Lecture notes to the 1-st year master course

Particle Physics 1

Nikhef - Autumn 2012

Lecture notes written by: Marcel Merk (marcel.merk@nikhef.nl)
Lecturer: Wouter Hulsbergen (wouter.hulsbergen@nikhef.nl)

Contents

Lecture 0

Introduction

The particle physics master course will be given in the autumn semester of 2011 and contains two parts: Particle Physics 1 (PP1) and Particle Physics 2 (PP2). The PP1 course consists of 12 lectures (Monday and Wednesday morning) and mainly follows the material as discussed in the books of Halzen and Martin and Griffiths.

These notes are my personal notes made in preparation of the lectures. They can be used by the students but should not be distributed. The original material is found in the books used to prepare the lectures (see below).

The contents of particle physics 1 is the following:

- Lecture 1: Concepts and History
- Lecture 2 - 5: Electrodynamics of spinless particles
- Lecture 6 - 8: Electrodynamics of spin 1/2 particles
- Lecture 9: The Weak interaction
- Lecture 10 - 12: Electroweak scattering: The Standard Model

Each lecture of 2×45 minutes is followed by a 1 hour problem solving session.

The particle physics 2 course contains the following topics:

- The Higgs Mechanism
- Quantum Chromodynamics

In addition the master offers in the next semester topical courses (not obligatory) on the particle physics subjects: CP Violation, Neutrino Physics and Physics Beyond the Standard Model

Examination

The examination consists of two parts: Homework (weight=1/3) and an Exam (weight=2/3).

Literature

The following literature is used in the preparation of this course (the comments reflect my personal opinion):

Halzen & Martin: "Quarks & Leptons: an Introductory Course in Modern Particle Physics ":
Although it is somewhat out of date (1984), I consider it to be the best book in the field for a master course. It is somewhat of a theoretical nature. It builds on the earlier work of Aitchison (see below). Most of the course follows this book.

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.

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.

Aitchison & Hey: "Gauge Theories in Particle Physics"
(1982) 2nd edition: An updated version of the book of Aitchison; a bit more theoretical.
(2003) 3rd edition (2 volumes): major rewrite in two volumes; very good but even more theoretical. It includes an introduction to quantum field theory.

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. Although the name Nikhef is kept, the acronym "Nationaal Instituut voor Kern en Hoge Energie Fysica" is no longer used. The name Nikhef is used to indicate simultaneously two overlapping organisations:

- Nikhef is a national research lab 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 Rijks Universiteit Groningen (the former FOM nuclear physics institute KVI) and there are contacts with the Universities of Twente, Leiden and Eindhoven.
For more information go to 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 started in 2009 and is currently producing data at increasing luminosity. The first results came out at the ICHEP 2010 conference in Paris, while the latest news of this summer on the search for the Higgs boson and "New Physics" have been discussed in the EPS conference in Grenoble and the lepton-photon conference in Mumbai. So far no convincing evidence for the Higgs particle or for New Physics have been observed.

In preparation of these LHC experiments Nikhef is/was also 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 huge datastreams from the (LHC-) experiments ("The Grid").

Nikhef program leaders/contact persons:

	Name	office	phone	email
Nikhef director	Frank Linde	H232	5001	z66@nikhef.nl
Theory departement:	Eric Laenen	H323	5127	t45@nikhef.nl
Atlas departement:	Stan Bentvelsen	H241	5150	stanb@nikhef.nl
B-physics departement:	Marcel Merk	N243	5107	marcel.merk@nikhef.nl
Alice departement:	Thomas Peitzmann	N325	5050	t.peitzmann@uu.nl
Antares experiment:	Maarten de Jong	H354	2121	mjg@nikhef.nl
Pierre Auger experiment:	Charles Timmermans	-	-	c.timmermans@hef.ru.nl
Virgo and ET experiment:	Jo van den Brand	N247	2015	jo@nikhef.nl
Xenon experiment:	Patrick Decowski	H349	2145	p.decowski@nikhef.nl
Detector R&D Departement:	Frank Linde	H232	5001	z66@nikhef.nl
Scientific Computing:	Jeff Templon	H158	2092	templon@nikhef.nl

History of Particle Physics

The book of Griffiths starts with a nice historical overview of particle physics in the previous century. Here's a summary:

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 existance 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.

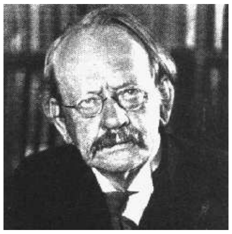
The Standard Model

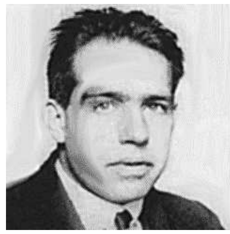
1978 *Glashow*, *Weinberg*, *Salam* formulate Standard Model for electroweak interactions

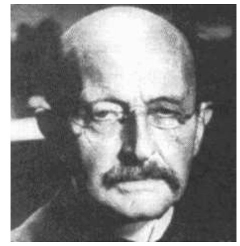
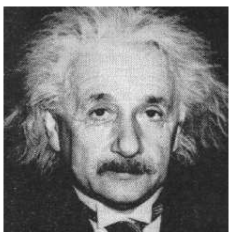
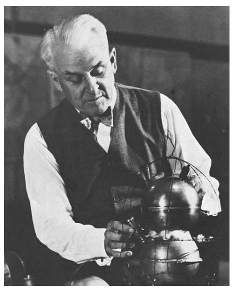
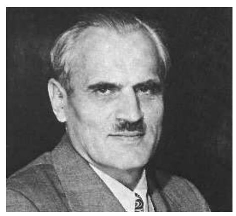
1983 W-boson has been found at CERN.

1984 Z-boson has been found at CERN.

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

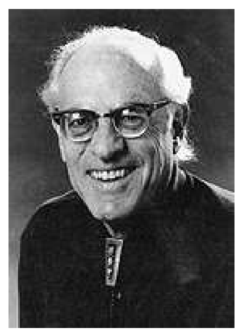




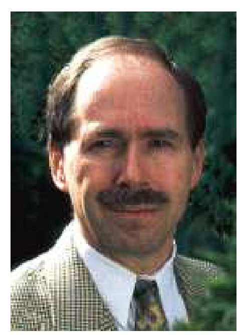




© Copyright California Institute of Technology. All rights reserved.
Commercial use or modification of this material is prohibited.











Natural Units

We will often make use of *natural units*. This means that we work in a system where the action is expressed in units of Planck's constant:

$$\hbar \approx 1.055 \times 10^{-34}\,\mathrm{Js}$$

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

$$c = 2.998 \times 10^{8}\mathrm{m/s}.$$

In other words we often use $\hbar = c = 1$.

This implies, however, that the results of calculations must be translated back to measureable quantities in the end. Conversion factors are the following:

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}	$\hbar/\,\mathrm{GeV}$
unit charge	$\mathrm{e} = \sqrt{4\pi\alpha}$	1	$\sqrt{\hbar c}$

Cross sections are expressed in *barn*, which is equal to $10^{-24}\mathrm{cm}^2$. Energy is expressed in GeV, or 10^9 eV, where 1 eV is the kinetic energy an electron obtains when it is accelerated over a voltage of 1V.

Exercise -1:
Derive the conversion factors for mass, length and time in the table above.

Exercise 0:
The Z-boson particle is a carrier of the weak force. It has a mass of 91.1 GeV. It can be produced experimentally 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) *Can you guess what the Feynman interaction diagram for this process is? Try to draw it.*

(b) *Assume that an electron and a positron are accelerated in opposite directions and collide head-on to produce a Z-boson in the lab frame. Calculate the beam energy required for the electron and the positron in order 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 in order to produce Z-bosons.*

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

Lecture 1

Particles and Forces

Introduction

After Chadwick had discovered the neutron in 1932, the elementary constituents of matter were the *proton* and the *neutron* inside the atomic nucleus, and the *electron* circling around it. With these constituents the atomic elements could be described as well as the chemistry with them. The answer to the question: "What is the world made of?" was indeed rather simple. The force responsible for interactions was the electromagnetic force, which was carried by the *photon*.

There were already some signs that there was more to it:

- Dirac had postulated in 1927 the existence of *anti-matter* as a consequence of his relativistic version of the Schrodinger equation in quantum mechanics. (We will come back to the Dirac theory later on.) The anti-matter partner of the electron, the positron, was actually discovered in 1932 by Anderson (see Fig. 1.1).

- 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. 1.2), 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.

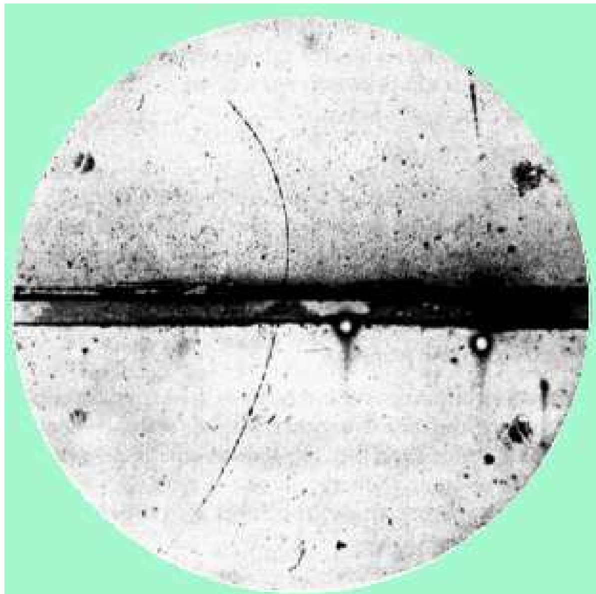


Figure 1.1: The discovery of the positron as reported by Anderson in 1932. Knowing the direction of the B field Anderson deduced that the trace was originating from an anti electron. *Question: how?*

1.1 The Yukawa Potential and the Pi meson

The year 1935 is a turning point in particle physics. Yukawa studied the strong interaction in atomic nuclei and proposed a new particle, a π-meson as the carrier of the nuclear force. His idea was that the nuclear force was carried by a **massive** particle (in contrast to the massless photon) such that the range of this force is limited to the nuclei.

The qualitative idea is that a virtual particle, the force carrier, can be created for a time $\Delta t < \hbar/2mc^2$. Electromagnetism is transmitted by the massless photon and has an infinite range, the strong force is transmitted by a massive meson and has a limited range, depending on the mass of the meson.

The Yukawa potential (also called the OPEP: One Pion Exchange Potential) is of the form:

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

where R is called the *range* of the force.

For comparison, the electrostatic potential of a point charge e as seen by a test charge e is given by:

$$V(r) = -e^2 \, \frac{1}{r}$$

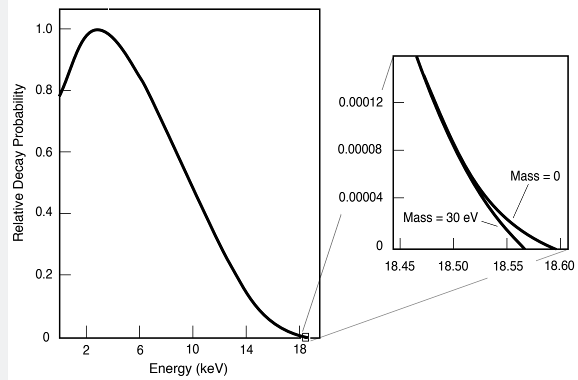


Figure 1. The Beta Decay Spectrum for Molecular Tritium
The plot on the left shows the probability that the emerging electron has a particular energy. If the electron were neutral, the spectrum would peak at higher energy and would be centered roughly on that peak. But because the electron is negatively 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 1.2: 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?*

The electrostatic potential is obtained in the limit that the range of the force is infinite: $R = \infty$. The constant g is referred to as the *coupling constant* of the interaction.

Exercise 1.

(a) *The wave equation for an electromagnetic wave in vaccuum 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 vaccuum. 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* $1.5 \times 10^{-15}\,\mathrm{m} = 1.5\,\mathrm{fm}$.

Yukawa called this particle a *meson* since it is expected to have an intermediate mass between 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 two things were not right:

(1) This particle did not interact strongly, which was very strange for a carrier of the strong force.

(2) Its mass was somewhat too low.

In fact this particle turned out to be the *muon*, the heavier brother of the electron.

In 1947 Powell (as well as Perkins) found the pion to be present in cosmic rays. They took their photographic emulsions to mountain tops to study the contents of cosmic rays (see Fig. 1.3). (In a cosmic ray event a cosmic proton scatters with high energy on an atmospheric nucleon and produces many secondary particles.) Pions produced in the atmosphere decay long before they reach sea level, which is why they were not observed before.

1.2 Strange Particles

After the pion had been identified as Yukawa's strong force carrier and the anti-electron was observed to confirm Dirac's theory, things seemed reasonably under control. The muon was a bit of a mystery. It lead to a famous quote of Isidore Rabi at the conference: "Who ordered *that*?"

But in December 1947 things went all wrong after Rochester and Butler published so-called V^0 events in cloud chamber photographs. What happened was that charged cosmic particles hit a lead target plate and as a result many different types of particles were produced. They were classified as:

baryons: particles whose decay product ultimately includes a proton.

mesons: particles whose decay product ultimately include only leptons or photons.

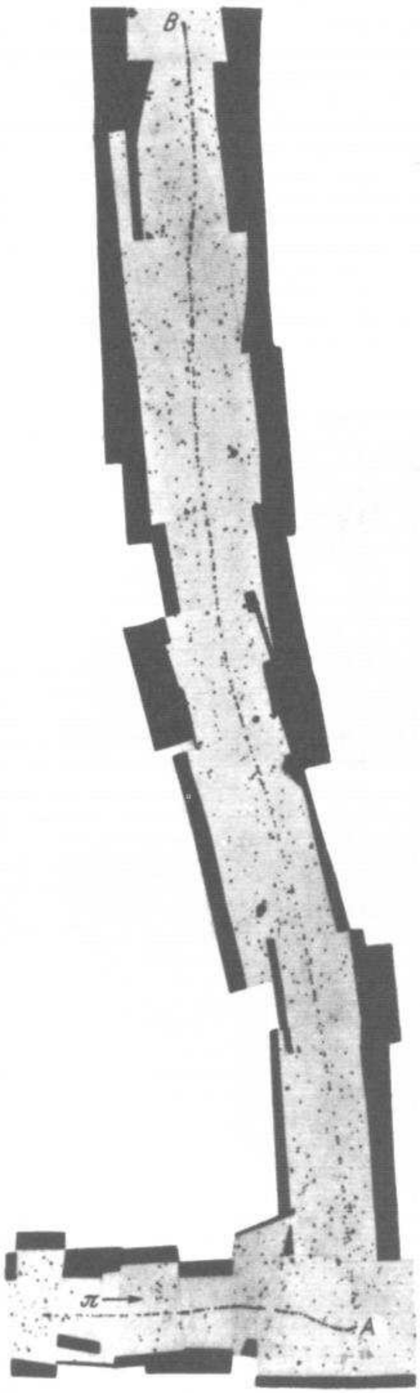


Figure 1.4 One of Powell's earliest pictures showing the track of a pion in a photographic emulsion exposed to cosmic rays at high altitude. The pion (entering from the left) decays into a muon and a neutrino (the latter is electrically neutral, and leaves no track). Reprinted by permission from C. F. 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 1.3: A pion entering from the left decays into a muon and an invisible neutrino.

Why were these events called *strange*? The mystery lies in the fact that certain (neutral) particles were produced (the "V^0's") with a large cross section ($\sim 10^{-27} cm^2$), while they decay according to a process with a small cross section ($\sim 10^{-40} cm^2$). The explanation to this riddle was given by Abraham Pais in 1952 and is called *associated production*. This means that strange particles are always *produced* in pairs by the strong interaction. It was suggested that strange particle carries a *strangeness* quantum number. In the strong interaction one then has the conservation rule $\Delta S = 0$, such that a particle with S=+1 (e.g. a K meson) is simultaneously produced with a particle with S=-1 (e.g. a Λ baryon). These particles then individually *decay* through the weak interaction, which does not conserve strangeness. An example of an associated production event is seen in Fig. 1.4.

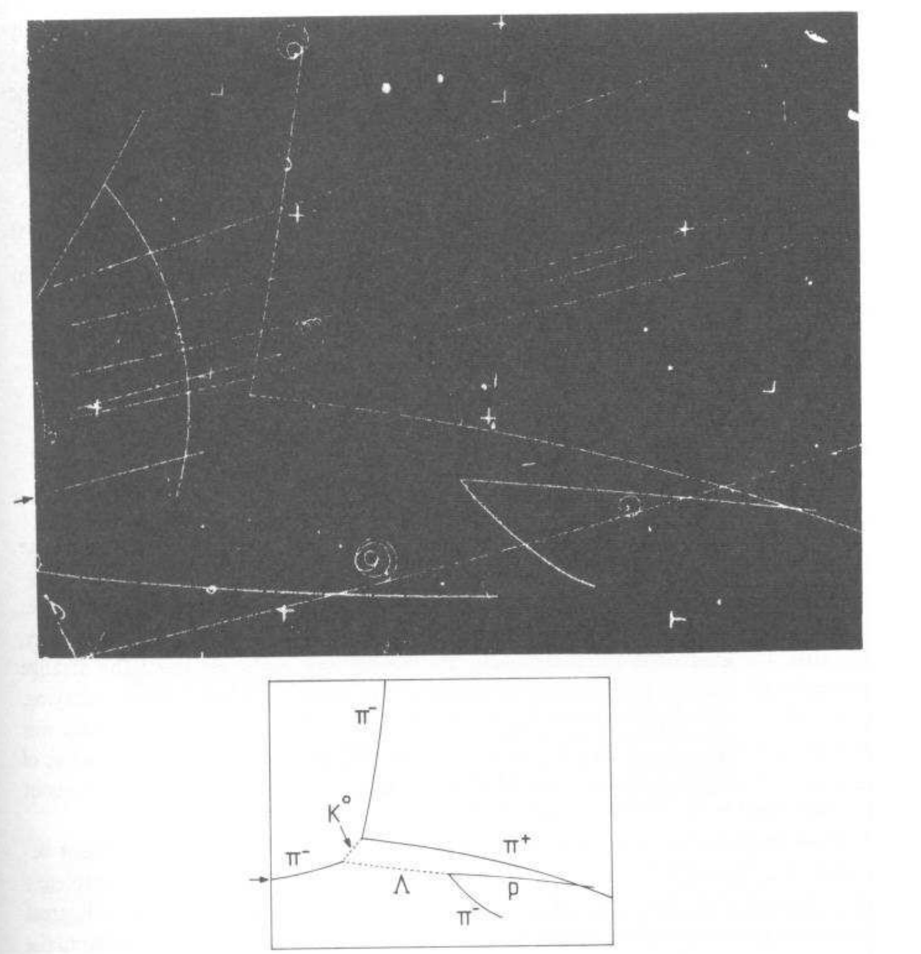


Fig. 2.7 A bubble chamber picture of the associated production reaction $\pi^- + p \rightarrow K^0 + \Lambda$. The incoming pion is indicated by the arrow, and the unseen neutrals are detected by their decays $K^0 \rightarrow \pi^+ + \pi^-$ and $\Lambda \rightarrow \pi^- + p$. This picture was taken in the 10 inch (25 cm) bubble chamber at the Lawrence Berkeley Radiation Laboratory. (Photograph courtesy of the Lawrence Berkeley Radiation Laboratory.)

Figure 1.4: A bubble chamber picture of associated production.

In the years 1950 - 1960 many elementary particles were discovered and one started to speak of the particle zoo. A quote: "The finder of a new particle used to be awarded the Nobel prize, but such a discovery now ought to be punished by a $10.000 fine."

1.3 The Eightfold Way

In the early 60's Murray Gell-Mann (at the same time also Yuvan Ne'eman) observed patterns of symmetry in the discovered mesons and baryons. He plotted the spin 1/2 baryons in a so-called octet (the "eightfold way" after the eighfold way to Nirvana in Buddhism). There is a similarity between Mendeleev's periodic table of elements and the supermultiplets of particles of Gell Mann. Both pointed out a deeper structure of matter. The eightfold way of the lightest baryons and mesons is displayed in Fig. 1.5 and Fig. 1.6. In these graphs the Strangeness quantum number is plotted vertically.

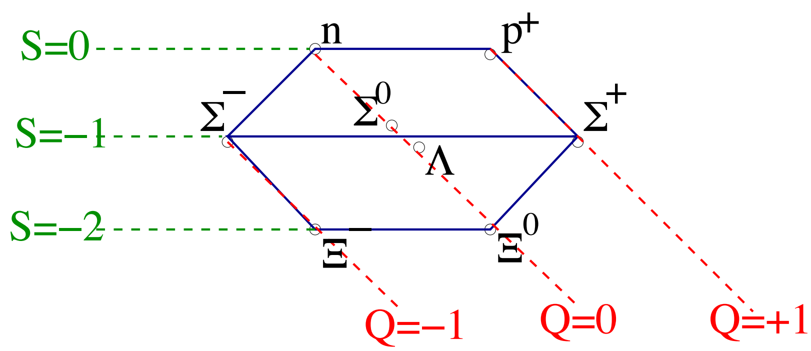


Figure 1.5: Octet of lightest baryons with spin=1/2.

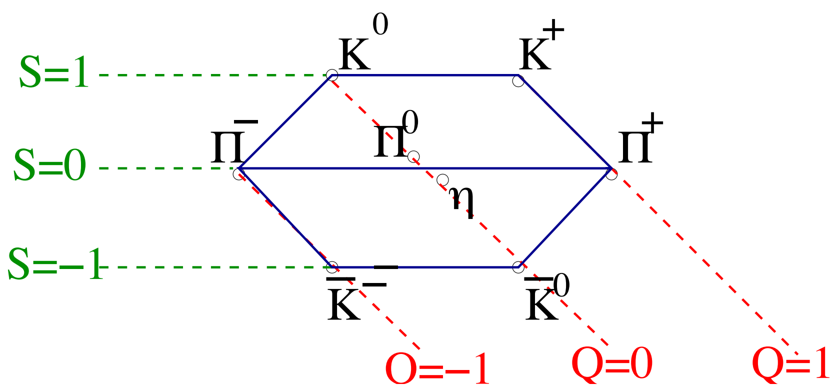


Figure 1.6: Octet with lightest mesons of spin=0

Also heavier hadrons could be given a place in *multiplets*. The baryons with spin=3/2 were seen to form a decuplet, see Fig. 1.7. The particle at the bottom (at S=-3) had not been observed. Not only was it found later on, but also its predicted mass was found to be correct! The discovery of the Ω^- particle is shown in Fig. 1.8.

