find

$$\begin{aligned}
-gJ_3^\mu b_\mu^3 - \frac{g'}{2}J_Y^\mu a_\mu &= -\left(g\sin\theta_W J_3^\mu + g'\cos\theta_W\frac{J_Y^\mu}{2}\right)A_\mu \\
&\quad -\left(g\cos\theta_W J_3^\mu - g'\sin\theta_W\frac{J_Y^\mu}{2}\right)Z_\mu \\
&\equiv -eJ_{EM}^\mu A_\mu - g_Z J_{NC}^\mu Z_\mu
\end{aligned} \tag{9.41}$$

where in the last line we defined the currents and coupling constants associated to the physical fields.

A direct consequence of Eq. (9.40) is that also the currents are related, namely by

$$\boxed{J_{EM}^\mu = J_3^\mu + \frac{1}{2}J_Y^\mu} \ . \tag{9.42}$$

Comparing this to Eq. (9.41) we find that

$$e = g\sin\theta_W = g'\cos\theta_W \tag{9.43}$$

This relation implies that there are only two independent parameters. In particular, the weak mixing angle is related to the $SU(2)_L$ and $U(1)_Y$ coupling constants by

$$g'/g = \tan\theta_W \tag{9.44}$$

Analogously, we find for the Z^0 interaction term,

$$\begin{aligned}
&-\left(g\cos\theta_W J_3^\mu - \frac{g'}{2}\sin\theta_W\cdot 2\,(J_{EM}^\mu - J_3^\mu)\right)Z_\mu \\
=\ &\ldots \\
=\ &-\frac{e}{\cos\theta_W\sin\theta_W}\left(J_3^\mu - \sin^2\theta_W J_{EM}^\mu\right)Z_\mu
\end{aligned}$$

Consequently, expressed in J_3 and J_{EM} the neutral current is

$$\boxed{J_{NC}^\mu = J_3^\mu - \sin^2\theta_W\ J_{EM}^\mu} \tag{9.45}$$

and its coupling constant becomes

$$g_Z = \frac{e}{\cos\theta_W\sin\theta_W} = \frac{g}{\cos\theta_W}\ . \tag{9.46}$$

9.4 Couplings for $Z \to ff$

Let us now take a closer look at the neutral current in Eq. (9.45). The J_3 current only involves the left-handed doublets, while the electromagnetic current couples both to left- and right-handed fields. Therefore, we conclude that the neutral current cannot be purely left-handed.

We defined the left-handed current in terms of a left-handed neutrino-electron (or up-down quark) douplet Ψ_L as

$$J_3^\mu \;=\; \overline{\Psi}_L\, \gamma^\mu\, T_3\, \Psi_L\;. \tag{9.47}$$

In analogy with the procedure that we applied for the charged current, we can now rewrite this in terms of the two fermion fields in the douplet using the explicit representation of T_3 as

$$J_3^\mu \;=\; \overline{\nu}_L\, \gamma^\mu\, T_3^\nu\, \nu_L \;+\; \overline{e}_L\, \gamma^\mu\, T_3^e\, e_L\;, \tag{9.48}$$

where the neutrino and electron weak isospin charges are $T_3^\nu = 1/2$ and $T_3^e = -1/2$. Since we scatter particles rather than douplets, we consider the currents for the neutrino and electron separately.

We can generalize this and write for any fermion field ψ_f

$$\begin{aligned} J^\mu_{3,f} &= \overline{\psi}_L\, \gamma^\mu\, T_3^f \psi_L \\ &= \frac{1}{2}\overline{\psi}\, \gamma^\mu\, (1-\gamma^5) T_3^f \psi \end{aligned} \tag{9.49}$$

Adding the contribution from the electromagnetic current, the neutral current for fermion f then becomes

$$J^\mu_{\mathrm{NC},f} = \overline{\psi}\, \gamma^\mu\, \left(\tfrac{1}{2}(1-\gamma^5) T_{3,f} - \sin^2\theta_W Q_f\right)\psi \tag{9.50}$$

It is customary to write the term on the right in terms of a vector and an axial vector coupling such that

$$J^\mu_{\mathrm{NC},f} \;=\; \overline{\psi}\, \gamma^\mu\, \frac{1}{2}\left(C_V^f - C_A^f\gamma^5\right)\psi \tag{9.51}$$

which implies that

$$\boxed{\begin{aligned} C_V^f &= T_3^f - 2Q^f \sin^2\theta_W \\ C_A^f &= T_3^f \end{aligned}} \tag{9.52}$$

Alternatively, we can write the current in terms of left- and right-handed fields as

$$J^\mu_{NC,f} \;=\; \frac{1}{2}\left(C_L^f\, \overline{\psi}_L^f\, \gamma^\mu\, \psi_L^f \;+\; C_R^f\, \overline{\psi}_R^f\, \gamma^\mu\, \psi_R^f\right) \tag{9.53}$$

with the left- and right-handed couplings given by

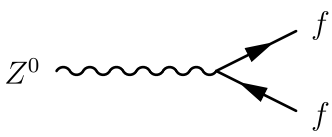
$$\begin{aligned} C_L^f \equiv\; C_V^f + C_A^f \;&=\; -\,2\,Q^f \sin^2\theta_W \;+\; 2\,T_3^f \\ C_R^f \equiv\; C_V^f - C_A^f \;&=\; -\,2\,Q^f \sin^2\theta_W \end{aligned} \tag{9.54}$$

As stated before the values of the charge of the different fermion fields is not predicted. Table 9.1 lists the quantum numbers and resulting vector and axial-vector couplings for all fermions in the Standard Model. The model can be experimentally tested by measuring these couplings in different processes.

fermion	T_3	Y	T_3	Y	Q	C_A^f	C_V^f
	left-handed		right-handed				
ν_e, ν_μ, ν_τ	$+\frac{1}{2}$	-1	0	0	0	$+\frac{1}{2}$	$+\frac{1}{2}$
e, μ, τ	$-\frac{1}{2}$	-1	0	-2	-1	$-\frac{1}{2}$	$-\frac{1}{2} + 2\sin^2\theta_W$
u, c, t	$+\frac{1}{2}$	$+\frac{1}{3}$	0	$+\frac{4}{3}$	$+\frac{2}{3}$	$+\frac{1}{2}$	$+\frac{1}{2} - \frac{4}{3}\sin^2\theta_W$
d, s, b	$-\frac{1}{2}$	$+\frac{1}{3}$	0	$-\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{2}$	$-\frac{1}{2} + \frac{2}{3}\sin^2\theta_W$

Table 9.1: Gauge interaction quantum numbers and corresponding vector and axial vector couplings for the fermions in the Standard Model.

Finally, expressed in terms of the left- and right-handed couplings, the Feynman rule corresponding to $Z\overline{\psi}\psi$ vertex becomes



$$-i\frac{g}{\cos\theta_W}\gamma^\mu\frac{1}{2}\left(C_V^f - C_A^f\gamma^5\right)$$

9.5 The mass of the W and Z bosons

In Lecture 7 we expressed the charged current coupling for processes with momentum transfer $q \ll M_W$ as a four-point interaction. Comparing the expressions to those in this lecture, we can show that the Fermi coupling constant is related to the gauge field couplings as

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2} \tag{9.55}$$

For neutral current processes we can also compute the coupling-constant of the four-point interaction. It is given by

$$\rho\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_Z^2\cos^2\theta_W} \tag{9.56}$$

The parameter ρ specifies the relative strength between the charged and neutral current weak interactions. Comparing the two expressions, we have

$$\rho = \frac{M_W^2}{M_z^2\cos^2\theta_W} \tag{9.57}$$

The masses of the $W^\pm$ and Z^0 can be precisely measured, for instance by reconstructing a 'two-jet' invariant mass distribution in high-energy e^+e^- collisions: Provided that the collision energy is large enough, the di-jet mass will show mass peaks for 'on-shell' produced $W^{\pm}$ and Z^0. The most precise measurement of the four-point coupling for the charged current comes from the measurement of the muon lifetime. The ratio of the charged and neutral current couplings was first measured by the Gargamelle experiment, which exploited an intense neutrino beam to measure the cross-section for a neutral current process $\nu_\mu + \text{nucleus} \to \nu_\mu + \text{hadrons}$ and a charged current process $\nu_\mu + \text{nucleus} \to \mu + \text{hadrons}$.

Upon combination of measurements for the couplings an the masses it is found that the experimental value for ρ is 1 within small uncertainties. This is actually a prediction of the Higgs mechanism. In the Higgs mechanism the mass generated for the W and Z are respectively

$$M_W \;=\; \frac{1}{2}\, v\, g \qquad \text{and} \qquad M_Z \;=\; \frac{1}{2}\, v \sqrt{g^2 + g'^2}\;, \tag{9.58}$$

where v is the so-called *vacuum expectation value* of the Higgs field. With $g'/g = \tan\theta_W$ we find that $\rho = 1$. Therefore, in the Standard Model the masses of the massive vectors bosons are related by

$$M_W = M_Z \cos\theta_W \;. \tag{9.59}$$

The best fit of the Standard Model to all experimental data gives approximately

$$\begin{aligned} \sin^2\theta_W &= 0.231 && (9.60)\\ M_W &= \sqrt{\frac{\sqrt{2}}{8G_F}}\,\frac{e}{\sin\theta_W} = 80.4\,\text{GeV} && (9.61)\\ M_Z &= M_W\;(g_z/g) = M_W/\cos\theta = 91.2\,\text{GeV} && (9.62)\end{aligned}$$

Summary

We have introduced a local gauge symmetry $SU(2)_L \otimes U(1)_Y$ to obtain a Lagrangian for electroweak interactions,

$$-\left(g\,\boldsymbol{J}_L^\mu \cdot \boldsymbol{b}_\mu + \frac{g'}{2} J_Y^\mu \cdot a_\mu\right) \tag{9.63}$$

The coupling constants g and g' are free parameters. We can also take e and $\sin^2\theta_W$. The electromagnetic and neutral weak currents are then given by:

$$\begin{aligned} J^\mu_{EM} &= J_3^\mu + \frac{1}{2} J_Y^\mu \\ J^\mu_{NC} &= J_3^\mu - \sin^2\theta_W\, J^\mu_{EM} = \cos^2\theta_W\, J_3^\mu - \sin^2\theta_W \frac{J_Y^\mu}{2}\end{aligned}$$

and the interaction term in the Lagrangian becomes:

$$-\left(e J^\mu_{EM} \cdot A_\mu + \frac{e}{\cos\theta_W \sin\theta_W} J^\mu_{NC} \cdot Z_\mu\right) \tag{9.64}$$

in terms of the physical fields A_μ and Z_μ.

Note that we still have two independent coupling constants (be it e and θ_W or g and g'). Therefore, it is sometimes said that we have not really 'unified' the electromagnetic and the weak interaction, but rather just put them under the same umbrella. Still, there are clear predictions in this model, such as all the relations between the couplings to the different fermion field and the W to Z mass ratio. Up to now the Glashow-Salam-Weinberg theory, which predicted the W and Z more than a decade before their experimental confirmation, has gloriously passed all experimental tests.

Exercises

Exercise 9.1 (Currents for left and right-handed chirality)
We define the chiral projection operators as $P_L \equiv \frac{1}{2}(1-\gamma^5)$ and $P_R \equiv \frac{1}{2}(1+\gamma^5) = 1 - P_L$ and the left- and right-handed chirality bi-spinor states as $\psi_L \equiv P_L\psi$ and $\psi_R \equiv P_R\psi$.

(a) Show that the left-handed adjoint spinor, defined as $\overline{\psi}_L \equiv \overline{(\psi_L)}$, is given by

$$\overline{\psi}_L = \overline{\psi} P_R$$

(b) Show that the vector current can be decomposed as

$$\overline{\psi}\,\gamma^\mu\,\psi = \overline{\psi}_L\,\gamma^\mu\,\psi_L + \overline{\psi}_R\,\gamma^\mu\,\psi_R$$

(c) Show that the scalar current can be decomposed as

$$\overline{\psi}\,\psi = \overline{\psi}_R\,\psi_L + \overline{\psi}_L\,\psi_R$$

Exercise 9.2 (Charged current interaction)
Show how we get from Eq. (9.22) to Eq. (9.24).

Exercise 9.3 (Symmetries (optional))
For each of the symmetries below indicate what the symmetry is about, and motivate whether it is an exact or a broken symmetry in nature:

(a) U1(Q) symmetry

(b) SU2(u-d-flavour) symmetry

(c) SU3(u-d-s-flavour) symmetry

(d) SU3(colour) symmetry

(e) SU2(weak-isospin) symmetry

(f) SU5(Grand unified) symmetry

(g) Super-symmetry (SUSY)

Lecture 10

The Process $e^- e^+ \to \mu^- \mu^+$

10.1 Helicity conservation

In Lecture 6 we performed computations of the cross-section of several scattering processes in QED. In Lecture 7 we also looked at the decay of the μ and the charged π via the weak interaction. The results of these computations must of course obey conservation of total angular momentum. In the processes that we look at, orbital angular momentum plays no role, such that we really talk about conservation of spin. (See also section 5.11.) The conservation rule is built into the computations, but sometimes it is still useful to understand how it comes about.

In the previous lecture we have defined the *chiral* projections,

$$\begin{aligned} \psi_L &= \tfrac{1}{2}(\mathbb{1}_4 - \gamma^5)\,\psi & \qquad \psi_R &= \tfrac{1}{2}(\mathbb{1}_4 + \gamma^5)\,\psi \\ \overline{\psi}_L &= \overline{\psi}\,\tfrac{1}{2}(\mathbb{1}_4 + \gamma^5) & \qquad \overline{\psi}_R &= \overline{\psi}\,\tfrac{1}{2}(\mathbb{1}_4 - \gamma^5) \end{aligned} \tag{10.1}$$

The states ψ_L and $\overline{\psi}_L$ are respectively the in- and outgoing left-handed particle chiral states, while ψ_R and $\overline{\psi}_R$ are the in- and outgoing right-handed particle chiral states.

We have argued before that in the ultra-relativistic limit there is a correspondence between helicity states and chiral states. You can formalize this by defining the helicity projection operators as

$$P_\uparrow = \tfrac{1}{2}\left(\mathbb{1}_4 + \boldsymbol{\Sigma}\cdot\hat{\boldsymbol{p}}\right) \qquad P_\downarrow = \tfrac{1}{2}\left(\mathbb{1}_4 - \boldsymbol{\Sigma}\cdot\hat{\boldsymbol{p}}\right) \tag{10.2}$$

and subsequently show that for any positive energy spinor with momentum p_z along the z-axis

$$\begin{aligned} u_R &= x\,u_\uparrow \;+\; (1-x)\,u_\downarrow \\ u_L &= x\,u_\downarrow \;+\; (1-x)\,u_\uparrow \end{aligned} \tag{10.3}$$

where we defined

$$x \equiv \frac{(1+\alpha)^2}{2(1+\alpha^2)} \qquad \text{with} \qquad \alpha \equiv \frac{|\boldsymbol{p}|}{E+m} = \sqrt{\frac{\gamma-1}{\gamma+1}}\,. \tag{10.4}$$

In the ultra-relativistic limit, we have $x \to 1$ and the correspondence between the chiral states and helicity states is obtained.

As a consequence, for any process that would violate helicity conservation in the ultra-relativistic limit, such as the $\pi^+ \to e^+\nu_e$ decay via the weak interaction, a helicity suppressing factor $(1-x)$ appears in the amplitude. Simply said, this is because the interaction can only couple to a 'fraction' $(1-x)$ of the lepton wave function.

You have shown in the previous lecture that for a vector coupling one can decompose the current in the vertex factor as follows

$$\overline{\psi}\gamma^\mu\psi = \overline{\psi}_R\gamma^\mu\psi_R + \overline{\psi}_L\gamma^\mu\psi_L \tag{10.5}$$

This means that a right-handed state only couples to a right-handed state, and a left-handed state only to a left-handed state. This results holds equally well for an axial vector coupling ($\gamma^\mu\gamma^5$). It is graphically illustrated in Fig. 10.1. Note that 'crossing' a particle flips its chirality.

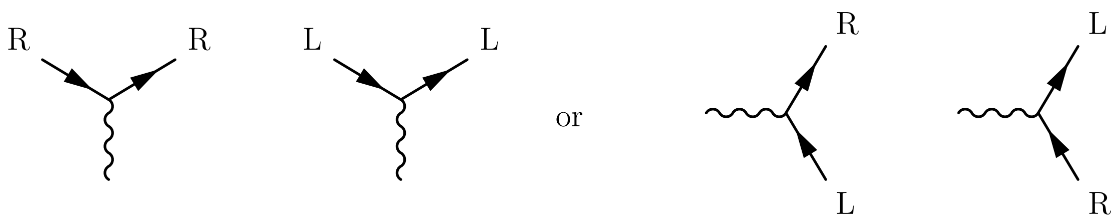


Figure 10.1: Helicity conservation in vector and axial-vector couplings. *left:* A right-handed incoming electron scatters into a right-handed outgoing electron and vice versa in a vector or axial vector interaction . *right:* In the crossed reaction the energy and momentum of one electron is reversed: i.e. in the e^+e^- pair production a right-handed electron and a left-handed positron (or vice versa) are produced. This is the consequence of a spin=1 force carrier. (In all diagrams time increases from left to right.)

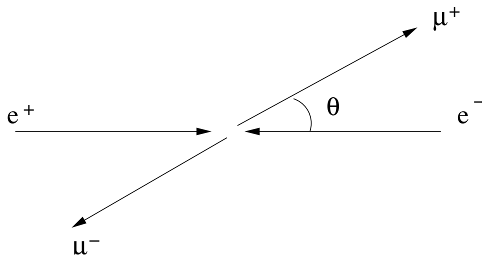
For scalar couplings the situation is exactly opposite, as its decomposition would read

$$\overline{\psi}\psi = \overline{\psi}_R\psi_L + \overline{\psi}_L\psi_R. \tag{10.6}$$

As we shall see in the remainder of this chapter, conservation of helicity has interesting consequences for the $e^-e^+ \to \mu^-\mu^+$ process as well. For example, it allows us to understand the angular dependence of a the polarized cross-sections without going in detail through the kinematics.

10.2 The cross-section of $e^-e^+ \to \mu^-\mu^+$

Equipped with the Feynman rules of the electroweak theory we now proceed with the calculation of the cross-section of the electroweak process $e^-e^+ \to (\gamma, Z) \to \mu^-\mu^+$. We study the process in the centre-of-momentum frame,



with p_i the momentum of an incoming electron, p_f the momentum of an outgoing muon and $\cos\theta$ the angle between the e^+ and the μ^+.

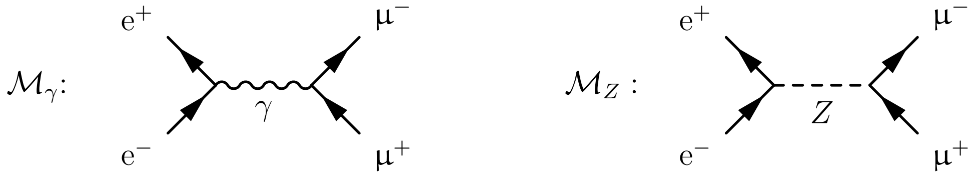


Figure 10.2: Leading order Feynman diagrams contributing to $e^-e^+ \to \mu^-\mu^+$.

In lecture 6 we considered this process in QED. At leading order there was only one contribution to the amplitude, namely via an intermediate photon. In the electroweak theory also the amplitude with an intermediate Z^0 boson contributes. The corresponding Feynman diagrams are shown in Fig. 10.2.

Once we have computed the relevant amplitudes, the differential cross-section follows as usual from the golden rule,

$$\frac{d\sigma(e^-e^+ \to \mu^-\mu^+)}{d\Omega} = \frac{1}{64\pi^2}\,\frac{1}{s}\,\frac{p_f}{p_i}\,\overline{|\mathcal{M}|}^2 \tag{10.7}$$

where the invariant amplitude is the sum of the photon and Z^0 contributions.

In Lecture 6 we computed the spin-averaged amplitude via a rather lengthy procedure, involving Casimir's track and the trace theorems. Because it is actually a nice illustration of the concept of helicity conservation, we will here follow a different approach.

10.2.1 Photon contribution

Consider first only the matrix element of the photon contribution (evaluated using the Feynman rules, see *e.g.* appendix C),

$$\mathcal{M}_\gamma \;=\; -e^2\left(\overline{\psi}_m\gamma^\mu\psi_m\right)\cdot\frac{g_{\mu\nu}}{q^2}\cdot\left(\overline{\psi}_e\gamma^\nu\psi_e\right) \tag{10.8}$$

where the subscript 'm' referes to the muon and the subscript 'e' to the electron. We now decompose the spinors in left- and right-handed chirality states, as we did in lecture 11,

$$\begin{aligned}\left(\overline{\psi}_m\gamma^\mu\psi_m\right) &= \left(\overline{\psi}_{Lm}\,\gamma^\mu\,\psi_{Lm}\right) + \left(\overline{\psi}_{Rm}\,\gamma^\mu\,\psi_{Rm}\right)\\ \left(\overline{\psi}_e\gamma_\mu\psi_e\right) &= \left(\overline{\psi}_{Le}\,\gamma_\mu\,\psi_{Le}\right) + \left(\overline{\psi}_{Re}\,\gamma_\mu\,\psi_{Re}\right)\;.\end{aligned}$$

The total amplitude then becomes

$$\mathcal{M}_\gamma = -\frac{e^2}{s} \left[\left(\overline{\psi}_{Lm}\, \gamma^\mu\, \psi_{Lm}\right) + \left(\overline{\psi}_{Rm}\, \gamma^\mu\, \psi_{Rm}\right) \right] \cdot \left[\left(\overline{\psi}_{L\mathrm{e}}\, \gamma_\mu\, \psi_{L\mathrm{e}}\right) + \left(\overline{\psi}_{R\mathrm{e}}\, \gamma_\mu\, \psi_{R\mathrm{e}}\right) \right] \tag{10.9}$$

where we have also substituted $q^2 = s$.

The matrix element thus consists of four contributions with definite chirality for the four particles in the process. For high energies the chiral projections are helicity states and the four contributions correspond exactly to four polarized amplitudes. Since we can choose a basis with helicity eigenstates, the four contributions do not interfere, and the total cross-section becomes the sum of the amplitudes squared:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}^{\text{unpolarized}} = \frac{1}{4} \left[\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\left(\mathrm{e}_L^-\mathrm{e}_R^+ \to \mu_L^-\mu_R^+\right) + \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\left(\mathrm{e}_L^-\mathrm{e}_R^+ \to \mu_R^-\mu_L^+\right) + \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\left(\mathrm{e}_R^-\mathrm{e}_L^+ \to \mu_L^-\mu_R^+\right) + \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\left(\mathrm{e}_R^-\mathrm{e}_L^+ \to \mu_R^-\mu_L^+\right) \right] \tag{10.10}$$

where we *average* over the incoming spins and *sum* over the final state spins. Note that $\mathrm{e}_R^+ \equiv \overline{\psi}_{L\mathrm{e}}$ etc.