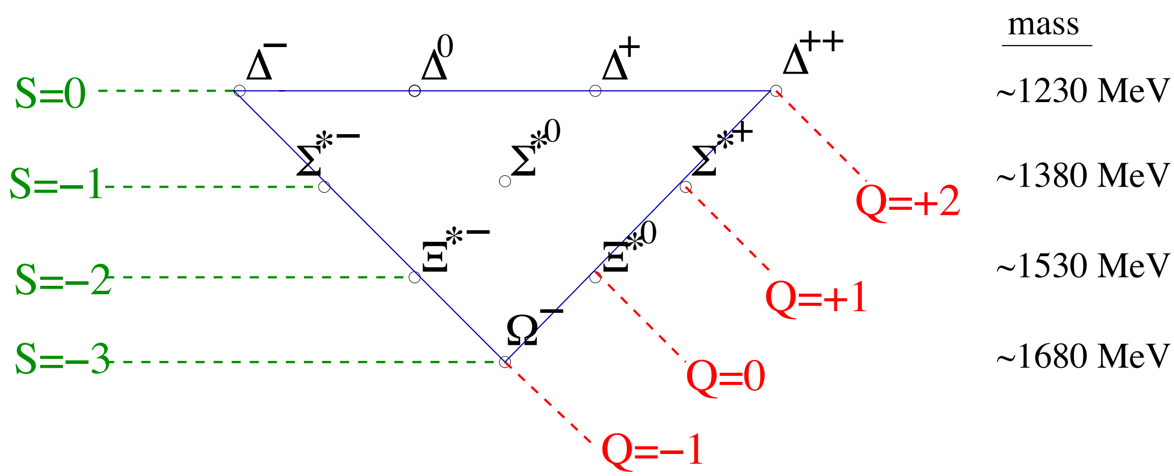


Figure 1.7: Decuplet of baryons with spin=3/2. The Ω^- was not yet observed when this model was introduced. Its mass was predicted.

1.4 The Quark Model

The observed structure of hadrons in multiplets hinted at an underlying structure. Gell-Mann and Zweig postulated indeed that hadrons consist of more fundamental partons: the quarks. Initially three quarks and their anti-particle were assumed to exist (see Fig. 1.9). A baryon consists of 3 quarks: (q, q, q), while a meson consists of a quark and an antiquark: $(q, \overline{q})$. Mesons can be their own anti-particle, baryons cannot.

Exercise 2:
Assign the quark contents of the baryon decuplet and the meson octet.

How does this explain that baryons and mesons appear in the form of octets, decuplets, nonets etc.? For example a baryon, consisting of 3 quarks with 3 flavours (u,d,s) could in principle lead to 3x3x3=27 combinations. The answer lies in the fact that the wave function of fermions is subject to a symmetry under exchange of fermions. The total wave function must be anti-symmetric with respect to the interchange of two fermions.

$$\psi\,(baryon) = \psi\,(space) \cdot \phi\,(spin) \cdot \chi\,(flavour) \cdot \zeta\,(color)$$

These symmetry aspects are reflected in group theory where one encounters expressions as: $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}$ and $\mathbf{3} \otimes \overline{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1}$.

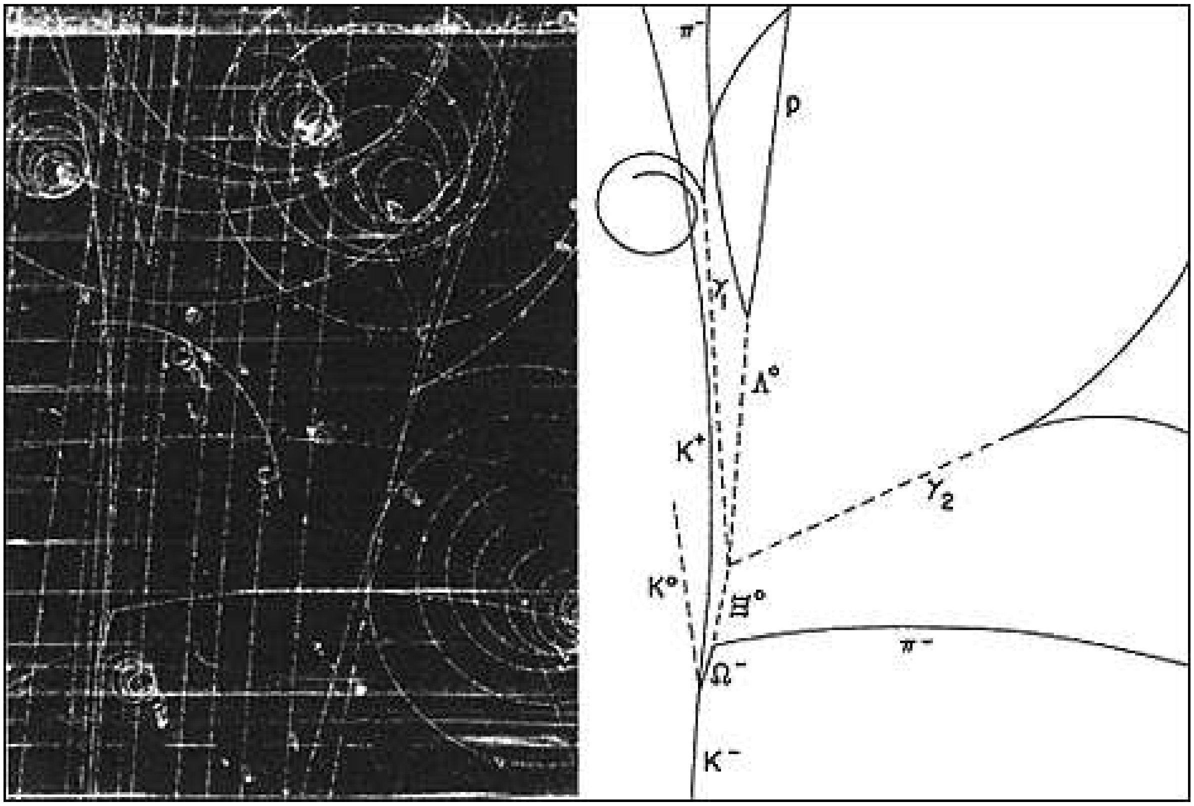


Figure 1.8: Discovery of the omega particle.

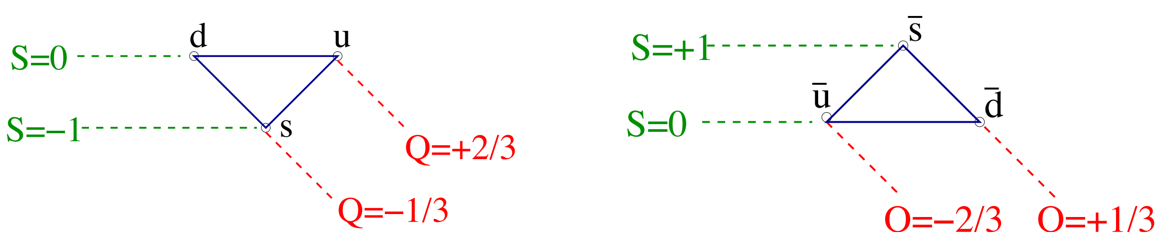


Figure 1.9: The fundamental quarks: u,d,s.

For more information on the static quark model read §2.10 and §2.11 in H&M, §5.5 and §5.6 in Griffiths, or chapter 5 in the book of Perkins.

1.4.1 Color

As indicated in the wave function above, a quark has another internal degree of freedom. In addition to electric charge a quark has a different charge, of which there are 3 types. This charge is referred to as the color quantum number, labelled as *r, g, b*. Evidence for the existence of color comes from the ratio of the cross section:

$$R \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} = N_C \sum_i Q_i^2$$

where the sum runs over the quark types that can be produced at the available energy. The plot in Fig. 1.10 shows this ratio, from which the result $N_C = 3$ is obtained.

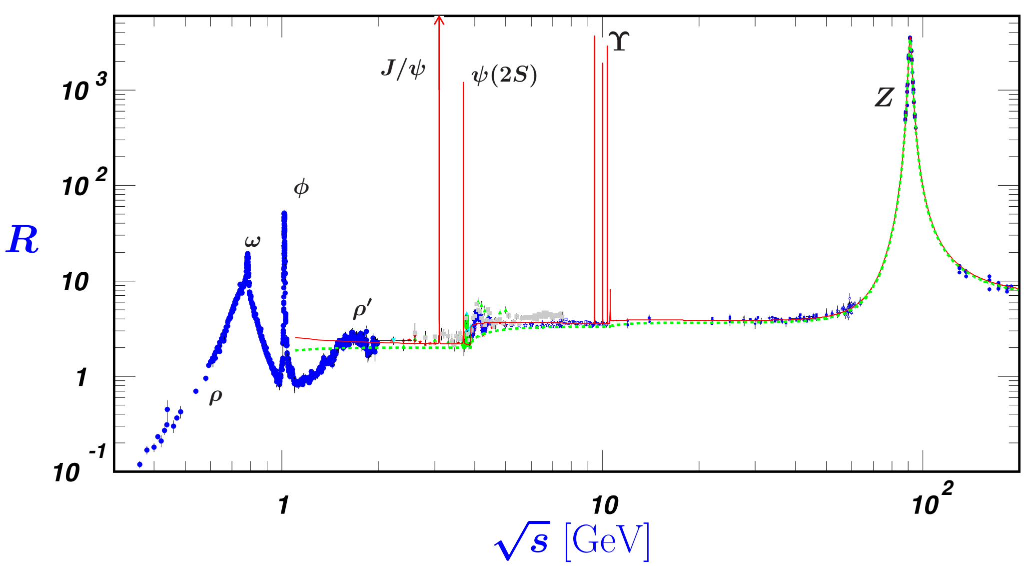


Figure 1.10: The R ratio.

<u>**Exercise 3: The Quark Model**</u>

(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 Color part of the wave function. Write down the color part of the wave-function taking into account that the particle is color 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?*

1.5 The Standard Model

The fundamental constituents of matter and the force carriers in the Standard Model can be represented as follows:

The fundamental particles:

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

The forces, their mediating bosons and their relative strength:

Force	Boson	Relative strength
Strong	g (8 gluons)	$\alpha_s \sim \mathcal{O}(1)$
Electromagnetic	γ (photon)	$\alpha \sim \mathcal{O}(10^{-2})$
Weak	Z^0,$W^{\pm}$ (weak bosons)	$\alpha_W \sim \mathcal{O}(10^{-6})$

Some definitions:

hadron (greek: strong)	particle that feels the strong interaction
lepton (greek: light, weak)	particle that feels only weak interaction
baryon (greek: heavy)	particle consisting of three quarks
meson (greek: middle)	particle consisting of a quark and an anti-quark
pentaquark	a hypothetical particle consisting of 4 quarks and an anti-quark
fermion	half-integer spin particle
boson	integer spin particle
gauge-boson	force carrier as predicted from local gauge invariance

In the Standard Model forces originate from a mechanism called local gauge invariance, which wil be discussed later on in the course. The strong force (or color force) is mediated by gluons, the weak force by intermediate vector bosons, and the electromagnetic force by photons. The fundamental diagrams are represented below.

a: b: c:

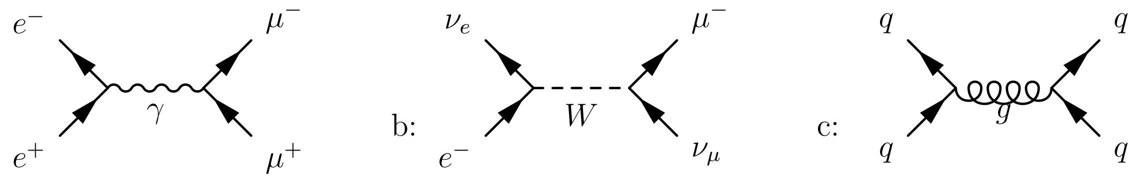


Figure 1.11: Feynman diagrams of fundamental lowest order perturbation theory processes in a: electromagnetic, b: weak and c: strong interaction.

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 charge and, therefore, does not interact with itself. The gluons, however, carry color and do interact amongst each other. Also, the weak vector bosons carry weak isospin and undergo a self coupling.

The strength of an interaction is determined by the coupling constant as well as the mass of the vector boson. Contrary to its name the couplings are not constant, but vary as a function of energy. At a momentum transfer of 10^{15} GeV the couplings of electromagnetic, weak and strong interaction all have the same value. In the quest of unification it is often assumed that the three forces unify to a grand unification force at this energy.

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 even diverges at momenta less than a few 100 MeV (the perturbative QCD description breaks down). This leads to confinement: the existence of colored objects (i.e. objects with net strong charge) is forbidden.

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. A new particle consistent with the Higgs particle was discovered in summer 2012 by the ATLAS and CMS collaborations.

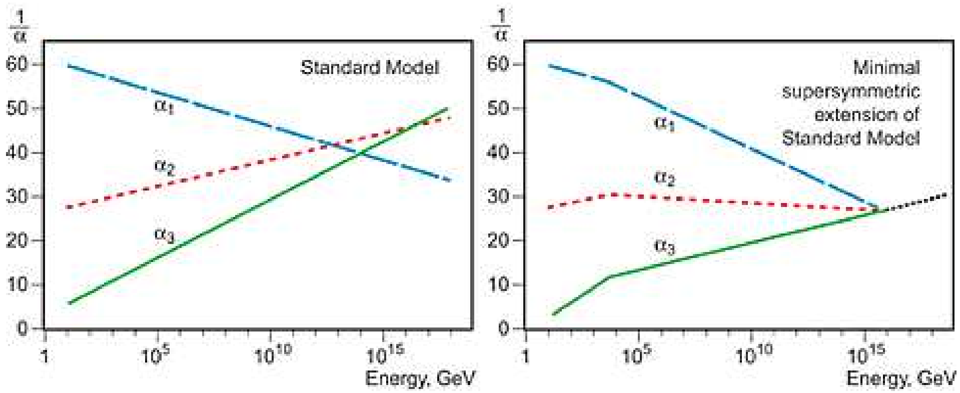


Figure 1.12: Running of the coupling constants and possible unification point. On the left: Standard Model. On the right: Supersymmetric Standard Model.

Open Questions

- Does the Higgs in fact exist? (Is the discovered particle the Higgs?)
- Why are the masses of the particles what they are?
- Why are there 3 generations of fermions?
- Are quarks and leptons truly fundamental?
- Why is the charge of the electron *exactly* opposite to that of the proton. Or: why is the total charge of leptons and quarks exactly equal to 0?
- Is a neutrino its own anti-particle?
- Can all forces be described in a single theory (unification)?
- Why is there no anti matter in the universe?
- What is the source of dark matter?
- What is the source of dark energy?

Lecture 2

Wave Equations and Anti-Particles

Introduction

In the course we develop a quantum mechanical framework to describe electromagnetic scattering, in short Quantum Electrodynamics (QED). The way we build it up is that we first derive a framework for non-relativistic scattering of spinless particles, which we then extend to the relativistic case. Also we will start with the wave equations for particles without spin, and address the spin 1/2 particles later on in the lectures ("the Dirac equation').

What is a spinless particle? There are two ways that you can think of it: either as charged mesons (e.g. pions or kaons) for which the strong interaction has been "switched off" or for electrons or muons for which the fact that they are spin-1/2 particles is ignored. In short: it not a very realistic case.

2.1 Non Relativistic Wave Equations

If we start with the non relativistic relation between kinetic energy and momentum

$$E = \frac{\vec{p}^2}{2m}$$

and make the quantum mechanical substitution:

$$E \to i\frac{\partial}{\partial t} \qquad \text{and} \qquad \vec{p} \to -i\vec{\nabla}$$

then we end up with Schrödinger's equation:

$$\boxed{i\frac{\partial}{\partial t}\,\psi = \frac{-1}{2m}\,\nabla^2\psi}$$

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

$$P(\text{particle in volume } V, t) = \int_V \rho(\vec{x}, t) d^3x$$

where the density is

$$\rho(\vec{x}, t) = |\psi(\vec{x}, t)|^2$$

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

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{j} = 0$$

where j is the *current density*. 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

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

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

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

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

The solution to the Schrödinger equation for a free particle is given by

$$\psi = N \, e^{i(\vec{p}\vec{x} - Et)}$$

where the energy and momentum are related as $E = p^2/2m$. (We leave the definition of the normalization constant N for the next lecture.) We usually call this the *plane wave* solution. For the density of the plane wave we obtain

$$\begin{aligned} \rho &= \psi^* \, \psi = |N|^2 \\ \vec{j} &= \frac{i}{2m} \left(\psi \, \vec{\nabla} \psi^* - \psi^* \, \vec{\nabla} \psi \right) = \frac{|N|^2}{m} \, \vec{p} \end{aligned}$$

Note that, as expected, the current is equal to the current density times the velocity $v = p/m$.

Exercise 4.

(a) *Show that the plane wave* $\psi = N e^{-i(Et - \vec{p}\cdot\vec{x})}$ *is a solution to the Schrödinger equation.*

(b) *Derive the expressions for the conserved current* $\rho = \psi^* \psi$ *and* $\vec{j} = \frac{i}{2m}(\psi \nabla \psi^* - \psi^* \nabla \psi)$

(c) *Substitute the plane solution in order to derive the expressions for the conserved current of the plane wave*

Intermezzo: 4-vector notation

We define the coordinate 4-vector x^μ as

$$x^\mu = (x^0, x^1, x^2, x^3)$$

where the first component $x^0 = ct$ is the time coordinate and the latter three components are the spatial coordinates $(x^1, x^2, x^3) = \vec{x}$. Under a Lorentz transformation along the x^1 axis with velocity $\beta = v/c$, x^μ transforms as

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

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

A general 'contravariant 4-vector' is defined to be any set of four quantities $A^\mu = (A^0, A^1, A^2, A^3) = (A^0, \vec{A})$ which transforms under Lorentz transformations exactly as the corresponding components of the coordinate 4-vector x^μ. Note that it is the *transformation property* that defines what a 'contravariant' vector is.

Lorentz transformations leave the quantity

$$|A|^2 = A^{0^2} - |\vec{A}|^2$$

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

$$A \cdot A \equiv |A|^2 = \sum_\mu A^\mu A_\mu.$$

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

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

we have $A_\mu = g_{\mu\nu} A^\nu$ and $A^\mu = g^{\mu\nu} A_\nu$. A scalar product of two 4-vectors A^μ and B^μ can then be written as

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

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

Exercise Extra 1 (from A&H, chapter 3. see also Griffiths, exercise 7.1).

Write down the inverse of the Lorentz transformation along the x^1 axis above (i.e. express (x^0, x^1) in $(x^{0'}, x^{1'})$) and use the 'chain rule' of partial differentiation to show that, under the Lorentz transformation, the quantities $(\partial/\partial x^0, -\partial/\partial x^1)$ transform in the same way as (x^0, x^1).

Consequently, if the contravariant and covariant 4-vectors for the coordinates are defined as above, then the 4-vectors of their derivatives are given by

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

Note that the position of the minus sign is opposite to that of the usual 4-vector.

2.2 Relativistic Wave Equations

If we start with the relativistic equation

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

and again make the substitution

$$E \to i\frac{\partial}{\partial t} \qquad \text{and} \qquad \vec{p} \to -i\vec{\nabla},$$

then we end up with the Klein Gordon equation for a wavefunction ϕ,

$$\boxed{-\frac{\partial^2}{\partial t^2}\phi = -\nabla^2\phi + m^2\,\phi}$$

This can be efficiently written in 4-vector notation as

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

where

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

is the so-called d'Alembert operator. The solution to the Klein-Gordon equation is the plane wave

$$\phi(x) \;=\; N\,e^{-ip_\mu x^\mu}$$

with $p^\mu = (E, \vec{p})$. For a given momentum the energy must satisfy the constraint $E^2 = p^2 + m^2$.

In analogy to the procedure applied above for the non-relativistic free particle, we now derive a continuity equation. We multiply the Klein Gorden equation for ϕ from the left by $-i\phi^*$, then add to the complext conjugate equation:

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

where we can recognize again the continuity equation. In 4-vector notation it becomes:

$$\begin{aligned}
j^\mu &= \left(\rho, \vec{j}\right) = i\left[\phi^*\left(\partial^\mu\phi\right) - \left(\partial^\mu\phi^*\right)\phi\right] \\
\partial_\mu j^\mu &= 0
\end{aligned}$$

You may wonder why we introduced the factor i: this is in order to make the density real. Substituting the plane wave solution gives

$$\begin{aligned}
\rho &= 2\,|N|^2\,E \\
\vec{j} &= 2\,|N|^2\,\vec{p} \\
\text{or}:\ \to\ j^\mu &= 2\,|N|^2\,p^\mu
\end{aligned}$$

Note that, in contrast to the solution for the classical Schrödinger equation, the density of the Klein-Gordon wave is proportional to the energy. This is a direct consequence of the Klein-Gordon equation being 2nd order in the time derivative.

The constraint $E^2 = p^2 + m^2$ 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.

Exercise 5.

Repeat exercise 4 but now for the Klein Gordon equation.

Exercise 6.

The relativitic energy-momentum relation can be written as:

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

This is linear in $E = \partial/\partial t$, but we do not know what to do with the square root of the momentum operator. However, for small $\vec{p}$ we can expand the expression in powers of $\vec{p}/m$. Write down the expansion terms up to and including $\vec{p}^2/m^2$. Write down the resulting wave equation. Argue that it is equivalent to the Schrödinger equation.

2.3 Interpretation of negative energy solutions

2.3.1 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. 2.1.

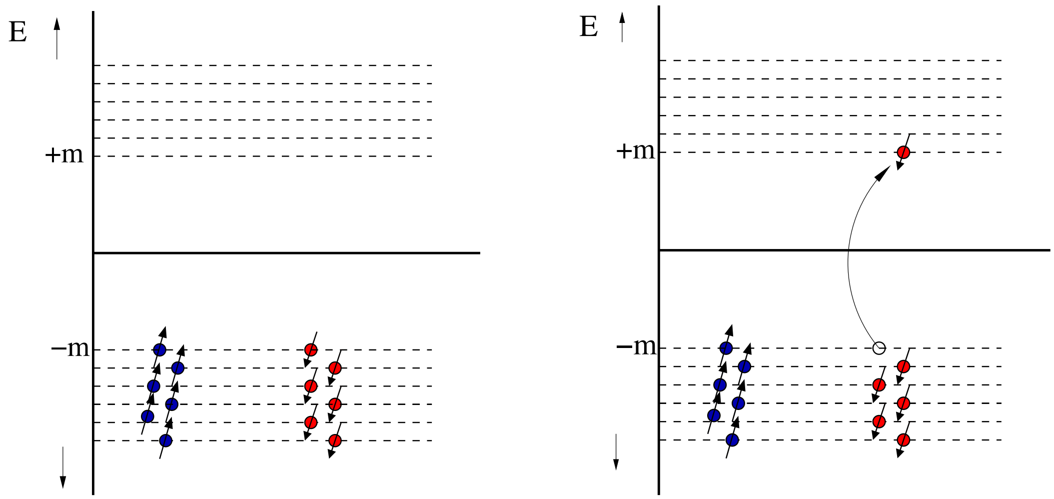


Figure 2.1: Dirac's interpretation of negative energy solutions: "holes"

The plot shows all the avaliable energy levels of an electron. It's 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!

2.3.2 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

$$j^\mu = -ie(\phi^* \partial^\mu \phi - \phi \partial^\mu \phi^*).$$

To describe electro-magnetic 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 answer the issue of negative energies.

2.3.3 Feynman-Stückelberg interpretation

Stückelberg and later Feynmann took this approach one step further. The current density for a particle with charge $-e$ and momentum $(E, \vec{p})$ is:

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

The current density for a particle with charge $+e$ and momentum $(E, \vec{p})$ is:

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

This means that the positive energy solution for a positron **is** the negative energy solution for an electron.

Note that indeed the wave function $Ne^{ip_\mu x^\mu} = Ne^{ip_\mu x^\mu}$ is invariant under: $p^\mu \to -p^\mu$ and $x_\mu \to -x_\mu$. So the wave functions that describe particles also describe anti-particles. The negative energy solutions give particles travelling backwards in time. They are the same as the positive energy solutions of anti-particles travelling forward in time. This is indicated in Fig. 2.2.

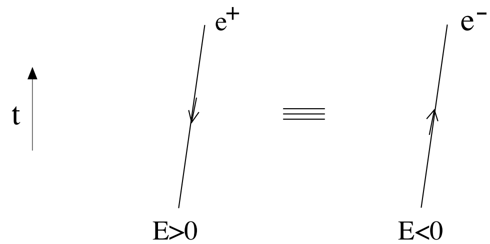


Figure 2.2: A positron travelling forward in time **is** an electron travelling backwards in time.

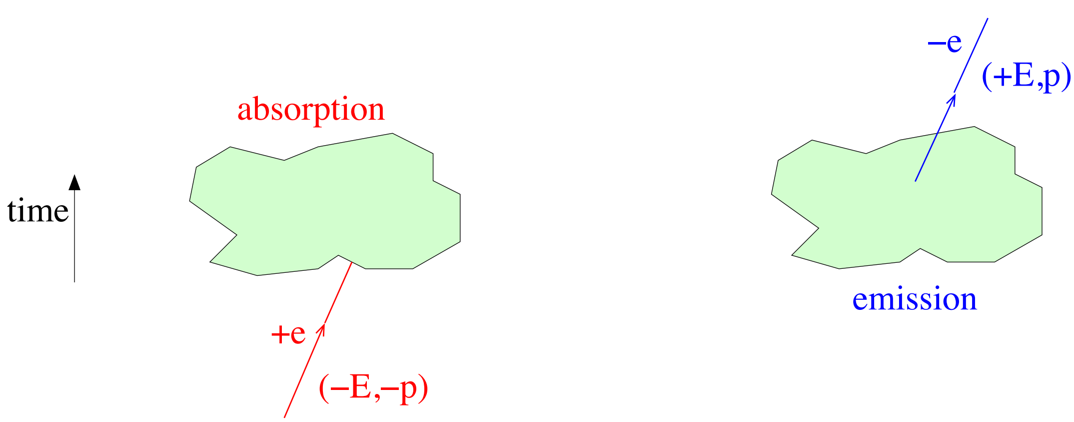


Figure 2.3: There is no difference between the process of an absorption of a positron with $p^\mu = (-E, -\vec{p})$ and the emission of an electron with $p^\mu = (e, \vec{p})$.

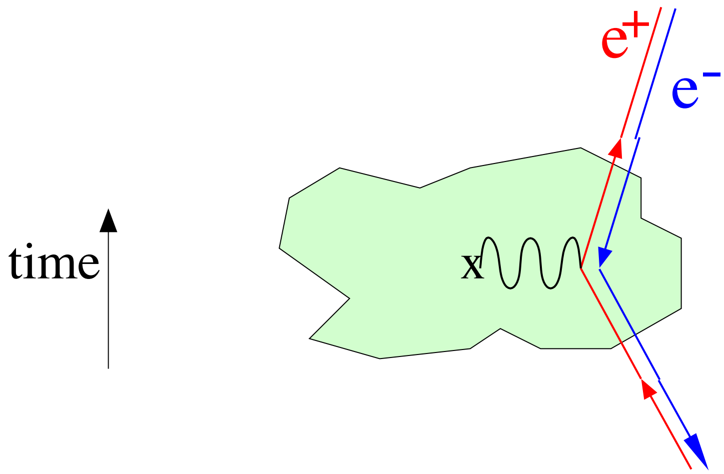


Figure 2.4: In terms of the charge current density $j^{\mu}_{+(E,\vec{p})}(+e) \equiv j^{\mu}_{-(E,\vec{p})}(-e)$

As a consequence of the Feynman-Stückelberg interpretation 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.2.3). In the calculations with Feynman diagrams we have made the convention that all scattering processes are calculated in terms of *particles* and **not** antiparticles. As an example, the process of an incoming positron scattering off a potential will be calculated as that of an electron travelling back in time (see Fig. 2.4).

Let us consider the scattering of an electron in a potential. The probability of a process is calculated in perturbation theory in terms of basic scattering processes (i.e. Feynman diagrams). In Fig. 2.5 the first and second order scattering of the electron is illustrated. To first order a single photon carries the interaction between the electron and the potential. When the calculation is extended to second order the electron interacts twice with the field. It is interesting to note that this scattering can occur in two time orderings as indicated in the figure. Note that the observable path of the electron before and after the scattering process is identical in the two processes. Because of our anti-particle interpretation, the second picture is also possible. It 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.

In quantum mechanics both time ordered processes (the left and the right picture) must be included in the calculation of the cross section. We realize that the vacuum has become a complex environment since particle pairs can spontaneously emerge from it and dissolve into it!

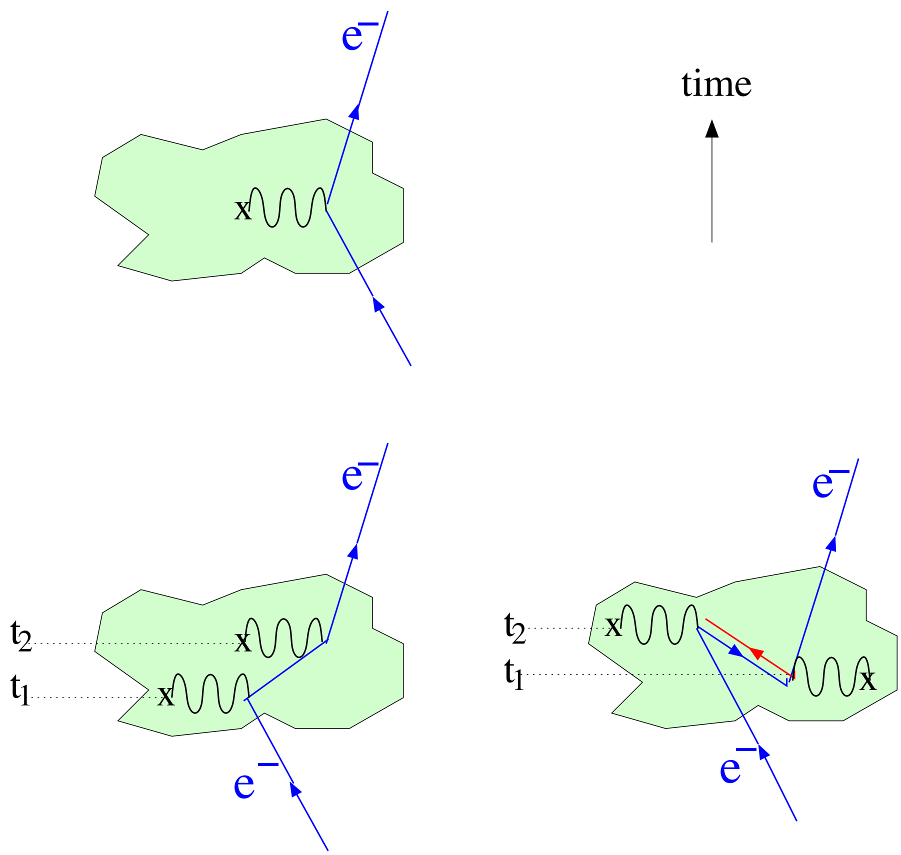
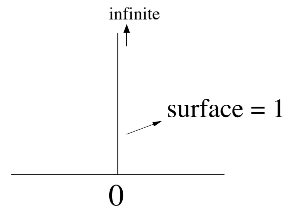


Figure 2.5: First and second order scattering.

2.4 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

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1$$

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

$$\int_{-\infty}^{\infty} f(x)\,\delta(x)\, dx = f(0).$$

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.

Exercise 7.

Starting from the properties of the δ-function,

(a) *Prove that*

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

(b) *Prove that*

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

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 8.

(a) *Calculate* $\int_0^3 \ln(1+x)\,\delta(\pi - x)\,dx$

(b) *Calculate* $\int_0^3 (2x^2 + 7x + 4)\,\delta(x-1)\,dx$

(c) *Calculate* $\int_0^3 \ln(x^3)\, \delta(x/e-1)\, dx$

(d) *Simplify* $\delta\left(\sqrt{(5x-1)} - x - 1\right)$

(e) *Simplify* $\delta(\sin x)$ *and draw the function*

(Note: just writing a number is not enough!)

Lecture 3

The Electromagnetic Field

3.1 Maxwell Equations

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

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

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

(1)	$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0}$	Gauss' law
(2)	$\vec{\nabla} \cdot \vec{B} = 0$	No magnetic charges
(3)	$\vec{\nabla} \times \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0$	Faraday's law of induction
(4)	$c^2\,\vec{\nabla} \times \vec{B} - \frac{\partial \vec{E}}{\partial t} = \frac{\vec{j}}{\epsilon_0}$	Modified Ampère's law

where ϵ_0 is the vacuum permittivity. From the first and the fourth equation we can 'derive' the continuity equation for electric charges, $\vec{\nabla} \cdot \vec{j} = -\frac{\partial \rho}{\partial t}$. Historically, 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 yet anticipate this. He did realize that c is the velocity of a propagating electromagnetic wave. The value of c^2 can be computed c^2 from measurements of ϵ_0 (e.g. with the force between static charges) and measurements of $c^2\epsilon_0$ (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 $c = 1$. (That is, we use so-called 'Heaviside-Lorentz rationalised units'. See appendix C of Aichison and Hey.)

In scattering with particles we need to work with relativistic velocities. Therefore, it is convenient if we 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 $A^\mu = (V, \vec{A})$. As suggested by our notation, the components of this field will turn out to transform as a Lorentz vector.

Remember that the following identities are valid for any vector field $\vec{A}$ and scalar field V:

$$\vec{\nabla} \cdot \left(\vec{\nabla} \times \vec{A}\right) = 0 \qquad \text{(the divergence of a rotation is 0)}$$
$$\vec{\nabla} \times \left(\vec{\nabla} V\right) = 0 \qquad \text{(the rotation of a gradient is 0)}$$

From electrostatics you may remember that, because the rotation of E is zero (which is the same as saying that 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, $E = -\nabla V$. The potential V is not unique: we can add an arbitrary constant and the physics will not change. Likewise, you may have seen in your electrodynamics course that, because the divergence of the B field is zero, we can always find a vector field $\vec{A}$ such that B is the rotation of $\vec{A}$.

Let us now choose a scalar and vector field such that

1. $\vec{B} = \vec{\nabla} \times \vec{A}$
   Then, automatically it follows that: $\vec{\nabla} \cdot \vec{B} = 0$.

2. $\vec{E} = -\frac{\partial \vec{A}}{\partial t} - \vec{\nabla} V$
   Then, automatically it follows that: $\vec{\nabla} \times \vec{E} = -\frac{\partial \left(\vec{\nabla} \times \vec{A}\right)}{\partial t} - 0 = -\frac{\partial \vec{B}}{\partial t}$.

So, by a suitable definition of how the potential A^μ is related to the physical fields, automatically Maxwell equations (2) and (3) are fulfilled.

Exercise 9:

(a) *Show that*

$$\vec{\nabla} \times \left(\vec{\nabla} \times \vec{A}\right) = -\nabla^2 \vec{A} + \vec{\nabla}\left(\vec{\nabla} \cdot \vec{A}\right)$$

Hint: use that $\varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$

(b) *Show that (setting* $c = 1$ *and* $\epsilon_0 = 1$*) Maxwell's equations can be written as:*

$$\partial_\mu \partial^\mu A^\nu - \partial^\nu \partial_\mu A^\mu = j^\nu$$

Hint: Derive the expressions for ρ *and* $\vec{j}$ *explicitly and use the expression in (a).*

(c) *The expressions can be made even more compact by introducing the tensor*

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

Show that with this definition Maxwell's equations reduce to

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

Let us take a closer look at our reformulation of the Maxwell equations,

$$\partial_\mu \partial^\mu A^\nu - \partial^\nu \partial_\mu A^\mu = j^\nu.$$

We have argued in the previous lecture that a conserved current j^μ transforms as a Lorentz vector. (It is easy to work this out for yourself. See also Feynman, Vol.2, section 13.6.) The derivative ∂^μ also transforms as a Lorentz vector. Therefore, it follows that 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, but you probably do remember that it was exactly the transformation properties of the fields that led Einstein to formulate his theory of special relativity.

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 freedom in more detail. Written out in terms of the components $\vec{E}$ and $\vec{B}$ the (4×4) matrix for the electromagnetic field tensor $F^{\mu\nu}$ is given by

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

Note that $F^{\mu\nu}$ *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 Invariance

As stated above the choice of 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*. After re-examining the equations that express the physical fields in terms A we realize that for any scalar field $\lambda(t, \vec{x})$, the transformations

$$\begin{aligned} A^\mu \to A'^\mu &= A^\mu + \partial^\mu \lambda && \text{or} \\ A_\mu \to A'_\mu &= A_\mu + \partial_\mu \lambda && \text{for any scalar field } \lambda \end{aligned}$$

do not change E and B. In terms of the potential V and vector potential $\vec{A}$ we have

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

Exercise 10:
Show explicitly that in such gauge transformations the $\vec{E}$ and $\vec{B}$ fields do not change:

$$\begin{aligned} \vec{B}' &= \vec{\nabla} \times \vec{A}' = ... = \vec{B} \\ \vec{E}' &= -\frac{\partial \vec{A}'}{\partial t} - \vec{\nabla} V' = ... = \vec{E} \end{aligned}$$

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: we can make any convenient choice for the function λ to calculate physics quantities. It is most elegant if we can perform all calculations in a way that is manifestly gauge invariant. However, sometimes we choose a particular gauge in order to make the expressions in calculations simpler. In other cases, we *require* gauge invariance to impose constraints on a solution, as with the photon below.

A gauge choice that is often made is called the *Lorentz condition*[1]: we choose A^μ such that

$$\boxed{\partial_\mu A^\mu = 0.}$$

[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.

which just means that we require A^μ to satisfy the continuity equation.

<u>**Exercise 11:**</u>
Show that it is always possible to define a A^μ field according to the Lorentz gauge. To do this assume that for a given A^μ field one has: $\partial_\mu A^\mu \neq 0$. Give then the equation for the gauge field λ by which that A^μ field must be transformed to obtain the Lorentz gauge.

In the Lorentz gauge the Maxwell equations simplify further:

$$\begin{aligned} \partial_\mu \partial^\mu A^\nu - \partial^\nu \partial_\mu A^\mu &= j^\nu \qquad \text{now becomes :} \\ \partial_\mu \partial^\mu A^\nu &= j^\nu \end{aligned}$$

However, A^μ still has some freedom since we have fixed: $\partial_\mu (\partial^\mu \lambda)$, but we have not yet fixed $\partial^\mu \lambda$! In other words a gauge transformation of the form

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

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

$$\boxed{A^0 = 0 \qquad \text{or equivalently :} \qquad \vec{\nabla} \cdot \vec{A} = 0}$$

At the same time we realize, however, that this is not elegant as we give the "0-th component" or "time-component" of the 4-vector a special treatment. Therefore the choice of this gauge is not Lorentz invariant. This means that one has to choose a different gauge condition if one goes from one reference frame to a different reference frame. This is allowed since the choice of the gauge is irrelevant for the physics observables, but it sometimes considered "not elegant".

3.3 The photon

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

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

We recognize in each component the Klein-Gordon equation of a particle with mass $m = 0$: $(\Box + m^2)\,\phi(x) = 0$. (See previous Lecture). 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 E and 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,

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

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,

$$A^\mu(x) = a^\mu(p)e^{-ipx} + a^\mu(p)^* e^{ipx}.$$

(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

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

which leads to

$$p_\mu a^\mu = 0.$$

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

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

with α a complex constant, then its derivative is

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

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

$$a^{\mu\prime} = a^\mu + \alpha p^\mu$$

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

$$\vec{a} \cdot \vec{p} = 0$$

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

At this point it is customary to uniquely factorize $\vec{a}(p)$ as follows

$$\vec{a}(\vec{p}) \equiv N(p)\,\vec{\varepsilon}(\vec{p})$$

such that the vector $\vec{\varepsilon}$ has unit length and $N(p)$ is real. The normalization $N(p)$ depends only on the magnitude of the momentum and is essentially just the energy density of the wave. The vector $\vec{\varepsilon}$ depends only on the direction of $\vec{p}$ 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

$$\vec{\varepsilon} = (c_1 e^{i\phi_1}, c_2 e^{i\phi_2}, 0) \; .$$

where c_i and ϕ_i are all real and $c_1^2 + c_2^2 = 1$. Note that 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*

$$\vec{\varepsilon}_1 = (1,0,0) \qquad \vec{\varepsilon}_2 = (0,1,0)\,.$$

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

$$\vec{\varepsilon}_+ = \frac{-\vec{\varepsilon}_1 - i\vec{\varepsilon}_2}{\sqrt{2}} \qquad \vec{\varepsilon}_- = \frac{+\vec{\varepsilon}_1 - i\vec{\varepsilon}_2}{\sqrt{2}}$$

<u>Exercise 12</u>

Show that the circular polarization vectors ε_+ and ε_- transform under a rotation of angle ϕ around the z-axis as:

$$\begin{aligned} \vec{\varepsilon}_+ \to \vec{\varepsilon}_+^{\,\prime} &= e^{-i\phi}\vec{\varepsilon}_+ \\ \vec{\varepsilon}_- \to \vec{\varepsilon}_-^{\,\prime} &= e^{i\phi}\vec{\varepsilon}_- \\ \text{or } \vec{\varepsilon}_i^{\,\prime} &= e^{-im\phi}\,\vec{\varepsilon}_i \end{aligned}$$

Hence $\vec{\varepsilon}_+$ and $\vec{\varepsilon}_-$ describe a photon of helicity +1 and -1 respectively.

Hint: The 3rd component of the angular momentum operator, J_3 is the generator *of rotations. That means that for an infinitesimal rotation ϵ around the z axis, we have for any function ψ of the coordinates*

$$U(\epsilon)\psi = (1 - i\epsilon J_3)\psi.$$

A rotation θ may be built up from infinitesimal rotations:

$$\begin{aligned} U(\theta) &= \lim_{n\to\infty} \left(U(\theta/n)\right)^n \\ &= \lim_{n\to infty} \left(1 - i\left(\frac{\theta}{n}\right) J_3)\right)^n = e^{-i\theta J_3} \end{aligned}$$

So, if ψ is an eigen vector of J_3 with eigenvalue m, then for any rotation around the z-axis we have

$$U(\theta)\psi = e^{-im\theta}\psi$$

Use this to translate the behaviour of the polarization vectors under rotations into a statement on spin or helicity.

Since the photon is a spin-1 particle 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 $p^2 =$, leading to $\vec{A} \cdot p = 0$. For massive vector fields (or virtual photons!), there is no corresponding gauge freedom and a component parallel to the momentum p may exist.

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. We can talk only about helicity (spin projection on the momentum) and not about spin. The equivalent of the m_z is zero spin state does not exist for the photon.

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

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

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 The Bohm Aharanov Effect

Later on in the course we will see that the presence of a vector field $\vec{A}$ affects the phase of a wave function of the particle. The phase factor is affected by the presence of the field in the following way:

$$\psi' = e^{i\frac{q}{\hbar}\alpha(\vec{r},t)}\psi$$

where q is the charge of the particle, $\hbar$ is Plancks constant, and α is given by:

$$\alpha(\vec{r},t) = \int_r d\vec{r}' \cdot A(\vec{r}',t)$$

Let us now go back to the famous two-slit experiment of Feynman in which he considers the interference between two possible electron trajectories. From quantum mechanics we know that the intensity at a detection plate positioned behind the two slits shows an interference pattern depending on the relative phases of the wave functions ψ_1 and ψ_2 that travel different paths. For a beautiful description of this 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"). The idea is schematically depicted in Fig. 3.1.