The amplitudes for each of the spinconfigurations can be computed using the expressions for the spinors in Eq. 5.83 and 5.84. We start with the outgoing $\mu^-\mu^+$ current, for the configuration in which the μ^- has positive helicity. In terms of the spinors, the transition current then takes the form

$$J^\mu(\mu_L^-\mu_R^+) = \overline{u_\uparrow(p_{\mu^-})}\, \gamma^\mu\, v_\downarrow(p_{\mu^+}) = \left[u_\uparrow(p_{\mu^-})\right]^\dagger\, \gamma^0\gamma^\mu\, v_\downarrow(p_{\mu^+}) \tag{10.11}$$

Before we fill in the expressions from Eq. 5.83 and 5.84 we need to worry about the momentum vectors. We denote the spherical angles of the μ^- with (θ, ϕ). Since the μ^+ is going exactly in the opposite direction, its spherical angles are $(\pi - \theta, \phi + \pi)$. For this transformation, we have for the various factors that appear in the anti-particle spinors

$$\begin{aligned} \phi &\to \phi + \pi \\ \theta &\to \pi - \theta \\ \sin(\theta/2) &\to \cos(\theta/2) \\ \cos(\theta/2) &\to \sin(\theta/2) \\ e^{i\phi} &\to -e^{i\phi} \end{aligned}$$

In the relativistic limit ($m \to 0$, $p \to E$) the current then becomes

$$J^\mu(\mu_L^-\mu_R^+) = E \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) \\ e^{-i\phi}\sin\left(\frac{\theta}{2}\right) \\ \cos\left(\frac{\theta}{2}\right) \\ e^{-i\phi}\sin\left(\frac{\theta}{2}\right) \end{pmatrix}^T \gamma^0\gamma^\mu \begin{pmatrix} \sin\left(\frac{\theta}{2}\right) \\ -e^{i\phi}\cos\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) \\ -e^{i\phi}\cos\left(\frac{\theta}{2}\right) \end{pmatrix} \tag{10.12}$$

where the factor E is the product of the two wave function normalization factors and we have used that $u^\dagger = (u^*)^T$.

The next step is to work out this result for all of the four components of γ^μ. For the zeroeth component we use that $(\gamma^0)^2 = \mathbb{1}$ and find $J^0 = 0$. For the spatial components we use that $\gamma^0\gamma^k = \alpha_k$ with the three matrices α given by

$$\alpha_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad \alpha_2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \quad \alpha_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \tag{10.13}$$

Inserting these in the expression above, gives for the spatial components of the current

$$\begin{aligned} J^k(\mu^-_L \mu^+_R) \;=\; E\,\big(\, & 2e^{-i\phi}\sin^2(\theta/2) - 2e^{i\phi}\cos^2(\theta/2), \\ & 2ie^{-i\phi}\sin^2(\theta/2) + 2ie^{i\phi}\cos^2(\theta/2), \\ & 4\cos(\theta/2)\sin(\theta/2)\,\big) \end{aligned} \tag{10.14}$$

We can re-use this expression to obtain the electron current for this spin configuration as well. The ingoing and outgoing particles are reversed, but as shown in Eq.6.22 that is just a matter of taking the complex conjugate. The electron and position travel along the z-axis, so $\theta = 0$. The value of ϕ is arbitrary: it will just lead to an overall phase which drops out when we take the amplitude squared. We choose $\phi = \pi$ since that gives the correct correspondence for earlier expressions for spinors with spin along the z-axis. Inserting these values for the polar angles and taking the complex conjugate to change the order of the ingoing and outgoing spinors, the electron transition current then becomes

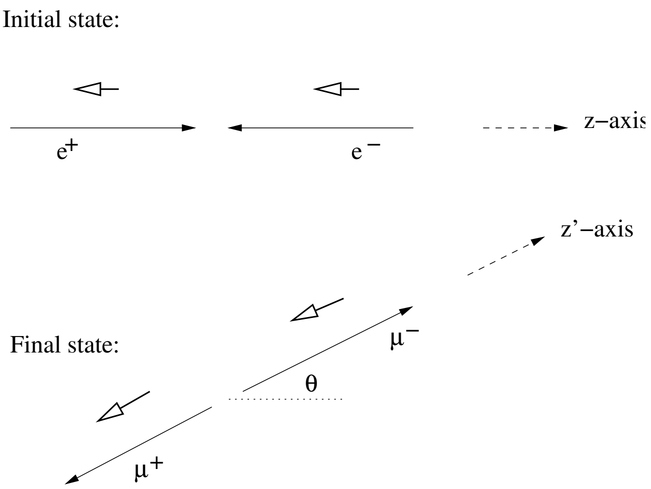
$$J^\mu(\mathrm{e}^-_L \mathrm{e}^+_R) \;=\; \overline{v_\downarrow}\gamma^\mu u_\uparrow \;=\; E\big(0, 2, 2i, 0\big) \tag{10.15}$$

Taking the metric into account, we then obtain for the matrix element

$$\mathcal{M} \;=\; -\frac{e^2}{s}\, J^\mu_e\, g_{\mu\nu}\, J^\nu_m \;=\; \frac{e^2}{s}\, 8E^2\, e^{i\phi}\, \cos^2(\theta/2) \;=\; e^2\, e^{i\phi}\, (1+\cos\theta) \tag{10.16}$$

where we used that $s = 4E^2$. The factor $e^{i\phi}$ drops out when we take the square of the amplitude.

The other three amplitudes can be obtained in a similar way. However, because it is more instructive, we will show you an alternative method to obtain the angular dependence, which just uses quantum mechanics. Let us look in more detail at the helicity dependence (H&M §6.6):



In the initial state the e^- and e^+ have opposite helicity (as they produce a spin 1 γ).

The same is true for the final state μ^- and μ^+.

In the center of mass frame, scattering proceeds from an initial state with $J_Z = +1$ or -1 along axis $\hat{z}$ into a final state with $J'_Z = +1$ or -1 along axis $\hat{z}'$. Since the interaction proceeds via a photon with spin $J = 1$ the amplitude for scattering over an angle θ is given by the rotation matrices[1]

$$d^j_{m'm}(\theta) \equiv \big\langle jm' | e^{-i\theta J_y} | jm \big\rangle \tag{10.18}$$

where J_y is the y component of the angular momentum operator (which is also the generator for rotations around the y-axis). The coefficients $d^j_{m,m'}$ are sometimes called 'Wigner d-matrices'. Computing them for the spin-1 system is not so hard (see *e.g.* exercise 10.2, or H&M exercise 2.6) and gives

$$\begin{aligned} d^1_{+1,+1}(\theta) &= d^1_{-1,-1}(\theta) = \frac{1}{2}\left(1+\cos\theta\right) \\ d^1_{+1,-1}(\theta) &= d^1_{-1,+1}(\theta) = \frac{1}{2}\left(1-\cos\theta\right) \end{aligned} \tag{10.19}$$

For the kinematic factors in the four amplitudes we then get

$$\begin{aligned} \left(\overline{\psi}_{Lm}\gamma^\mu\psi_{Lm}\right) g_{\mu\nu} \left(\overline{\psi}_{Le}\gamma^\nu\psi_{Le}\right) &\propto d_{-1,-1}(\theta) = \tfrac{1}{2}(1+\cos\theta) \\ \left(\overline{\psi}_{Rm}\gamma^\mu\psi_{Rm}\right) g_{\mu\nu} \left(\overline{\psi}_{Re}\gamma^\nu\psi_{Re}\right) &\propto d_{+1,+1}(\theta) = \tfrac{1}{2}(1+\cos\theta) \\ \left(\overline{\psi}_{Lm}\gamma^\mu\psi_{Lm}\right) g_{\mu\nu} \left(\overline{\psi}_{Re}\gamma^\nu\psi_{Re}\right) &\propto d_{+1,-1}(\theta) = \tfrac{1}{2}(1-\cos\theta) \\ \left(\overline{\psi}_{Rm}\gamma^\mu\psi_{Rm}\right) g_{\mu\nu} \left(\overline{\psi}_{Le}\gamma^\nu\psi_{Le}\right) &\propto d_{+1,+1}(\theta) = \tfrac{1}{2}(1-\cos\theta) \end{aligned} \tag{10.20}$$

As a consequence of the helicity conservation, scattering is suppressed in the direction in which the spin of the μ^- is not aligned with the spin of the e^-.

[1] See H&M§2.2:

$$e^{-i\theta J_2}\,|j\ m\rangle = \sum_{m'} d^j_{m\,m'}(\theta)\,|j\ m'\rangle \tag{10.17}$$

and also appendix H in Burcham & Jobes

When we compare the amplitudes in Eq. 10.20 to the full equation derived using the spinors, we find that we still miss a proportionality factor that comes from the normalization of the wave functions. For each of the amplitudes this factor is $|N|^2 = E = \sqrt{s}/2$.

After combining all of this, the polarized cross-sections become

$$\begin{aligned} \frac{d\sigma}{d\Omega}\left(e_L^- e_R^+ \to \mu_L^- \mu_R^+\right) &= \frac{d\sigma}{d\Omega}\left(e_R^- e_L^+ \to \mu_R^- \mu_L^+\right) &= \frac{\alpha^2}{4s}\left(1+\cos\theta\right)^2 \\ \frac{d\sigma}{d\Omega}\left(e_L^- e_R^+ \to \mu_R^- \mu_L^+\right) &= \frac{d\sigma}{d\Omega}\left(e_R^- e_L^+ \to \mu_L^- \mu_R^+\right) &= \frac{\alpha^2}{4s}\left(1-\cos\theta\right)^2 \end{aligned} \tag{10.21}$$

The unpolarised cross-section is obtained as the spin-averaged sum,

$$\frac{d\sigma}{d\Omega}^{\text{unpol}} = \frac{1}{4}\,\frac{\alpha^2}{4s}\,2\,\left[(1+\cos\theta)^2 + (1-\cos\theta)^2\right] = \frac{\alpha^2}{4s}\left(1+\cos^2\theta\right) \quad . \tag{10.22}$$

in agreement with our computation in Lecture 6.

10.2.2 Z^0 contribution

Having written the total cross-section as a sum of polarized amplitudes we are ready to include the contribution from the Z^0 boson amplitude. Using the Feynman rules (see *e.g.* appendix C) we find for the invariant amplitude

$$\mathcal{M}_Z = -\frac{g^2}{4\cos^2\theta_w}\left[\overline{\psi}_m\gamma^\mu\left(C_V^m - C_A^m\gamma^5\right)\psi_m\right]\cdot\frac{g_{\mu\nu} - q_\mu q_\nu/M_Z^2}{q^2 - M_Z^2}\cdot\left[\overline{\psi}_e\gamma^\nu\left(C_V^e - C_A^e\gamma^5\right)\psi_e\right] \tag{10.23}$$

We can simplify the Z^0 propagator if we ignore the lepton masses ($m_\ell \ll \sqrt{s}$). In that case the Dirac equation becomes:

$$\overline{\psi}_e\,\left(i\partial_\mu\gamma^\mu - m\right) = 0 \quad \Rightarrow \quad \overline{\psi}_e\,\left(\gamma^\mu p_{\mu,e}\right) = 0 \tag{10.24}$$

Since $p_e = \frac{1}{2}q$ we also have:

$$\frac{1}{2}\,\overline{\psi}_e\,\left(\gamma^\mu q_\mu\right) = 0 \tag{10.25}$$

As a result the $q_\mu q_\nu$ term in the propagator vanishes and we obtain for the matrix element

$$\mathcal{M}_Z = \frac{-g^2}{4\cos^2\theta_w}\,\frac{1}{q^2 - M_Z^2}\cdot\left[\overline{\psi}_m\,\gamma^\mu\left(C_V^m - C_A^m\gamma^5\right)\,\psi_m\right]\left[\overline{\psi}_e\,\gamma_\mu\left(C_V^e - C_A^e\gamma^5\right)\,\psi_e\right] \quad . \tag{10.26}$$

We have shown in the previous lecture how the $(C_V^e - C_A^e\gamma^5)$ terms can be written in terms of left- and right-handed couplings. Defining

$$C_R \equiv C_V - C_A \qquad \text{and} \qquad C_L \equiv C_V + C_A \tag{10.27}$$

one finds

$$\left(C_V - C_A\gamma^5\right)\psi \;=\; C_R\psi_R \;+\; C_L\psi_L \;. \tag{10.28}$$

Consequently, the Z^0 amplitude can be written as

$$\begin{aligned}\mathcal{M}_Z \;=\; -\frac{g^2}{4\cos^2\theta_w}\,\frac{1}{s-M_Z^2}\;&\left[\,C_L^m\left(\overline{\psi}_{Lm}\,\gamma^\mu\,\psi_{Lm}\right) + C_R^m\left(\overline{\psi}_{Rm}\,\gamma^\mu\,\psi_{Rm}\right)\,\right]\cdot \\ &\left[\,C_L^e\left(\overline{\psi}_{Le}\,\gamma_\mu\,\psi_{Le}\right) + C_R^e\left(\overline{\psi}_{Re}\,\gamma_\mu\,\psi_{Re}\right)\,\right]\end{aligned} \tag{10.29}$$

Comparing this the expression to Eq. (10.9) we realize that we can obtain the polarized cross-sections directly from the results obtained for the QED process. For two of the four contributions we then obtain

$$\begin{aligned}\frac{d\sigma}{d\Omega}_{\gamma,Z}\left(e_L^-e_R^+\to\mu_L^-\mu_R^+\right) &= \frac{\alpha^2}{4s}\,\left(1+\cos\theta\right)^2\cdot\left|1+r\,C_L^mC_L^e\right|^2 \\ \frac{d\sigma}{d\Omega}_{\gamma,Z}\left(e_L^-e_R^+\to\mu_R^-\mu_L^+\right) &= \frac{\alpha^2}{4s}\,\left(1-\cos\theta\right)^2\cdot\left|1+r\,C_R^mC_L^e\right|^2\end{aligned} \tag{10.30}$$

with the relative contribution of the Z^0 and γ parameterized as

$$r = \frac{g^2}{e^2}\,\frac{1}{4\cos^2\theta_w}\,\frac{s}{s-M_z^2}\;. \tag{10.31}$$

The other two helicity configuration follow using the relation in Eq. (10.21) and replacing C_L by C_R etc . Using the relation between the coupling constants

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2} = \frac{g^2}{8M_Z^2\cos^2\theta_w}\;. \tag{10.32}$$

we can also write r in terms of G_F as

$$r = \frac{\sqrt{2}G_FM_Z^2}{e^2}\,\frac{s}{s-M_Z^2}\;. \tag{10.33}$$

10.2.3 Correcting for the finite width of the Z^0

The propagator for the massive vectors bosons has a ‘pole’ at the boson mass: it becomes infinitely large for an ‘on-shell’ ($p^2 = m^2$) boson. As you can readily see from the expression above, this would lead to an infinite cross-section when we tune the beam energies to $\sqrt{s} = M_Z$. The problem is that the propagator does not take into account the finite decay width of the Z^0. The Z^0 boson is not a stable particle and hence the ’on-shell’ Z^0 is actually something with a rather broad mass distribution. We can account for the width by replacing the mass in the propagator with

$$M_Z \;\to\; M_Z - \frac{i}{2}\Gamma_Z \tag{10.34}$$

where Γ is the total decay width of 'on-shell' (*i.e.* not virtual) Z^0-bosons.

A heuristic explanation (Halzen and Martin, §2.10) is as follows. The decay of an unstable particle follows the exponential law

$$|\psi(t)|^2 = |\psi(0)|^2 \, e^{-\Gamma t} \tag{10.35}$$

where $|\psi(0)|$ is the probability (density) at $t = 0$ and $1/\Gamma$ is the lifetime. Therefore, the time-dependence of the wave function, which already involves the rest mass, must also include a factor $\sqrt{e^{-\Gamma t/2}}$, or

$$\psi(t) = \psi(0) e^{-imt} e^{-\Gamma t/2} \tag{10.36}$$

Consequently, with the substitution above we can 'correct' the propagator mass for the finite decay width. The lineshape that results from such a propagator is usually called a (spin-1) Breit-Wigner.

With the replacement $M_Z \to M_Z - i\Gamma_Z/2$ in the propagator, the expression for r becomes

$$\begin{aligned} r &= \frac{\sqrt{2} G_F M_Z^2}{e^2} \frac{s}{s - \left(M_Z - i\frac{\Gamma_Z}{2}\right)^2} \\ &= \frac{\sqrt{2} G_F M_Z^2}{e^2} \frac{s}{s - \left(M_z^2 - \frac{\Gamma_Z^2}{4}\right) + iM_Z\Gamma_Z} \end{aligned} \tag{10.37}$$

Note that r is now complex and that the phase of r depends on $\sqrt{s}$.

10.2.4 Total unpolarized cross-section

As a final step we add up the unpolarized cross-sections and find

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4s} \left[A_0 \left(1 + \cos^2\theta\right) + A_1 \cos\theta \right] \tag{10.38}$$

with A_0 and A_1 given by

$$\begin{aligned} A_0 &= 1 + 2Re(r)\, C_V^e C_V^f + |r|^2 \left({C_V^e}^2 + {C_A^e}^2\right) \left({C_V^f}^2 + {C_A^f}^2\right) \\ A_1 &= 4Re(r)\, C_A^e C_A^f + 8|r|^2\, C_V^e C_V^f C_A^e C_A^f \end{aligned} \tag{10.39}$$

In this expression we replaced C_R and C_L by C_V and C_A using the definitions given earlier.

Since we have not actually used that the final state fermions are muons, our result is valid for any process $e^-e^+ \to \gamma, Z^0 \to f\bar{f}$, provided that we insert the correct weak couplings.[2]

[2]This is not entirely true: For quarks we also need to insert the charge in the coupling to the photon.

Assuming "lepton universality", we have $C_V^e = C_V^\mu$ and $C_A^e = C_A^\mu$. The expressions for A_0 and A_1 then become

$$\begin{aligned} A_0 &= 1 + 2\,Re(r)\,C_V^2 + |r|^2\left(C_V^2 + C_A^2\right)^2 \\ A_1 &= 4\,Re(r)\,C_A^2 + 8|r|^2 C_V^2 C_A^2 \end{aligned}$$

In the Standard Model the coefficients for leptons are $C_A = -\frac{1}{2}$ and $C_V = -\frac{1}{2} + 2\sin^2\theta_w$.

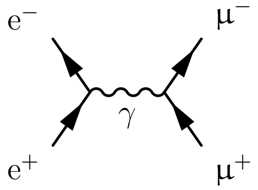
Finally, we consider the total cross-section. In the integrated cross-section the term proportional to $\cos\theta$ vanishes and for the other term we use

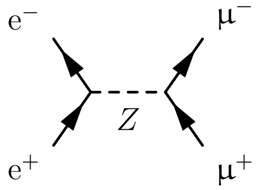
$$\int (1 + \cos^2\theta)\,\mathrm{d}\Omega \;=\; \frac{16\pi}{3} \tag{10.40}$$

Consequently, the total cross-section is

$$\sigma(s) \;=\; \frac{4\pi\alpha^2}{3s}\,A_0(s) \tag{10.41}$$

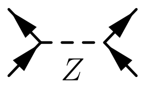
To summarize, on the *amplitude level* there are two diagrams that contribute:

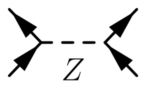
$\mathcal{M}_\gamma$: 

$\mathcal{M}_Z$: 

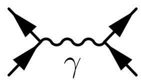
(10.42)

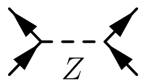
Using the following notation

$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}[Z,Z] =$ 

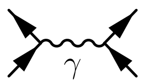
$\cdot$ 

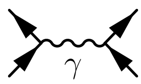
$\propto |r|^2$

$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}[\gamma Z] =$ 

$\cdot$ 

$\propto Re\,(r)$

$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}[\gamma,\gamma] =$ 

$\cdot$ 

$\propto 1$

the expression for the differential cross-section is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \;=\; \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}[\gamma,\gamma] \;+\; \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}[Z,Z] \;+\; \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}[\gamma,Z] \tag{10.43}$$

with

$$\begin{aligned} \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}[\gamma,\gamma] &= \frac{\alpha^2}{4s}\left(1 + \cos^2\theta\right) \\ \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}[Z,Z] &= \frac{\alpha^2}{4s}\,|r|^2\left[\left({C_V^e}^2 + {C_A^e}^2\right)\left({C_V^f}^2 + {C_A^f}^2\right)\left(1 + \cos^2\theta\right) + 8C_V^e C_V^f C_A^e C_A^f\cos\theta\right] \\ \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}[\gamma,Z] &= \frac{\alpha^2}{4s}\,Re|r|\left[C_V^e C_V^f\left(1 + \cos^2\theta\right) + 2C_A^e C_A^f\cos\theta\right] \end{aligned}$$

10.3 Near the resonance

Let us take a closer look at the cross-section for beam energies close to the Z^0 mass. One can see from Eq. (10.37) that $|r|$ is maximal for

$$s_0 = M_Z^2 - \frac{\Gamma_Z^2}{4} \,. \tag{10.44}$$

Since the denominator in Eq. (10.37) is purely imaginary for $s = s_0$, the interference term, which is proportional to $Re(r)$, vanishes at the peak. Defining the kinematic factor in r as

$$\kappa(s) \;=\; \frac{s}{s - \left(M_Z^2 - \frac{\Gamma_Z^2}{4}\right) + iM_Z\Gamma_Z} \tag{10.45}$$

you will show in exercise 10.1(b) that

$$Re(\kappa) = \left(1 - \frac{s_0}{s}\right)|\kappa|^2 \quad \text{with} \quad |\kappa|^2 = \frac{s^2}{(s-s_0)^2 + M_Z^2\Gamma_Z^2} \,. \tag{10.46}$$

You will also show that at the resonance $|r|^2 \gg 1$ such that we can ignore the photon contribution entirely. With neither the interference nor the photon contribution, we have

$$A_0(s) \approx \left(\frac{\sqrt{2}G_F M_Z^2}{e^2}\right)^2 \frac{s^2}{(s-s_0)^2 + M_z^2\Gamma_Z^2} \left({C_V^e}^2 + {C_A^e}^2\right)\left({C_V^f}^2 + {C_A^f}^2\right) \tag{10.47}$$

Exactly at the resonance, this gives for the total cross-section to the final state $f\bar{f}$:

$$\sigma(\mathrm{e}^+\mathrm{e}^- \to f\bar{f})|_{s=s_0} \;=\; \frac{G_F^2}{6\pi}\frac{s_0 M_Z^2}{\Gamma_Z^2}\left({C_V^e}^2 + {C_A^e}^2\right)\left({C_V^f}^2 + {C_A^f}^2\right) \tag{10.48}$$

For quark-antiquark final states ($f = q$) we need to take into account that there are three distinct colour configurations, namely blue-anti-blue, red-anti-red and green-anti-green. Therefore, for a quark-anti-quark pair, the cross-section involves another factor $N_c = 3$,

$$\sigma(\mathrm{e}^+\mathrm{e}^- \to q\bar{q})|_{s=s_0} \;=\; N_c \cdot \frac{G_F^2}{6\pi}\frac{s_0 M_Z^2}{\Gamma_Z^2}\left({C_V^e}^2 + {C_A^e}^2\right)\left({C_V^q}^2 + {C_A^q}^2\right) \tag{10.49}$$

Figure 10.3 shows the measured cross-section in hadronic final states as function of the collision energy.

At collision energies well above the typical QCD binding energy ($\sqrt{s} \gg 2m_\pi$), the $q\bar{q}$ state is observed as two 'jets', collimated showers of light mesons. The ratio between the hadronic and leptonic event yields at the Z^0 resonance,

$$R_l \;=\; \frac{\sigma(\mathrm{e}^+\mathrm{e}^- \to \text{hadrons})}{\sigma(\mathrm{e}^+\mathrm{e}^- \to \mu^+\mu^-)} \tag{10.50}$$

provides an important test of the standard model, as shown in Fig. 10.4.

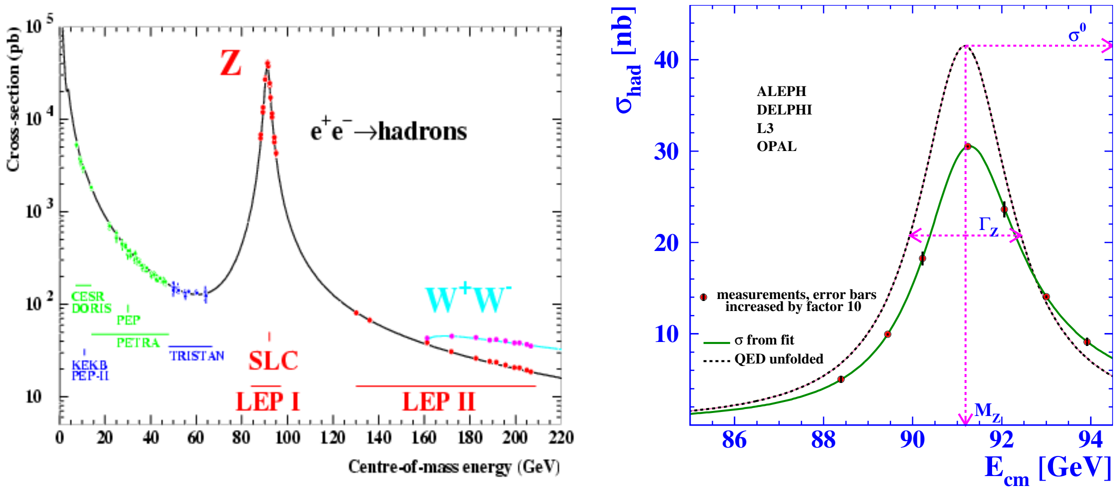


Figure 10.3: *left:* The Z^0-lineshape: the cross-section for $e^+e^- \to$ hadrons as a function of $\sqrt{s}$. *right:* Same but now near the resonance. The dashed line represents the leading order computation,while the continuous gray line includes higher order corrections.

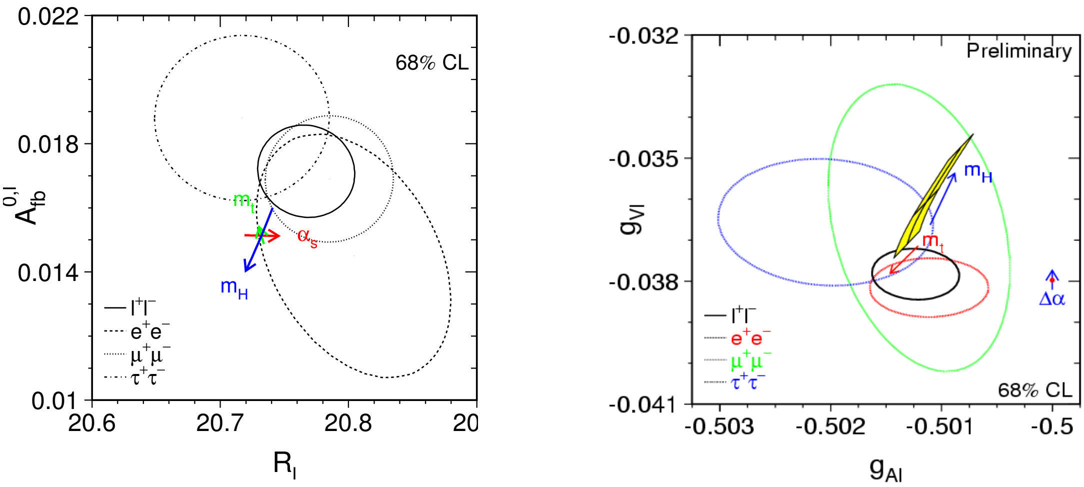


Figure 10.4: *left:* Tests of the standard model. The leptonic A_{fb} vs. R_l. The contours show the measurements while the arrows show the dependency on Standard Model parameters. *right:* Determination of the vector and axial vector couplings.

10.4 The forward-backward asymmetry

A direct consequence of the photon-Z^0 interference is that the angular distribution is not symmetric. Figure 10.5 shows the $\cos\theta$ distribution observed at the Jade experiment, which operated at the PETRA collider in Hamburg. The beam energy in this experiment was not yet sufficient to directly produce Z^0 bosons. Still, the effect of the interference was clearly visible long before the direct discovery of the Z^0 resonance.

At the peak and ignoring the pure photon exchange (because it is negligibly small), the

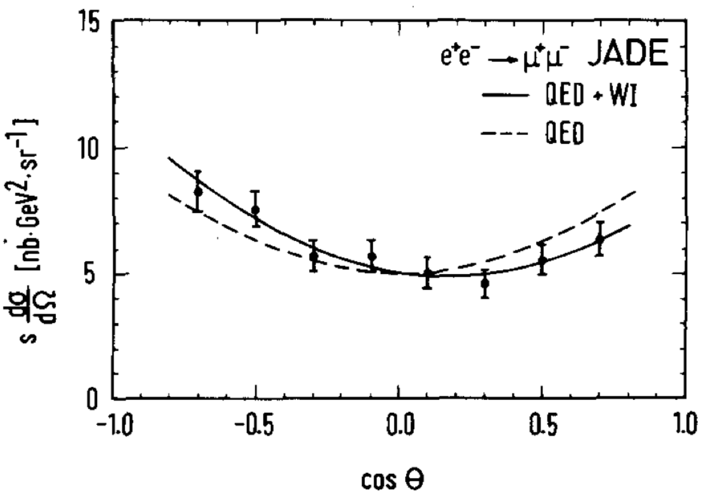


Figure 10.5: Angular distribution for $\mathrm{e}^+\mathrm{e}^- \to \mu^+\mu^-$ for $25 < \sqrt{s} < 37$ GeV at the JADE experiment. θ is the angle between the outgoing μ^+ and the incoming e^+. The curves show fits to the data $p(1+\cos^2\theta) + q\cos\theta$ (full curve) and $p(1+\cos^2\theta)$ (dashed curve). (Source: JADE collaboration, Phys. Lett. Vol108B, p140-144, 1982.)

polar angle distribution is given by

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\cos\theta} \propto 1 + \cos^2\theta + \frac{8}{3} A_{\mathrm{FB}} \cos\theta \,, \tag{10.51}$$

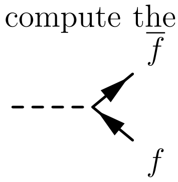
where we defined the 'forward-backward asymmetry',

$$A_{\mathrm{FB}}^{0,f} = \frac{3}{4} A_{\mathrm{e}} A_f \qquad \text{with} \qquad A_f = \frac{2C_V^f C_A^f}{C_V^2 + C_A^2} \,. \tag{10.52}$$

Precise measurements of the forward-backward asymmetry can be used to determine the couplings C_V and C_A.

10.5 The Z^0 decay width and the number of light neutrinos

Using the Feynman rules we can also compute the $Z^0 \to f\bar{f}$ decay width, represented by the diagram



You cannot easily do this computation yourself, since we have not discussed the external line for the Z^0 in this course. (The computation needs to take into account the three polarization states of the massive vector boson.) The result of the computation is

$$\begin{aligned}\Gamma\left(Z \to f\bar{f}\right) &= \frac{1}{16\pi}\frac{1}{M_Z}\left|\overline{\mathcal{M}}\right|^2 \\ &= \frac{g^2}{48\pi}\frac{M_z}{\cos^2\theta_w}\left({C_V^f}^2 + {C_A^f}^2\right) \\ &= \frac{G_F}{6\sqrt{2}}\frac{M_Z^3}{\pi}\left({C_V^f}^2 + {C_A^f}^2\right)\end{aligned} \tag{10.53}$$

For quark-antiquark final states ($f = q$) we again need to multiply by the colour factor $N_c = 3$,

$$\Gamma(Z \to \overline{q}q) = \frac{G_F}{6\sqrt{2}}\frac{M_Z^3}{\pi}\left({C_V^q}^2 + {C_A^q}^2\right) \cdot N_C \,. \tag{10.54}$$

The total decay width of the Z^0 is the sum of all partial widths to all accessible final states,

$$\Gamma_Z = \Gamma_{ee} + \Gamma_{\mu\mu} + \Gamma_{\tau\tau} + \Gamma_{uu} + \Gamma_{dd} + \Gamma_{ss} + \Gamma_{cc} + \Gamma_{bb} + N_\nu \cdot \Gamma_{\nu\nu} \,, \tag{10.55}$$

where N_ν is the number of neutrino species, which is equal to three in the standard model.

Using all available data to extract information on the couplings we can compute the decay widths to all final states within the standard model,

$$\begin{array}{lllll}\Gamma_{ee} & \approx & \Gamma_{\mu\mu} \approx \Gamma_{\tau\tau} = 84 \text{ MeV} & C_V \approx 0 & C_A = -\frac{1}{2} \\ \Gamma_{\nu\nu} & = & 167 \text{ MeV} & C_V = \frac{1}{2} & C_A = \frac{1}{2} \\ \Gamma_{uu} & \approx & \Gamma_{cc} = 276 \text{ MeV} & C_V \approx 0.19 & C_A = \frac{1}{2} \\ \Gamma_{dd} & \approx & \Gamma_{ss} \approx \Gamma_{bb} = 360 \text{ MeV} & C_V \approx -0.35 & C_A = -\frac{1}{2}\end{array}$$

A measurement of the lineshape (the cross-section as function of $\sqrt{s}$) gives for the total decay width of the Z^0,

$$\Gamma_Z \;\approx\; 2490 \text{ MeV}$$

So, even though we cannot see the neutrino contribution, we can estimate the number of neutrinos from the total width of the Z^0. The result is

$$N_\nu = \frac{\Gamma_Z - 3\Gamma_l - \Gamma_{had}}{\Gamma_{\nu\nu}} = 2.984 \pm 0.008 \,. \tag{10.56}$$

Figure 10.6 shows the predicted lineshape for different values of N_ν. This results put strong constraints on extra generations: if there is a fourth generation, then either it has a very heavy neutrino, or its neutrino does not couple to the Z^0. In either case, this generation would be very different from the known generations of quarks and leptons.

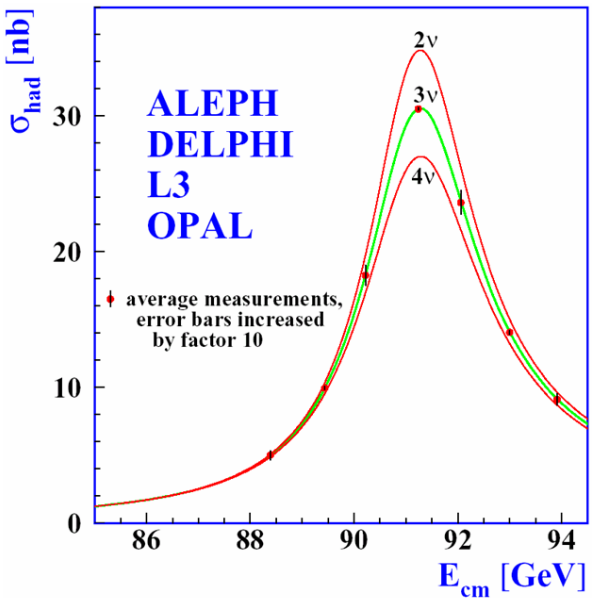


Figure 10.6: The Z lineshape fit for $N_\nu = 2,\ 3,\ 4$.

Exercises

Exercise 10.1 (Z^0 production and decay)

(a) Derive the expression for $Re(\kappa)$ in Eq. (10.46).

(b) Calculate the relative contribution of the Z^0-exchange and the γ exchange to the cross-section at the Z^0 peak. Use $\sin^2\theta_W = 0.23$, $M_z = 91GeV$ and $\Gamma_Z = 2.5GeV$.

(c) Show also that at the peak

$$\sigma_{peak}(e^-e^+ \to \mu^-\mu^+) \approx \frac{12\pi}{M_z^2}\frac{\Gamma_e\Gamma_\mu}{\Gamma_Z^2} \tag{10.57}$$

(d) Why does the top quark not contribute to the decay width of the Z^0?

(e) Calculate the value of $R_l = \Gamma_{\text{had}}/\Gamma_{\text{lep}}$ at the resonance $s = s_0$. Ignore the masses of the fermions, as we did in the lecture. You may also ignore the contribution of the photon, as it is very small at the resonance.

(f) The actual line shape of the Z^0-boson is not a pure Breit Wigner: at the high $\sqrt{s}$ side of the peak the cross-section is higher then expected from the formula derived in the lectures. Can you think of a reason why this would be the case?

(g) The number of light neutrino generations is determined from the "invisible width" of the Z^0-boson as follows:

$$N_\nu = \frac{\Gamma_Z - 3\Gamma_l - \Gamma_{had}}{\Gamma_\nu}$$

Can you think of another way to determine the decay rate of $Z^0 \to \nu\bar{\nu}$ directly? Do you think this method is more precise or less precise?

Exercise 10.2 (Spin projections (optional!) (From Thomson, ex. 6.6))
Consider a spin-1 system. The eigenstates of the operator $S_n = \boldsymbol{n}\cdot\boldsymbol{S}$ correspond to the spin projections in the direction $\boldsymbol{n}$. These can be written in terms of the eigenstates of the operator S_z, for instance as

$$|1,+1\rangle_n = \alpha\,|1,+1\rangle + \beta\,|1,0\rangle + \gamma\,|1,-1\rangle$$

Taking $n = (\sin\theta, 0, \cos\theta)$ show that

$$|1,+1\rangle_n = \tfrac{1}{2}(1-\cos\theta)\,|1,+1\rangle + \tfrac{1}{\sqrt{2}}\sin\theta\,|1,0\rangle + \tfrac{1}{2}(1+\cos\theta)\,|1,-1\rangle \tag{10.58}$$

Hint: Write S_x in terms of the spin ladder operators and use that all states are normalized to 1.

Lecture 11

Symmetry breaking

After a review of the shortcomings of the model of electroweak interactions in the Standard Model, in this section we study the consequences of spontaneous symmetry breaking of (gauge) symmetries. We will do this in three steps of increasing complexity and focus on the principles of how symmetry breaking can be used to obtain massive gauge bosons by working out in full detail the breaking of a local U(1) gauge invariant model (QED) and giving the photon a mass.

11.1 Problems in the Electroweak Model

The electroweak model, beautiful as it is, has some serious shortcomings.

1] Local $\mathbf{SU(2)_L \times U(1)_Y}$ gauge invariance forbids massive gauge bosons

In the theory of Quantum Electro Dynamics (QED) the requirement of local gauge invariance, i.e. the invariance of the Lagrangian under the transformation $\phi' \rightarrow e^{i\alpha(x)}\phi$ plays a fundamental role. Invariance was achieved by replacing the partial derivative by a covariant derivative, $\partial_\mu \rightarrow \mathcal{D}_\mu = \partial_\mu - ieA_\mu$ and the introduction of a new vector field A with very specific transformation properties: $A'_\mu \rightarrow A_\mu + \frac{1}{e}\partial_\mu\alpha$. The Lagrangian for a free particle then changed to:

$$\mathcal{L}_{\rm QED} \;=\; \mathcal{L}_{\rm free} + \mathcal{L}_{\rm int} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \qquad (11.1)$$

which not only 'explained' the presence of a vector field in nature (the photon), but also automatically yielded an interaction term $\mathcal{L}_{\rm int} = eJ^\mu A_\mu$ between the vector field and the particle, as explained in detail in the lectures on the electroweak model. Under these symmetry requirements it is unfortunately not possible for a gauge boson to acquire a mass. In QED for example, a mass term for the photon, would not be allowed as such a term breaks gauge invariance:

$$\frac{1}{2}m_\gamma^2 A_\mu A^\mu = \frac{1}{2}m_\gamma^2 (A_\mu + \frac{1}{e}\partial_\mu \alpha)(A^\mu + \frac{1}{e}\partial^\mu \alpha) \neq \frac{1}{2}m_\gamma^2 A_\mu A^\mu \tag{11.2}$$

The example using only $U(1)$ and the mass of the photon might sound strange as the photon is actually massless, but a similar argument holds in the electroweak model for the W and Z bosons, particles that we *know* are massive and make the weak force only present at very small distances.

2] Local $\mathbf{SU(2)_L \times U(1)_Y}$ gauge invariance forbids massive fermions

Just like in QED, invariance under local gauge transformations in the electroweak model requires introducing a covariant derivative of the form $D_\mu = \partial_\mu + ig\frac{1}{2}\vec{\tau}\cdot\vec{W}_\mu + ig'\frac{1}{2}YB_\mu$ introducing a weak current, J^{weak} and a different transformation for isospin singlets and doublets. A mass term for a fermion in the Lagrangian would be of the form $-m_f\bar{\psi}\psi$, but such terms in the Lagrangian are not allowed as they are *not* gauge invariant. This is clear when we decompose the expression in helicity states:

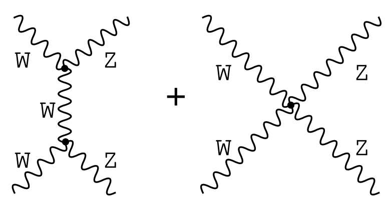
$$\begin{aligned} -m_f\bar{\psi}\psi &= -m_f\left(\bar{\psi_R}+\bar{\psi_L}\right)(\psi_L+\psi_R) \\ &= -m_f\left[\bar{\psi_R}\psi_L+\bar{\psi_L}\psi_R\right] \qquad \text{, since } \bar{\psi_R}\psi_R=\bar{\psi_L}\psi_L=0 \end{aligned}$$

Since ψ_L (left-handed, member of an isospin doublet, I $= \frac{1}{2}$) and ψ_R (right-handed, isospin singlet, I $=$ 0) behave differently under rotations these terms are not gauge invariant:

$$\begin{aligned} \psi_L{}' \to \psi_L &= e^{i\alpha(x)T+i\beta(x)Y}\psi_L \\ \psi_R{}' \to \psi_R &= e^{i\beta(x)Y}\psi_R \end{aligned}$$

3] Violating unitarity

Several Standard Model scattering cross-sections, like WW-scattering (some Feynman graphs are shown in the picture on the right) violate unitarity at high energy as $\sigma(\text{WW}\to\text{ZZ}) \propto \text{E}^2$. This energy dependency clearly makes the theory non-renormalizable.



How to solve the problems: a way out

To keep the theory renormalizable, we need a very high degree of symmetry (local gauge invariance) in the model. Dropping the requirement of the local $\mathrm{SU(2)_L \times U(1)_Y}$ gauge invariance is therefore not a wise decision. Fortunately there is a way out of this situation:

> Introduce a new field with a very specific potential that keeps the full Lagrangian invariant under $\mathrm{SU(2)_L \times U(1)_Y}$, but will make the vacuum *not* invariant under this symmetry. We will explore this idea, spontaneous symmetry breaking of a local gauge invariant theory (or Higgs mechanism), in detail in this section.

The Higgs mechanism: - Solves all the above problems
- Introduces a fundamental scalar $\rightarrow$ the Higgs boson !

11.2 A few basics on Lagrangians

A short recap of the basics on Lagrangians we will be using later.

$$\mathcal{L} = \mathrm{T(kinetic)} - \mathrm{V(potential)}$$

The Euler-Lagrange equation then give you the equations of motion:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0$$

For a real scalar field for example:

$$\mathcal{L}_{\mathrm{scalar}} = \frac{1}{2}\left(\partial_\mu \phi\right)\left(\partial^\mu \phi\right) - \frac{1}{2}m^2\phi^2 \rightarrow \text{Euler-Lagrange} \rightarrow \underbrace{(\partial_\mu\partial^\mu + m^2)\phi = 0}_{\text{Klein-Gordon equation}}$$

In electroweak theory, kinematics of fermions, i.e. spin-1/2 particles is described by:

$$\mathcal{L}_{\mathrm{fermion}} = i\bar{\psi}\gamma_\mu\partial^\mu\psi - m\bar{\psi}\psi \rightarrow \text{Euler-Lagrange} \rightarrow \underbrace{\left(i\gamma_\mu\partial^\mu - m\right)\psi = 0}_{\text{Dirac equation}}$$

In general, the Lagrangian for a real scalar particle (ϕ) is given by:

$$\mathcal{L} = \underbrace{\left(\partial_\mu\phi\right)^2}_{\text{kinetic term}} + \underbrace{C}_{\text{constant}} + \underbrace{\alpha\phi}_{?} + \underbrace{\beta\phi^2}_{\text{mass term}} + \underbrace{\gamma\phi^3}_{\text{3-point int.}} + \underbrace{\delta\phi^4}_{\text{4-point int.}} + \;... \tag{11.3}$$

We can interpret the particle spectrum of the theory when studying the Lagrangian under small perturbations. In expression (11.3), the constant (potential) term is for most purposes of no importance as it does not appear in the equation of motion, the term linear in the field has no direct interpretation (and should not be present as we will explain later), the quadratic term in the fields represents the mass of the field/particle and higher order terms describe interaction terms.

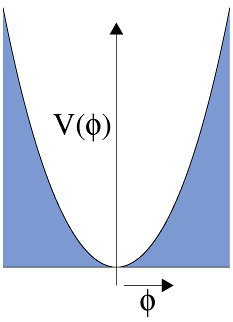
11.3 Simple example of symmetry breaking

To describe the main idea of symmetry breaking we start with a simple model for a real scalar field ϕ (or a theory to which we add a new field ϕ), with a specific potential term:

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}(\partial_\mu\phi)^2 - \mathrm{V}(\phi) \\ &= \frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}\mu^2\phi^2 - \frac{1}{4}\lambda\phi^4 \qquad (11.4)\end{aligned}$$

Note that $\mathcal{L}$ is symmetric under $\phi \to -\phi$ and that λ is positive to ensure an absolute minimum in the Lagrangian. We can investigate in some detail the two possibilities for the sign of μ^2: positive or negative.

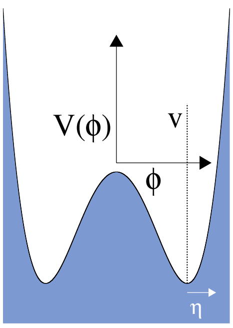
11.3.1 $\mu^2 > 0$: Free particle with additional interactions



To investigate the particle spectrum we look at the Lagrangian for small perturbations around the minimum (vacuum). The vacuum is at $\phi = 0$ and is symmetric in ϕ. Using expression (11.3) we see that the Lagrangian describes a free particle with mass μ that has an additional four-point self-interaction:

$$\mathcal{L} = \underbrace{\frac{1}{2}(\partial_\mu\phi)^2 - \frac{1}{2}\mu^2\phi^2}_{\text{free particle, mass }\mu} \qquad \underbrace{-\frac{1}{4}\lambda\phi^4}_{\text{interaction}}$$

11.3.2 $\mu^2 < 0$: Introducing a particle with imaginary mass ?



The situation with $\mu^2 < 0$ looks strange since at first glance it would appear to describe a particle ϕ with an imaginary mass. However, if we take a closer look at the potential, we see that it does not make sense to interpret the particle spectrum using the field ϕ since perturbation theory around $\phi = 0$ will not converge (not a stable minimum) as the vacuum is located at:

$$\phi_0 = \sqrt{-\frac{\mu^2}{\lambda}} = v \quad \text{or} \quad \mu^2 = -\lambda v^2 \qquad (11.5)$$

As before, to investigate the particle spectrum in the theory, we have to look at small perturbations around this minimum. To do this it is more natural to introduce a field η (simply a shift of the ϕ field) that is centered at the vacuum: $\eta = \phi - v$.

Rewriting the Lagrangian in terms of η

Expressing the Lagrangian in terms of the shifted field η is done by replacing ϕ by $\eta + v$ in the original Lagrangian from equation (11.4):

$$\begin{aligned}\text{Kinetic term:}\quad \mathcal{L}_{\text{kin}}(\eta) &= \frac{1}{2}(\partial_\mu(\eta+v)\partial^\mu(\eta+v)) \\ &= \frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) \qquad\qquad \text{, since } \partial_\mu v = 0.\end{aligned}$$

$$\begin{aligned}\text{Potential term:}\quad \mathrm{V}(\eta) &= +\frac{1}{2}\mu^2(\eta+v)^2 + \frac{1}{4}\lambda(\eta+v)^4 \\ &= \lambda v^2\eta^2 + \lambda v\eta^3 + \frac{1}{4}\lambda\eta^4 - \frac{1}{4}\lambda v^4,\end{aligned}$$

where we used $\mu^2 = -\lambda v^2$ from equation (11.5). Although the Lagrangian is still symmetric in ϕ, the perturbations around the minimum are *not* symmetric in η, i.e. $\mathrm{V}(-\eta) \neq \mathrm{V}(\eta)$. Neglecting the irrelevant $\frac{1}{4}\lambda v^4$ constant term and neglecting terms or order η^2 we have as Lagrangian:

$$\begin{aligned}\text{Full Lagrangian:}\quad \mathcal{L}(\eta) &= \frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \lambda v^2\eta^2 - \lambda v\eta^3 - \frac{1}{4}\lambda\eta^4 - \frac{1}{4}\lambda v^4 \\ &= \frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta) - \lambda v^2\eta^2\end{aligned}$$

From section 11.2 we see that this describes the kinematics for a massive scalar particle:

$$\frac{1}{2}m_\eta^2 = \lambda v^2 \to m_\eta = \sqrt{2\lambda v^2} \quad \left(= \sqrt{-2\mu^2}\right) \qquad \text{Note: } m_\eta > 0.$$

Executive summary on $\mu^2 < 0$ scenario

At first glance, adding a $V(\phi)$ term as in equation (11.4) to the Lagrangian implies adding a particle with imaginary mass with a four-point self-interaction. However, when examining the particle spectrum using perturbations around the vacuum, we see that it actually describes a massive scalar particle (real, positive mass) with three- and four-point self-interactions. Although the Lagrangian retains its original symmetry (symmetric in ϕ), the vacuum is not symmetric in the field η: spontaneous symmetry breaking. Note that we have added a single degree of freedom to the theory: a scalar particle.