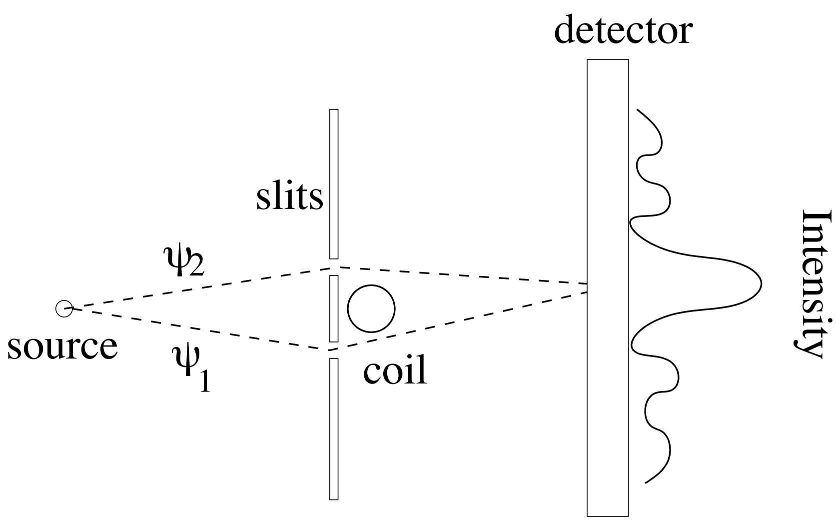


Figure 3.1: The schematical setup of an experiment that investigates the effect of the presence of an A field on the phase factor of the electron wave functions.

In case a field $\vec{A}$ is present the phases of the wave functions are affected, such that the wave function on the detector is:

$$\psi = \psi_1 \, e^{iq\alpha_1(\vec{r},t)} + \psi_2 \, e^{iq\alpha_2(\vec{r},t)} = \left(\psi_1 \, e^{iq(\alpha_1-\alpha_2)} + \psi_2\right) \, e^{iq\alpha_2}$$

We note that the interference between the two amplitudes depends on the relative phase:

$$\begin{aligned}\alpha_1 - \alpha_2 &= \int_{r_1} d\vec{r}'_1 A_1 - \int_{r_2} d\vec{r}'_2 A_2 = \oint d\vec{r}' \cdot \vec{A}(\vec{r}', t) \\ &= \int_S \vec{\nabla} \times \vec{A}(\vec{r}', t) \cdot d\vec{S} = \int_S \vec{B} \cdot d\vec{S} = \Phi\end{aligned}$$

where we have used Stokes theorem to relate the integral around a closed loop to the magnetic flux through the surface. In this way the presence of a magnetic field can affect, (i.e. *shift*) the interference pattern on the screen.

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 B-field is homogeneous and outside it is $B = 0$ (or sufficiently small), see Fig. 3.2. However, from electrodynamics we recall the $\vec{A}$ field is **not** zero outside the coil. There is a lot of $\vec{A}$ circulation around the thin coil.

The electrons in the experiment pass through this $\vec{A}$ field which quantum mechanically affects the phase of their wave function and therefore also the interference pattern on the detector. On the other hand, there is no B field in the region, so classically there is no effect. Experimentally it has been verified (in a technically difficult experiment) that the interference pattern will indeed shift.

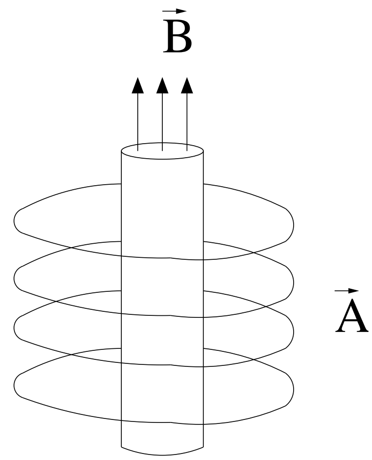


Figure 3.2: Magnetic field and vector potential of a long solenoid.

<u>*Discussion:*</u>
We have introduced the vector potential as a mathematical tool to write Maxwells equations in a Lorentz covariant form. In this formulation we noticed that the A-field has some arbitrariness due to gauge invariance. Quantummechanically we observe, however, that the A field is *not* just a mathematical tool, but gives a more fundamental description of "forces". The aspect of gauge invariance seems an unwanted (*"not nice"*) aspect now, but later on it will turn out to be a fundamental concept in our description of interactions.

Exercise 13 *The delta function*

(a) *Show that*

$$\frac{d^3p}{(2\pi)^3 2E} \tag{3.1}$$

is Lorentz invariant ($d^3p = dp_x dp_y dp_z$). Do this by showing that

$$\int M(p)\; 2d^4p\; \delta(p^2 - m^2)\; \theta(p^0) = \int M(\vec{p}) \frac{d^3p}{E}. \tag{3.2}$$

The 4-vector p is (E, p_x, p_y, p_z), and $M(p)$ is a Lorentzinvariant function of p and $\theta(p^0)$ is the Heavyside function.

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

$$\delta(x) = \lim_{\alpha\to\infty} \frac{1}{\pi} \frac{\sin^2 \alpha x}{\alpha x^2} \tag{3.3}$$

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

(c) *Remember the Fourier transform,*

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} g(k)\, e^{ikx}\, dk$$
$$g(k) = \int_{-\infty}^{+\infty} f(x)\, e^{-ikx}\, dx$$

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

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

Lecture 4

Perturbation Theory and Fermi's Golden Rule

4.1 Scattering cross-section

Consider a scattering process $A+B \to C+D$. As an example, take two spinless electrons scatter in their mutual electromagnetic potential A^μ, as depicted in Fig. 4.1.

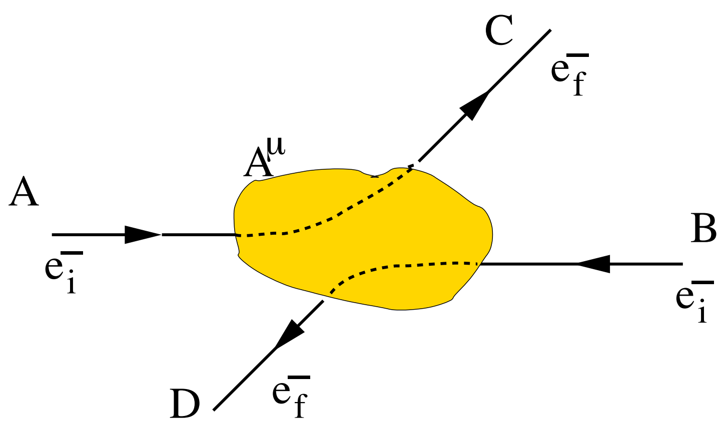


Figure 4.1: Scattering of two electrons in a electromagnetic potential.

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

$$\boxed{\mathrm{d}\sigma = \frac{W_{fi}}{\text{flux}}\,\mathrm{d}\Phi} \tag{4.1}$$

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 usually can not observe individual states but integrate over a number of states with near equal momenta.)

Note that 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 a measurements, such as a cross-section or decay rate.

The computation of the transition rate requires a full treatment in quantum electrodynamics (QED), which is outside the scope of this course. However, to illustrate the concepts we can limit the discussion to non-relativistic scattering of a single particle in a time-dependent potential. In the end we will formulate the result in a Lorentz covariant way. Only in the next lecture we will introduce the concept of Feynman diagrams to compute a transition rate in a realistic theory. This is roughly the approach followed in Halzen and Martin. Griffiths takes an alternative approach: he simply present the full expressions for scattering and decay rates and then guides you through the interpretation of the components in the formulas.

4.2 Non-relativistic scattering

In this section we discuss the ingredients to the general expression for the scattering cross-section in the non-relativistic limit, that is, in classical quantum mechanics.

4.2.1 The transition rate

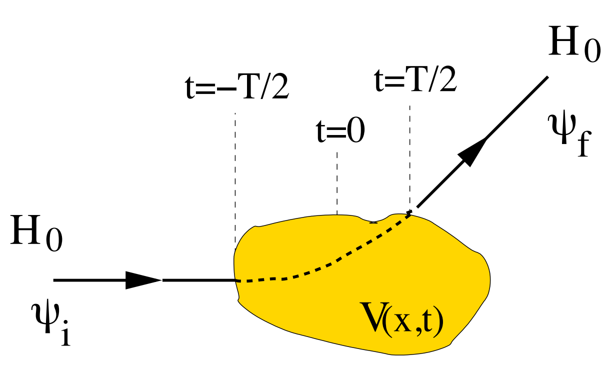


Figure 4.2: Scattering of a single particle in a potential.

Consider the scattering of a particle in a potential as depicted in Fig. 4.2 Assume that before the interaction takes place, as well as after, the system is described by the Schrödinger equation:

$$i\,\frac{\partial\psi}{\partial t} = H_0\,\psi \tag{4.2}$$

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

$$H_0\phi_m(\vec{x}) = E_m\phi_m(\vec{x}). \tag{4.3}$$

The states ϕ_m form an orthonormal basis,

$$\int \phi^*_m(\vec{x})\,\phi_n(\vec{x})\,d^3x = \delta_{mn}. \tag{4.4}$$

The time-dependent wave function

$$\psi_m(\vec{x},t) = \phi_m(\vec{x})\,e^{-iE_mt}. \tag{4.5}$$

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,

$$i\frac{\partial\psi}{\partial t} = \left(H_0 + V\left(\vec{x},t\right)\right)\,\psi\;. \tag{4.6}$$

Any solution ψ can be written as

$$\psi = \sum_{n=0}^{\infty} a_n(t)\,\phi_n(\vec{x})\,e^{-iE_nt}. \tag{4.7}$$

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 coëfficients $a_n(t)$ we substitute 4.7 in 4.6 and find

$$i\sum_{n=0}^{\infty}\frac{da_n(t)}{dt}\,\phi_n(\vec{x})\,e^{-iE_nt} \;=\; \sum_{n=0}^{\infty} V(\vec{x},t)\,a_n(t)\,\phi_n(\vec{x})\,e^{-iE_nt}\;,$$

where we have used that the ψ_m are solutions of the free Schrödinger equation. Multiply the resulting equation from the left with $\psi^*_f = \phi^*_f(\vec{x})\,e^{iE_ft}$ and integrate over $\vec{x}$ to obtain

$$i\sum_{n=0}^{\infty}\frac{da_n(t)}{dt}\underbrace{\int d^3x\,\phi^*_f(\vec{x})\,\phi_n(\vec{x})}_{\delta_{fn}}\,e^{-i\left(E_n-E_f\right)t} \;=$$

$$\sum_{n=0}^{\infty} a_n(t)\int d^3x\,\phi^*_f(\vec{x})\,V(\vec{x},t)\,\phi_n(\vec{x})\,e^{-i\left(E_n-E_f\right)t}$$

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

$$\frac{da_f(t)}{dt} = -i \sum_{n=0}^{\infty} a_n(t) \int d^3x\, \phi_f^*(\vec{x})\, V(\vec{x},t)\, \phi_n(\vec{x})\, e^{-i\left(E_n - E_f\right)t} . \tag{4.8}$$

We now make a few simplications (without actual loss in generality):

- we assume that the interaction is turned on at time $-T/2$ and turned off at time $T/2$;
- we prepare the incoming wave in an eigenstate i of the free Hamiltonian. The incoming wave is: $\psi_i = \phi_i(\vec{x})\, e^{-iE_i t}$. In other words: $a_i(-\infty) = 1$ and $a_n(-\infty) = 0$ for $(n \neq i)$.
- we assume that the interaction is sufficiently weak that for all times t each coefficients for $f \neq i$ is $a_f(t) \ll 1$. We then ignore all terms with $n \neq i$ in the right hand side of the differential equation, which gives us the lowest order in perturbation theory.

With that last approximation, the right hand side of the differential equation for a_f reduces to

$$\frac{da_f(t)}{dt} = -i \int d^3x\, \phi_f^*(\vec{x})\, V(\vec{x},t)\, \phi_i(\vec{x})\, e^{-i\left(E_i - E_f\right)t} \tag{4.9}$$

Using that $a_f(-T/2) = 0$ and integrating this equation we obtain for the coefficient $a_f(t)$ at time t',

$$a_f(t') = \int_{-T/2}^{t'} \frac{da_f(t)}{dt}\, dt = -i \int_{-T/2}^{t'} dt \int d^3x\, \left[\phi_f(\vec{x})\, e^{-iE_f t}\right]^* V(\vec{x},t)\, \left[\phi_i(\vec{x})\, e^{-iE_i t}\right]$$

We define the *transition amplitude* T_{fi} as the amplitude to go from a state i to a final state f at the end of the interaction:

$$T_{fi} \equiv a_f(T/2) = -i \int_{-T/2}^{T/2} dt \int d^3x\, \psi_f^*(\vec{x},t)\, V(\vec{x},t)\, \psi_i(\vec{x},t)$$

Finally we take the limit $T \to \infty$ and write the result in "4-vector" notation,

$$\boxed{T_{fi} = -i \int d^4x\, \psi_f^*(x)\, V(x)\, \psi_i(x)} \tag{4.10}$$

Note that the expression for T_{fi} has a Lorentzinvariant form: since it is a scalar, and any dependence on $\vec{x}$ and t is ingrated out, it is independent of the Lorentz frame. Indeed, the expression does not change when we consider relativistic scattering.

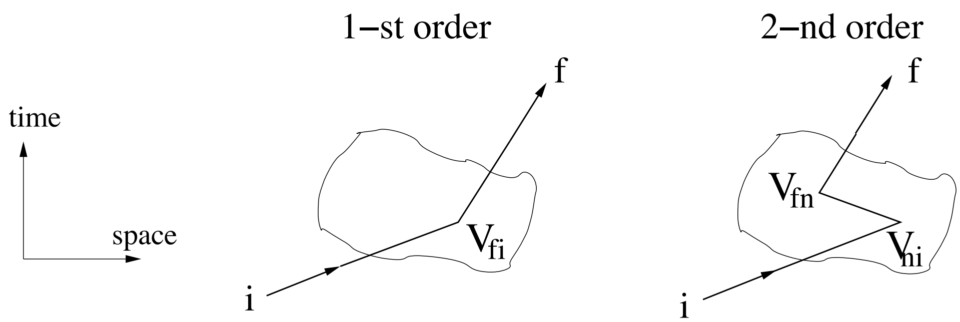


Figure 4.3: First and Second order approximation in scattering.

What is the meaning of the initial conditions: $a_i(t) = 1, a_n(t) = 0$? It implies that the potential can only make **one** quantum perturbation from the initial state i to the final state f. For example the perturbation: $i \to n \to f$ is not included in this approximation. (It is called a 2$^{\text{nd}}$ order perturbation).

If we want to improve the calculation to second order in perturbation theory we replace the approximation $a_n(t) = 0$ by the first order result:

$$\begin{aligned} \frac{da_f(t)}{dt} &= -i\, V_{fi}\, e^{i\left(E_f - E_i\right)t} \\ &+ (-i)^2 \left[\sum_{n \neq i} V_{ni} \int_{-T/2}^{t} dt'\, e^{i(E_n - E_i)t'} \right] V_{fn}\, e^{i\left(E_f - E_n\right)t} \end{aligned}$$

where we have assumed that the perturbation $V(\vec{x}, t)$ is constant over the time interval $[-T/2, T/2]$ and defined

$$V_{fi} \equiv \int d^3x \phi_f^*(\vec{x})\, V(\vec{x})\, \phi_i(\vec{x})$$

See the book of Halzen and Martin how to work out the second order calculation. A graphical illustration of the first and second order perturbation is given in Fig. 4.3.

Can we interpret $|T_{fi}|^2$ as the probability that a particle has scattered from state i to state f? Consider the case where the perturbation is time **in**dependent. Then:

$$T_{fi} = -i\, V_{fi} \int_{-\infty}^{\infty} dt\, e^{i\left(E_f - E_i\right)t} = -2\pi i\, V_{fi}\, \delta\left(E_f - E_i\right)$$

The δ-function expresses energy conservation in $i \to f$. From the uncertainty princple it can then be inferred that the transition between two exactly defined energy states E_i and E_f must be infinitely seperated in time. Therefore the quantity $|T_{fi}|^2$ is not a meaningfull quantity. Instead, we define the transition probability per unit time, or *transition rate*, as

$$W_{fi} \equiv \lim_{T \to \infty} \frac{|T_{fi}|^2}{T}. \tag{4.11}$$

The calculation of the transition probability is non-trivial as it involves the square of a δ-function. A proper treatment is rather lengthy[1] and relies on the use of wave packets. Instead we will apply a "trick". If we assume that the interaction occurs during a time period T from $t = -T/2$ until $t = +T/2$ we can write:

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

The δ-function in the first integral implies that there is only contribution for E_f equal to E_i in the second integral.

Then we note that the arbitrary chosen time period T drops out of the formula such that the transition probability per unit time becomes:

$$W_{fi} = \lim_{T\to\infty} \frac{|T_{fi}|^2}{T} = \frac{2\pi}{\hbar} \, |V_{fi}|^2 \, \delta\left(E_f - E_i\right) \ ,$$

where we temporarily revived $\hbar$ to show that W_{fi} is indeed a *rate*. (It has units of "energy/$\hbar$" or "inverse time".) This is the transition rate for a given initial state into a specific final state with energy $E_f = E_i$.

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. The number of final states with energy between E_f and $E_f + \mathrm{d}E_f$ is given by

$$\mathrm{d}n \ = \ \rho(E_f)\mathrm{d}E_f$$

where $\rho(E_f)$ is the density of states per unit energy near E_f. Integrating the expression for the transition rate over all final state energies, we obtain Fermi's (Second) Golden Rule,

$$\boxed{\begin{aligned} \overline{W}_{fi} &\equiv \int W_{fi} \, \rho\left(E_f\right) \, \mathrm{d}E_f \\ &= \frac{2\pi}{\hbar} \, |V_{fi}|^2 \, \rho\left(E_i\right) \ . \end{aligned}} \tag{4.11}$$

In the following exercise we show an alternative way of arriving at this result, without explicitly taking the limit $T \to \infty$.

Exercise 14.

Assume that there is a constant perturbation potential between $t = -T/2$ and $t = T/2$.

[1] see e.g. the book by K.Gottfried, "Quantum Mechanics" (1966), Volume 1, sections 12, 56.

(a) *Write down the expression for T_{fi} and do the integral over t.*

(b) *Write down the expression for W_{fi}. Use the result of exercise 13b to rewrite $\sin^2 x/x^2$ in the limit of large T as a δ-function.*

(c) *Assume that density for final states is given by $\rho(E_f)$ and perform the integral over all final states dE_f. Compare it to the expression of Fermi's Golden rule.*

4.2.2 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

$$\psi(\vec{x}, t) \;=\; N e^{-i(Et - \vec{x}\cdot\vec{p})} \;.$$

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

$$\int_V \psi^*(\vec{x}, t)\psi(\vec{x}, t)\mathrm{d}^3x = 1 \;.$$

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 we had a density $\rho = |\psi|^2$ so that the normalization gives one particle per unit volume. For the plane wave solutions of the Klein-Gordon equation, we had $\rho = 2|N|^2 E$, which with the box normalization becomes

$$\rho = 2E/V. \tag{4.12}$$

In other words, in the relativistic case we have $2E$ particles per unit volume. Remember that the density ρ is proportional to E in order to compensate for the Lorentz contraction of the volume element d^3x: we need the number of particles in a finite volume, $\rho\,\mathrm{d}^3x$, to remain remain constant in Lorentz transformations. With this result we are ready to compute the flux factor.

4.2.3 The Flux Factor

The flux factor or the initial flux corresponds to the number of particles that pass each other per unit area and per unit time. It can be most easily computed in the lab frame. Consider the case that a beam of particles (A) is shot on a target (B), see Fig. 4.4.

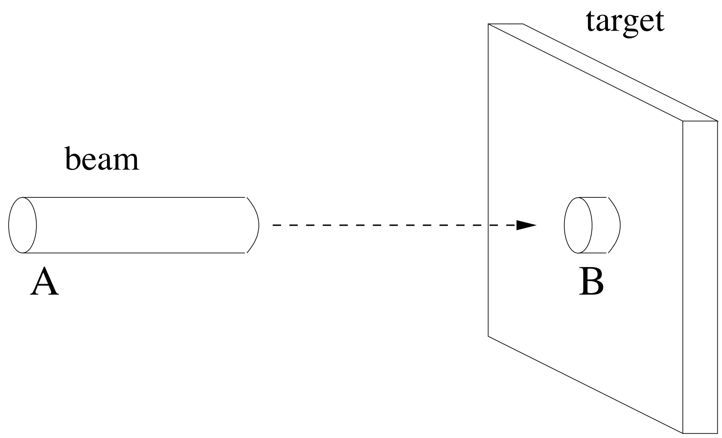


Figure 4.4: A beam incident on a target.

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

$$\text{flux} = |\vec{v}_A|\, n_a\, n_b = |\vec{v}_A|\, \frac{2E_A}{V}\, \frac{2E_B}{V}. \tag{4.13}$$

In the following exercise we show that the kinematic factor is Lorentz invariant. (The volume factor is not, but it will drop out later.)

Exercise 15.

Prove that the following Lorentz invariant expression for the flux (ignoring transformation of the volume) is equal to the one in Eq. 4.13:

$$\boxed{flux = 4\sqrt{(p_A \cdot p_B)^2 - m_A^2 m_B^2}\ /\ V^2} \tag{4.14}$$

Hint: Since the kinematic factor is Lorentz invariant, you can compute it in any frame. Now simply choose the frame in which Eq. 4.13 is defined. Note that (ignoring factors c) $\vec{v} = \vec{p}/E$. Also note that $p_A \cdot p_B$ is an inner product of the four-vectors (not the three-vectors).

4.2.4 The Phase Space Factor

In the final step to Fermi's golden rule we introduced the density of final states $\rho(E)$. The phase space factor accounts for the number of accessible 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. It is customary to express the cross-section as a differential cross-section to the final state momenta,

$$\mathrm{d}\sigma = \ldots \prod_f \mathrm{d}^3 p_f$$

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' $V\mathrm{d}^3p$".