11.4 Breaking a global symmetry

In an existing theory we are free to introduce an additional complex scalar field: $\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$ (two degrees of freedom):

$$\mathcal{L} = (\partial_\mu \phi)^*(\partial^\mu \phi) - \mathrm{V}(\phi) \qquad \text{, with } \mathrm{V}(\phi) = \mu^2(\phi^*\phi) + \lambda(\phi^*\phi)^2$$

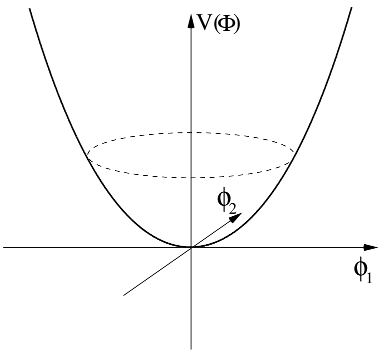
Note that the Lagrangian is invariant under a U(1) global symmetry, i.e. under $\phi' \to e^{i\alpha}\phi$ since $\phi'^*\phi' \to \phi^*\phi e^{-i\alpha}e^{+i\alpha} = \phi^*\phi$.

The Lagrangian in terms of ϕ_1 and ϕ_2 is given by:

$$\mathcal{L}(\phi_1,\phi_2) = \frac{1}{2}(\partial_\mu\phi_1)^2 + \frac{1}{2}(\partial_\mu\phi_2)^2 - \frac{1}{2}\mu^2(\phi_1^2+\phi_2^2) - \frac{1}{4}\lambda(\phi_1^2+\phi_2^2)^2$$

There are again two distinct cases: $\mu^2 > 0$ and $\mu^2 < 0$. As in the previous section, we investigate the particle spectrum by studying the Lagrangian under small perturbations around the vacuum.

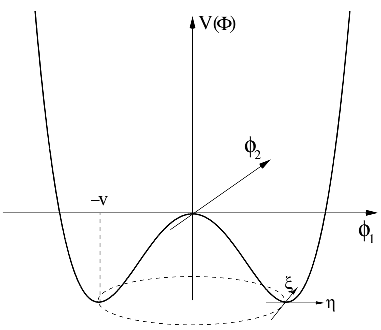
11.4.1 $\mu^2 > 0$



This situation simply describes two massive scalar particles, each with a mass μ with additional interactions:

$$\begin{aligned}\mathcal{L}(\phi_1,\phi_2) &= \underbrace{\frac{1}{2}(\partial_\mu\phi_1)^2 - \frac{1}{2}\mu^2\phi_1^2}_{\text{particle } \phi_1,\text{ mass } \mu} + \underbrace{\frac{1}{2}(\partial_\mu\phi_2)^2 - \frac{1}{2}\mu^2\phi_2^2}_{\text{particle } \phi_2,\text{ mass } \mu} \\ &+ \text{ interaction terms}\end{aligned}$$

11.4.2 $\mu^2 < 0$

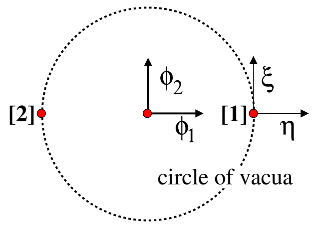


When $\mu^2 < 0$ there is not a single vacuum located at $\binom{0}{0}$, but an infinite number of vacua that satisfy:

$$\sqrt{\phi_1^2+\phi_2^2} = \sqrt{\frac{-\mu^2}{\lambda}} = v$$

From the infinite number we choose ϕ_0 as $\phi_1 = v$ and $\phi_2 = 0$. To see what particles are present in this model, the behaviour of the Lagrangian is studied under small oscillations around the vacuum.

Looking at the symmetry we would use a $\alpha e^{i\beta}$. When looking at perturbations around this minimum it is natural to define the shifted fields η and ξ, with: $\eta = \phi_1 - v$ and $\xi = \phi_2$, which means that the (perturbations around the) vacuum are described by (see section 11.5.2):



$$\phi_0 = \frac{1}{\sqrt{2}}(\eta + v + i\xi)$$

Using $\phi^2 = \phi^*\phi = \frac{1}{2}[(v+\eta)^2 + \xi^2]$ and $\mu^2 = -\lambda v^2$ we can rewrite the Lagrangian in terms of the shifted fields.

$$\begin{aligned}\text{Kinetic term:} \quad \mathcal{L}_{\text{kin}}(\eta,\xi) &= \frac{1}{2}\partial_\mu(\eta + v - i\xi)\partial^\mu(\eta + v + i\xi) \\ &= \frac{1}{2}(\partial_\mu\eta)^2 + \frac{1}{2}(\partial_\mu\xi)^2 \quad \text{, since } \partial_\mu v = 0.\end{aligned}$$

$$\begin{aligned}\text{Potential term:} \quad \mathrm{V}(\eta,\xi) &= \mu^2\phi^2 + \lambda\phi^4 \\ &= -\frac{1}{2}\lambda v^2[(v+\eta)^2 + \xi^2] + \frac{1}{4}\lambda[(v+\eta)^2 + \xi^2]^2 \\ &= -\frac{1}{4}\lambda v^4 + \lambda v^2\eta^2 + \lambda v\eta^3 + \frac{1}{4}\lambda\eta^4 + \frac{1}{4}\lambda\xi^4 + \lambda v\eta\xi^2 + \frac{1}{2}\lambda\eta^2\xi^2\end{aligned}$$

Neglecting the constant and higher order terms, the full Lagrangian can be written as:

$$\mathcal{L}(\eta,\xi) = \underbrace{\frac{1}{2}(\partial_\mu\eta)^2 - (\lambda v^2)\eta^2}_{\text{massive scalar particle } \eta} + \underbrace{\frac{1}{2}(\partial_\mu\xi)^2 + 0\cdot\xi^2}_{\text{massless scalar particle } \xi} + \text{higher order terms}$$

We can identify this as a massive η particle and a massless ξ particle:

$$m_\eta = \sqrt{2\lambda v^2} = \sqrt{-2\mu^2} > 0 \quad \text{and} \quad m_\xi = 0$$

Unlike the η-field, describing radial excitations, there is no 'force' acting on oscillations along the ξ-field. This is a direct consequence of the U(1) symmetry of the Lagrangian and the massless particle ξ is the so-called Goldstone boson.

Goldstone theorem:
For each broken generator of the original symmetry group, i.e. for each generator that connects the vacuum states, one massless spin-zero particle will appear.

Executive summary on breaking a global gauge invariant symmetry

Spontaneously breaking a continuous global symmetry gives rise to a massless (Goldstone) boson. When we break a *local* gauge invariance something special happens and the Goldstone boson will disappear.

11.5 Breaking a local gauge invariant symmetry: the Higgs mechanism

In this section we will take the final step and study what happens if we break a *local* gauge invariant theory. As promised in the introduction, we will explore its consequences using a local U(1) gauge invariant theory we know (QED). As we will see, this will allow to add a mass-term for the gauge boson (the photon).

Local U(1) gauge invariance is the requirement that the Lagrangian is invariant under $\phi' \rightarrow e^{i\alpha(x)}\phi$. From the lectures on electroweak theory we know that this can be achieved by switching to a covariant derivative with a special transformation rule for the vector field. In QED:

$$\begin{aligned} \partial_\mu &\rightarrow D_\mu = \partial_\mu - ieA_\mu && \text{[covariant derivatives]} \\ A'_\mu &= A_\mu + \frac{1}{e}\partial_\mu \alpha && [A_\mu \text{ transformation}] \end{aligned} \tag{11.6}$$

The local U(1) gauge invariant Lagrangian for a complex scalar field is then given by:

$$\mathcal{L} = (D^\mu \phi)^\dagger (D_\mu \phi) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - V(\phi)$$

The term $\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$ is the kinetic term for the gauge field (photon) and $V(\phi)$ is the extra term in the Lagrangian we have seen before: $V(\phi^*\phi) = \mu^2(\phi^*\phi) + \lambda(\phi^*\phi)^2$.

11.5.1 Lagrangian under small perturbations

The situation $\mu^2 > 0$: we have a vacuum at $\binom{0}{0}$. The exact symmetry of the Lagrangian is preserved in the vacuum: we have QED with a massless photon and two massive scalar particles ϕ_1 and ϕ_2 each with a mass μ.

In the situation $\mu^2 < 0$ we have an infinite number of vacua, each satisfying $\phi_1^2 + \phi_2^2 = -\mu^2/\lambda = v^2$. The particle spectrum is obtained by studying the Lagrangian under small oscillations using the same procedure as for the continuous global symmetry from section (11.4.2). Because of <u>local</u> gauge invariance some important differences appear. Extra terms will appear in the kinetic part of the Lagrangian through the covariant derivatives. Using again the shifted fields η and ξ we define the vacuum as $\phi_0 = \frac{1}{\sqrt{2}}[(v + \eta) + i\xi]$.

$$\text{Kinetic term:} \quad \mathcal{L}_{\text{kin}}(\eta,\xi) = (D^\mu\phi)^\dagger (D_\mu\phi) = (\partial^\mu + ieA^\mu)\phi^*(\partial_\mu - ieA_\mu)\phi = \text{... see Exercise 1}$$

$$\text{Potential term:} \quad V(\eta,\xi) = \lambda v^2\eta^2 \text{ , up to second order in the fields. See section 11.4.2.}$$

The full Lagrangian can be written as:

$$\mathcal{L}(\eta,\xi) = \underbrace{\frac{1}{2}(\partial_\mu\eta)^2 - \lambda v^2\eta^2}_{\eta\text{-particle}} + \underbrace{\frac{1}{2}(\partial_\mu\xi)^2}_{\xi\text{-particle}} \underbrace{- \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}e^2v^2A_\mu^2}_{\text{photon field}} - \underbrace{evA_\mu(\partial^\mu\xi)}_{?} + \text{int.-terms} \tag{11.7}$$

At first glance: massive η, massless ξ (as before) and also a mass term for the photon. However, the Lagrangian also contains strange terms that we cannot easily interpret: $-evA_\mu(\partial^\mu\xi)$. This prevents making an easy interpretation.

11.5.2 Rewriting the Lagrangian in the unitary gauge

In a local gauge invariance theory we see that A_μ is fixed up to a term $\partial_\mu\alpha$ as can be seen from equation (11.6). In general, A_μ and ϕ change simultaneously. We can exploit this freedom, to redefine A_μ and remove all terms involving the ξ field.

Looking at the terms involving the ξ-field, we see that we can rewrite them as:

$$\frac{1}{2}(\partial_\mu\xi)^2 - evA^\mu(\partial_\mu\xi) + \frac{1}{2}e^2v^2A_\mu^2 = \frac{1}{2}e^2v^2\left[A_\mu - \frac{1}{ev}(\partial_\mu\xi)\right]^2 = \frac{1}{2}e^2v^2(A_\mu^{'})^2$$

This specific choice, i.e. taking $\alpha = -\xi/v$, is called the *unitary gauge*. Of course, when choosing this gauge (phase of rotation α) the field ϕ changes accordingly (see first part of section 11.1 and dropping terms of $\mathcal{O}(\xi^2,\eta^2,\xi\eta)$):

$$\phi^{'} \rightarrow e^{-i\ \xi/v}\phi = e^{-i\ \xi/v}\frac{1}{\sqrt{2}}(v+\eta+i\xi) = e^{-i\ \xi/v}\frac{1}{\sqrt{2}}(v+\eta)e^{+i\ \xi/v} = \frac{1}{\sqrt{2}}(v+h)$$

Here we have introduced the real h-field. When writing down the full Lagrangian in this specific gauge, we will see that all terms involving the ξ-field will disappear and that the additional degree of freedom will appear as the mass term for the gauge boson associated to the broken symmetry.

11.5.3 Lagrangian in the unitary gauge: particle spectrum

$$
\begin{aligned}
\mathcal{L}_{\text{scalar}} &= (D^\mu\phi)^\dagger (D_\mu\phi) - V(\phi^\dagger\phi) \\
&= (\partial^\mu + ieA^\mu)\frac{1}{\sqrt{2}}(v+h) \;\; (\partial_\mu - ieA_\mu)\frac{1}{\sqrt{2}}(v+h) - V(\phi^\dagger\phi) \\
&= \frac{1}{2}(\partial_\mu h)^2 + \frac{1}{2}e^2A_\mu^2(v+h)^2 - \lambda v^2h^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4 + \frac{1}{4}\lambda v^4
\end{aligned}
$$

Expanding $(v+h)^2$ into 3 terms (and ignoring $\frac{1}{4}\lambda v^4$) we end up with:

$$
= \underbrace{\frac{1}{2}(\partial_\mu h)^2 - \lambda v^2 h^2}_{\substack{\text{massive scalar} \\ \text{particle h}}} + \underbrace{\frac{1}{2}e^2v^2A_\mu^2}_{\substack{\text{gauge field } (\gamma) \\ \text{with mass}}} + \underbrace{e^2vA_\mu^2h + \frac{1}{2}e^2A_\mu^2h^2}_{\substack{\text{interaction Higgs} \\ \text{and gauge fields}}} \underbrace{- \lambda v h^3 - \frac{1}{4}\lambda h^4}_{\substack{\text{Higgs self-} \\ \text{interactions}}}
$$

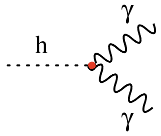
11.5.4 A few words on expanding the terms with $(\mathbf{v}+\mathbf{h})^2$

Expanding the terms in the Lagrangian associated to the vector field we see that we do not only get terms proportional to A_μ^2, i.e. a mass term for the gauge field (photon), but also automatically terms that describe the interaction of the Higgs field with the gauge field. These interactions, related to the mass of the gauge boson, are a consequence of the Higgs mechanism.

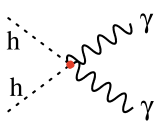
In our model, QED with a massive photon, when expanding $\frac{1}{2}e^2A_\mu^2(v+h)^2$ we get:

1] $\frac{1}{2}e^2v^2A_\mu^2$: the mass term for the gauge field (photon)
Given equation (11.3) we see that $m_\gamma = ev$.

2] $e^2vA_\mu^2h$: photon-Higgs three-point interaction



3] $\frac{1}{2}e^2A_\mu^2h^2$: photon-Higgs four-point interaction



Executive summary: breaking a local gauge invariant symmetry

We added a complex scalar field (2 degrees of freedom) to our existing theory and broke the original symmetry by using a 'strange' potential that yielded a large number of vacua. The additional degrees of freedom appear in the theory as a mass term for the gauge boson connected to the broken symmetry (m_γ) and a massive scalar particle (m_h).

Exercises

Exercise 11.1 (interaction terms)

(a) Compute the 'interaction terms' as given in equation (11.7).

(b) Are the interaction terms symmetric in η and ξ ?

Exercise 11.2 (Toy-model with a massive photon)

(a) Derive expression (14.58) in Halzen & Martin.
Hint: you can either do the full computation or, much less work, just insert $\phi = \frac{1}{\sqrt{2}}(v+h)$ in the Lagrangian and keep A_μ unchanged.

(b) Show that in this model the Higgs boson can decay into two photons and that the coupling $h \to \gamma\gamma$ is proportional to m_γ.

(c) Draw all Feynman vertices that are present in this model and show that Higgs three-point (self-)coupling, or $h \to hh$, is proportional to $\mathrm{m_h}$.

(d) Higgs boson properties: how can you see from the Lagrangian that the Higgs boson is a scalar (spin 0) particle ? What defines the 'charge' of the Higgs boson ?

Exercise 11.3 (the potential part: $\mathbf{V}(\phi^\dagger\phi)$)

Use in this exercise $\phi = \frac{1}{\sqrt{2}}(v+h)$ and that ϕ is real (1 dimension).

(a) The normal Higgs potential: $V(\phi^\dagger\phi) = \mu^2\phi^2 + \lambda\phi^4$.
Show that $\frac{1}{2}m_h^2 = \frac{1}{2}\lambda v^2$, where $(\phi_0 = v)$. How many vacua are there?

(b) Why is $V(\phi^\dagger\phi) = \mu^2\phi^2 + \beta\phi^3$ not possible ?
How many vacua are there?

Terms $\propto \phi^6$ are allowed since they introduce additional interactions that are not cancelled by gauge boson interactions, making the model non-renormalizable. Just ignore this little detail for the moment and compute the 'prediction' for the Higgs boson mass.

(c) Use $V(\phi^\dagger\phi) = \mu^2\phi^2 + \lambda\phi^4 + \frac{4}{3}\delta\phi^6$, with $\mu^2 < 0$, $\lambda > 0$ and $\delta = -\frac{2\lambda^2}{\mu^2}$.
Show that $\mathrm{m}_h(\text{new}) = \sqrt{\frac{3}{2}}\mathrm{m}_h(\text{old})$, with 'old': m_h for the normal Higgs potential.

Lecture 12

The Higgs mechanism in the Standard Model

In this section we will apply the idea of spontaneous symmetry breaking from section 11 to the model of electroweak interactions. With a specific choice of parameters we can obtain massive Z and W bosons while keeping the photon massless.

12.1 Breaking the local gauge invariant $\mathbf{SU(2)_L \times U(1)_Y}$ symmetry

To break the $\mathrm{SU(2)_L \times U(1)_Y}$ symmetry we follow the ingredients of the Higgs mechanism:

1) Add an isospin doublet:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}$$

Since we would like the Lagrangian to retain all its symmetries, we can only add $\mathrm{SU(2)_L \times U(1)_Y}$ multiplets. Here we add a left-handed doublet (like the electron neutrino doublet) with weak Isospin $\frac{1}{2}$. The electric charges of the upper and lower component of the doublet are chosen to ensure that the hypercharge Y=+1. This requirement is vital for reasons that will become more evident later.

2) Add a potential V(ϕ) for the field that will break (spontaneously) the symmetry:

$$V(\phi) = \mu^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2 \quad \text{, with } \mu^2 < 0$$

The part added to the Lagrangian for the scalar field

$$\mathcal{L}_{\text{scalar}} = (D^\mu\phi)^\dagger(D_\mu\phi) - V(\phi),$$

where D_μ is the covariant derivative associated to $\mathrm{SU(2)_L \times U(1)_Y}$:

$$D_\mu = \partial_\mu + ig\frac{1}{2}\vec{\tau}\cdot\vec{W}_\mu + ig'\frac{1}{2}YB_\mu$$

3) Choose a vacuum:
We have seen that any choice of the vacuum that breaks a symmetry will generate a mass for the corresponding gauge boson. The vacuum we choose has ϕ_1=ϕ_2=ϕ_4=0 and $\phi_3 = v$:

$$\text{Vacuum} = \phi_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+h \end{pmatrix}$$

This vacuum as defined above is neutral since $I = \frac{1}{2}$, $I_3 = -\frac{1}{2}$ and with our choice of $Y = +1$ we have $Q = I_3 + \frac{1}{2}Y$=0. We will see that this choice of the vacuum breaks $\mathrm{SU(2)_L \times U(1)_Y}$,but leaves $\mathrm{U(1)_{EM}}$ invariant, leaving only the photon massless. In writing down this vacuum we immediately went to the unitary gauge (see section 11.5).

12.2 Checking which symmetries are broken in a given vacuum

How do we check if the symmetries associated to the gauge bosons are broken ? Invariance implies that $e^{i\alpha Z}\phi_0 = \phi_0$, with Z the associated 'rotation'. Under infinitesimal rotations this means $(1 + i\alpha Z)\phi_0 = \phi_0 \rightarrow Z\phi_0 = 0$.

What about the $\mathrm{SU(2)_L}$, $\mathrm{U(1)_Y}$ and $\mathrm{U(1)_{EM}}$ generators:

$$\begin{aligned}
\mathrm{SU(2)_L}: \quad \tau_1\phi_0 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+h \end{pmatrix} = +\frac{1}{\sqrt{2}}\begin{pmatrix} v+h \\ 0 \end{pmatrix} \neq 0 \rightarrow \text{broken} \\
\tau_2\phi_0 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+h \end{pmatrix} = -\frac{i}{\sqrt{2}}\begin{pmatrix} v+h \\ 0 \end{pmatrix} \neq 0 \rightarrow \text{broken} \\
\tau_3\phi_0 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+h \end{pmatrix} = -\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+h \end{pmatrix} \neq 0 \rightarrow \text{broken} \\
\mathrm{U(1)_Y}: \quad Y\phi_0 &= Y_{\phi_0}\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+h \end{pmatrix} = +\frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v+h \end{pmatrix} \neq 0 \rightarrow \text{broken}
\end{aligned}$$

This means that all 4 gauge bosons (W_1, W_2, W_3 and B) acquire a mass through the Higgs mechanism. In the lecture on electroweak theory we have seen that the W_1 and W_2 fields mix to form the charged W^+ and W^- bosons and that the W_3 and B field will mix to form the neutral Z-boson and photon.

$$\underbrace{W_1 \quad W_2}_{\text{W}^+ \text{ and W}^- \text{ bosons}} \qquad \underbrace{W_3 \quad B}_{\text{Z-boson and } \gamma}$$

When computing the masses of these mixed physical states in the next sections, we will see that one of these combinations (the photon) remains massless. Looking at the symmetries we can already predict this is the case. For the photon to remain massless the $\mathrm{U(1)_{EM}}$ symmetry should leave the vacuum invariant. And indeed:

$$\mathrm{U(1)_{EM}}: \quad Q\phi_0 \quad = \quad \frac{1}{2}(\tau_3 + Y)\phi_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+h \end{pmatrix} = 0 \rightarrow \text{unbroken}$$

It is not so strange that $\mathrm{U(1)_{EM}}$ is conserved as the vacuum is neutral and we have:

$$\phi_0' \rightarrow e^{i\alpha Q_{\phi_0}} \phi_0 = \phi_0$$

Breaking of $\mathbf{SU(2)_L \times U(1)_Y}$: looking a bit ahead

1) W_1 and W_2 mix and will form the massive a W^+ and W^- bosons.
2) W_3 and B mix to form massive Z and massless γ.
3) Remaining degree of freedom will form the mass of the scalar particle (Higgs boson).

12.3 Scalar part of the Lagrangian: gauge boson mass terms

Studying the scalar part of the Lagrangian

To obtain the masses for the gauge bosons we will only need to study the scalar part of the Lagrangian:

$$\mathcal{L}_{\text{scalar}} = (D^\mu\phi)^\dagger (D_\mu\phi) - V(\phi) \tag{12.1}$$

The $V(\phi)$ term will again give the mass term for the Higgs boson and the Higgs self-interactions. The $(D^\mu\phi)^\dagger(D_\mu\phi)$ terms:

$$D_\mu\phi = \left[\partial_\mu + ig\frac{1}{2}\vec{\tau}\cdot\vec{W}_\mu + ig'\frac{1}{2}YB_\mu\right] \; \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+h \end{pmatrix}$$

will give rise to the masses of the gauge bosons (and the interaction of the gauge bosons with the Higgs boson) since, as we discussed in section 11.5.4, working out the $(v+h)^2$-terms from equation (12.1) will give us three terms:

1) Masses for the gauge bosons ($\propto v^2$)
2) Interactions gauge bosons and the Higgs ($\propto vh$) and ($\propto h^2$)

In the exercises we will study the interactions of the Higgs boson and the gauge boson (the terms in 2)) in detail, but since we are here primarily interested in the masses of the vector bosons we will only focus on 1):

$$
\begin{aligned}
(D_\mu \phi) &= \frac{1}{\sqrt{2}}\left[ig\frac{1}{2}\vec{\tau}\cdot\vec{W}_\mu + ig'\frac{1}{2}YB_\mu\right]\begin{pmatrix}0\\ v\end{pmatrix}\\
&= \frac{i}{\sqrt{8}}\left[g(\tau_1 W_1 + \tau_2 W_2 + \tau_3 W_3) + g^{'}YB_\mu\right]\begin{pmatrix}0\\ v\end{pmatrix}\\
&= \frac{i}{\sqrt{8}}\left[g\left(\begin{pmatrix}0 & W_1\\ W_1 & 0\end{pmatrix} + \begin{pmatrix}0 & -iW_2\\ iW_2 & 0\end{pmatrix} + \begin{pmatrix}W_3 & 0\\ 0 & -W_3\end{pmatrix}\right) + g^{'}\begin{pmatrix}Y_{\phi_0}B_\mu & 0\\ 0 & Y_{\phi_0}B_\mu\end{pmatrix}\right]\begin{pmatrix}0\\ v\end{pmatrix}\\
&= \frac{i}{\sqrt{8}}\begin{pmatrix}gW_3 + g^{'}Y_{\phi_0}B_\mu & g(W_1 - iW_2)\\ g(W_1 + iW_2) & -gW_3 + g^{'}Y_{\phi_0}B_\mu\end{pmatrix}\begin{pmatrix}0\\ v\end{pmatrix}\\
&= \frac{iv}{\sqrt{8}}\begin{pmatrix}g(W_1 - iW_2)\\ -gW_3 + g^{'}Y_{\phi_0}B_\mu\end{pmatrix}
\end{aligned}
$$

We can then also easily compute $(D^\mu\phi)^\dagger$: $(D^\mu\phi)^\dagger = -\frac{iv}{\sqrt{8}}\left(g(W_1 + iW_2)\ ,\ (-gW_3 + g^{'}Y_{\phi_0}B_\mu)\right)$ and we get the following expression for the kinetic part of the Lagrangian:

$$
(D^\mu\phi)^\dagger\,(D_\mu\phi) = \frac{1}{8}v^2\left[g^2(W_1^2 + W_2^2) + (-gW_3 + g^{'}Y_{\phi_0}B_\mu)^2\right] \tag{12.2}
$$

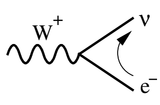
12.3.1 Rewriting $(\mathbf{D}^\mu\phi)^\dagger\,(\mathbf{D}_\mu\phi)$ in terms of physical gauge bosons

Before we can interpret this we need to rewrite this in terms of W$^+$, W$^-$, Z and γ since that are the gauge bosons that are observed in nature.

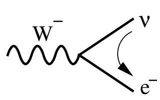
1] Rewriting terms with $\mathbf{W_1}$ and $\mathbf{W_2}$ terms: charged gauge bosons $\mathbf{W^+}$ and $\mathbf{W^-}$

When discussing the charged current interaction on SU(2)$_\mathrm{L}$ doublets we saw that the charge raising and lowering operators connecting the members of isospin doublets were τ_+ and τ_-, linear combinations of τ_1 and τ_2 and that each had an associated gauge boson: the W$^+$ and W$^-$.

$$
\tau_+ = \frac{1}{2}(\tau_1 + i\tau_2) = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}
$$



$$
\tau_- = \frac{1}{2}(\tau_1 - i\tau_2) = \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}
$$



We can rewrite W_1, W_2 terms as W^+, W^- using $W^\pm = \frac{1}{\sqrt{2}}(W_1 \mp iW_2)$. In particular, $\frac{1}{2}(\tau_1 W_1 + \tau_2 W_2) = \frac{1}{\sqrt{2}}(\tau_+ W^+ + \tau_- W^-)$.

Looking at the terms involving W_1 and W_2 in the Lagrangian in equation (12.2), we see that:

$$g^2(W_1^2 + W_2^2) = g^2(W^{+2} + W^{-2}) \text{ or, alternatively, } 2g^2W^+W^- \tag{12.3}$$

2] Rewriting terms with W$_3$ and B$_\mu$ terms: neutral gauge bosons Z and γ

$$(-gW_3 + g'Y_{\phi_0}B_\mu)^2 = (W_3, B_\mu)\begin{pmatrix} g^2 & -gg'Y_{\phi_0} \\ -gg'Y_{\phi_0} & g'^2 \end{pmatrix}\begin{pmatrix} W_3 \\ B_\mu \end{pmatrix}$$

When looking at this expression there are some important things to note, especially related to the role of the hypercharge of the vacuum, Y_{ϕ_0}:

1 Only if $Y_{\phi_0} \neq 0$, the W_3 and B_μ fields mix.

2 If $Y_{\phi_0} = \pm\,1$, the determinant of the mixing matrix vanishes and one of the combinations will be massless (the coefficient for that gauge field squared is 0). In our choice of vacuum we have $Y_{\phi_0} = +1$ (see Exercise 4 why that is a good idea). In the rest of our discussion we will drop the term Y_{ϕ_0} and simply use its value of 1.

The two eigenvalues and eigenvectors are given by [see Exercise 3]:

eigenvalue *eigenvector*

$$\lambda = 0 \quad \rightarrow \quad \frac{1}{\sqrt{g^2+g'^2}}\begin{pmatrix} g' \\ g \end{pmatrix} = \frac{1}{\sqrt{g^2+g'^2}}(g'W_3 + gB_\mu) = A_\mu \quad \text{photon}(\gamma)$$

$$\lambda = (g^2+g'^2) \quad \rightarrow \quad \frac{1}{\sqrt{g^2+g'^2}}\begin{pmatrix} g \\ -g' \end{pmatrix} = \frac{1}{\sqrt{g^2+g'^2}}(gW_3 - g'B_\mu) = Z_\mu \quad \text{Z-boson (Z)}$$

Looking at the terms involving W_3 and B in the Lagrangian we see that:

$$(-gW_3 + g'Y_{\phi_0}B_\mu)^2 = (g^2+g'^2)Z_\mu^2 + 0\cdot A_\mu^2 \tag{12.4}$$

3] Rewriting Lagrangian in terms of physical fields: masses of the gauge bosons

Finally, by combining equation (12.3) and (12.4) we can rewrite the Lagrangian from equation (12.2) in terms of the physical gauge bosons:

$$(D^\mu\phi)^\dagger (D_\mu\phi) = \frac{1}{8}v^2[g^2(W^+)^2 + g^2(W^-)^2 + (g^2+g'^2)Z_\mu^2 + 0\cdot A_\mu^2] \tag{12.5}$$

12.4 Masses of the gauge bosons

12.4.1 Massive charged and neutral gauge bosons

As a general mass term for a massive gauge boson V has the form $\frac{1}{2}M_V^2 V_\mu^2$, from equation (12.5) we see that:

$$\begin{aligned} M_{W^+} = M_{W^-} &= \frac{1}{2}vg \\ M_Z &= \frac{1}{2}v\sqrt{(g^2 + g'^2)} \end{aligned}$$

Although since g and g' are free parameters, the SM makes no absolute predictions for $\mathrm{M_W}$ and $\mathrm{M_Z}$, it has been possible to set a lower limit before the W- and Z-boson were discovered (see Exercise 2). The measured values are $M_W = 80.4$ GeV and $M_Z = 91.2$ GeV.

Mass relation W and Z boson:

Although there is no absolute prediction for the mass of the W- and Z-boson, there is a clear prediction on the ratio between the two masses. From discussions in QED we know the photon couples to charge, which allowed us to relate e, g and g' (see Exercise 3):

$$e = g\sin(\theta_\mathrm{W}) = g'\cos(\theta_\mathrm{W}) \tag{12.6}$$

In this expression θ_W is the Weinberg angle, often used to describe the mixing of the W_3 and B_μ -fields to form the physical Z boson and photon. From equation (12.6) we see that $g'/g = \tan(\theta_\mathrm{W})$ and therefore:

$$\frac{M_W}{M_Z} = \frac{\frac{1}{2}vg}{\frac{1}{2}v\sqrt{g^2 + g'^2}} = \cos(\theta_\mathrm{W})$$

This predicted ratio is often expressed as the so-called ρ-(Veltman) parameter:

$$\rho = \frac{M_W^2}{M_Z^2 \cos^2(\theta_W)} = 1$$

The current measurements of the M_W, M_Z and θ_W confirm this relation.

12.4.2 Massless neutral gauge boson (γ):

Similar to the Z boson we have now a mass for the photon: $\frac{1}{2}M_\gamma^2 = 0$, so:

$$M_\gamma = 0. \tag{12.7}$$

12.5 Mass of the Higgs boson

Looking at the mass term for the scalar particle, the mass of the Higgs boson is given by:

$$m_h = \sqrt{2\lambda v^2}$$

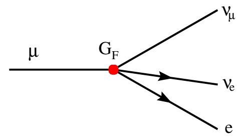
Although v is known ($v \approx 246$ GeV, see below), since λ is a free parameter, the mass of the Higgs boson is **not** predicted in the Standard Model.

Extra: how do we know v ?:

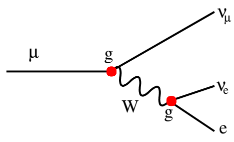
$$\text{Muon decay: } \frac{g^2}{8M_W^2} = \frac{G_F}{\sqrt{2}} \rightarrow v = \sqrt{\frac{1}{\sqrt{2}G_F}}$$

We used $M_W = \frac{1}{2}vg$. Given $G_F = 1.166 \cdot 10^{-5}$, we see that $v = 246$ GeV. This energy scale is known as the electroweak scale.

Fermi: $\propto \frac{G_F}{\sqrt{2}}$



EW: $\propto \frac{g^2}{8M_W^2}$



Exercises

Exercise 12.1 (Higgs - Vector boson couplings)
In the lecture notes we focussed on the masses of the gauge bosons, i.e. part 1) when expanding the $((v+h)^2)$-terms as discussed in Section 11.5.4 and 12.3. Looking now at the terms in the Lagrangian that describe the interaction between the gauge fields and the Higgs field, show that the four vertex factors describing the interaction between the Higgs boson and gauge bosons: hWW, hhWW, hZZ, hhZZ are given by:

3-point: $2i\frac{M_V^2}{v}g^{\mu\nu}$ and 4-point: $2i\frac{M_V^2}{v^2}g^{\mu\nu}$, with (V= W,Z).

Note: A vertex factor is obtained by multiplying the term involving the interacting fields in the Lagrangian by a factor i and a factor $n!$ for n identical particles in the vertex.

Exercise 12.2 (History: lower limits on $\mathrm{M_W}$ and $\mathrm{M_Z}$)
Use the relations $e = g\sin\theta_W$ and $G_F = (v^2\sqrt{2})^{-1}$ to obtain lower limits for the masses of the W and Z boson assuming that you do not know the value of the weak mixing angle.

Exercise 12.3 (Electroweak mixing: $(\mathbf{W^3_\mu}, \mathbf{B_\mu}) \rightarrow (\mathbf{A_\mu}, \mathbf{Z_\mu})$)
The mix between the W^3_μ and B_μ fields in the lagrangian can be written in a matrix

notation:

$$(W_\mu^3, B_\mu)\begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix}\begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}$$

(a) Show that the eigenvalues of the matrix are $\lambda_1 = 0$ and $\lambda_2 = (g^2 + g'^2)$.

(b) Show that these eigenvalues correspond to the two eigenvectors:

$$V_1 = \frac{1}{\sqrt{g^2+g'^2}}(g'W_\mu^3 + gB_\mu) \equiv \mathrm{A}_\mu \quad \text{and} \quad V_2 = \frac{1}{\sqrt{g^2+g'^2}}(gW_\mu^3 - g'B_\mu) \equiv \mathrm{Z}_\mu$$

(c) **bonus:** Imagine that we would have chosen $Y_{\phi_0'} = -1$. What, in that scenario, would be the (mass-)eigenvectors A_μ' and Z_μ', the 'photon' and 'Z-boson' ? In such a model, what would be their masses ? Compare them to those in the Standard Model.

Exercise 12.4 (A closer look at the covariant derivative)
The covariant derivative in the electroweak theory is given by:

$$D_\mu = \partial_\mu + ig'\frac{Y}{2}B_\mu + ig\vec{T}\cdot\vec{W}_\mu$$

(a) Looking only at the part involving W_μ^3 and B_μ show that:

$$D_\mu = \partial_\mu + i\mathrm{A}_\mu\frac{gg'}{\sqrt{g'^2+g^2}}\left(T_3 + \frac{Y}{2}\right) + i\mathrm{Z}_\mu\frac{1}{\sqrt{g'^2+g^2}}\left(g^2T_3 - g'^2\frac{Y}{2}\right)$$

(b) Make also a final interpretation step for the A_μ part and show that:

$$\frac{gg'}{\sqrt{g'^2+g^2}} = e \quad \text{and} \quad T_3 + \frac{Y}{2} = Q, \text{ the electric charge.}$$

(c) **bonus:** Imagine that we would have chosen $Y_{\phi_0'} = -1$. Show explicitly that in that case the photon does not couple to the electric charge.

Exercise 12.5 (Gauge bosons in a model with an SU(2)$_\mathrm{L}$ symmetry)
Imagine a system described by a local SU(2)$_\mathrm{L}$ gauge symmetry (iso-spin only) in which all gauge bosons are be massive. Note that this is different from the SU(2)$_\mathrm{L}\times$U(1)$_\mathrm{Y}$ symmetry of the SM involving also hypercharge. In this alternative model:

(a) Explain why the Higgs field ϕ needs to be an SU(2)$_\mathrm{L}$ doublet.

(b) How many gauge bosons are there and how many degrees of freedom does ϕ have?

(c) Determine the masses of the gauge bosons in this model.

(d) What property of the particles do the gauge bosons couple to and what defines the 'charge' of the gauge bosons themselves?

Lecture 13

Fermion masses, Higgs decay and limits on $\mathbf{m_h}$

In this section we discuss how fermions acquire a mass and use our knowledge on the Higgs coupling to fermions and gauge bosons to predict how the Higgs boson decays as a function of its mass. Even though the Higgs boson has been discovered, we also discuss what theoretical information we have on the mass of the Higgs boson as it reveals the impact on the Higgs boson at higher energy scales (evolution of the universe).

13.1 Fermion masses

In section 11 we saw that terms like $\frac{1}{2}B_\mu B^\mu$ and $m\bar{\psi}\psi$ were not gauge invariant. Since these terms are not allowed in the Lagrangian, both gauge bosons and fermions are massless. In the previous section we have seen how the Higgs mechanism can be used to accommodate massive gauge bosons in our theory while keeping the local gauge invariance. As we will now see, the Higgs mechanism can also give fermions a mass: 'twee vliegen in een klap'.

Chirality and a closer look at terms like $-m\bar{\psi}\psi$

A term like $-m\bar{\psi}\psi = -m[\bar{\psi_L}\psi_R + \bar{\psi_R}\psi_L]$, i.e. a decomposition in chiral states (see exercise 1). Such a term in the Lagrangian is <u>not</u> gauge invariant since the left handed fermions form an isospin doublet (for example $\begin{pmatrix} \nu \\ e \end{pmatrix}_L$) and the right handed fermions form isospin singlets like e_R. They transform differently under $\mathrm{SU(2)_L \times U(1)_Y}$.

$$\begin{aligned} \text{left handed doublet} &= \chi_L \to \chi_L^{'} = \chi_L e^{i\vec{W}\cdot\vec{T}+i\alpha Y} \\ \text{right handed singlet} &= \psi_R \to \psi_R^{'} = \psi_R e^{i\alpha Y} \end{aligned}$$

This means that the term is not invariant under all $\mathrm{SU(2)_L \times U(1)_Y}$ 'rotations'.

Constructing an $\mathrm{SU(2)_L \times U(1)_Y}$ invariant term for fermions

If we could make a term in the Lagrangian that is a *singlet* under $\mathrm{SU(2)_L}$ and $\mathrm{U(1)_Y}$, it would remain invariant. This can be done using the complex (Higgs) doublet we introduced in the previous section. It can be shown that the Higgs has exactly the right quantum numbers to form an $SU(2)_L$ and $U(1)_Y$ singlet in the vertex: $-\lambda_f \bar{\psi}_L \phi \psi_R$, where λ_f is a so-called Yukawa coupling.

Executive summary: - a term: $\propto \bar{\psi}_L \psi_R$ is **not** invariant under $\mathrm{SU(2)_L \times U(1)_Y}$
- a term: $\propto \bar{\psi}_L \phi \psi_R$ is invariant under $\mathrm{SU(2)_L \times U(1)_Y}$

We have constructed a term in the Lagrangian that couples the Higgs doublet to the fermion fields:

$$\mathcal{L}_{\text{fermion-mass}} = -\lambda_f [\bar{\psi}_L \phi \psi_R + \bar{\psi}_R \bar{\phi} \psi_L] \tag{13.1}$$

When we write out this term we'll see that this does not only describe an interaction between the Higgs field and fermion, but that the fermions will acquire a finite mass if the ϕ-doublet has a non-zero expectation value. This is the case as $\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+h \end{pmatrix}$ as before.

13.1.1 Lepton masses

$$\begin{aligned}
\mathcal{L}_e &= -\lambda_e \frac{1}{\sqrt{2}} \left[(\bar{\nu}, \bar{e})_L \begin{pmatrix} 0 \\ v+h \end{pmatrix} e_R + \bar{e}_R (0, v+h) \begin{pmatrix} \nu \\ e \end{pmatrix}_L \right] \\
&= -\frac{\lambda_e (v+h)}{\sqrt{2}} [\bar{e}_L e_R + \bar{e}_R e_L] \\
&= -\frac{\lambda_e (v+h)}{\sqrt{2}} \bar{e} e \\
&= - \underbrace{\frac{\lambda_e v}{\sqrt{2}} \bar{e} e}_{\text{electron mass term}} - \underbrace{\frac{\lambda_e}{\sqrt{2}} h \bar{e} e}_{\text{electron-higgs interaction}}
\end{aligned}$$

$$m_e = \frac{\lambda_e v}{\sqrt{2}} \qquad \frac{\lambda_e}{\sqrt{2}} \propto m_e$$

A few side-remarks:

1) The Yukawa coupling is often expressed as $\lambda_f = \sqrt{2} \left(\frac{m_f}{v} \right)$ and the coupling of the fermion to the Higgs field is $\frac{\lambda_f}{\sqrt{2}} = \frac{m_f}{v}$, so proportional to the mass of the fermion.
2) The mass of the electron is **not** predicted since λ_e is a free parameter. In that sense the Higgs mechanism does not say anything about the electron mass itself.

3) The coupling of the Higgs boson to electrons is **very** small:
The coupling of the Higgs boson to an electron-pair ($\propto \frac{m_e}{v} = \frac{gm_e}{2M_W}$) is very small compared to the coupling of the Higgs boson to a pair of W-bosons ($\propto gM_W$).

$$\frac{\Gamma(h \to ee)}{\Gamma(h \to WW)} \propto \frac{\lambda_{eeh}^2}{\lambda_{WWh}^2} = \left(\frac{gm_e/2M_W}{gM_W}\right)^2 = \frac{m_e^2}{4M_W^4} \approx 1.5 \cdot 10^{-21}$$

13.1.2 Quark masses

The fermion mass term $\mathcal{L}_{\text{down}} = \lambda_f \bar{\psi}_L \phi \psi_R$ (leaving out the hermitian conjugate term $\bar{\psi}_R \bar{\phi} \psi_L$ for clarity) only gives mass to 'down' type fermions, i.e. only to one of the isospin doublet components. To give the neutrino a mass and give mass to the 'up' type quarks (u, c, t), we need another term in the Lagrangian. Luckily it is possible to compose a new term in the Lagrangian, using again the complex (Higgs) doublet in combination with the fermion fields, that is gauge invariant under $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ and gives a mass to the up-type quarks. The mass-term for the up-type fermions takes the form:

$$\mathcal{L}_{\text{up}} = \bar{\chi}_L \tilde{\phi}^c \phi_R + \text{h.c., with}$$

$$\tilde{\phi}^c = -i\tau_2 \phi^* = -\frac{1}{\sqrt{2}} \begin{pmatrix} (v+h) \\ 0 \end{pmatrix} \tag{13.2}$$

Mass terms for fermions (leaving out h.c. term):

$$\begin{aligned} \text{down-type:} \quad \lambda_d(\bar{u}_L, \bar{d}_L)\phi d_R &= \lambda_d(\bar{u}_L, \bar{d}_L) \begin{pmatrix} 0 \\ v \end{pmatrix} d_R = \lambda_d v \; \bar{d}_L d_R \\ \text{up-type:} \quad \lambda_u(\bar{u}_L, \bar{d}_L)\tilde{\phi}^c d_R &= \lambda_u(\bar{u}_L, \bar{d}_L) \begin{pmatrix} v \\ 0 \end{pmatrix} u_R = \lambda_u v \; \bar{u}_L u_R \end{aligned}$$

As we will discuss now, this is not the whole story. If we look more closely we'll see that we can construct more fermion-mass-type terms in the Lagrangian that cannot easily be interpreted. Getting rid of these terms is at the origin of quark mixing.

13.2 Yukawa couplings and the origin of Quark Mixing

This section will discuss in full detail the consequences of all possible allowed quark 'mass-like' terms and study the link between the Yukawa couplings and quark mixing in the Standard Model: the difference between *mass eigenstates* and *flavour eigenstates*.

If we focus on the part of the SM Lagrangian that describes the dynamics of spinor (fermion) fields ψ, the kinetic terms, we see that:

$$\mathcal{L}_{\text{kinetic}} = i\bar{\psi}(\partial^\mu \gamma_\mu)\psi,$$

where $\bar{\psi} \equiv \psi^\dagger \gamma^0$ and the spinor fields ψ. It is instructive to realise that the spinor fields ψ are the three fermion generations can be written in the following five (interaction) representations:

general spinor field	Ψ^I(color, weak iso-spin, hypercharge)
1) left handed quarks	$Q^I_{Li}(3, 2, +1/3)$
2) right handed up-type quarks	$u^I_{Ri}(3, 1, +4/3)$
3) right handed down-type quarks	$d^I_{Ri}(3, 1, +1/3)$
4) left handed fermions	$L^I_{Li}(1, 2, -1)$
5) right handed fermions	$l^I_{Ri}(1, 1, -2)$

In this notation, $Q^I_{Li}(3, 2, +1/3)$ describes an $\mathrm{SU(3)_C}$ triplet, $\mathrm{SU(2)_L}$ doublet, with hypercharge $Y = 1/3$. The superscript I implies that the fermion fields are expressed in the *interaction (flavour)* basis. The subscript i stands for the three generations (families). Explicitly, $Q^I_{Li}(3, 2, +1/3)$ is therefor a shorthand notation for:

$$Q^I_{Li}(3, 2, +1/3) = \begin{pmatrix} u^I_g, u^I_r, u^I_b \\ d^I_g, d^I_r, d^I_b \end{pmatrix}_i = \begin{pmatrix} u^I_g, u^I_r, u^I_b \\ d^I_g, d^I_r, d^I_b \end{pmatrix}, \begin{pmatrix} c^I_g, c^I_r, c^I_b \\ s^I_g, s^I_r, s^I_b \end{pmatrix}, \begin{pmatrix} t^I_g, t^I_r, t^I_b \\ b^I_g, b^I_r, b^I_b \end{pmatrix}.$$

We saw that using the Higgs field ϕ we could construct terms in the Lagrangian of the form given in equation (13.1). For up and down type fermions (leaving out the hermitian conjugate term) that would allow us to write for example:

$$\begin{aligned} \mathcal{L}_{\text{quarks}} &= -\Lambda_{\text{down}}\bar{\chi}_L \phi \psi_R - \Lambda_{\text{up}}\bar{\chi}_L \tilde{\phi}^c \psi_R \\ &= -\Lambda_{\text{down}}\frac{v}{\sqrt{2}}\,\bar{d}^I d^I - \Lambda_{\text{up}}\frac{v}{\sqrt{2}}\,\bar{u}^I u^I, \\ &= -m_d\,\bar{d}^I d^I - m_u\,\bar{u}^I u^I, \end{aligned}$$

where the strength of the interactions between the Higgs and the fermions, the so-called Yukawa couplings, had again to be added by hand.