Assume that our volume V is rectangular with sides L_x, L_y, L_z. A particle with momentum p has a "size" given by: $\lambda = 2\pi\hbar/p$. Using periodic boundary conditions to ensure no net particle flow out of the volume we see that the number of states with a momentum between $\vec{p} = (0,0,0)$ and $\vec{p} = (p_x, p_y, p_z)$ is

$$N = n_x\, n_y\, n_z = \frac{L_x}{\lambda_x}\frac{L_y}{\lambda_y}\frac{L_z}{\lambda_z} = \frac{L_x p_x}{2\pi\hbar}\frac{L_y p_y}{2\pi\hbar}\frac{L_z p_z}{2\pi\hbar} = \frac{V}{(2\pi\hbar)^3}\, p_x\, p_y\, p_z\ .$$

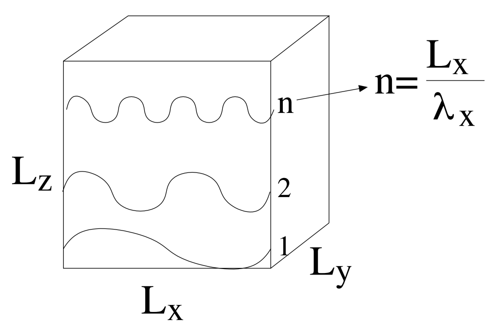


Figure 4.5: Schematic calculation of the number of states in a box of volume V.

An alternative view is given by Burcham & Jobes on page 305. The number of final states is given by the total size of the available phase space for the final state divided by the volume of the elementary cell, h^3 (within an elementary cell states cannot be distinguished):

$$N = \frac{1}{h^3}\int dx\, dy\, dz\, dp_x\, dp_y\, dp_z = \frac{V}{(2\pi)^3}\, p_x\, p_y\, p_z\ .$$

As a consequence, the number of states with momentum between $\vec{p}$ and $\vec{p} + \mathrm{d}\vec{p}$ (i.e. between (p_x, p_y, p_z) and $(p_x + dp_x, p_y + dp_y, p_z + dp_z)$) is:

$$\mathrm{d}N = \frac{V\mathrm{d}^3p}{(2\pi\hbar)^3}\ .$$

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:

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

where we dropped $\hbar$ again. For a process in the form $A + B \to C + D + E + ...$ with N final state particles the Lorentz invariant phase space factor then becomes

$$\mathrm{d}\Phi = \mathrm{dLIPS} = \prod_{i=1}^{N} \frac{V}{(2\pi)^3}\frac{\mathrm{d}^3 p_i}{2E_i}\,. \tag{4.15}$$

4.2.5 Summary

Finally we arrive at the formula to calculate a cross section for the process $A_i + B_i \to C_f + D_f + ...$:

$$\boxed{\begin{aligned} \mathrm{d}\sigma_{fi} &= \frac{1}{\text{flux}} W_{fi}\,\mathrm{d}\Phi \\ W_{fi} &= \lim_{T\to\infty}\frac{|T_{fi}|^2}{T} \\ T_{fi} &= -i\int d^4x\,\psi_f^*(x)\,V(x)\,\psi_i(x) \\ \mathrm{d}\Phi &= \prod_{f=1}^{N}\frac{V}{(2\pi)^3}\frac{\mathrm{d}^3\vec{p}_f}{2E_f} \\ \text{flux} &= 4\sqrt{(p_A\cdot p_B)^2 - m_A^2 m_B^2}\;/\;V^2 \end{aligned}} \tag{4.16}$$

Exercise 16
Show that the cross section does not depend on the arbitrary volume V.

Exercise 17
Why is the phase space factor indeed Lorentz invariant? (Hint: You may just refer to a previous exercise.)

4.3 Extension to Relativistic Scattering

The attentive reader will have realized that something is odd in the discussion above. Although we talked about a 'scattering' cross-section, we never actually considered two particles A and B in the computation of the transition rate. We just looked at scattering of a single particle in a time-dependent potential, as if one of the particles could just be considered as a static source.

In a more realistic computation we need to deal with two particles. As explained in the 1st lecture, such scattering processes can be described by the exchange of virtual particles, Yukawa's force carriers. A full computation requires quantum field theory, but even without that we can identify one place where it should differ from the formalism above: the result must somehow encode four-momentum conservation, and not just conservation of energy.

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. 4.10. Performing the integral using incoming and outgoing plane waves $\phi = N e^{-ipx}$ the result can be written as

$$T_{fi} = -i\, N_A N_B N_C N_D\, (2\pi)^4\, \delta^4(p_A + p_B - p_C - p_D)\, \mathcal{M}\ . \tag{4.17}$$

The quantity $\mathcal{M}$ is usually called the *matrix element*. It is computed using Feynman diagrams. The δ-function takes care of energy and momentum conservation in the process. (Note that the momentum vectors are four-vectors). We rely again on box-normalization on a finite volume V such that for each particle $N = 1/\sqrt{V}$.

To find the transition probability we square this expression

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

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 these, we consider a transition probability per unit time *and* per unit volume:

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

With this new definition the cross section is still given by[2]

$$\mathrm{d}\sigma = \frac{W_{fi}}{\text{flux}}\, \mathrm{d}\Phi$$

[2] Usually we will write this as:

$$\mathrm{d}\sigma = \frac{|\mathcal{M}|^2}{\text{flux}}\, \mathrm{d}\Phi$$

and absorb the delta function in the phase space factor.

and also the expressions for phase space and flux factors to not change. Inserting the plane wave normalization the result for the differential cross-section becomes

$$\boxed{\mathrm{d}\sigma = \frac{(2\pi)^4\ \delta^4\left(p_A + p_B - p_C - p_D\right)}{4\sqrt{\left(p_A \cdot p_B\right)^2 - m_A^2 m_B^2}} \cdot |\mathcal{M}|^2 \cdot \frac{\mathrm{d}^3 p_C}{(2\pi)^3\ 2E_C} \frac{\mathrm{d}^3 p_D}{(2\pi)^3\ 2E_D}} \tag{4.18}$$

Note that 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.

The computation of a decay rate for the process $A \to C + D$ follows a similar strategy. Here we just quote the result:

$$\boxed{d\Gamma = \frac{(2\pi)^4\ \delta^4\left(p_A - p_C - p_D\right)}{2E_A} \cdot |\mathcal{M}|^2 \cdot \frac{d^3 p_C}{(2\pi)^3\ 2E_C} \frac{d^3 p_D}{(2\pi)^3\ 2E_D}} \tag{4.19}$$

Exercise 18. (See also Halzen and Martin, 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:*

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

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

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

where we have made use spherical coordinates (i.e.: $d^3 p_C = |p_C|^2 dp_C\, d\Omega$) and defined $p_f \equiv |p_C|$.

(b) *In the C.M. frame we have $|\vec{p}_A| = |\vec{p}_B| = p_i$ and $|\vec{p}_C| = |\vec{p}_D| = p_f$. 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):*

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

So that we finally get:

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

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

$$F = 4p_i\sqrt{s}$$

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

$$\boxed{\left.\frac{d\sigma}{d\Omega}\right|_{cm} = \frac{1}{64\pi^2 s}\,\frac{p_f}{p_i}\,|\mathcal{M}|^2} \tag{4.20}$$

For the decay rate $A \to B + C$ one finds $(4p_i\sqrt{s} \to 2E_A = 2m_A)$

$$\boxed{\left.\frac{d\Gamma}{d\Omega}\right|_{cm} = \frac{1}{32\pi^2 m_A^2}\,p_f\,|\mathcal{M}|^2} \tag{4.21}$$

Lecture 5

Electromagnetic Scattering of Spinless Particles

In this lecture we discuss electromagnetic scattering of spinles particles. First we describe an example of a charged particle scattering in an external electric field. Second we derive the cross section for two particles that scatter in each-others field. We end the lecture with a prescription how to treat antiparticles.

Intermezzo: The principle of least action

In classical mechanics the equations of motion can be derived using the variational principle of Hamilton which states that the action integral S should be stationary under arbitrary variations of the so-called generalized coordinates $q_i, \dot{q}_i$. 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. For example, take 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 would say that the movement of the pendulum can be described by one generic coordinate. We could choose x, 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*.

Defining the Lagrangian of the system as the kinetic energy minus the potential energy,

$$L = T(\dot{q}) - V(q) \tag{5.1}$$

(where the potential energy only depends on q and the kinetic energy only on $\dot{q}$), we denote the action (or 'action integral') of a path that starts at t_1 and ends at t_2 with

$$S(q) \;=\; \int_{t_0}^{t_1} L(q,\dot{q})\mathrm{d}t. \tag{5.2}$$

Hamilton's principle now states that the actual trajectory $q(t)$ followed by the system is the trajectory *that minimizes the action.*

You will derive later in this course that for each of the coordinates q_i, this leads to the so-called Euler-Lagrange equation of motion

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} \;=\; \frac{\partial L}{\partial q_i}. \tag{5.3}$$

This may also be written more symmetrically as Hamilton's equations,

$$\dot{p}_i = \frac{\partial \mathcal{L}}{\partial q_i} \qquad \text{with} \qquad p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \quad , \tag{5.4}$$

where p_i is called the generalized momentum, or sometimes *the momentum canonical to* q_i. In terms of these coordinates, the Hamiltonian takes the form

$$H \;=\; \sum_i p_i \dot{q}_i - L. \tag{5.5}$$

Finally, the classical system can be quantized by imposing the fundamental postulate of quantum mechanics,

$$[q_i, p_j] \;=\; i\hbar\delta_{ij}. \tag{5.6}$$

The formulation of physics laws in terms of Euler-Lagrange equations derived from a Lagrangian (or Lagrangian *density*, for fields) plays a fundamental role in particle physics. As we shall see in chapter 10, any symmetry of the Lagrangian leads to a conserved quantity. (This is Noether's theorem. As an example, $U(1)$ symmetry leads to the conserved current for the Klein-Gordon and Dirac fields.) Furthermore, all interactions in the standard model can be constructed by requiring so-called local gauge-symmetries of the Lagrangian.

5.1 Electrodynamics

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

$$p^\mu \rightarrow p^\mu - qA^\mu. \tag{5.7}$$

in the equations of motion of the free particle. Written out in terms of the energy Φ and momentum $\vec{A}$ components of the field, the free Hamiltonian is then replaced by

$$H\,\psi = \left[\frac{1}{2m}\left(\vec{p} - q\vec{A}(\vec{r},t)\right)^2 + q\Phi(\vec{r},t)\right]\,\psi$$

It can be shown (see e.g. Jackson §12.1, page 575) that this indeeds leads to the known equation of motion,

$$\frac{d\vec{p}}{dt} = q\left(\vec{E} + \vec{v}\times\vec{B}\right).$$

Exercise 19 (from Kibble, p. 243).

The Lagrangian for a charged particle moving in a electromagnetic field is

$$L \;=\; \frac{1}{2}m\vec{v}^2 + q\vec{v}\cdot\vec{A}(\vec{r},t) - q\Phi(\vec{r},t)$$

(a) *Show that for a uniform magnetic field, we may take:*

$$V = 0, \qquad \vec{A} = \frac{1}{2}\vec{B}\times\vec{r}$$

If we choose the z-axis in the direction of $\vec{B}$ we have in cylindrical coordinates (r,ϕ,z):

$$V = 0, \quad A_r = 0, \quad A_\phi = \frac{1}{2}Br, \quad A_z = 0$$

Hint: In cylindrical coordinates the cross product is defined as:

$$\vec{\nabla}\times\vec{A} = \left(\frac{1}{r}\frac{\partial A_z}{\partial\phi} - \frac{\partial A_\phi}{\partial z}, \frac{\partial A_z}{\partial z} - \frac{\partial A_z}{\partial r}, \frac{1}{r}\left[\frac{\partial\,(rA_\phi)}{\partial r} - \frac{\partial A_r}{\partial\phi}\right]\right)$$

(b) *Write down the Lagrangian in cylindrical coordinates*

(c) *Write out the Lagrangian equations:*

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_\alpha}\right) = \frac{\partial\mathcal{L}}{\partial q_\alpha}$$

in the cylindrical coordinates.

(d) *Show that the equation of motion in terms of the coordinate $\dot{\phi}$ yields (assume r=constant):*

$$\dot{\phi} = 0 \quad \text{or} \quad \dot{\phi} = -\frac{qB}{m}$$

i.e. that it is in agreement with the law:

$$\vec{F} = \frac{d\vec{p}}{dt} = q\left(\vec{E} + \vec{v}\times\vec{B}\right)$$

In quantum mechanics we make the replacement $p^\mu \to i\partial^\mu$, such that the method of minimal substitution takes the form

$$\boxed{\partial^\mu \to \partial^\mu + iqA^\mu} \tag{5.8}$$

This is the essence of quantum electrodynamics. As we will see later in the lectures this substitution is equivalent to requiring QED to be locally $U(1)$ (phase) gauge invariant: We can introduce the field by requiring a symmetry!

We now proceed as follows. Start with the Klein-Gordon equation, the wave-equation for a spinless particle with mass m,

$$\left(\partial_\mu \partial^\mu + m^2\right)\ \phi = 0$$

and substitute $\partial^\mu \to \partial^\mu - ieA^\mu$ for a particle with charge $-e$

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

This is of the form

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

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

$$V(x) \equiv -ie\left(\partial_\mu A^\mu + A_\mu \partial^\mu\right) - e^2 A^2\ . \tag{5.9}$$

The sign of V is chosen such that compared to the kinetic energy it gets the same sign as in the Schrödinger equation. Since e^2 is small ($\alpha = e^2/4\pi = 1/137$) we can neglect the second order term, $e^2A^2 \approx 0$. Note that the derivatives as they appear here act on anything to their right. In particular, they still act on the field in terms $V(x)\phi(x)$.

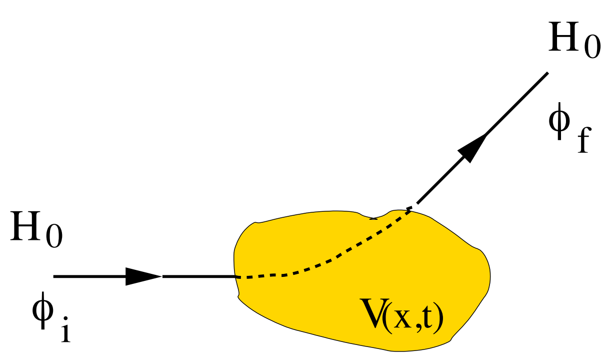


Figure 5.1: Scattering potential

From the previous lecture we take the general expression for the transition amplitude

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

We can use integration by parts to write

$$\int \mathrm{d}^4x\, \phi_f^*\, \partial_\mu \left(A^\mu\, \phi_i\right) = \underbrace{\left[\phi_f^*\, A^\mu\, \phi_i\right]_{-\infty}^{\infty}}_{=0} - \int \partial_\mu \left(\phi_f^*\right)\, A^\mu\, \phi_i\, \mathrm{d}^4x$$

Requiring the field to be zero at $t = \pm\infty$, the first term on the left vanishes, leading to

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

Remember the definition of the current density for the complex scalar field,

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

In complete analogy we define the "transition current density" to go from initial state i to final state f as

$$j_\mu^{fi} \equiv -ie\, \left[\phi_f^* \left(\partial_\mu \phi_i\right) - \left(\partial_\mu \phi_f^*\right) \phi_i\right]\, . \tag{5.10}$$

The transition amplitude can then be written as

$$\boxed{T_{fi} = -i \int j_\mu^{fi}\, A^\mu\, \mathrm{d}^4x} \tag{5.11}$$

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.

If we substitute the free particle solutions of the unperturbed Klein-Gordon equation in initial and final states we find for the transition current of spinless particles

$$\boxed{\begin{aligned} \phi_i &= N_i\, e^{-ip_i x} \quad ; \quad \phi_f^* = N_f^* e^{ip_f x} \\ j_\mu^{fi} &= -eN_i N_f^* \left(p_\mu^i + p_\mu^f\right)\, e^{i\left(p_f - p_i\right)x} \quad . \end{aligned}} \tag{5.12}$$

Verify that the conservation law $\partial^\mu j_\mu^{fi} = 0$ holds. From this equation it can be derived that the charge is conserved in the interaction.

5.2 Scattering in an External Field

Consider the case that the external field is a static field of a point charge Z located in the origin:

$$A_\mu = \left(V, \vec{A}\right) = \left(V, \vec{0}\right) \qquad \text{with} \qquad V(x) = \frac{Ze}{4\pi|\vec{x}|}$$

The transition amplitude is:

$$\begin{aligned} T_{fi} &= -i \int j^{\mu}_{fi}\, A_\mu \, d^4x \\ &= -i \int (-e)\, N_i N_f^* \, \left(p_i^\mu + p_f^\mu\right)\, A_\mu \, e^{i\left(p_f - p_i\right)x}\, d^4x \end{aligned}$$

Insert that $A_\mu = \left(V, \vec{0}\right)$ and thus: $p^\mu A_\mu = E\, V$:

$$T_{fi} = i \int eN_i N_f^* \, \left(E_i + E_f\right)\, V(x)\, e^{i\left(p_f - p_i\right)x}\, d^4x$$

Split the integral in a part over time and in a part over space and note that $V(\vec{x})$ is not time dependent. Use also again: $\int e^{i\left(E_f - E_i\right)t} dt = 2\pi\, \delta\left(E_f - E_i\right)$ to find that:

$$T_{fi} = ieN_i N_f^* \, \left(E_i + E_f\right)\, 2\pi\, \delta\left(E_f - E_i\right) \int \frac{Ze}{4\pi |\vec{x}|}\, e^{-i\left(\vec{p}_f - \vec{p}_i\right)\vec{x}}\, d^3x$$

Now we make use of the Fourier transform:

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

Using this with $\vec{q} \equiv \left(\vec{p}_f - \vec{p}_i\right)$ we obtain:

$$T_{fi} = ieN_i N_f^* \, \left(E_i + E_f\right)\, 2\pi\, \delta\left(E_f - E_i\right)\, \frac{Ze}{\left|\vec{p}_f - \vec{p}_i\right|^2}$$

The next step is to calculate the transition probability:

$$\begin{aligned} W_{fi} &= \lim_{T\to\infty} \frac{\left|T_{fi}\right|^2}{T} \\ &= \lim_{T\to\infty} \frac{1}{T}\, \left|N_i N_f^*\right|\, \left[2\pi\, \delta\left(E_f - E_i\right)\right]^2 \left(\frac{Ze^2\left(E_i + E_f\right)}{\left|\vec{p}_f - \vec{p}_i\right|^2}\right)^2 \end{aligned}$$

We apply again our "trick" (or calculate the integral explicitly and let $T \to \infty$):

$$\begin{aligned} \lim_{T\to\infty} \left[2\pi\, \delta\left(E_f - E_i\right)\right]^2 &= 2\pi\, \delta\left(E_f - E_i\right) \cdot \lim_{T\to\infty} \int_{-T/2}^{T/2} dt\, e^{i\left(E_f - E_i\right)t} \\ &= 2\pi\, \delta\left(E_f - E_i\right) \cdot \lim_{T\to\infty} \underbrace{\int_{-T/2}^{T/2} e^{i0t} dt}_{T} \\ &= \lim_{T\to\infty} 2\pi\, \delta\left(E_f - E_i\right) \cdot T \end{aligned}$$

Putting this back into W_{fi} we obtain:

$$W_{fi} = \lim_{T\to\infty} \frac{1}{T} \cdot T \; |N_i N_f|^2 \; 2\pi\,\delta\left(E_f - E_i\right) \left(\frac{Ze^2\left(E_i + E_f\right)}{|\vec{p}_f - \vec{p}_i|^2}\right)^2$$

The cross section is given by[1]:

$$\begin{aligned}
d\sigma &= \frac{W_{fi}}{\text{Flux}}\,\text{dLips} \\
\text{with :} & \\
\text{Flux} &= \vec{v}\,\frac{2E_i}{V} = \frac{\vec{p}_i}{E_i}\frac{2E_i}{V} = \frac{2\vec{p}_i}{V} \\
\text{dLips} &= \frac{V}{(2\pi)^3}\frac{d^3p_f}{2E_f} \\
\text{Normalization : } N &= \frac{1}{\sqrt{V}} \qquad\qquad \rightarrow \qquad \int_V \phi^*\phi dV = 1
\end{aligned}$$

In addition, from energy and momentum conservation we write $E = E_i = E_f$ and $p = |\vec{p}_f| = |\vec{p}_i|$

Putting everything together:

$$d\sigma = \frac{1}{V^2}\,2\pi\,\delta\left(E_f - E_i\right) \cdot \left(\frac{Ze^2\left(E_i + E_f\right)}{|\vec{p}_f - \vec{p}_i|^2}\right)^2 \cdot \frac{V}{2\,|\vec{p}_i|}\,\frac{V}{(2\pi)^3}\,\frac{d^3p_f}{2E_f}$$

Note that the arbitrary volume V drops from the expression!
Use now $d^3p_f = p_f^2\,dp_f\,d\Omega$ and $|p_f| = |p_i| = p$ to get:

$$\begin{aligned}
d\sigma &= \frac{1}{(2\pi)^2}\delta\left(E_f - E_i\right) \left(\frac{Ze^2\left(E_i + E_f\right)}{|\vec{p}_f - \vec{p}_i|^2}\right)^2 \frac{p_f^2\,dp_f\,d\Omega}{2|\vec{p}_i|\,2E_f} \\
&= \frac{1}{(2\pi)^2}\delta\left(E_f - E_i\right) \left(\frac{Ze^2\left(E_i + E_f\right)}{\underbrace{2p^2\left(1 - \cos\theta\right)}_{4p^2\sin^2\theta/2}}\right)^2 \frac{p\,dp\,d\Omega}{4E}
\end{aligned}$$

now, since $E^2 = m^2 + \vec{p}^2$, use $p\,\mathrm{d}p = E\,\mathrm{d}E$ such that:

$$\frac{p\,\mathrm{d}p\,\mathrm{d}\Omega}{4E}\,\delta\left(E_f - E_i\right) = \frac{\mathrm{d}E\,\delta\left(E_f - E_i\right)\,\mathrm{d}\Omega}{4} = \frac{\mathrm{d}\Omega}{4}\;,$$

---

[1]Note that $E = m_0\gamma$ and $\vec{p} = m_0\gamma\vec{v}$ so that $\vec{v} = \vec{p}/E$.

where we integrated over all E_f in the second step. We thus arrive at the following expression for the differential cross section,

$$\mathrm{d}\sigma = \left(\frac{Ze^2 E}{4\pi p^2 \sin^2 \theta/2} \right)^2 \mathrm{d}\Omega \tag{5.13}$$

or

$$\boxed{\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{Z^2 E^2 e^4}{16\pi^2 p^4 \sin^4 \theta/2} = \frac{Z^2 E^2 \alpha^2}{p^4 \sin^4 \theta/2}} \tag{5.14}$$

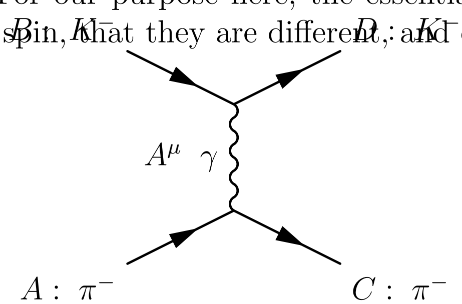
where we defined $\alpha \equiv e^2/4\pi$. In the classical (i.e. non-relativistic) limit we can take $E \to m$ and $E_{kin} = \frac{p^2}{2m}$ such that:

$$\boxed{\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{Z^2 m^2 \alpha^2}{4m^2 E_{kin}^2 \sin^4 \theta/2} = \frac{Z^2 \alpha^2}{4E_{kin}^2 \sin^4 \theta/2}\ ,} \tag{5.15}$$

which is the well-known Rutherford scattering formula.