This looks straightforward, but there is an additional complication when you realize that in the most general realization the Λ's are matrices. This will introduce mixing between different flavours as we will see a little bit later. In the most general case, again leaving out the *h.c.*, the expression for the fermion masses is written as:

$$\begin{aligned} -\mathcal{L}_{\text{Yukawa}} &= Y_{ij}\overline{\psi_{Li}}\phi\psi_{Rj} \\ &= Y^d_{ij}\overline{Q^I_{Li}}\,\phi\,d^I_{Rj} + Y^u_{ij}\overline{Q^I_{Li}}\,\tilde{\phi}^c\,u^I_{Rj} + Y^l_{ij}\overline{L^I_{Li}}\,\phi\,l^I_{Rj}, \end{aligned} \tag{13.3}$$

where the last term is the mass term for the charged leptons. The matrices Y^d_{ij}, Y^u_{ij} and Y^l_{ij} are arbitrary complex matrices that connect the flavour eigenstate since also terms like $Y_{\rm uc}$ will appear. These terms have no easy interpretation:

$$-\mathcal{L}_{\rm Yukawa} = \quad \ldots + \quad \underbrace{-\Lambda_{dd}\frac{v}{\sqrt{2}}\bar{u}^I u^I}_{\text{mass−term down quark}} \quad \underbrace{-\Lambda_{us}\frac{v}{\sqrt{2}}\bar{u}^I s^I}_{??} \quad \underbrace{-\Lambda_{ss}\frac{v}{\sqrt{2}}\bar{s}^I s^I}_{\text{mass−term strange quark}} \quad + \ \ldots \tag{13.4}$$

To interpret the fields in the theory as physical particles, the fields in our model should have a well-defined mass. This is not the case in equation (13.4). If we write out all Yukawa terms in the Lagrangian we realize that it is possible to re-write them in terms of mixed fields that *do* have a well-defined mass. These states are the physical particles in the theory

Writing out the full Yukawa terms:

Since this is the crucial part of flavour physics, we spell out the term $Y^d_{ij}\overline{Q^I_{Li}}\ \phi\ d^I_{Rj}$ explicitly and forget about the other 2 terms in expression (13.3):

$$Y^d_{ij}\overline{Q^I_{Li}}\ \phi\ d^I_{Rj} \quad = \quad Y^d_{ij}\overline{(\text{up-type}\ \ \text{down-type})^I_{iL}}\begin{pmatrix}\phi^+\\ \phi\end{pmatrix}(\text{down-type})^I_{Rj} =$$

$$\begin{pmatrix} Y_{11}\overline{(u\ \ d)^I_L}\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} & Y_{12}\overline{(u\ \ d)^I_L}\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} & Y_{13}\overline{(u\ \ d)^I_L}\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} \\ Y_{21}\overline{(c\ \ s)^I_L}\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} & Y_{22}\overline{(c\ \ s)^I_L}\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} & Y_{23}\overline{(c\ \ s)^I_L}\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} \\ Y_{31}\overline{(t\ \ b)^I_L}\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} & Y_{32}\overline{(t\ \ b)^I_L}\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} & Y_{33}\overline{(t\ \ b)^I_L}\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} \end{pmatrix} \cdot \begin{pmatrix} d^I_R \\ s^I_R \\ b^I_R \end{pmatrix}$$

After symmetry breaking we get the following mass terms for the fermion fields:

$$\begin{aligned} -\mathcal{L}^{\rm quarks}_{\rm Yukawa} &= Y^d_{ij}\overline{Q^I_{Li}}\ \phi\ d^I_{Rj} + Y^u_{ij}\overline{Q^I_{Li}}\ \tilde{\phi}\ u^I_{Rj} \\ &= Y^d_{ij}\overline{d^I_{Li}}\ \frac{v}{\sqrt{2}}\ d^I_{Rj} + Y^u_{ij}\overline{u^I_{Li}}\ \frac{v}{\sqrt{2}}\ u^I_{Rj} + ... \\ &= M^d_{ij}\overline{d^I_{Li}}d^I_{Rj} + M^u_{ij}\overline{u^I_{Li}}u^I_{Rj} + , \end{aligned} \tag{13.5}$$

where we omitted the corresponding interaction terms of the fermion fields to the Higgs field, $\bar{q}qh(x)$ and the hermitian conjugate terms. Note that the d's and u's in equation (13.5) still each represent the three down-type and up-type quarks respectively, so the 'mixed'-terms are still there. To obtain mass eigenstates, i.e. states with proper mass terms, we should diagonalize the matrices M^d and M^u. We do this with unitary matrices V^d as follows:

$$M^d_{\rm diag} = V^d_L M^d V^{d\dagger}_R$$
$$M^u_{\rm diag} = V^u_L M^d V^{u\dagger}_R$$

Using the requirement that the matrices V are unitary ($V_L^{d\dagger} V_L^d = \mathbb{1}$) and leaving out again the hermitian conjugate terms the Lagrangian can now be expressed as follows:

$$
\begin{aligned}
-\mathcal{L}_{\text{Yukawa}}^{\text{quarks}} &= \overline{d_{Li}^I}\; M_{ij}^d\; d_{Rj}^I + \overline{u_{Li}^I}\; M_{ij}^u\; u_{Rj}^I + ... \\
&= \overline{d_{Li}^I}\; V_L^{d\dagger} V_L^d M_{ij}^d V_R^{d\dagger} V_R^d\; d_{Rj}^I + \overline{u_{Li}^I}\; V_L^{u\dagger} V_L^u M_{ij}^u V_R^{u\dagger} V_R^u\; u_{Rj}^I + ... \\
&= \overline{d_{Li}}\; (M_{ij}^d)_{\text{diag}}\; d_{Rj} + \overline{u_{Li}}\; (M_{ij}^u)_{\text{diag}}\; u_{Rj} + ...,
\end{aligned}
$$

where in the last line the matrices V have been absorbed in the quark states. Note that the up-type and down-type fields are now no longer the interaction states u^I and d^I, but are now 'simply' u and d. A bit more explicit, we now have the following quark mass eigenstates:

$$
\begin{aligned}
d_{Li} = (V_L^d)_{ij} d_{Lj}^I \quad & d_{Ri} = (V_R^d)_{ij} d_{Rj}^I \\
u_{Li} = (V_L^u)_{ij} u_{Lj}^I \quad & u_{Ri} = (V_R^u)_{ij} u_{Rj}^I,
\end{aligned}
$$

which allowed us to express the quark interaction eigenstates d^I, u^I as quark mass eigenstates d, u. It is now interesting to see how various parts of the Standard Model Langrangian change when you write them either in the mass or the interaction eigenstates.

Rewriting interaction terms using quark mass eigenstates

The interaction terms are obtained by imposing gauge invariance by replacing the partial derivative by the covariant derivate

$$
\mathcal{L}_{\text{kinetic}} = i\bar{\psi}(D^\mu \gamma_\mu)\psi, \tag{13.6}
$$

with the covariant derivative defined as $D^\mu = \partial^\mu + ig\frac{1}{2}\vec{\tau} \cdot \vec{W}_\mu$. The τ's are the Pauli matrices and W_i^μ and B^μ are the three weak interaction bosons and the single hypercharge boson, respectively. It is very natural to write the charged current interaction between the (left-handed) iso-spin doublet interaction eigenstates that are connected by W-bosons:

$$
\begin{aligned}
\mathcal{L}_{\text{kinetic, weak}}(Q_L) &= i\overline{Q_{Li}^I}\gamma_\mu\big(\partial^\mu + \frac{i}{2} g W_i^\mu \tau_i\big) Q_{Li}^I \\
&= i\overline{(u\;\; d)_{iL}^I}\gamma_\mu\big(\partial^\mu + \frac{i}{2} g W_i^\mu \tau_i\big) \begin{pmatrix} u \\ d \end{pmatrix}_{iL}^I \\
&= i\overline{u_{iL}^I}\gamma_\mu \partial^\mu u_{iL}^I + i\overline{d_{iL}^I}\gamma_\mu \partial^\mu d_{iL}^I - \frac{g}{\sqrt{2}}\overline{u_{iL}^I}\gamma_\mu W^{-\mu} d_{iL}^I - \frac{g}{\sqrt{2}}\overline{d_{iL}^I}\gamma_\mu W^{+\mu} u_{iL}^I + ...
\end{aligned}
$$

, where we used $W^{\pm} = \frac{1}{\sqrt{2}}(W_1 \mp i W_2)$, see Section 12.

If we now express the Lagrangian in terms of the quark mass eigenstates d, u instead of the weak interaction eigenstates d^I, u^I, the 'price' to pay is that the quark mixing between families (i.e. the off-diagonal elements) appear in the charged current interaction as each of the interaction fields is now replaced by a combination of the mass eigenstates:

$$\begin{aligned}\mathcal{L}_{\text{kinetic, cc}}(Q_L) &= \frac{g}{\sqrt{2}}\overline{u^I_{iL}}\gamma_\mu W^{-\mu}d^I_{iL} + \frac{g}{\sqrt{2}}\overline{d^I_{iL}}\gamma_\mu W^{+\mu}u^I_{iL} + ... \\ &= \frac{g}{\sqrt{2}}\overline{u_{iL}}(V^u_L V^{d\dagger}_L)_{ij}\gamma_\mu W^{-\mu}d_{iL} + \frac{g}{\sqrt{2}}\overline{d_{iL}}(V^d_L V^{u\dagger}_L)_{ij}\gamma_\mu W^{+\mu}u_{iL} + ...\end{aligned}$$

The CKM matrix

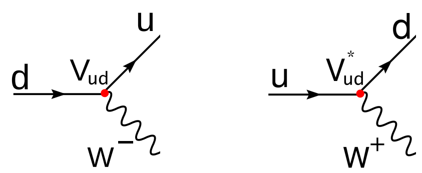
The combination of matrices $(V^d_L V^{u\dagger}_L)_{ij}$, a unitary 3×3 matrix is known under the shorthand notation V_{CKM}, the famous Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix. By convention, the interaction eigenstates and the mass eigenstates are chosen to be equal for the up-type quarks, whereas the down-type quarks are chosen to be rotated, going from the interaction basis to the mass basis:

$$\begin{aligned} u^I_i &= u_j \\ d^I_i &= V_{\text{CKM}} d_j \end{aligned}$$

or explicitly:

$$\begin{pmatrix} d^I \\ s^I \\ b^I \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{13.7}$$

From the definition of V_{CKM} it follows that the transition from a down-type quark to an up-type quark is described by V_{ud}, whereas the transition from an up type quark to a down-type quark is described by V^*_{ud}. A separate lecture describes in detail how V_{CKM} allows for CP-violation in the SM.



Note on lepton masses

We should note here that in principle a similar matrix exists that connects the lepton *flavour* and *mass* eigenstates. In this case, contrary to the quarks, the down-type interaction doublet-states (charged leptons) are chosen to be the same as the mass eigenstates. The rotation between mass and interaction eigenstates is in the neutrino sector. This matrix is known as the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix and has a completely different structure than the one for quarks. Just like for the CMS matrix, the origin of the observed patterns are completely unknown. A last thing to remember: neutrino interaction eigenstates are known as ν_e, ν_μ and ν_τ, whereas the physical particles, the mass eigenstates, are ν_1, ν_2 and ν_3.

13.3 Higgs boson decay

It is interesting to study details of the Higgs boson properties like its coupling to fermions and gauge bosons as that determines if and how the Higgs boson is produced in experiments and what the event topology will be. In Section 13.3.3 we list all couplings and as an example we'll compute the decay rate fractions of a Higgs boson into fermions as a function of it's unknown mass in Section 13.3.1.

13.3.1 Higgs boson decay to fermions

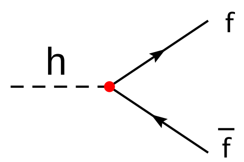
Now that we have derived the coupling of fermions and gauge bosons to the Higgs field, we can look in more detail at the decay of the Higgs boson.

The general expression for the two-body decay rate:

$$\frac{d\Gamma}{d\Omega} = \frac{|\mathcal{M}|^2}{32\pi^2 s}\,|p_f|\,S, \tag{13.8}$$

with $\mathcal{M}$ the matrix element, $|p_f|$ the momentum of the produced particles and $S = \frac{1}{n!}$ for n identical particles. In a two-body decay we have $\sqrt{s} = m_h$ and $|p_f| = \frac{1}{2}\beta\sqrt{s}$ (see exercise 2). Since the Higgs boson is a scalar particle, the Matrix element takes a simple form:

$$\begin{aligned} -i\mathcal{M} &= \bar{u}(p_1)\frac{im_f}{v}v(p_2) \\ i\mathcal{M}^\dagger &= \bar{v}(p_2)\frac{-im_f}{v}u(p_1) \end{aligned}$$



Since there are no polarizations for the scalar Higgs boson, computing the Matrix element squared is 'easy':

$$
\begin{aligned}
\mathcal{M}^2 &= \left(\frac{m_f}{v}\right)^2 \sum_{s_1,s_2} \bar{v}^{(s_2)}(p_2)u^{(s_1)}(p_1)\bar{u}^{(s_1)}(p_1)v^{(s_2)}(p2) \\
&\quad \text{write out the matrix multiplications (and omit momenta, temporarily):} \\
&= \left(\frac{m_f}{v}\right)^2 \sum_{m,n}\sum_{s_1,s_2} \bar{v}_m^{(s_2)}u_m^{(s_1)}\bar{u}_n^{(s_1)}v_n^{(s_2)} \\
&\quad \text{reorder and apply completeness relations (Eq. 5.77):} \\
&= \left(\frac{m_f}{v}\right)^2 \sum_{m,n} \left(\sum_{s_1} u_m^{(s_1)}\bar{u}_n^{(s_1)}\right)\left(\sum_{s_2} v_n^{(s_2)}\bar{v}_m^{(s_2)}\right) \\
&= \left(\frac{m_f}{v}\right)^2 \sum_{m,n} (\not{p}_1 + m_f)_{mn}(\not{p}_2 - m_f)_{nm} \\
&= \left(\frac{m_f}{v}\right)^2 \mathrm{Tr}\left((\not{p}_1 + m_f)(\not{p}_2 - m_f)\right) \\
&\quad \text{use the trace theorems (exercise 5.5):} \\
&= \left(\frac{m_f}{v}\right)^2 \left[4p_1 \cdot p_2 - 4m_f^2\right] \\
&\quad \text{use } m_h^2 = (p_1+p_2)^2 = p_1^2 + p_2^2 + 2p_1\cdot p_2 = 2m_f^2 + 2p_1 \cdot p_2\text{:} \\
&= \left(\frac{m_f}{v}\right)^2 \left[2m_h^2 - 8m_f^2\right] \\
&= \left(\frac{m_f}{v}\right)^2 2m_h^2\beta^2\,, \qquad \text{with } \beta = \sqrt{1 - \frac{4m_f^2}{m_h^2}}
\end{aligned}
$$

Including the number of colours (for quarks) we finally have:

$$\mathcal{M}^2 = \left(\frac{m_f}{v}\right)^2 2m_h^2\beta^2 N_c$$

Decay rate:

Starting from equation (13.8) and using $\mathcal{M}^2$ (above), $|p_f| = \frac{1}{2}\beta\sqrt{s}$, S=1 and $\sqrt{s} = m_h$ we get:

$$\frac{d\Gamma}{d\Omega} = \frac{|\mathcal{M}|^2}{32\pi^2 s}\,|p_f|\;S \quad = \quad \frac{N_c m_h}{32\pi^2}\left(\frac{m_f}{v}\right)^2\beta^3$$

Doing the angular integration $\int d\Omega = 4\pi$ we finally end up with:

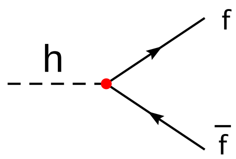
$$\Gamma(h \to f\bar{f}) = \frac{N_c}{8\pi v^2}m_f^2 m_h \beta_f^3.$$

13.3.2 Higgs boson decay to gauge bosons

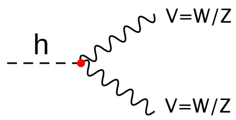
The decay ratio to gauge bosons is a bit more tricky, but is explained in great detail in Exercise 5.

13.3.3 Review Higgs boson couplings to fermions and gauge bosons

A summary of the Higgs boson couplings to fermions and gauge bosons.

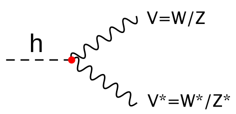


$$\Gamma(h \to f\bar{f}) = \frac{N_c}{8\pi v^2}\, m_f^2\, m_h\, \sqrt{1-x} \qquad \text{, with } \; x = \frac{4m_f^2}{m_h^2}$$



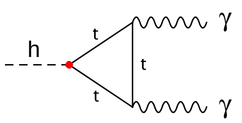
$$\Gamma(h \to VV) = \frac{g^2}{64\pi M_W^2}\, m_h^3\, \mathcal{S}_{VV}\, (1 - x + \tfrac{3}{4}x^2)\sqrt{1-x}$$
$$\text{, with } x = \frac{4M_V^2}{m_h^2} \text{ and } \mathcal{S}_{WW,ZZ} = 1, \tfrac{1}{2}.$$

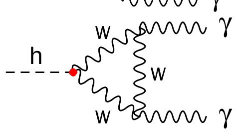
The decay of the Higgs boson to two off-shell gauge bosons is given by:



$$\Gamma(h \to VV^*) = \frac{3M_V^4}{32\pi^2 v^4}\, m_h\, \delta_V'\, \mathcal{R}(x) \qquad \text{, with}$$
$$\delta_W' = 1,\; \delta_Z' = \frac{7}{12} - \frac{10}{9}\sin^2\theta_W + \frac{40}{27}\sin^4\theta_W \qquad \text{, with}$$
$$\mathcal{R}(x) = \frac{3(1-8x+20x^2)}{\sqrt{4x-1}}\,\mathrm{acos}\left(\frac{3x-1}{2x^{3/2}}\right) - \frac{1-x}{2x}(2 - 13x + 47x^2) - \frac{3}{2}(1 - 6x + 4x^2)\ln(x)$$

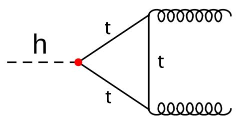
Since the coupling of the Higgs boson to gauge bosons is so much larger than that to fermions, the Higgs boson decays to off-shell gauge bosons even though $M_{V^*} + M_V < 2M_V$. The increase in coupling 'wins' from the Breit-Wigner suppression. For example: at m_h= 140 GeV, the $h \to WW^*$ is already larger than $h \to b\bar{b}$.





$$\Gamma(h \to \gamma\gamma) = \frac{\alpha^2}{256\pi^3 v^2}\, m_h^3\, \left|\frac{4}{3}\sum_f N_c^{(f)} e_f^2 - 7\right|^2$$

, where e_f is the fermion's electromagnetic charge.
Note: - WW contribution ≈ 5 times top contribution
- Some computation also gives $h \to \gamma Z$

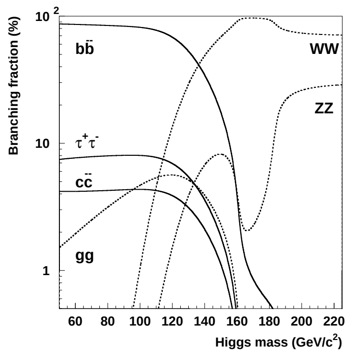


$$\Gamma(h \to \text{gluons}) = \frac{\alpha_s^2}{72\pi^3 v^2}\, m_h^3 \left[1 + \left(\frac{95}{4} - \frac{7N_f}{6}\right)\frac{\alpha_s}{\pi} + ...\right]^2$$

Note: - The QCD higher order terms are large.
- Reading the diagram from right to left you see the dominant production mechanism of the Higgs boson at the LHC.

13.3.4 Higgs branching fractions

Having computed the branching ratios to fermions and gauge bosons in Section 13.3.1 and Section 13.3.2 we can compute the relative branching fractions for the decay of a Higgs boson as a function of its mass. The distribution is shown here.



13.4 Theoretical bounds on the mass of the Higgs boson

Although the Higgs mass is not predicted within the minimal SM, there are theoretical upper and lower bounds on the mass of the Higgs boson if we assume there is no new physics between the electroweak scale and some higher scale called Λ. In this section we present a quick sketch of the various arguments and present the obtained limits.

As the Higgs boson mass is now known to quite some precision this section might feel strange and unnecessary to revisit. Since similar arguments are used to obtain theoretical limits on the mass of hypothetical particles that are predicted in models that go beyond the Standard Model it is good to understand the various elements that enter in such a discussion.

13.4.1 Unitarity

In the absence of a scalar field the amplitude for elastic scattering of longitudinally polarised massive gauge bosons (e.g. $\mathrm{W_L^+W_L^-} \to \mathrm{W_L^+W_L^-}$) diverges quadratically with the centre-of-mass energy when calculated in perturbation theory and at an energy of 1.2 TeV this process violates unitarity. In the Standard Model, the Higgs boson plays an important role in the cancellation of these high-energy divergences. Once diagrams involving a scalar particle (the Higgs boson) are introduced in the gauge boson scattering mentioned above, these divergences are no longer present and the theory remains unitary and renormalizable. Focusing on solving these divergences alone also yields most of the Higgs bosons properties. This cancellation only works however if the Higgs boson is not too heavy. By requiring that perturbation theory remains valid an upper limit on the Higgs mass can be extracted. With the requirement of unitarity and using all (coupled) gauge boson scattering processes it can be shown that:

$$m_h \ < \ \sqrt{\frac{4\pi\sqrt{2}}{3G_F}} \ \sim 700 \ \ \mathrm{GeV/c^2}.$$

It is important to note that this does not mean that the Higgs boson can not be heavier than 700 GeV/c^2. It only means that for heavier Higgs masses, perturbation theory is not valid and the theory is not renormalisable.

This number comes from an analysis that uses a partial wave decomposition for the matrix element $\mathcal{M}$, i.e.:

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi s}\mathcal{M}^2 \ \ , \text{with} \quad \mathcal{M} = 16\pi \sum_{l=0}^{l=\infty}(2l+1)\mathrm{P}_l(\cos\theta)a_l,$$

where P_l are Legendre polynomials and a_l are spin-l partial waves. Since $(W_L^+W_L^- + Z_L + Z_L + HH)^2$ is well behaved, it must respect unitarity, i.e. $|a_i| < 1$ or $|Re(a_i)| \leq 0.5$. As the largest amplitude is given by:

$$a_0^{\max} = -\frac{G_F m_h^2}{4\pi\sqrt{2}} \ \cdot \ \frac{3}{2}$$

This can then be transformed into an upper limit on m_h:

$$\begin{aligned} |a_0| < \frac{1}{2} \rightarrow m_h^2 \ &< \ \frac{8\pi\sqrt{2}}{6G_F} \ \left(= \ \frac{8}{3}\pi v^2 \text{ using } G_F = \frac{1}{\sqrt{2}v^2}\right) \\ m_h \ &< \ 700 \text{ GeV} \quad \text{using } v\text{= 246 GeV.} \end{aligned}$$

This limit is soft, i.e. it means that for Higgs boson masses > 700 GeV perturbation theory breaks down.

13.4.2 Triviality and Vacuum stability

In this section, the running of the Higgs self-coupling λ with the renormalisation scale μ is used to put both a theoretical upper and a lower limit on the mass of the Higgs boson as a function of the energy scale Λ.

Running Higgs coupling constant

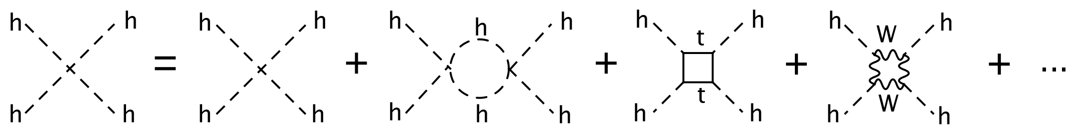
Similar to the gauge coupling constants, the coupling λ 'runs' with energy.

$$\frac{d\lambda}{dt} = \beta_\lambda \qquad \text{, where } t = \ln(Q^2).$$

Although these evolution functions (called β-functions) have been calculated for all SM couplings up to two loops, to focus on the physics, we sketch the arguments to obtain these mass limits by using only the one-loop results. At one-loop the quartic coupling runs with the renormalisation scale as:

$$\frac{d\lambda}{dt} \;\equiv\; \beta_\lambda \;=\; \frac{3}{4\pi^2}\left[\lambda^2 + \frac{1}{2}\lambda h_t^2 - \frac{1}{4}h_t^4 + \mathcal{B}(g,g')\right] \tag{13.9}$$

, where h_t is the top-Higgs Yukawa coupling as given in equation (13.1). The dominant terms in the expression are the terms involving the Higgs self-coupling λ and the top quark Yukawa coupling h_t. The contribution from the gauge bosons is small and explicitly given by $\mathcal{B}(g,g') = -\frac{1}{8}\lambda(3g^2 + g'^2) + \frac{1}{64}(3g^4 + 2g^2g'^2 + g'^4)$. The terms involving the mass of the Higgs boson, top quark and gauge bosons can be understood from looking in more detail at the effective coupling at higher energy scales, where contributions from higher order diagrams enter:



This expression allows to evaluate the value of $\lambda(\Lambda)$ relative to the coupling at a reference scale which is taken to be $\lambda(v)$.

If we study the β-function in 2 special regimes: $\lambda \gg g, g', h_t$ or $\lambda \ll g, g', h_t$, we'll see that we can set both a lower *and* an upper limit on the mass of the Higgs boson as a function of the energy-scale cut-off in our theory (Λ):

$$\underbrace{\text{Triviality}}_{\text{upper bound on } m_h} \qquad \text{and} \qquad \underbrace{\text{Vacuum stability}}_{\text{lower bound on } m_h}$$

$$m_h^{\max}(\Lambda) \qquad\qquad\qquad m_h^{\min}(\Lambda)$$

13.4.3 Triviality: $\lambda \gg \mathbf{g}, \mathbf{g}', \mathbf{h_t}$ heavy Higgs boson $\rightarrow$ upper limit on $\mathbf{m_h}$

For large values of λ (heavy Higgs boson since $m_h^2 = 2\lambda v^2$) and neglecting the effects from gauge interactions and the top quark, the evolution of λ is given by the dominant term in equation (13.9) that can be easily solved for $\lambda(\Lambda)$:

$$\frac{d\lambda}{dt} = \frac{3}{4\pi^2}\lambda^2 \qquad \Rightarrow \qquad \lambda(\Lambda) = \frac{\lambda(v)}{1 - \frac{3\lambda(v)}{4\pi^2}\ln\left(\frac{\Lambda^2}{v^2}\right)} \tag{13.10}$$

Note:

- We now have related λ at a scale v to λ at a higher scale Λ. We see that as Λ grows, $\lambda(\Lambda)$ grows. We should remember that $\lambda(v)$ is related to m_h: $m_h = \sqrt{-2\lambda v^2}$.
- There is a scale Λ at which $\lambda(\Lambda)$ is infinite. As Λ increases, $\lambda(\Lambda)$ increases until at $\Lambda = v\exp(2\pi^2/3\lambda(v))$ there is a singularity, known as the Landau pole.

$$\frac{3\lambda(v)}{4\pi^2}\ln\left(\frac{\Lambda^2}{v^2}\right) = 1 \rightarrow \text{At a scale } \Lambda = ve^{2\pi^2/3\lambda(v)} \quad \lambda(\Lambda) \text{ is infinite.}$$

If the SM is required to remain valid up to some cut-off scale Λ, i.e. if we require $\lambda(Q) < \infty$ for all $Q < \Lambda$ this puts a constraint (a maximum value) on the value of the Higgs self-coupling at the electroweak scale (v): $\lambda(v)^{\max}$ and therefore on the maximum Higgs mass since $m_h^{\max} = \sqrt{2\lambda(v)^{\max}v^2}$. Taking $\lambda(\Lambda) = \infty$ and 'evolving the coupling downwards', i.e. find $\lambda(v)$ for which $\lambda(\Lambda) = \infty$ (the Landau pole) we find:

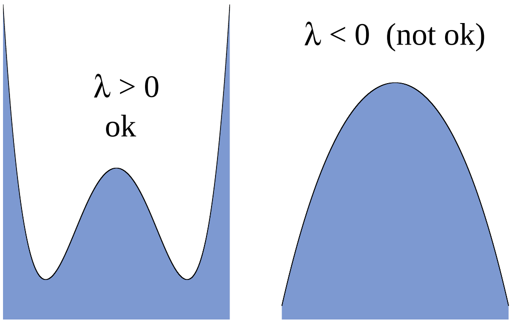
$$\lambda^{\max}(v) = \frac{4\pi^2}{3\ln\left(\frac{\Lambda^2}{v^2}\right)} \qquad \Rightarrow \quad m_h < \sqrt{\frac{8\pi^2 v^2}{3\ln\left(\frac{\Lambda^2}{v^2}\right)}} \tag{13.11}$$

For $\Lambda = 10^{16}$ GeV the upper limit on the Higgs mass is 160 GeV/c^2. This limit gets less restrictive as Λ decreases. The upper limit on the Higgs mass as a function of Λ from a computation that uses the two-loop β function and takes into account the contributions from top-quark and gauge couplings is shown in the Figure at the end of Section 13.4.4.

13.4.4 Vacuum stability $\lambda \ll \mathrm{g}, \mathrm{g}', \mathrm{h_t}$ light Higgs boson $\rightarrow$ lower limit on $\mathrm{m_h}$

For small λ (light Higgs boson since $m_h^2 = 2\lambda v^2$), a lower limit on the Higgs mass is found by the requirement that the minimum of the potential be lower than that of the unbroken theory and that the electroweak vacuum is stable. In equation (13.9) it is clear that for small λ the dominant contribution comes from the top quark through the Yukawa coupling ($-h_t^4$).

$$\begin{aligned}
\beta_\lambda &= \frac{1}{16\pi^2}\left[-3h_t^4 + \frac{3}{16}(2g^4 + (g^2 + g'^2)^2)\right] \\
&= \frac{3}{16\pi^2 v^4}\left[2M_W^4 + M_Z^4 - 4m_t^4\right] \\
&< 0.
\end{aligned}$$



Since this contribution is negative, there is a scale Λ for which $\lambda(\Lambda)$ becomes negative. If this happens, i.e. when $\lambda(\mu) < 0$ the potential is unbounded from below. As there is no minimum, no consistent theory can be constructed.

The requirement that λ remains positive up to a scale Λ, such that the Higgs vacuum is the global minimum below some cut-off scale, puts a lower limit on $\lambda(v)$ and therefore on the Higgs mass:

$$\frac{d\lambda}{dt} = \beta_\lambda \to \lambda(\Lambda) - \lambda(v) = \beta_\lambda \ln\left(\frac{\Lambda^2}{v^2}\right) \quad \text{and require } \lambda(\Lambda) > 0.$$

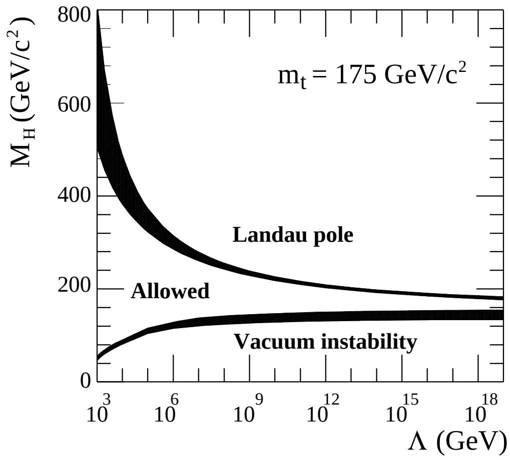
$$\begin{aligned}
\lambda(v) &> \beta_\lambda \ln\left(\frac{\Lambda^2}{v^2}\right) \text{ and } \lambda^{\min}(v) \to (m_h^{\min})^2 > 2\lambda^{\min}(v)v^2 \text{ , so} \\
m_h^2 &> 2v^2\beta_\lambda \ln\left(\frac{\Lambda^2}{v^2}\right) \\
(m_h^{\min})^2 &= \frac{3}{8\pi^2 v^2}\left[2M_W^4 + M_Z^4 - 4m_t^4\right] \\
&> -493 \ln\left(\frac{\Lambda^2}{v^2}\right)
\end{aligned}$$

Note: This result makes no sense, but is meant to describe the logic. If we go to the 2-loop beta-function we get a new limit: $m_h > 130 - 140$ GeV if $\Lambda = 10^{19}$ GeV. A detailed evaluation taking into account these considerations has been performed. The region of excluded Higgs masses as a function of the scale Λ from this analysis is also shown in the Figure at the end of Section 13.4.4 by the lower excluded region.

Summary of the theoretical bounds on the Higgs mass

In the Figure on the right the theoretically allowed range of Higgs masses is shown as a function of Λ.

For a small window of Higgs masses around 160 GeV/c^2 the Standard Model is valid up to the Planck scale ($\sim 10^{19}$ GeV). For other values of the Higgs mass the Standard Model is only an effective theory at low energy and new physics has to set in at some scale Λ.



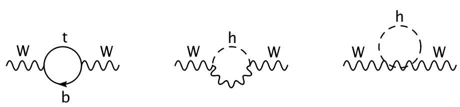
13.5 Experimental limits on the mass of the Higgs boson

13.5.1 Indirect measurements

The electroweak gauge sector of the SM is described by only three independent parameters: g, g' and v. The predictions for electroweak observables, are often presented using three (related) variables that are known to high precision: G_F, M_Z and α_{QED}. To obtain predictions to a precision better than the experimental uncertainties (often at the per mill level) higher order loop corrections have to be computed. These higher order radiative corrections contain, among others, contributions from the mass of the top quark and the Higgs boson. Via the precision measurements one is sensitive to these small contributions and thereby to the masses of these particles.

Radiative corrections

An illustration of the possibility to estimate the mass of a heavy particle entering loop corrections is the very good agreement between the estimate of the top quark mass using only indirect measurements and the direct observation.



$$
\begin{array}{lll}
\text{Estimate:} & m_t & = 177.2^{+2.9}_{-3.1}\ \mathrm{GeV/c^2} \\
\text{Measurement:} & m_t & = 173.2 \pm 0.9\ \mathrm{GeV/c^2}
\end{array}
$$

Sensitivity to Higgs boson mass through loop corrections

Apart from the mass of the W-boson, there are more measurements that provide sensitivity to the mass of the Higgs boson. A summary of the measurements of several SM measurements is given in the left plot of Figures 13.1.

While the corrections connected to the top quark behave as m_t^2, the sensitivity to the mass of the Higgs boson is unfortunately only logarithmic ($\sim \ln m_h$):

$$\begin{aligned}
\rho &= \frac{\mathrm{M_W^2}}{\mathrm{M_Z^2 cos\theta_W}}\left[1+\Delta_\rho^{\text{quarks}}+\Delta_\rho^{\text{higgs}}+...\right] \\
&= \frac{\mathrm{M_W^2}}{\mathrm{M_Z^2}\cos\theta_\mathrm{W}}\left[1+\frac{3}{16\pi^2}\left(\frac{m_t}{v}\right)^2+1-\frac{11\tan\theta_W}{96\pi^2}g^2\ln\left(\frac{m_h}{\mathrm{M_W}}\right)+...\right]
\end{aligned}$$

The results from a global fit to the electroweak data with only the Higgs mass as a free parameter is shown in the right plot of Figure 13.1. The plot shows the $\Delta\chi^2$ distribution as a function of m_h. The green band indicates the remaining theoretical uncertainty in the fit. The result of the fit suggested a rather light Higgs boson and it could be

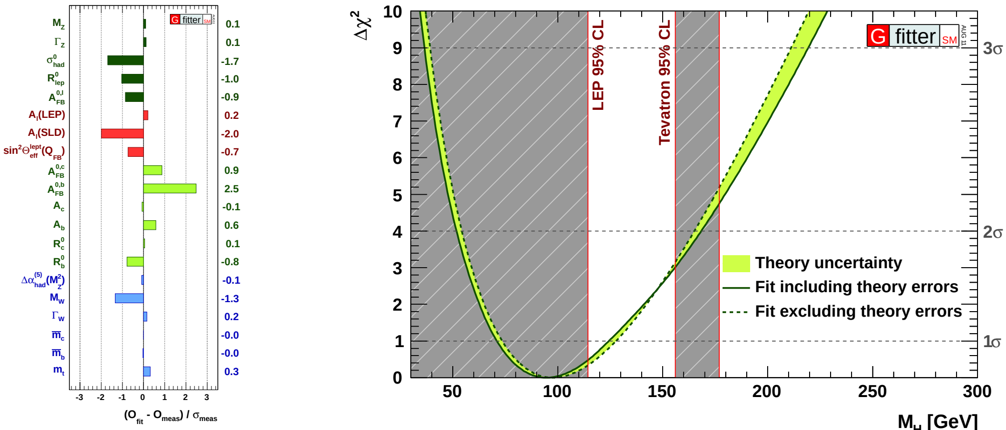


Figure 13.1: *Status of various SM measurements (left) and the $\Delta\chi^2$ distribution as a function of m_h from a global fit with only m_h as a free parameter (right). Before the discovery.*

summarised by the central value with its one standard deviation and the one-sided (95% CL) upper limit:

$$m_\mathrm{h} = 95^{+30}_{-24}\,{}^{+74}_{-43}\ \ \mathrm{GeV/c^2} \qquad \text{and} \qquad m_\mathrm{h} < 162\ \ \mathrm{GeV/c^2} \quad (\text{at } 95\%\ \text{CL}).$$

13.5.2 Direct measurements

In July 2012 the ATLAS and CMS experiments at the Large Hadron Collider at CERN announced the discovery of the Higgs boson. We will discuss the details of the search for the Higgs boson and its discovery in a separate lecture, but we since we cannot have

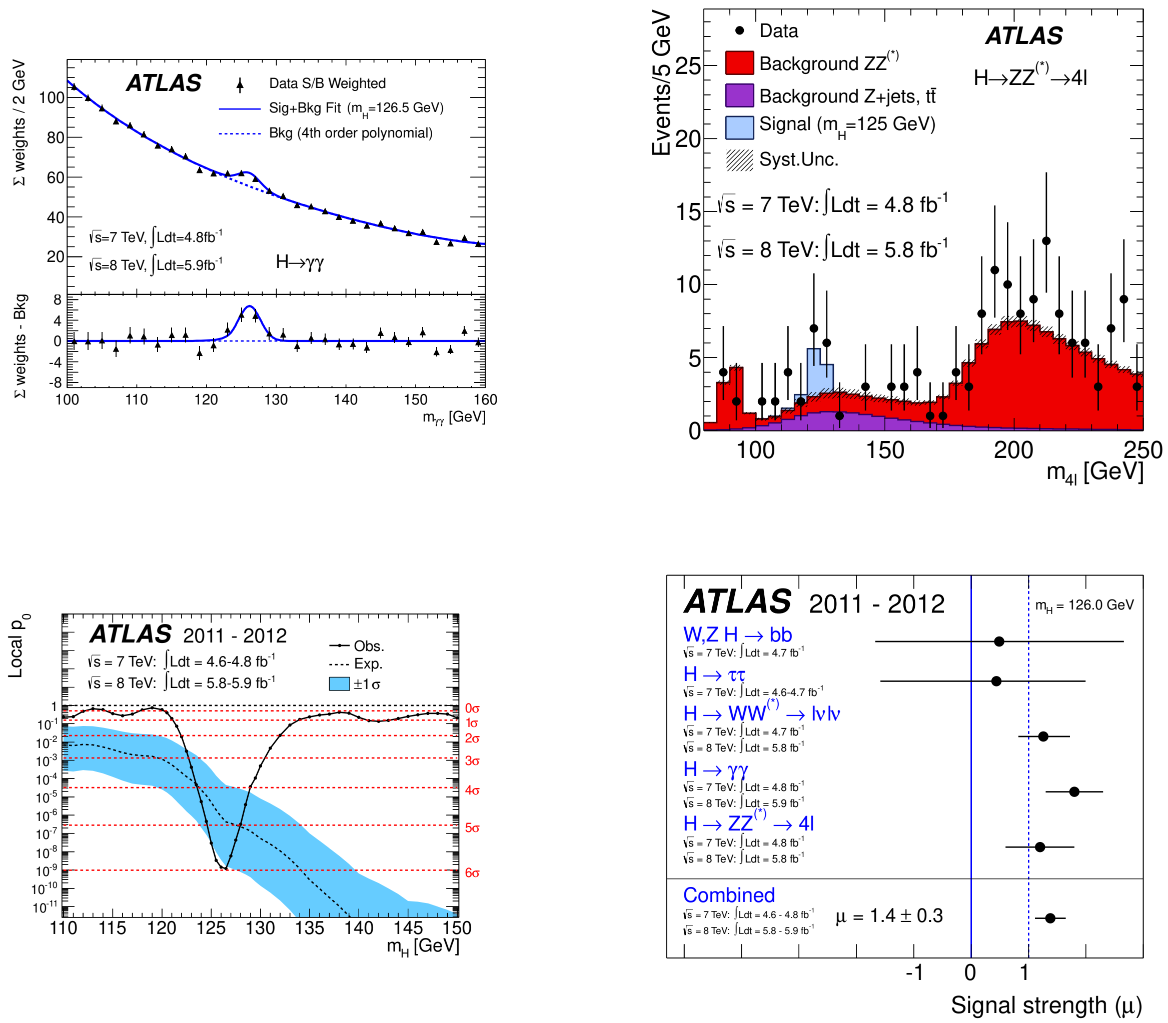


Figure 13.2: *Plots from the Higgs discovery paper from ATLAS. Two-photon invariant mass distribution (top left), the 4-lepton invariant mass distribution (top right), the p-value as a function of the Higgs mass (bottom left) and the measurement of the coupling strength of the Higgs boson to gauge bosons and fermions (bottom right).*

a lecture note on the Higgs boson without proof of its discovery I include here 4 plots that were in the discovery paper of the ATLAS experiment.

All results on the Higgs boson from the ATLAS and CMS experiments at the LHC can be found on these locations:

ATLAS: `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/HiggsPublicResults`
CMS: `http://cms.web.cern.ch/org/cms-higgs-results`

Exercises

Exercise 13.1
Show that $\bar{\mathrm{u}}\mathrm{u} = (\bar{\mathrm{u}}_\mathrm{L}\mathrm{u}_\mathrm{R} + \bar{\mathrm{u}}_\mathrm{R}\mathrm{u}_\mathrm{L})$

Exercise 13.2
Show that in a two body decay (a heavy particle M decaying into two particles with mass m) the momentum of the decay particles can be written as:

$$|\mathrm{p_f}| = \frac{\sqrt{s}}{2}\beta, \text{ with } \beta = \sqrt{1-x} \text{ and } x = \frac{4m^2}{M^2}$$

Exercise 13.3
Higgs decay into fermions for $\mathrm{m_h} = 100$ GeV.
Use $\mathrm{m_b} = 4.5$ GeV, $\mathrm{m}_\tau = 1.8$ GeV, $\mathrm{m_c} = 1.25$ GeV

(a) Compute $\Gamma(\mathrm{H}\rightarrow \mathrm{b\bar{b}})$.

(b) Compute $\Gamma(\mathrm{H}\rightarrow$ all) assuming only decay into the three heaviest fermions.

(c) What is the lifetime of the Higgs boson. Compare it to that of the Z boson.

Exercise 13.4 (H&M exercise 6.16)
The helicity states λ of a massive vector particle can be described by polarization vectors. Show that:

$$\sum_\lambda \epsilon_\mu^{(\lambda)*}\epsilon_\nu^{(\lambda)} = -g_{\mu\nu} + \frac{p_\mu p_\nu}{M^2}$$

Exercise 13.5 (Higgs decay to vector bosons)
Computing the Higgs boson decay into gauge bosons (W/Z = V), with boson momenta p, q and helicities λ, δ is a bit more tricky. Let's go through it step by step.

(a) Draw the Feynman diagram and use the vertex factor you computed last week to show that the matrix element squared is given by:

$$M^2 = \left(\frac{gM_V^2}{M_W}\right)^2 \sum_{\lambda,\delta} g_{\mu\nu}(\epsilon_\lambda^\mu)^*(\epsilon_\delta^\nu)^* g_{\alpha\beta}(\epsilon_\lambda^\alpha)(\epsilon_\delta^\beta),$$

where λ and δ are the helicity states of the Z bosons.

(b) Use your results of exercise 13.4 and work out to show that:

$$M^2 = \left(\frac{gM_V^2}{M_W}\right)^2 \left[2 + \frac{(p\cdot q)^2}{M_V^4}\right],$$

where p and q are are the momenta of the two Z bosons.

(c) Show that the matrix element can finally be written as:

$$M^2 = \frac{g^2}{4\mathrm{M_W^2}} m_h^4 (1 - x + \frac{3}{4}x^2), \text{ with x} = \frac{4\mathrm{M_V^2}}{m_h^2}$$

(d) Show that the Higgs decay into vector bosons can be written as:

$$\Gamma(h \to VV) = \frac{g^2 S_{VV}}{64\pi \mathrm{M_W^2}} m_h^3 (1 - x + \frac{3}{4}x^2)\sqrt{1-x},$$

with x $= \frac{4M_V^2}{m_h^2}$ and $S_{\mathrm{WW,ZZ}} = 1, \frac{1}{2}$.

(e) Compute Γ(h$\rightarrow$ WW) for $m_h = 200$ GeV.
What is the total width (only WW and ZZ decays)? And the lifetime ?

Lecture 14

Problems with the Higgs mechanism and Higgs searches

Although the Higgs mechanism cures many of the problems in the Standard Model, there are also several 'problems' associated to the Higgs mechanism. We will explore these problems in this section and very briefly discuss the properties of non-SM Higgs bosons.

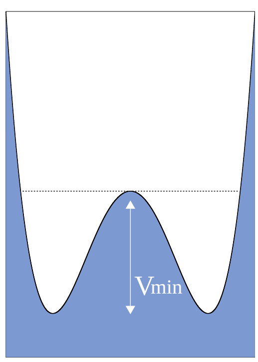
14.1 Problems with the Higgs boson

14.1.1 Problems with the Higgs boson: Higgs self-energy

Since the Higgs field occupies all of space, the non-zero vacuum expectation value of the Higgs field (v) will contribute to the vacuum energy, i.e. it will contribute to the cosmological constant in Einstein's equations: $\Lambda = \frac{8\pi G_N}{c^4}\rho_{\text{vac}}$.