5.3 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 (e.g. we could consider scattering at large distances). We could equally well consider a process like spinless $e^- \mu^-$ scattering. For our purpose here, the essential restrictions are that the incident particles carry no spin, that they are different, and of equal charge.



We know from the previous calculation 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. 5.10)

$$j^\mu_{BD} = -eN_B N_D^* \ (p^\mu_B + p^\mu_D) \ e^{i(p_D - p_B)x}$$

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.*

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

where we have adopted the Lorentz condition. (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

$$A^\mu = -\frac{1}{q^2} \, j^\mu_{BD} \quad , \tag{5.16}$$

where we defined $q = p_D - p_B$. Note that the latter corresponds to the *four-momentum* transfered by the photon from the kaon to the pion.

$$T_{fi} = -i \int j^\mu_{AC} \, A_\mu \, d^4x = -i \int j^\mu_{AC} \, \frac{-1}{q^2} j^{BD}_\mu \, d^4x = -i \int j^\mu_{AC} \, \frac{-g_{\mu\nu}}{q^2} \, j^\nu_{BD} \, d^4x \tag{5.17}$$

Four-momentum conservation (which appears as a result of the integral) 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.

Also note that 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 real scattering comes from photons that are "off the mass-shell". We call these *virtual* photons.

Inserting the plane wave solutions

$$T_{fi} = -ie^2 \int (N_A N_C^*) \, (p^\mu_A + p^\mu_C) \, e^{i(p_C - p_A)x} \cdot \frac{-1}{q^2} \cdot (N_B N_D^*) \, (p^\mu_B + p^\mu_D) \, e^{i(p_D - p_B)x} \, d^4x$$

and performing the integral over x we obtain

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

where the δ-function that takes care of four-momentum conservation appears. Usually this is written in terms of the *matrix element* $\mathcal{M}$ as

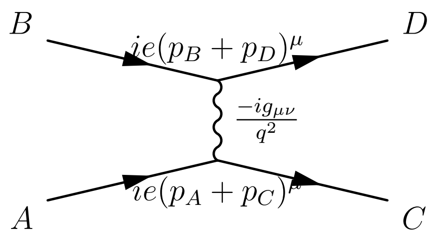
$$T_{fi} \;=\; -i \, N_A N_B N_C^* N_D^* \, (2\pi)^4 \, \delta^4 \left(p_A + p_B - p_C - p_D\right) \cdot \mathcal{M} \tag{5.18}$$

with the matrix element given by

$$-i\mathcal{M} \;=\; \underbrace{ie \, (p_A + p_C)^\mu}_{\text{vertex factor}} \cdot \underbrace{\frac{-ig_{\mu\nu}}{q^2}}_{\text{propagator}} \cdot \underbrace{ie \, (p_B + p_D)^\nu}_{\text{vertex factor}} \; . \tag{5.19}$$

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!



The matrix element $\mathcal{M}$ 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 exchanged photon quantum.

Using Fermi's golden rule we now proceed to calculate the relativistic transition probability:

$$W_{fi} = \lim_{T,V\to\infty} \frac{|T_{fi}|^2}{TV} = \lim_{T,V\to\infty} \frac{1}{TV} \; |N_A N_B N_C^* N_D^*|^2 \; |\mathcal{M}|^2 \; \left|(2\pi)^4 \delta^4 \left(p_A + p_B - p_C - p_D\right)\right|^2$$

Again we use the "trick" :

$$\delta(p) = \lim_{T,V\to\infty} \frac{1}{(2\pi)^4} \int_{-T/2}^{+T/2} dt \int_{-V/2}^{+V/2} d^3x \; e^{ipx}$$

such that

$$\lim_{T,V\to\infty} \frac{1}{TV} \; |\delta^4(p)|^2 = \frac{1}{TV} \; TV \; \delta(p)$$

We get for the transition amplitude:

$$W_{fi} = |N_A N_B N_C N_D|^2 \; |\mathcal{M}|^2 \; (2\pi)^4 \; \delta^4 \left(p_A + p_B - p_C - p_D\right) \tag{5.20}$$

We remember from the previous lecture that the cross-section for the scattering process $A + B \to D + D$ was obtained from

$$\begin{aligned} d\sigma &= \frac{W_{fi}}{\text{Flux}} \, \text{dLips} \\ \text{Flux} &= 4\sqrt{(p_A \cdot p_B)^2 - m_A^2 m_B^2}/V^2 \\ \text{dLips} &= \frac{V}{(2\pi)^3} \frac{d^3 p_C}{2E_C} \; \frac{V}{(2\pi)^3} \frac{d^3 p_D}{2E_D} \end{aligned}$$

The volume V cancels again and we obtain

$$\mathrm{d}\sigma = \frac{(2\pi)^4\,\delta^4\,(p_A + p_B - p_C - p_D)}{4\sqrt{(p_A \cdot p_B)^2 - m_A^2 m_B^2}}\ |\mathcal{M}|^2\ \frac{d^3p_C}{(2\pi)^3\,2E_C}\ \frac{d^3p_D}{(2\pi)^3\,2E_D}\ . \tag{5.21}$$

In the center-of-momentum frame ($\vec{p}_A = -\vec{p}_B$) this expression can be written as (see exercise 18)

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{64\pi^2}\,\frac{1}{s}\,\left|\frac{\vec{p}_f}{\vec{p}_i}\right|\ |\mathcal{M}|^2\ . \tag{5.22}$$

where $s = (p_A + p_B)^2$.

Finally, consider the limit of mass-less particles. Define $\vec{p} \equiv \vec{p}_A$ and $\vec{p\prime} \equiv \vec{p}_C$. In the CM frame the four-vectors are given by

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

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

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

Inserting this in Eq. 5.22 gives (with $\alpha = e^2/4\pi$),

$$\boxed{\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{64\pi^2}\,\frac{1}{s}\,\left(e^2\right)^2\left(\frac{3 + \cos\theta}{1 - \cos\theta}\right)^2 = \frac{\alpha^2}{4s}\left(\frac{3 + \cos\theta}{1 - \cos\theta}\right)^2} \tag{5.23}$$

This is the QED cross section for spinless scattering.

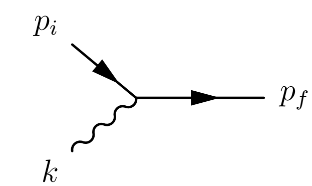
5.4 Particles and Anti-Particles

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 Matrix element has the form of:

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

Let us examine four cases:

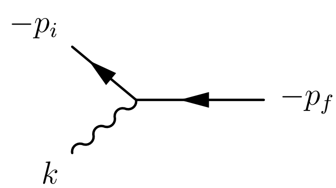
- Scattering of an electron and a photon:



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

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

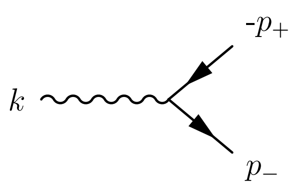
- Scattering of a positron and a photon:



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

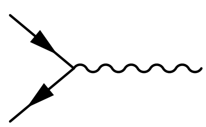
$$\begin{aligned}\mathcal{M} &\propto \int \left(e^{-i(-p_i)x}\right)^* e^{-ikx}\, e^{-i(-p_f)x}\, dx \\ &= \int e^{-i\left(p_i - p_f + k\right)x}\, dx \\ &= (2\pi)^4\, \delta\left(E_i + \omega - E_f\right)\, \delta^3\left(\vec{p_i} + \vec{k} - \vec{p_f}\right)\end{aligned}$$

- Electron positron pair production:



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

- Electron positron annihilation:



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

<u>**Exercise 20 (see also Griffiths, ex.6.6)**</u>
Decay rate of $\pi^0 \to \gamma\gamma$:

(a) Write down the expression for the total decay rate 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) For the Matrix element we have: $\mathcal{M} \sim f_\pi \, e^2$, where for the decay constant we insert $f_\pi = m_\pi$.

(i) Where does the factor e^2 come from?

(ii) What do you think is the meaning of the factor f_π? Describe it qualitatively.

(d) The π^0 is actually a $u\bar{u} + d\bar{d}$ wave with 3 colour degrees of freedom.

(i) Give the expression for the decay rate.

(ii) Calculate the decay rate expressed in GeV.

(iii) Convert the rate into seconds using the conversion table of the introduction lecture.

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

Lecture 6

The Dirac Equation

Introduction

It is sometimes said that Schrödinger discovered the Klein-Gordon equation before the equation carrying his own name, but that he rejected it because it was quadratic in $\partial/\partial t$. In Lecture 2 we have seen how the Klein-Gordon equation leads to solutions with negative energy and negative 'probability density'.

In 1928 in an attempt to avoid this problem Dirac tried to develop a relativistic wave equation that is linear in $\partial/\partial t$. Lorentz invariance requires such an wave equation to be also linear in in $\vec{\nabla}$. What Dirac found, to his own surprise, was an equation that describes particles with spin $\frac{1}{2}$, just what was needed for electrons. We now know that all fundamental fermions should be described by Dirac's wave equation. At the same time Dirac also predicted the existence of anti-particles, an idea that was not taken seriously until 1932, when Anderson discovered the positron.

6.1 Dirac Equation

In order to appreciate what Dirac discovered we follow (a modern interpretation of) Dirac's approach that led to a linear wave equation. (For a different approach, which may be closer to what Dirac actually did, see Griffiths, §7.1.) Consider the usual form of the Schrödinger equation,

$$i\frac{\partial}{\partial t}\psi \;=\; H\psi \;. \tag{6.1}$$

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 [1]:

$$H \; = \; (\vec{\alpha} \cdot \vec{p} + \beta m) \tag{6.2}$$

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 with momentum p of H that

$$H^2 \psi = \left(\vec{p}^2 + m^2 \right) \; \psi \tag{6.3}$$

where $\vec{p}^2 + m^2$ is the eigenvalue. What should H look like such that these eigenvectors exist? Squaring Dirac's ansatz for the Hamiltonian gives

$$\begin{aligned} H^2 &= \left(\sum_i \alpha_i p_i + \beta m \right) \left(\sum_j \alpha_j p_j + \beta m \right) \\ &= \left(\sum_{i,j} \alpha_i \alpha_j p_i p_j + \sum_i \alpha_i \beta p_i m + \sum_i \beta \alpha_i p_i m + \beta^2 m^2 \right) \\ &= \left(\sum_i \alpha_i^2 p_i^2 + \sum_{i>j} (\alpha_i \alpha_j + \alpha_j \alpha_i) p_i p_j + \sum_i (\alpha_i \beta + \beta \alpha_i) p_i m + \beta^2 m^2 \right) \end{aligned} \tag{6.4}$$

where we on purpose did not impose that the coefficients (α_i, β) commute. In fact, comparing to equation 6.3 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

$$\{A, B\} \; = \; AB + BA. \tag{6.5}$$

we can also write these requirements as

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij} \qquad \{\alpha_i, \beta\} = 0 \qquad \beta^2 = 1 \; . \tag{6.6}$$

Immediately we conclude that the α_i and β cannot be ordinary numbers. At this point Dirac has a brilliant idea: what if we represent them as matrices that act on a wave function that is a column vector? As we require the Hamiltonian to be hermitian (such that its eigenvalues are real), the matrices α_i and β must be hermitian as well,

$$\alpha_i^\dagger = \alpha_i \qquad \text{and} \qquad \beta^\dagger = \beta \; . \tag{6.7}$$

[1] We take $\vec{\alpha} \cdot \vec{p} = \alpha_x p_x + \alpha_y p_y + \alpha_z p_z$.

Furthermore, we can show using just the anti-commutation relations and normalization above that they all have eigenvalues ± 1 and zero trace. It then also follows that they must have even dimension.

It can be shown that the lowest dimensional matrices that have the desired behaviour are 4×4 matrices. (See exercise 21 below and also Aitchison (1972) §8.1). The choice of the matrices α_i and β is not unique. Here we choose the Dirac-Pauli representations,

$$\vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{6.8}$$

where $\vec{\sigma}$ are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{6.9}$$

Note that the Pauli matrices are also traceless and hermitian and that they satisfy

$$\sigma_i \sigma_j = \epsilon_{ijk} \sigma_k \tag{6.10}$$

which implies that they anticommute as well.

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,

$$\vec{\alpha} = \begin{pmatrix} -\vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} . \tag{6.11}$$

Exercise 21.

(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 $b = s + \text{it}$ 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 $\vec{\alpha}$ and β cannot be 2×2 matrices?

(b) *Show that the $\vec{\alpha}$ and β matrices in both the Dirac-Pauli as well as in the Weyl representation have the required anti-commutation behaviour.*

6.2 Covariant form of the Dirac Equation

With Dirac's Hamiltonian and the substitution $\vec{p} = -i\hbar\vec{\nabla}$ we arrive at the following relativistic Schrödinger-like wave equation,

$$i\frac{\partial}{\partial t}\psi = \left(-i\vec{\alpha}\cdot\vec{\nabla} + \beta m\right)\,\psi\,. \tag{6.12}$$

Multiplying on the left by β and using $\beta^2 =$ we can write this equation as

$$\left(i\beta\frac{\partial}{\partial t}\psi + i\beta\vec{\alpha}\cdot\vec{\nabla} - m\right)\psi = 0\,. \tag{6.13}$$

We now define the four Dirac γ-matrices by

$$\gamma^\mu \equiv (\beta, \beta\vec{\alpha}) \tag{6.14}$$

The wave equation then takes the simple form

$$\boxed{(i\gamma^\mu\partial_\mu - m)\ \psi\ =\ 0} \tag{6.15}$$

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 in the next lecture 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,

$$\text{for each j=1,2,3,4} \quad : \quad \sum_{k=1}^{4}\left[\sum_{\mu=0}^{3} i\left(\gamma^\mu\right)_{jk}\ \partial_\mu - m\delta_{jk}\right](\psi_k) = 0$$

$$\text{or} \quad : \quad \left[i\underbrace{\begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}}_{\gamma^\mu}\cdot\partial_\mu - \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\cdot m\right]\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix}$$

or even more explicit

$$\left[\begin{pmatrix} 1\!\!1 & 0 \\ 0 & -1\!\!1 \end{pmatrix}\frac{i\partial}{\partial t} + \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}\frac{i\partial}{\partial x} + \begin{pmatrix} 0 & \sigma_2 \\ -\sigma_2 & 0 \end{pmatrix}\frac{i\partial}{\partial y} + \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix}\frac{i\partial}{\partial z} - \begin{pmatrix} 1\!\!1 & 0 \\ 0 & 1\!\!1 \end{pmatrix} m\right]\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix} = \begin{pmatrix}0\\0\\0\\0\end{pmatrix}$$

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

$$\not{a} \; = \; \gamma_\mu a^\mu \; . \tag{6.16}$$

The wave equation for spin-$\frac{1}{2}$ particles can then be written very concisely as

$$(i \not{\partial} - m)\, \psi = 0 \; . \tag{6.17}$$

Note that, 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 may still depend on the frame. The behaviour of Dirac spinors under Lorentz transformations is not entirely trivial. See also Griffiths §7.3 and Halzen and Martin §5.6.

6.3 The Dirac Algebra

From the definitions of $\vec{\alpha}$ and β we can derive the following relation for the anti-commutator of two γ-matrices

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu \; = \; 2\, g^{\mu\nu}\, \mathbb{1}_4 \tag{6.18}$$

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

$$\left(\gamma^0\right)^2 = \mathbb{1} \quad ; \quad \left(\gamma^1\right)^2 = \left(\gamma^2\right)^2 = \left(\gamma^3\right)^2 = -\mathbb{1} \tag{6.19}$$

Also we have the Hermitian conjugates:

$$\gamma^{0\dagger} \; = \; \gamma^0 \tag{6.20}$$

$$\gamma^{i\dagger} \; = \; \left(\beta \alpha^i\right)^\dagger = \alpha^{i\dagger} \beta^\dagger = \alpha^i \beta = -\gamma^i \tag{6.21}$$

Using the anticommutator relation once more then gives

$$\boxed{\gamma^{\mu\dagger} = \gamma^0 \gamma^\mu \gamma^0} \tag{6.22}$$

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

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \tag{6.23}$$

which has the characteristics

$$\gamma^{5\dagger} = \gamma^5 \qquad \left(\gamma^5\right)^2 = \mathbb{1} \qquad \left\{\gamma^5, \gamma^\mu\right\} = 0 \tag{6.24}$$

6.4 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

$$i\gamma^0 \frac{\partial \psi}{\partial t} + i\gamma^k \frac{\partial \psi}{\partial x^k} - m\psi = 0 \qquad k = 1,2,3 \tag{6.25}$$

As we now work with matrices, we use hermitian conjugates rather than complex conjugates and find for the conjugate equation

$$-i\frac{\partial \psi^\dagger}{\partial t}\gamma^0 - i\frac{\partial \psi^\dagger}{\partial x^k}\left(-\gamma^k\right) - m\psi^\dagger = 0 \tag{6.26}$$

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 then define the *adjoint Dirac spinor*

$$\boxed{\overline{\psi} = \psi^\dagger \gamma^0} \; . \tag{6.27}$$

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

$$-i\frac{\partial \overline{\psi}}{\partial t}\gamma^0 - i\frac{\partial \overline{\psi}}{\partial x^k}\gamma^k - m\overline{\psi} = 0 \qquad k = 1,2,3 \tag{6.28}$$

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

$$\boxed{j^\mu \;=\; \overline{\psi}\gamma^\mu \psi} \tag{6.29}$$

then this current satisfies a continuity equation, $\partial_\mu j^\mu = 0$. The first component of this current is simply

$$j^0 = \overline{\psi}\gamma^0\psi = \psi^\dagger\psi = \sum_{i=1}^{4} |\psi_i|^2 \tag{6.30}$$

which is always positive. This property was the original motivation of Dirac's work. Note that the current density can also be written as

$$\vec{j} = \psi^\dagger \vec{\alpha} \psi. \tag{6.31}$$

The form Eq. 6.29 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.

6.5 Bilinear covariants

The Dirac probability current in Eq. 6.29 is an example of a so-called *bilinear covariant*: a quantity that is a product of components of ϕ^* and ϕ 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 ϕ^* and ϕ each have four components, we have 16 independent combinations. Requiring the currents to be covariant, then leads to the following types of currents:

		# of components
scalar	$\overline{\psi}\,\psi$	1
vector	$\overline{\psi}\,\gamma^\mu\,\psi$	4
tensor	$\overline{\psi}\,\sigma^{\mu\nu}\,\psi$	6
axial vector	$\overline{\psi}\,\gamma^5\,\gamma^\mu\,\psi$	4
pseudo scalar	$\overline{\psi}\,\gamma^5\,\psi$	1

(6.32)

where the (anti-symmetric) tensor is defined as

$$\sigma^{\mu\nu} \equiv \frac{i}{2}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right) \tag{6.33}$$

The names 'axial' and 'pseudo' refer to the behaviour of these objects under the parity transformation, $\vec{x} \to -\vec{x}$. 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 9.

6.6 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 $-e$,

$$j^\mu_{\rm em} \;=\; -e\,\overline{\psi}\,\gamma^\mu\,\psi\;. \tag{6.34}$$

Using the ansatz $\psi = u(p)e^{-ipx}$ (motivated in the next lecture) the interaction current density 4-vector takes the form

$$j^\mu_{fi} \;=\; -e\;\left(\quad \overline{u}_f \quad\right)\left(\quad \gamma^\mu \quad\right)\left(\; u_i \;\right)\; e^{i(p_f - p_i)x} \tag{6.35}$$

In the next lecture we will see that although the probability density of the Dirac fields is now positive, the negative energy solutions just remain. Following the Feynman-Stückelberg interpretation the solution with negative energy is again seen as the antiparticle 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 antiparticle ($j^\mu = 2|N|^2 p^\mu$) 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-$\frac{1}{2}$ antiparticle.

This additional minus sign between particles and antiparticles 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[2] 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.

Exercise 22. Traces and products of γ matrices.

Use the anticommutator relation for Dirac γ-matrices in Eq. 6.18 (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) $\gamma_\mu\;\not{a}\,\gamma^\mu = -2\;\not{a}$

[2]See Aitchison & Hey, 3rd edition §7.2

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)* $\mathrm{Tr}\,(\textit{odd number of } \gamma\textit{-matrices}) = 0$

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 excercise for "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.*

Lecture 7

Solutions of the Dirac Equation

The notes of this Lecture follow very closely sections 5.2 - 5.5 of Halzen & Martin. The same material is covered in Griffiths section 7.4

7.1 Solutions for plane waves with $\vec{p} = 0$

We look for free particle solutions of the Dirac equation, which was given by

$$\boxed{(i\gamma^\mu \partial_\mu - m)\ \psi = 0} \ . \tag{7.1}$$

Assuming the particle has non-zero mass, we could first look for the solution of a wave with zero momentum, $\vec{p} = 0$. A *quick* way to get wave solutions is to realize that zero momentum this implies $-i\vec{\nabla}\ \psi = 0$, or that the wave function ψ has no explicit space dependence. In that case the Dirac equation reduces to $i\gamma^0 \frac{\partial \psi}{\partial t} = m\psi$, or written in the Dirac-Pauli representation,

$$\begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix} \begin{pmatrix} \partial\psi_A/\partial t \\ \partial\psi_B\ \partial t \end{pmatrix} = -i\, m \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix} \tag{7.2}$$

For the solution we find that

$$\psi = \begin{pmatrix} \psi_A \\ \psi_B \end{pmatrix} = \begin{pmatrix} e^{-imt}\ \psi_A(0) \\ e^{+imt}\ \psi_B(0) \end{pmatrix} \ . \tag{7.3}$$

Note that ψ_A represents a two-component spinor with positive energy and ψ_B a two-component spinor with negative energy. In the following, however, we will follow the standard textbook method to derive the Dirac solutions.

In exercise 23 you will show that each of the components of the Dirac wave satisfies the Klein-Gordon equation. Therefore, we try the plane wave solutions

$$\psi(x) = u(p)\, e^{-ipx} \tag{7.4}$$

where $u(p)$ is a 4-component column-vector that does not depend on x. After substitution in the Dirac equation we find what is called the Dirac equation in the momentum representation,

$$\boxed{(\gamma^\mu p_\mu - m)\ u(p)\ =\ 0} \tag{7.5}$$

which, using the 'slash notation' can also be written as

$$(\not p - m)\ u(p)\ =\ 0\ . \tag{7.6}$$

In momentum-space the coupled differential equations reduce to a set of coupled linear equations, *e.g.* in the Pauli-Dirac representation,

$$\left[\begin{pmatrix} 1\!\!1 & 0 \\ 0 & -1\!\!1 \end{pmatrix} E - \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} p^i - \begin{pmatrix} 1\!\!1 & 0 \\ 0 & 1\!\!1 \end{pmatrix} m\right] \begin{pmatrix} u_A \\ u_B \end{pmatrix} = 0 \tag{7.7}$$

We can rewrite this as a set of two equations

$$\begin{aligned} (\vec\sigma\cdot\vec p)\ u_B &= (E-m)\, u_A \\ (\vec\sigma\cdot\vec p)\ u_A &= (E+m)\, u_B\ , \end{aligned} \tag{7.8}$$

where u_A and u_B are (still) two-component spinors and $\vec\sigma \equiv (\sigma_1, \sigma_2, \sigma_3)$. Now consider again particles at rest ($p = 0$). In this case, the two equations decouple,

$$\begin{aligned} E\, u_A &= m\, u_A \\ E\, u_B &= -m\, u_B\ . \end{aligned} \tag{7.9}$$

There are four independent solutions, which we write as

$$u^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad , \quad u^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad u^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad , \quad u^{(4)} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{7.10}$$

The first two have positive energy eigenvalue $E = m$ and the second two a negative energy $E = -m$.

7.2 Solutions for moving particles $\vec p \neq 0$

To extend the solution to particles with non-zero momentum, we first define

$$\chi^{(1)}\ \equiv\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad \chi^{(2)}\ \equiv\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{7.11}$$

Now consider two Dirac spinors for which the two upper coordinates $u_A(p)$ of $u(p)$ are given by χ, *i.e.* ($s \in \{1, 2\}$)

$$u_A^{(1)}\ =\ \chi^{(1)} \qquad \text{and} \qquad u_A^{(2)}\ =\ \chi^{(2)}\ . \tag{7.12}$$

Substituting this into the second equation of 7.8 gives for the lower two components

$$u_B^{(1,2)} = \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\, u_A^{(1,2)} \;=\; \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\, \chi^{(1,2)} \; . \tag{7.13}$$

To prove that these are indeed solutions of the equations, one can use the identity

$$(\vec{\sigma}\cdot\vec{p})\,(\vec{\sigma}\cdot\vec{p}) \;=\; |\vec{p}|^2 \; \mathbb{1}_2 \tag{7.14}$$

such that $u_A^{(1,2)}$ and $u_B^{(1,2)}$ also satisfy the first equation in 7.8. See also exercise 24.

Furthermore, to prove that these are indeed a *positive* energy solutions, it is illuminating to operate with the Hamiltonian on the states $u^{(1,2)}$. In the Pauli-Dirac representation the Hamiltonian is given by

$$H \;=\; \vec{\alpha}\cdot\vec{p} + \beta m \;=\; \begin{pmatrix} 0 & \vec{\sigma}\cdot\vec{p} \\ \vec{\sigma}\cdot\vec{p} & 0 \end{pmatrix} + \begin{pmatrix} m\mathbb{1}_2 & 0 \\ 0 & -m\mathbb{1}_2 \end{pmatrix} \tag{7.15}$$

With a bit of algebra we obtain for our solution

$$Hu^{(1)} \;=\; \begin{pmatrix} \left[m + \frac{p^2}{E+m}\right] & u_A^{(1)} \\ E & u_B^{(1)} \end{pmatrix} , \tag{7.16}$$

which illustrates two things: In order that $u^{(1)}$ be a solution we need indeed that $E^2 = p^2 + m^2$. 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

$$\boxed{u^{(1)} = N \begin{pmatrix} \chi^{(1)} \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi^{(1)} \end{pmatrix} \qquad \text{and} \qquad u^{(2)} = N \begin{pmatrix} \chi^{(2)} \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi^{(2)} \end{pmatrix}} \tag{7.17}$$

We shall discuss the normalization constant N later.

In an exactly analogous manner, we can start for our $E < 0$ solutions with the *lower* component given by $\chi^{(s)}$,

$$u_B^{(3)} = \chi^{(1)} \qquad , \qquad u_B^{(4)} = \chi^{(2)} \tag{7.18}$$

which using the first of the equations in Eq. 7.8 gives for the upper coordinates

$$u_A^{(3,4)} \;=\; \frac{\vec{\sigma}\cdot\vec{p}}{E-m}\, u_B^{(3,4)} \;=\; -\frac{\vec{\sigma}\cdot\vec{p}}{(-E)+m}\, \chi^{(1,2)} \tag{7.19}$$

Note the difference in the enumerator: it has become $(E - m)$ rather than $(E + m)$. Evaluating the energy eigenvalue, we now find *e.g.*

$$Hu^{(3)} \;=\; \begin{pmatrix} E & u_A^{(3)} \\ \left[-m + \frac{p^2}{E-m}\right] & u_B^{(3)} \end{pmatrix} , \tag{7.20}$$

which again requires $E^2 = p^2 + m^2$ and in the limit $p \to 0$ gives $E = -m$, a negative energy solution. Consequently, two negative-energy orthogonal solutions are given by

$$\boxed{u^{(3)} = N \begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{E-m}\chi^{(1)} \\ \chi^{(1)} \end{pmatrix} \qquad \text{and} \qquad u^{(2)} = N \begin{pmatrix} \frac{\vec{\sigma}\cdot\vec{p}}{E-m}\chi^{(2)} \\ \chi^{(2)} \end{pmatrix}} \tag{7.21}$$

To gain slightly more insight, let's write them out in momentum components. Use the definition of the Pauli matrices we have

$$\vec{\sigma}\cdot\vec{p} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} p_x + \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} p_y + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} p_z \tag{7.22}$$

to find

$$(\vec{\sigma}\cdot\vec{p})\; u_A^{(1)} = \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} p_z \\ p_x + ip_y \end{pmatrix} \tag{7.23}$$

and similar for $u_A^{(2)}$, $u_B^{(3)}$, $u_B^{(4)}$. The solutions can then be written as

$$E>0 \text{ spinors} \qquad u^{(1)}(p) = N \begin{pmatrix} 1 \\ 0 \\ \frac{p_z}{E+m} \\ \frac{p_x+ip_y}{E+m} \end{pmatrix}, \qquad u^{(2)}(p) = N \begin{pmatrix} 0 \\ 1 \\ \frac{p_x-ip_y}{E+m} \\ \frac{-p_z}{E+m} \end{pmatrix}$$

$$E<0 \text{ spinors} \qquad u^{(3)}(p) = N \begin{pmatrix} \frac{-p_z}{-E+m} \\ \frac{-p_x-ip_y}{-E+m} \\ 1 \\ 0 \end{pmatrix}, \qquad u^{(4)}(p) = N \begin{pmatrix} \frac{-p_x+ip_y}{-E+m} \\ \frac{p_z}{-E+m} \\ 0 \\ 1 \end{pmatrix}$$

We shall discuss the normalization N later. We can verify that the $u^{(1)}$ - $u^{(4)}$ solutions are indeed orthogonal, *i.e.* that

$$u^{(i)\dagger}u^{(j)} = 0 \qquad \text{for } i \neq j. \tag{7.24}$$

7.3 Helicity

The Dirac spinors for a given momentum p have a two-fold degeneracy. This implies that there must be an additional observable that commutes with H and p and the eigenvalues of which distinguish between the degenerate states. Could the extra quantum number be spin? So, eg.: $u^{(1)}$ = spin "up", and $u^{(2)}$ =spin "down"?

The spin operator is defined as

$$\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}. \tag{7.25}$$

In exercise 26 you will show that it does not commute with the Hamiltonian in Eq. 7.15. 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,

$$\vec{\Sigma}\, u^{(i)} = s\, u^{(i)} \ \ ?$$

where s is the spin eigenvalue. Now insert one of the solutions, for example $u^{(1)}$,

$$\begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} p_z/\left(E+m\right) \\ \left(p_x+ip_y\right)/\left(E+m\right) \end{pmatrix} \end{pmatrix} \overset{?}{=} s \begin{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \begin{pmatrix} p_z/\left(E+m\right) \\ \left(p_x+ip_y\right)/\left(E+m\right) \end{pmatrix} \end{pmatrix}$$

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

$$\vec{L} = \vec{r} \times \vec{p} \tag{7.26}$$

You will also show in exercise 26 that the total angular momentum

$$\vec{J} \ = \ \vec{L} + \tfrac{1}{2}\vec{\Sigma} \tag{7.27}$$

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

$$\lambda = \frac{1}{2}\vec{\Sigma}\cdot\hat{p} \equiv \frac{1}{2}\begin{pmatrix} \vec{\sigma}\cdot\hat{p} & 0 \\ 0 & \vec{\sigma}\cdot\hat{p} \end{pmatrix} \tag{7.28}$$

where $\hat{p} \equiv \vec{p}/|\vec{p}|$. We could interpret the helicity as the "spin component in the direction of movement". One can verify that indeed λ commutes with the Hamiltonian in 7.15.

The fact that λ and H commutes means that 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 $\vec{p} = (0, 0, p)$. Applying the helicity operator on $u^{(i)}$ gives

$$\begin{aligned} \frac{1}{2}\left(\vec{\sigma}\cdot\hat{p}\right)\ u_A &= \frac{1}{2}\sigma_3 u_A = \pm\frac{1}{2}u_A \\ \frac{1}{2}\left(\vec{\sigma}\cdot\hat{p}\right)\ u_B &= \frac{1}{2}\sigma_3 u_B = \pm\frac{1}{2}u_B \end{aligned}$$

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.

7.4 Antiparticle 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 antiparticle (the positron). We define the antiparticle components of the wave function as

$$\begin{aligned} v^{(1)}(p) &\equiv u^{(4)}(-p) \\ v^{(2)}(p) &\equiv -u^{(3)}(-p) \;. \end{aligned} \tag{7.29}$$

The minus sign in $u^{(3)}$ is chosen such that the charge conjugation transformation implies $u^{(1)} \to v^{(1)}$ and $u^{(2)} \to v^{(2)}$. We will discuss that below.

Using this definition we can replace the two negative energy solutions by the following anti-particle spinors with positive energy, $E = +\sqrt{\vec{p}^2 + m^2}$,

$$v^{(1)}(p) \;=\; N \begin{pmatrix} \frac{p_x - i p_y}{E+m} \\ \frac{-p_z}{E+m} \\ 0 \\ 1 \end{pmatrix} \qquad , \qquad v^{(2)}(p) \;=\; N \begin{pmatrix} \frac{-p_z}{E+m} \\ \frac{-(p_x + i p_y)}{E+m} \\ -1 \\ 0 \end{pmatrix} .$$

The spinors $u(p)$ of matter waves are solutions of the Dirac equation in momentum space, Eq. 7.5. Replacing p with $-p$ in the Dirac equation we find that our positive energy anti-particle spinors satisfy another Dirac equation,

$$\boxed{(\not{p} + m)\; v(p) = 0} \tag{7.30}$$

7.5 Normalization of the wave function

As for the Klein-Gordon case we choose a normalization such that there are $2E$ particles per unit volume. Remember that we had in the previous lecture for the first component of the current of the Dirac wave

$$\rho(x) \;=\; \psi^\dagger(x)\, \psi(x) \;. \tag{7.31}$$

Substituting the plane wave solution $\psi = u(p)\, e^{-ipx}$, and integrating over a volume V we find

$$\int_V \rho \, \mathrm{d}^3 x \;=\; \int_V u^\dagger(p)\, e^{ipx}\, u(p)\, e^{-ipx}\, \mathrm{d}^3 x \;=\; u^\dagger(p)\, u(p)\, V \tag{7.32}$$

Consequently, to find $2E$ particles per unit volume we must normalize such that

$$u^\dagger(p)\, u(p) \;=\; 2E \tag{7.33}$$

Explicit calculation for the positive energy solutions ($s \in \{1,2\}$) gives

$$\begin{aligned} u^{(s)\dagger}\, u^{(s)} &= N^2 \left(\chi^{(s)T} \chi^{(s)} + \chi^{(s)T} \frac{(\vec{\sigma}\cdot\vec{p})^\dagger\, (\vec{\sigma}\cdot\vec{p})}{(E+m)^2} \chi^{(s)} \right) \\ &= N^2 \left(1 + \frac{\vec{p}^2}{(E+m)^2} \right) \;=\; N^2 \frac{2E}{E+m} \end{aligned}$$

Consequently, in order to have $2E$ particles per unit volume we choose

$$N \;=\; \sqrt{E+m}\,. \tag{7.34}$$

The computation for the positive energy antiparticle waves $v(p)$ leads to the same normalization. We can now write the orthogonality relations as (with $r, s \in \{1,2\}$)

$$\begin{aligned} u^{(r)\dagger}\, u^{(s)} &= 2E\, \delta_{rs} \\ v^{(r)\dagger}\, v^{(s)} &= 2E\, \delta_{rs} \end{aligned} \tag{7.35}$$

7.6 The completeness relation

We now consider the Dirac equation for the adjoint spinor $\overline{u}$, $\overline{v}$. Taking the hermitian conjugate of Eq. 7.5 and multiplying on the right by γ^0 we have

$$u^\dagger \gamma^{\mu\dagger} \gamma^0 p_\mu - u^\dagger \gamma^0 m = 0 \tag{7.36}$$

Using that $\gamma^{\mu\dagger}\gamma^0 = \gamma^0\gamma^\mu$ we then find for the Dirac equation of the adjoint spinor $\overline{u} = u^\dagger\gamma^0$,

$$\boxed{\overline{u}\ (\not p - m) = 0} \tag{7.37}$$

In the same manner we find for the adjoint antiparticle spinors

$$\boxed{\overline{v}\ (\not p + m) = 0} \tag{7.38}$$

Using these results you will derive in exercise 25 the so-called *completeness relations*

$$\boxed{\begin{aligned} \sum_{s=1,2} u^{(s)}(p)\, \overline{u}^{(s)}(p) &= (\not p + m) \\ \sum_{s=1,2} v^{(s)}(p)\, \overline{v}^{(s)}(p) &= (\not p - m) \end{aligned}} \tag{7.39}$$

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

$$\begin{pmatrix} \cdot \\ \cdot \\ \cdot \\ \cdot \end{pmatrix} \cdot (....) = \begin{pmatrix} & & \\ & \gamma^\mu & \\ & & \end{pmatrix} \cdot p_\mu + \begin{pmatrix} & & \\ & \mathbb{1} & \\ & & \end{pmatrix} \cdot m$$

(**Note**: $\sum_{s=3,4} u^{(s)}(p)\,\overline{u}^{(s)}(p) = \sum_{s=1,2} v^{(s)}(-p)\,\overline{v}^{(s)}(-p) = -(\not{p} + m)$)

7.7 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:

$$[\gamma^\mu (i\partial_\mu + eA_\mu) - m]\,\psi = 0\ . \tag{7.40}$$

Similarly, there must be a Dirac equation describing the positron ($q = +e$):

$$[\gamma^\mu (i\partial_\mu - eA_\mu) - m]\,\psi_C = 0\ , \tag{7.41}$$

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,

$$[-\gamma^{\mu *} (i\partial_\mu - eA_\mu) - m]\,\psi^* = 0\ . \tag{7.42}$$

Now suppose that there is a matrix M such that

$$\gamma^{\mu *} \;=\; M^{-1}\gamma^\mu M\ . \tag{7.43}$$

We can then rewrite the equation above as

$$M^{-1}\,[\gamma^\mu (i\partial_\mu - eA_\mu) - m]\,M\psi^* = 0\ . \tag{7.44}$$

and we obtain the relation

$$\psi_C = M\psi^* = M\gamma^0\overline{\psi}^T \equiv C\overline{\psi}^T \tag{7.45}$$

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

$$M = C\gamma^0 = i\gamma^2 = \begin{pmatrix} & & & 1 \\ & & -1 & \\ & -1 & & \\ 1 & & & \end{pmatrix}\ . \tag{7.46}$$

Exercise 23.

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* $\gamma^\nu \partial_\nu$ *or (easier) with* $(i\gamma^\nu \partial_\nu + m)$.

Hint: Use the anti-commutation relation of the γ-matrices.

Exercise 24.

(a) *Prove the identity in Eq. 7.14,* i.e.

$$(\vec{\sigma} \cdot \vec{p})\,(\vec{\sigma} \cdot \vec{p}) \;=\; |\vec{p}|^2 \; 1\!\!1_2$$

Hint: either use the explicit form of the Pauli matrices in Eq. 6.9, or exploit that they anti-commute, that is $\{\sigma_i, \sigma_j\} = 0$ for $i \neq j$.

(b) *Show explicitly that a solution to the Dirac equation satisfies the relativistic relation between energy and momentum by substituting*

$$u_B = \frac{\vec{\sigma} \cdot \vec{p}}{E + m} u_A \qquad \textit{into} \qquad u_A = \frac{\vec{\sigma} \cdot \vec{p}}{E - m} u_B$$

and using the identify derived in (a).

Exercise 25. (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. 7.17 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

$$\not{p} = \begin{pmatrix} E\,\mathbb{1}_2 & -\vec{\sigma}\cdot\vec{p} \\ \vec{\sigma}\cdot\vec{p} & -E\,\mathbb{1}_2 \end{pmatrix} \tag{7.47}$$

Exercise 26. (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 ($\vec{S}$) in addition to their orbital angular momentum ($\vec{L}$). We will see that $\vec{L}$ and $\vec{S} = \vec{\Sigma}/2$ are not conserved individually but that their sum *is.*

(a) *Consider the Hamiltonian that leads to the Dirac equation,*

$$H \; = \; \vec{\alpha} \cdot \vec{p} + \beta m$$

Use the fundamental commutator $[x_i, p_j] = i\hbar\delta_{ij}$ *(with* $\hbar = 1$*) to show that*

$$[H, \vec{L}] = -i \, \vec{\alpha} \times \vec{p} \tag{7.48}$$

where $\vec{L} = \vec{x} \times \vec{p}$.
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 spin operator $\vec{\Sigma}$ *(see also Eq. 7.25) and* $\vec{\alpha}$ *in the Pauli-Dirac representation were*

$$\vec{\Sigma} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix} \qquad \textit{and} \qquad \vec{\alpha} \; = \; \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix}$$

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*

$$\left[H, \vec{\Sigma} \right] = 2 \, i \, \vec{\alpha} \times \vec{p} \tag{7.49}$$

We see from (a) and (c) that the Hamiltonian commutes with $J = L + \frac{1}{2}\Sigma$.

Exercise 27. (Exercise 5.5 of H & M)

(a) *Use the equation*

$$(\vec{\sigma} \cdot \vec{p})\, u_A = (E + m)\, u_B \tag{7.50}$$

to show that, for a non-relativistic electron with velocity β*,* u_B *is a factor* $\frac{1}{2}\beta$ *smaller than* u_A*. In a non-relativistic description* ψ_A *and* ψ_B *are often called respectively the "large" and "small" components of the electron wave function.*

(b) *Show that the Dirac equation for an electron with charge* $-e$ *in the non-relativistic limit in an electromagnetic field* $A^\mu = (A^0, \mathbf{A})$ *reduces to the Schrödinger-Pauli equation*

$$\left(\frac{1}{2m} \left(\vec{p} + e\vec{A}\right)^2 + \frac{e}{2m}\vec{\sigma} \cdot B - eA^0 \right) \psi_A - E_{NR}\psi_A \,, \tag{7.51}$$

where the magnetic field $\vec{B} = \vec{\nabla} \times \vec{A}$*, and the non-relativistic energy* $E_{NR} = E - m$*. Assume that* $|eA^0| \ll m$*.*
Do this by substituting $p^\mu + eA^\mu$ *for* p^μ *in Eq. 7.50 and solve the equations for* ψ_A*. Use:*

$$\vec{p} \times \vec{A} + \vec{A} \times \vec{p} = -i\vec{\nabla} \times \vec{A} \,, \tag{7.52}$$

where $\vec{p} = -i\vec{\nabla}$*.*

The term with eA^0 in 7.51 is a constant potential energy that is of no further importance. The term with $\vec{B}$ arises due to the fact that $\vec{p}$ and $\vec{A}$ don't commute. In this term we recognise the magnetic field:

$$-\vec{\mu} \cdot \vec{B} = -g\frac{e}{2m}\vec{S} \cdot \vec{B} \,. \tag{7.53}$$

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 ($\vec{S} = \frac{1}{2}\vec{\sigma}$) one finds $g = 2$. The current value of $(g-2)/2$ is according to the Particle Data Book

$$(g-2)/2 = 0.001159652193 \pm 0.000000000010 \tag{7.54}$$

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 8

Spin-1/2 Electrodynamics

8.1 Feynman rules for fermion scattering

With the spinor solutions of the Dirac equation we finally have the tools to calculate cross section for fermions (spin-1/2 particles). Analogously to the case of spin 0 particles (K.G.-waves) we determine the solutions of the Dirac equations in the presence of a perturbation potential. So we work with the free spin-1/2 solutions $\psi = u(p)\, e^{-ipx}$ that satisfy the free Dirac equation: $(\gamma_\mu p^\mu - m)\,\psi = 0$.