Energy density Higgs field:
With $V(\phi^\dagger\phi) = \mu^2\phi^2 + \lambda\phi^4$, The 'depth' of the potential is:

$$\begin{aligned} V_{\text{min}} = V(v) &= \frac{1}{2}\mu^2 v^2 + \frac{1}{4}\lambda v^4 \quad \text{use } \mu^2 = -\lambda v^2 \\ &= -\frac{1}{4}\lambda v^4 \qquad\quad \text{use } m_h^2 = 2\lambda v^2 \\ &= -\frac{1}{8} m_h^2 v^2 \end{aligned}$$



Note that we cannot simply redefine V_{min} to be 0, or any arbitrary number since quantum corrections will always yield a value like the one (order of magnitude) given above. The Higgs mass is unknown, but since we have a lower limit on the (Standard Model) Higgs boson mass from direct searches at LEP ($m_h > 114.4$ GeV/c^2) we *can* compute the contribution of the Higgs field to ρ_{vac}.

$$\begin{aligned}\rho_{\mathrm{vac}}^{\mathrm{Higgs}} &= \frac{1}{8}m_h^2 v^2 \\ &> 1\cdot 10^8\ \mathrm{GeV}^4 \qquad \text{and since GeV} = \frac{1}{r} \\ &> 1\cdot 10^8\ \mathrm{GeV/r}^3 \qquad \text{(energy density)}\end{aligned}$$

Measured vacuum energy density:
An experiment to measure the energy density in vacuum and the energy density in matter has shown:

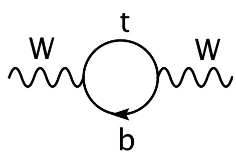
$$\Omega_m \approx 30\% \ \text{ and } \ \Omega_\Lambda \approx 70\% \sim 10^{-46}\mathrm{GeV}^4 \quad \rightarrow \quad \text{empty space is really quite empty.}$$

Problem:
- 10^{54} orders of magnitude mismatch.
- Why is the universe larger than a football ?

14.1.2 Problems with the Higgs boson: the hierarchy problem

In the electroweak theory of the SM, loop corrections are small. In the loops the integration is done over momenta up to a cut-off value Λ.

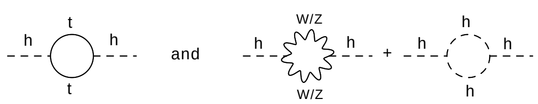
Success of radiative corrections:
When we discussed the sensitivity of the electroweak measurements to the mass of the Higgs boson through the radiative corrections, the example of the prediction of the top quark mass was mentioned:



Indirect estimate:	$m_t = 178^{+9.8}_{-4.2}$ GeV/c^2
Direct result:	$m_t = 172.4 \pm 1.2$ GeV/c^2

Failure of radiative corrections:
Also the Higgs propagator receives quantum corrections.

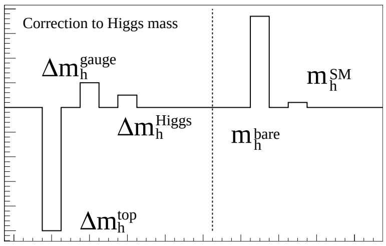
$$m_h = m_h^{\mathrm{bare}} + \Delta m_h^{\mathrm{ferm.}} + \Delta m_h^{\mathrm{gauge}} + \Delta m_h^{\mathrm{Higgs}} + \ldots$$



The corrections from the fermions (mainly from the top quark) are large. Expressed in terms of the loop-momentum cut-off Λ given by:

$$\left(\Delta m_h^2\right)^{\mathrm{top}} = -\frac{3}{8\pi^2}\lambda_t^2 \Lambda^2$$

The corrections from the top quark are not small at all, but huge and of order Λ. If Λ is chosen as 10^{16} (GUT) or 10^{19} (Planck), and taking the corrections into account (same order of magnitude), it is unnatural for m_h to be of order of $\mathrm{M_{EW}}(\approx v)$.



The hierarchy problem: why is $\mathrm{M_{EW}} \ll \mathrm{M_{PL}}$?

Most popular theoretical solution to the hierarchy problem is the concept of Supersymmetry, where for every fermion/boson there is a boson/fermion as partner. For example, the top and stop (supersymmetric bosonic partner of the top quark) contributions (almost) cancel. The quadratic divergences have disappeared and we are left with

$$\Delta m_h^2 \propto (m_f^2 - m_S^2) \ln\left(\frac{\Lambda}{m_S}\right).$$

14.2 Higgs bosons in models beyond the SM (SUSY)

When moving to a supersymmetric description of nature we can no longer use a single Higgs doublet, but will need to introduce at least two, because:

A) In the SM we used $\phi/\tilde{\phi}^c$ to give mass to down/up-type particles in $\mathrm{SU(2)_L}$ doublets. In susy models these two terms cannot appear together in the Lagrangian. We need an additional Higgs doublet to give mass to the up-type particles.

B) Anomalies disappear only if in a loop $\sum_f \mathrm{Y}_f = 0$. In SUSY there is an additional fermion in the model: the partner for the Higgs boson, the Higgsino. This will introduce an anomaly unless there is a second Higgsino with opposite hypercharge.

$$\phi_1 = \underbrace{\begin{pmatrix} \phi_1^+ \\ \phi_1^0 \end{pmatrix} \to \begin{pmatrix} 0 \\ v_1 \end{pmatrix}}_{\mathrm{Y}_{\phi_1}=+1} \quad \text{and} \quad \phi_2 = \underbrace{\begin{pmatrix} \phi_2^0 \\ \phi_2^- \end{pmatrix} \to \begin{pmatrix} v_2 \\ 0 \end{pmatrix}}_{\mathrm{Y}_{\phi_2}=-1}$$

Number of degrees of freedom in SUSY models:

SM: Add 4 degrees of freedom $\to$ 3 massive gauge bosons $\to$ 1 Higgs boson (h)
SUSY: Add 8 degrees of freedom $\to$ 3 massive gauge bosons $\to$ 5 Higgs boson (h, H, A, $\mathrm{H^+}$, $\mathrm{H^-}$)

parameters: $\tan(\beta) = \frac{v_2}{v_1}$ and $\mathrm{M_A}$.

Note: - Sometimes people choose α = mixing angle to give h,A, similar to W_3/B_μ-mixing to give Z-boson and photon.
- $\mathrm{M_W} = \frac{1}{2}\sqrt{\mathrm{v}_1^2 + \mathrm{v}_2^2}\,\mathrm{g} \to \mathrm{v}_1^2 + \mathrm{v}_2^2 = \mathrm{v}^2$ (246 GeV).

Differences SM and SUSY Higgses:

With the new parameters, all couplings to gauge bosons and fermions change:

$$\begin{aligned}
g_{hVV}^{\mathrm{SUSY}} &= g_{hVV}^{\mathrm{SM}} \ \sin(\beta-\alpha) \\
g_{hb\bar{b}}^{\mathrm{SUSY}} &= g_{hb\bar{b}}^{\mathrm{SM}} \ - \frac{\sin\alpha}{\cos\beta} \rightarrow \frac{\Gamma(h\rightarrow b\bar{b})^{\mathrm{SUSY}}}{\Gamma(h\rightarrow b\bar{b})^{\mathrm{SM}}} = \frac{\sin^2(\alpha)}{\cos^2(\beta)} \\
g_{ht\bar{t}}^{\mathrm{SUSY}} &= g_{ht\bar{t}}^{\mathrm{SM}} \ - \frac{\cos\alpha}{\sin\beta} \rightarrow \frac{\Gamma(h\rightarrow t\bar{t})^{\mathrm{SUSY}}}{\Gamma(h\rightarrow t\bar{t})^{\mathrm{SM}}} = \frac{\cos^2(\alpha)}{\sin^2(\beta)}
\end{aligned}$$

To determine if an observed Higgs sparticle is a SM or SUSY Higgs a detailed investigation of the branching fraction is required. Unfortunately, also SUSY does not give a prediction for the lightest Higgs boson mass:

$$\begin{aligned}
m_h^2 &< M_Z^2 + \delta^2 m_{top} + \delta^2 m_X + ... \\
&\leq 130 \text{ GeV}.
\end{aligned}$$

Exercises

Exercise 14.1 (b-tagging at LEP)
A Higgs boson of 100 GeV decays at LEP: given a lifetime of a B mesons of roughly 1.6 picoseconds, what distance does it travel in the detector before decaying ? What is the most likely decay distance ?

Exercise 14.2 (H $\rightarrow$ ZZ $\rightarrow$ 4 leptons at the LHC (lepton = e/μ))

(a) Why is there a 'dip' in te fraction of Higgs bosons that decays to 2 Z bosons (between 160 and 180 GeV)?

(b) How many events H $\rightarrow$ ZZ $\rightarrow e^+e^-\mu^+\mu^-$ muons are produced in 1 fb^{-1} of data for $\mathrm{m_h}$= 140, 160, 180 and 200 GeV ? The expected number of evets is the product of the luminosity and the cross-section: N = $\mathcal{L}\cdot\sigma$

On the LHC slides, one of the LHC experiments shows its expectation for an analysis aimed at trying to find the Higgs boson in the channel with 2 electrons and 2 muons. We concentrate on $\mathrm{m_h}$=140 GeV.

(c) What is the fraction of events in which all 4 leptons have been well reconstructed in the detector ? What is the single (high-energy) lepton detection efficiency ? Name reasons why not all leptons are detected.

We do a counting experiment using the two bins around the expected Higgs boson mass (we assume for the moment that the background is extremely well known and does not fluctuate). In a counting experiment a Poisson distribution describes the probabilities to observe x events when λ are expected:

$$\mathrm{P}(\mathrm{x}|\lambda) \;=\; \frac{\lambda^{\mathrm{x}}\,\mathrm{e}^{-\lambda}}{\mathrm{x}!}$$

(d) Does this experiment expect to be able to discover the $\mathrm{m_h}$=140 GeV hypothesis after 9.3 fb^{-1}.

(e) Imagine the data points was the actual measurement after 9.3 fb^{-1}. Can this experiment claim to have discovered the Higgs boson at $\mathrm{m_h}$=140 GeV?

Appendix A

Some properties of Dirac matrices α_i and β

This appendix lists some properties of the operators α_i and β in the Dirac Hamiltonian:

$$E\psi = i\frac{\partial}{\partial t}\psi = \left(-i\vec{\alpha}\cdot\vec{\nabla} + \beta m\right)\psi$$

1. α_i and β are hermitian.
   They have real eigenvalues because the operators E and $\vec{p}$ are hermitian. (Think of a plane wave equation: $\psi = Ne^{-ip_\mu x^\mu}$.)

2. $Tr(\alpha_i) = Tr(\beta) = 0$.
   Since $\alpha_i\beta = -\beta\alpha_i$, we have also: $\alpha_i\beta^2 = -\beta\alpha_i\beta$. Since $\beta^2 = 1$, this implies: $\alpha_i = -\beta\alpha_i\beta$ and therefore $Tr(\alpha_i) = -Tr(\beta\alpha_i\beta) = -Tr(\alpha_i\beta^2) = -Tr(\alpha_i)$, where we used that $Tr(A\cdot B) = Tr(B\cdot A)$.

3. The eigenvalues of α_i and β are ± 1.
   To find the eigenvalues bring α_i, β to diagonal form and since $(\alpha_i)^2 = 1$, the square of the diagonal elements are 1. Therefore the eigenvalues are ± 1. The same is true for β.

4. The dimension of α_i and β matrices is even.
   The $Tr(\alpha_i) = 0$. Make α_i diagonal with a unitary rotation: $U\alpha_i U^{-1}$. Then, using again $Tr(AB) = Tr(BA)$, we find: $Tr(U\alpha_i U^{-1}) = Tr(\alpha_i U^{-1}U) = Tr(\alpha_i)$. Since $U\alpha_i U^{-1}$ has only $+1$ and -1 on the diagonal (see 3.) we have: $Tr(U\alpha_i U^{-1}) = j(+1) + (n-j)(-1) = 0$. Therefore $j = n - j$ or $n = 2j$. In other words: n is even.

Appendix B

Propagators and vertex factors

In the lectures we expressed the invariant amplitude for particle-particle scattering in vertex factors for the 'particle-photon-particle' points in the Feynman diagram and one propagator for the photon-line connecting the two vertices. The expression of the amplitude in terms of propagators and vertex factors is part of what we refer to as the Feynman calculus. To derive the validity of the Feynman calculus approach required field theory. However, the attractiveness of this approach is that once you have established the recipe, you can derive the Feynman rules (the expressions for the propagators and vertices) directly from the Lagrangian density that specifies the dynamics of your favourite theory: if you insert a new type of particle or interaction in your Lagrangian, you do not really need field theory anymore to compute cross-sections.

To extract the propagators for a certain type of particle, one takes the part of the Lagrangian that specifies the 'free particle', derives its wave equation using the Euler-Lagrange equations and then takes the inverse of the wave equation in momentum space, multiplied by i. For instance, for a complex scalar (spin-0) field, the Lagrangian density is

$$\mathcal{L} \;=\; (\partial_\mu \phi^*)(\partial^\mu \phi) - m^2 \phi^* \phi \tag{B.1}$$

The corresponding Euler-Lagrange equations (obtained by requiring the action to be stationary, see Lecture 8) lead to the Klein-Gordon equation that we encountered in Lecture 1,

$$\left(\partial_\mu \partial^\mu + m^2\right)\; \phi(x) \;=\; 0 \tag{B.2}$$

In momentum space ($p^\mu \leftrightarrow i\hbar\partial^\mu$) this reads

$$\left(p_\mu p^\mu - m^2\right)\; \phi(p) \;=\; 0 \tag{B.3}$$

The resulting propagator for the field ϕ with momentum p is then

$$\text{propagator} = \frac{i}{p_\mu p^\mu - m^2} \tag{B.4}$$

You cannot directly use this result for a spin-1 particle like the photon, but you will notice that the result for the photon propagator ($-ig^{\mu\nu}/p^2$) resembles this result closely in the limit $m = 0$. (The propagator for spin-1 particles with non-zero mass is called the Proca propagator and we will see it in Lecture 9.)

Interactions in the Lagrangian density take the form of terms that have fields of different types. You cannot just write down arbitrary terms: the requirement that they be Lorentz invariant restricts their form. For instance, the interaction term for the photon with the spin-0 field that we have seen above, takes the form

$$\mathcal{L}_{\text{int}} = -g j^\mu A_\mu \tag{B.5}$$

where g is the coupling constant (e for the EM interaction) and j^μ is the current for the complex scalar field

$$j_\mu \equiv i\ \left[\phi^* \left(\partial_\mu \phi\right) - \left(\partial_\mu \phi\right)^* \phi\right]\ . \tag{B.6}$$

It is important that ϕ and ϕ^* are different fields here: in our discussion above one would be the incoming field and the other the outgoing field. The interaction term then has three fields and leads to a vertex with three lines. Labeling ϕ by A and ϕ^* by C, we now express the current in momentum space,

$$j^\mu_{AC} = i\left[\phi^*_C \left(p^\mu_A \phi_A\right) + \left(p^\mu_C \phi^*_C\right)\phi_A\right] \tag{B.7}$$

(Note how the plus sign appears: it is because $p^\mu \leftrightarrow i\hbar\partial^\mu$.) The rule to obtain the vertex factor is now to omit all the fields from the interaction term and write down what is left, multiplied by i,

$$\text{vertex factor} = ig(p^\mu_A + p^\mu_C) \tag{B.8}$$

This summarized the recipe to get the vertex factors and propagators. It directly leads to the 'Feynman rules' for a particular Lagrangian. Using the rules we can now compute the amplitude $\mathcal{M}$ for any Feynman diagram.

Appendix C

Summary of electroweak theory

Take the Lagrangian of free fermions (leptons and quarks)

$$\mathcal{L} \;=\; \sum_f \overline{\psi_f}(i\gamma^\mu\partial_\mu - m)\psi_f \tag{C.1}$$

Arrange the left-handed projections of the lepton and quark fields in doublets

$$\Psi_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \qquad \text{or} \qquad \Psi_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \tag{C.2}$$

Ignore their masses (or choose them equal within the doublet). Now consider that the Lagrangian remains invariant under

$U(1)_Y$:

$$\psi \to \psi' = e^{iY\beta(x)}\psi \tag{C.3}$$

$SU(2)_L$:

$$\Psi_L \to \Psi'_L \;=\; \exp\left[iY\vec{\alpha}(x)\cdot\vec{\tau}\right]\;\Psi_L \tag{C.4}$$

To keep the Lagrangian invariant compensating gauge fields must be introduced. These transform simultaneously with the Dirac spinors in the doublet:

$U(1)_Y$: hypercharge field a_μ

$$\partial_\mu \to D_\mu = \partial_\mu + ig'\frac{Y}{2}a_\mu \tag{C.5}$$

$SU(2)_L$: weak isospin fields b^1_μ, b^2_μ, b^3_μ (only couple to left-handed doublet):

$$\partial_\mu \to D_\mu = \partial_\mu + ig\vec{\tau}\cdot\vec{b}_\mu \tag{C.6}$$

Ignoring the kinetic and self-coupling terms of the gauge fields, the Lagrangian becomes

$$\mathcal{L} \;=\; \mathcal{L}^{\text{free}} - i\frac{g'}{2}J^\mu_Y a_\mu - ig\vec{J}^\mu_L\cdot\vec{b}_\mu \tag{C.7}$$

For the generators of $SU(2)$ we choose the Pauli spin matrices. The first field in a left-handed doublet has $T_3 = +1/2$ and the second field $T_3 = -1/2$. By construction the right-handed projections are singlets under $SU(3)_L$ and therefore have $T_3 = 0$.

The physical gauge fields (connecting the particle fields) become

"charged currents"

$$W_\mu^\pm = \frac{b_\mu^1 \mp i b_\mu^2}{\sqrt{2}} \tag{C.8}$$

"neutral currents"

$$\begin{aligned} Z_\mu &= -a_\mu \sin\theta_w + b_\mu^3 \cos\theta_w \\ A_\mu &= a_\mu \cos\theta_w + b_\mu^3 \sin\theta_w \end{aligned} \tag{C.9}$$

The Higgs mechanism takes care that 3 out of 4 gauge bosons get mass. For the field A_μ (the photon) to be massless, we need

$$\tan\theta_w = \frac{g'}{g} \tag{C.10}$$

The coupling of the massless field becomes proportional to a charge

$$Q = T_3 + \tfrac{1}{2}Y \tag{C.11}$$

Furthermore, the W and Z masses obey the relation

$$M_W = M_Z \cos\theta \tag{C.12}$$

The interaction Lagrangian for the doublet can now be written as

$$\begin{aligned} \mathcal{L}^{\text{int}} = &- \frac{g}{\sqrt{2}} \overline{\psi}_u \gamma^\mu \tfrac{1}{2}\left(1-\gamma^5\right)\psi_d\, W_\mu^+ \\ &- \frac{g}{\sqrt{2}} \overline{\psi}_d \gamma^\mu \tfrac{1}{2}\left(1-\gamma^5\right)\psi_u\, W_\mu^- \\ &- e\left[\sum_{f=u,d} Q_f \overline{\psi}_f \gamma^\mu \psi_f\right] A_\mu \\ &- g_z \left[\sum_{f=u,d} \overline{\psi}_f \gamma^\mu \tfrac{1}{2}\left(C_V^f - C_A^f \gamma^5\right)\psi_f\right] Z_\mu \end{aligned} \tag{C.13}$$

with

$$e = g\sin\theta_w \qquad g_z = g/\cos\theta_w \qquad \tan\theta_w = g'/g \tag{C.14}$$

and

$$C_V^f = T_3^f - 2Q^f \sin^2\theta_w \qquad C_A^f = T_3^f \tag{C.15}$$

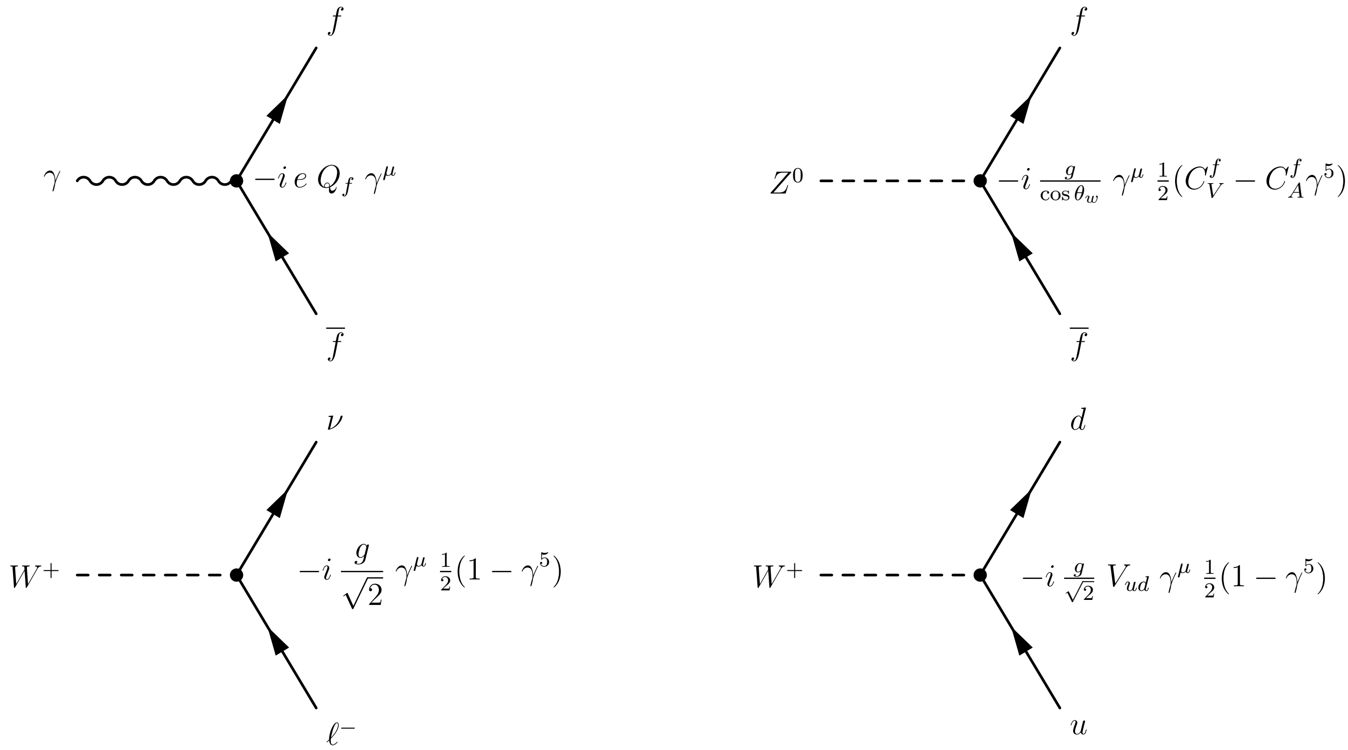
The relevant quantum numbers for our fields are

f	Q	T_L^3	T_R^3
u, c, t	$+\frac{2}{3}$	$+\frac{1}{2}$	0
d, s, b	$-\frac{1}{3}$	$-\frac{1}{2}$	0
ν_e, ν_μ, ν_τ	0	$+\frac{1}{2}$	–
e^-, μ^-, τ^-	-1	$-\frac{1}{2}$	0

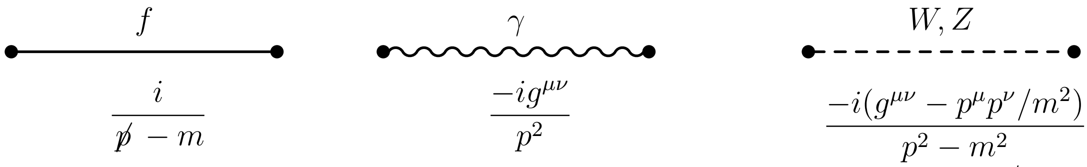
Till now we have ignored that the weak interaction mixes the quark fields. Inserting the CKM matrix we get for the charged currents,

$$\begin{aligned} \mathcal{L}^{\text{C.C.}} &= - i\frac{g}{\sqrt{2}}\, V_{ud}\, \overline{\psi}_u \gamma^\mu \tfrac{1}{2}\left(1-\gamma^5\right) \psi_d\, W_\mu^+ \\ &\quad - i\frac{g}{\sqrt{2}}\, V_{ud}^*\, \overline{\psi}_d \gamma^\mu \tfrac{1}{2}\left(1-\gamma^5\right) \psi_u\, W_\mu^- \end{aligned} \tag{C.16}$$

The Feynman rules for the vertex factors are then as follows



while those for the propagators are



The vector boson propagators are not unique and depend on the gauge. The Z/W boson propagator takes the form above in the so-called unitary gauge, while the expression for photon propagator is valid in the Feynman-'t Hooft gauge. In the Lorenz (or Landau) gauge the photon propagator becomes $-i(g^{\mu\nu} - p^\mu p^\nu/p^2)/p^2$.