In order to introduce an electromagnetic perturbation we make again the substitution for a particle with $q = -e$: $p^\mu \to p^\mu + eA^\mu$. The Dirac equation for an electron then becomes:

$$(\gamma_\mu p^\mu - m)\,\psi + e\gamma_\mu A^\mu \psi = 0 \tag{8.1}$$

Again, we would like to have a kind of Schrödinger equation, ie. an equation of the type:

$$(H_0\, +\, V)\,\psi \;=\; E\,\psi \tag{8.2}$$

In order to get to this form, we multiply Eq. 8.1 from the left by γ^0,

$$\begin{aligned}
\to\quad & \left(\gamma^0\gamma_\mu p^\mu - \gamma^0 m\right)\psi + e\gamma^0\gamma_\mu A^\mu \psi = 0 \\
\to\quad & \left(E - \gamma^0\gamma^k p^k - \gamma^0 m\right)\psi = -e\gamma^0\gamma_\mu A^\mu \psi \\
\to\quad & E\psi = \underbrace{\left(\gamma^0\gamma^k p^k + \gamma^0 m\right)}_{H_0 = \vec{\alpha}\cdot\vec{p} + \beta m}\psi - \underbrace{e\gamma^0\gamma_\mu A^\mu}_{V}\,\psi
\end{aligned}$$

Consequently, the perturbation potential is given by

$$V(x) \;=\; -e\gamma^0\gamma_\mu A^\mu \tag{8.3}$$

In analogy to spinless scattering we now write for the transition amplitude

$$T_{fi} = -i\int\; \psi_f^\dagger(x)\, V(x)\, \psi_i(x)\, \mathrm{d}^4x \tag{8.4}$$

Note the differences with the case of the KG solutions in spinless scattering: The wave function has four components and the perturbation $V(x)$ is now a (4×4) matrix. We take a hermitian conjugate of the wave ψ, rather than just its complex conjugate. The transition amplitude is still just a scalar.

Substituting the expression for $V(x)$ we obtain

$$\begin{aligned} T_{fi} &= -i \int \psi_f^\dagger(x) \; \left(-e\gamma^0\gamma_\mu A^\mu(x)\right) \; \psi_i(x) \mathrm{d}^4x \\ &= -i \int \overline{\psi}_f(x) \; (-e) \, \gamma_\mu \psi_i(x) A^\mu(x) \, \mathrm{d}^4x \end{aligned} \tag{8.5}$$

In lecture 6 we defined the charge current density of the Dirac wave as

$$j^\mu(x) \;=\; -e\, \overline{\psi}(x)\, \gamma^\mu \, \psi(x)$$

In complete analogy to the spinless particle case we define the electromagnetic *transition current* between states i and $f$as

$$j^\mu_{fi}(x) \;=\; -e\, \overline{\psi}_f(x)\, \gamma^\mu \, \psi_i(x) \; , \tag{8.6}$$

such that the transition amplitude can be written as

$$T_{fi} \;=\; -i \int j^{fi}_\mu \; A^\mu \; \mathrm{d}^4x \; . \tag{8.7}$$

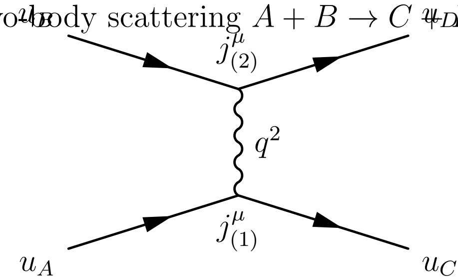
Using the plane wave decomposition $\psi(x) \;=\; u(p)e^{-ipx}$ we can still rewrite the transition current as

$$j^\mu_{fi} \;=\; -e\overline{u}_f \, \gamma^\mu \, u_i \, e^{i\left(p_f - p_i\right)x} \; . \tag{8.8}$$

Remember that the current is a 'scalar' in Dirac spinor space, or schematically,

$$j^\mu_{fi} = \left(\quad \overline{u}_f \quad \right) \left(\quad \gamma^\mu \quad \right) \left(\; u_i \; \right)$$

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 2') as a source. That is, we take

$$\Box A^\mu \;=\; j^\mu_{(2)} \; .$$

and obtain for the potential

$$A^\mu = -\frac{1}{q^2} j^\mu_{(2)} \, .$$

where $q \equiv p_i - p_f$ is the momentum transfer. The transition amplitude is then again:

$$T_{fi} \;=\; -i \int j^{(1)}_\mu \, \frac{-1}{q^2} \, j^\mu_{(2)} \, \mathrm{d}^4 x \;=\; -i \int j^\mu_{(1)} \, \frac{-g_{\mu\nu}}{q^2} \, j^\nu_{(2)} \, \mathrm{d}^4 x \tag{8.9}$$

which is symmetric in terms of particle (1) and (2). Inserting the expressions of the plane wave currents using Eq. 8.8 we obtain

$$T_{fi} = -i \int -e \overline{u}_C \gamma^\mu u_A \, e^{i(p_C - p_A)x} \cdot \frac{-g_{\mu\nu}}{q^2} \cdot -e \overline{u}_D \gamma^\nu u_B \, e^{i(p_D - p_B)x} \, d^4 x \tag{8.10}$$

Performing the integral (and realizing that nothing depends on x except the exponentials) leads us to the expression

$$T_{fi} \;=\; -i \, (2\pi)^4 \; \delta^4 \left(p_D + p_C - p_B - p_A\right) \cdot \mathcal{M} \tag{8.11}$$

with the matrix element given by

$$-i\mathcal{M} \;=\; \underbrace{ie \left(\overline{u}_C \gamma^\mu u_A\right)}_{\text{vertex}} \cdot \underbrace{\frac{-i g_{\mu\nu}}{q^2}}_{\text{propagator}} \cdot \underbrace{ie \left(\overline{u}_D \gamma^\nu u_B\right)}_{\text{vertex}} \tag{8.12}$$

From the matrix element we can now read of 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. Graphically, the difference between the spinless and spin-$\frac{1}{2}$ case is this:

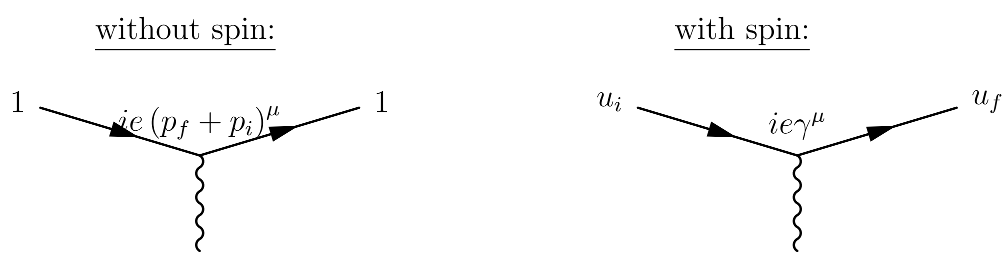


Figure 8.1: Vertex factors for *left:* spinless particles, *right:* spin 1/2 particles.

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. As you will prove in exercise 28, we can rewrite the Dirac current as

$$\overline{u}_f \gamma^\mu u_i = \frac{1}{2m} \overline{u}_f \left[\left(p_f + p_i\right)^\mu + i\sigma^{\mu\nu} \left(p_f - p_i\right)_\nu \right] u_i \tag{8.13}$$

where the tensor is defined as

$$\sigma^{\mu\nu} = \frac{i}{2}\left(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu\right) \ . \tag{8.14}$$

(We have seen this tensor before in lecture 6, when we discussed bilinear covariants.) This formulation of the current is called the 'Gordon decomposition'. We observe that in addition to the contribution that also 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.

8.2 Electron-muon scattering

We proceed to use the Feynman rules to calculate the cross section of the process $e^-\mu^- \to e^-\mu^-$.The Feynman diagram is drawn in Fig. 8.2.

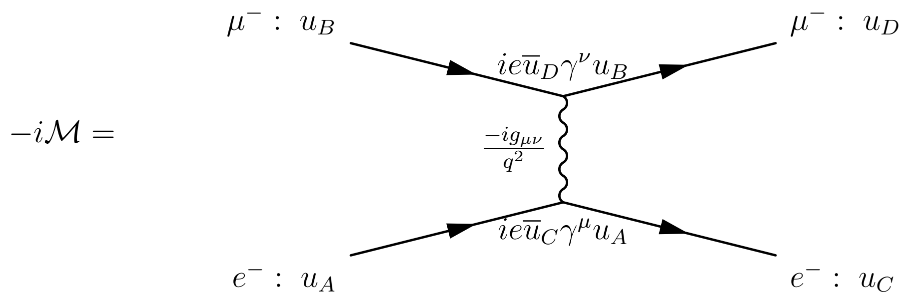


Figure 8.2: Lowest order Feynman diagram for $e^-\mu^-$ scattering.

Using the diagram we find for the lowest-order amplitude

$$-i\mathcal{M} \ = \ -e^2\,\overline{u}_C\gamma^\mu u_A\,\frac{-i}{q^2}\,\overline{u}_D\gamma_\mu u_B \tag{8.15}$$

and for its square

$$|\mathcal{M}|^2 \ = \ e^4\left[(\overline{u}_C\gamma^\mu u_A)\ \frac{1}{q^2}\ (\overline{u}_D\gamma_\mu u_B)\right]\left[(\overline{u}_C\gamma^\nu u_A)\ \frac{1}{q^2}\ (\overline{u}_D\gamma_\nu u_B)\right]^* \tag{8.16}$$

Note that 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

$$|\mathcal{M}|^2 \ = \ \frac{e^4}{q^4}\sum_{\mu\nu}\left[(\overline{u}_C\gamma^\mu u_A)\ (\overline{u}_C\gamma^\nu u_A)^*\right]\ \left[(\overline{u}_D\gamma_\mu u_B)\ (\overline{u}_D\gamma_\nu u_B)^*\right] \tag{8.17}$$

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

$$|\mathcal{M}|^2 \rightarrow \overline{|\mathcal{M}|^2} = \frac{1}{(2s_A+1)\,(2s_B+1)} \sum_{\text{spin}} |\mathcal{M}|^2 \tag{8.18}$$

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 cannot be interference between states with different helicity in the final state. 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

$$\overline{|\mathcal{M}|^2} \;=\; \frac{1}{4}\,\frac{e^4}{q^4}\, L^{\mu\nu}_{\text{electron}}\, L_{\mu\nu}^{\text{muon}} \tag{8.19}$$

where the unpolarized lepton tensors are defined as

$$\begin{aligned} L^{\mu\nu}_{\text{electron}} &= \sum_{e-\text{spin}} \left[\overline{u}_C \gamma^\mu u_A\right] \left[\overline{u}_C \gamma^\nu u_A\right]^* \\ L^{\mu\nu}_{\text{muon}} &= \sum_{\mu-\text{spin}} \left[\overline{u}_D \gamma^\mu u_B\right] \left[\overline{u}_D \gamma^\nu u_B\right]^* \;. \end{aligned} \tag{8.20}$$

The spin summation is in fact 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

$$\begin{aligned} \left[\overline{u}_C \gamma^\nu u_A\right]^* &= \left[\overline{u}_C \gamma^\nu u_A\right]^\dagger \\ &= \left[u_C^\dagger \gamma^0 \gamma^\nu u_A\right]^\dagger = \left[u_A^\dagger \gamma^{\nu\dagger} \gamma^0 u_C\right] \\ &= \left[\overline{u}_A \gamma^0 \gamma^{\nu\dagger} \gamma^0 u_C\right] = \left[\overline{u}_A \gamma^\nu u_C\right] \end{aligned} \tag{8.21}$$

In other words, by reversing the order of the spinors, we can get rid of the complex conjugation and find

$$L_{\mathrm{e}}^{\mu\nu} = \sum_{e \text{ spin}} (\overline{u}_C \gamma^\mu u_A)\ (\overline{u}_A \gamma^\nu u_C)$$

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',

$$L_{\mathrm{e}}^{\mu\nu} = \sum_{s'} \sum_{s} \sum_{klmn} \overline{u}_{C,k}^{(s')}\, \gamma_{kl}^{\mu}\, u_{A,l}^{(s)}\, \overline{u}_{A,m}^{(s)}\, \gamma_{mn}^{\nu}\, u_{C,n}^{(s')} \tag{8.22}$$

Note that all of the factors on the right are just complex numbers, so we can manipulate their order and write this as

$$L_{\mathrm{e}}^{\mu\nu} = \sum_{klmn} \sum_{s'} u_{C,n}^{(s')}\, \overline{u}_{C,k}^{(s')}\, \gamma_{kl}^{\mu} \sum_{s} u_{A,l}^{(s)}\, \overline{u}_{A,m}^{(s)}\, \gamma_{mn}^{\nu} \tag{8.23}$$

Now remember the completeness relation that we derived in the previous lecture[1],

$$\sum_s u^{(s)}\, \overline{u}^{(s)} = \not{p} + m$$

This is exactly what we need to sum over the spin states! Substituting this expression for the spin sums gives

$$L_{\mathrm{e}}^{\mu\nu} = \sum_{klmn} (\not{p}_C + m_{\mathrm{e}})_{nk}\ \gamma_{kl}^{\mu}\ (\not{p}_A + m_{\mathrm{e}})_{lm}\ \gamma_{mn}^{\nu} \tag{8.24}$$

where m_{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

$$L_{\mathrm{e}}^{\mu\nu} = \mathrm{Tr}\left[(\not{p}_C + m)\ \gamma^\mu\ (\not{p}_A + m)\ \gamma^\nu\right] \tag{8.25}$$

You will now realize why we made you compute the traces of products of γ-matrices in lecture 6. We briefly repeat here the properties that we need:

- In general, for matrices A, B and C and any complex number z
  - $Tr(zA) = z\, Tr(A)$

[1] for anti-fermions this gives an overall "−" sign in the tensor: $L_{\mathrm{e}}^{\mu\nu} \to -L_{\mathrm{\overline{e}}}^{\mu\nu}$ for each particle → anti-particle.

– $\mathrm{Tr}\,(A+B) \;=\; \mathrm{Tr}(A) + \mathrm{Tr}(B)$
– $\mathrm{Tr}\,(ABC) \;=\; \mathrm{Tr}\,(CAB) \;=\; \mathrm{Tr}\,(BCA)$

- For γ-matrices (from the anti-commutator $\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$):
  – Tr(odd number of γ-matrices $= 0$)
  – $\mathrm{Tr}\,(\gamma^\mu\gamma^\nu) = 4\,g^{\mu\nu}$
  – $\mathrm{Tr}(\gamma^\alpha\gamma^\beta\gamma^\mu\gamma^\nu) \;=\; 4\left(g^{\alpha\beta}g^{\mu\nu} - g^{\alpha\mu}g^{\beta\nu} + g^{\alpha\nu}g^{\beta\mu}\right)$

Using the first rule we can write out the tensor as a sum of traces,

$$\begin{aligned} L_{\mathrm{e}}^{\mu\nu} &= \mathrm{Tr}\left[(\not{p}_C + m)\;\gamma^\mu\;(\not{p}_A + m)\;\gamma^\nu\right] \\ &= \underbrace{\mathrm{Tr}\left[\not{p}_C\,\gamma^\mu\;\not{p}_A\,\gamma^\nu\right]}_{\text{case 1}} + \underbrace{\mathrm{Tr}\left[m\gamma^\mu m\gamma^\nu\right]}_{\text{case 2}} + \underbrace{\mathrm{Tr}\left[\not{p}_C\,\gamma^\mu m\gamma^\nu\right]}_{3\gamma' s\Rightarrow 0} + \underbrace{\mathrm{Tr}\left[m\gamma^\mu\;\not{p}_A\,\gamma^\nu\right]}_{3\gamma' s\Rightarrow 0} \end{aligned} \tag{8.26}$$

The last two terms vanish because they contain an odd number of γ-matrices. For the second term ('case 2') we find

$$\mathrm{Tr}\left[m\gamma^\mu m\gamma^\nu\right] = m^2\,\mathrm{Tr}\left[\gamma^\mu\gamma^\nu\right] \;=\; 4\,m^2\,g^{\mu\nu}\;. \tag{8.27}$$

Finally, for the first term ('case 1') we have

$$\begin{aligned} \mathrm{Tr}\left[\not{p}_C\,\gamma^\mu\;\not{p}_A\,\gamma^\nu\right] &\equiv \mathrm{Tr}\left[\gamma^\alpha p_{C,\alpha}\gamma^\mu\gamma^\beta p_{A,\beta}\gamma^\nu\right] \\ &= Tr\left[\gamma^\alpha\gamma^\mu\gamma^\beta\gamma^\nu\right]\;p_{C,\alpha}\,p_{A,\beta} \\ &= 4\;\left(g^{\alpha\mu}g^{\beta\nu} - g^{\alpha\beta}g^{\mu\nu} + g^{\alpha\nu}g^{\beta\mu}\right)\;p_{C,\alpha}\,p_{A,\beta} \\ &= 4\left(p_C^\mu p_A^\nu + p_C^\nu p_A^\mu - g^{\mu\nu}(p_A\cdot p_C)\right), \end{aligned} \tag{8.28}$$

where we used the trace formula for four γ-matrices in the third step. Adding the two contributions gives for the lepton tensor

$$L_{\mathrm{e}}^{\mu\nu} \;=\; 4\;\left[p_C^\mu p_A^\nu \;+\; p_C^\nu p_A^\mu \;+\; \left(m_{\mathrm{e}}^2 - p_C\cdot p_A\right)g^{\mu\nu}\right] \tag{8.29}$$

The expression for the muon tensor is obtained with the substitution $(p_A, p_C, m_{\mathrm{e}}) \to (p_B, p_D, m_{\mu})$,

$$L_{\mu}^{\mu\nu} \;=\; 4\;\left[p_D^\mu p_B^\nu \;+\; p_D^\nu p_B^\mu + : \left(m_{\mu}^2 - p_D\cdot p_B\right)g^{\mu\nu}\right] \tag{8.30}$$

To compute the contraction of the two tensors, which appears in the amplitude, we just apply brute force:

$$\begin{aligned} L_{\mathrm{e}}^{\mu\nu}\,L^{\mu}_{\mu\nu} \;&=\; 4\left[p_C^\mu p_A^\nu + p_C^\nu p_A^\mu + \left(m_{\mathrm{e}}^2 - p_C\cdot p_A\right)\;g^{\mu\nu}\right]\cdot 4\;\left[p_{D\mu}p_{B\nu} + p_{D\nu}p_{B\mu} + \left(m_{\mu}^2 - p_D\cdot p_B\right)\;g_{\mu\nu}\right] \\ &=\; 16\left[(p_C\cdot p_D)\,(p_A\cdot p_B) + (p_C\cdot p_B)\,(p_A\cdot p_D) - (p_C\cdot p_A)\,(p_D\cdot p_B) + (p_C\cdot p_A)\,m_{\mu}^2\right. \\ &\quad + (p_C\cdot p_B)\,(p_A\cdot p_D) + (p_C\cdot p_D)\,(p_A\cdot p_B) - (p_C\cdot p_A)\,(p_D\cdot p_B) + (p_C\cdot p_A)\,m_{\mu}^2 \\ &\quad - (p_C\cdot p_A)\,(p_D\cdot p_B) - (p_C\cdot p_A)\,(p_D\cdot p_B) + 4\,(p_C\cdot p_A)\,(p_D\cdot p_B) - 4\,(p_C\cdot p_A)\,m_{\mu}^2 \\ &\quad \left. + m_e^2\;(p_D\cdot p_B) + m_e^2\;(p_D\cdot p_B) - 4m_e^2\;(p_D\cdot p_B) + 4m_e^2 m_{\mu}^2\right] \\ &=\; 32\left[(p_A\cdot p_B)\,(p_C\cdot p_D) + (p_A\cdot p_D)\,(p_C\cdot p_B) - m_e^2\,(p_D\cdot p_B) - m_{\mu}^2\,(p_A\cdot p_C) + 2m_e^2 m_{\mu}^2\right] \end{aligned}$$

Combining everything we obtain for the square of the unpolarized amplitude for electron-muon scattering

$$\begin{aligned}\overline{|\mathcal{M}|^2} &= 8\,\frac{e^4}{q^4}\,\Big[\,(p_C\cdot p_D)\,(p_A\cdot p_B)+ \\ &\qquad (p_C\cdot p_B)\,(p_A\cdot p_D)-m_e^2\,(p_D\cdot p_B)-m_\mu^2\,(p_A\cdot p_C)+2m_e^2m_\mu^2\,\Big]\end{aligned} \tag{8.31}$$

We now consider the ultra-relativistic limit which allows us to ignore the rest masses of the particles. The amplitude squared then becomes

$$\overline{|\mathcal{M}|^2} \simeq 8\,\frac{e^4}{q^4}\,\Big[\,(p_C\cdot p_D)\,(p_A\cdot p_B)+(p_C\cdot p_B)\,(p_A\cdot p_D)\,\Big] \tag{8.32}$$

Furthermore, we define the Mandelstam variables

$$\begin{aligned} s &\equiv (p_A+p_B)^2 = p_A^2+p_B^2+2\,(p_A\cdot p_B) && \simeq 2\,(p_A\cdot p_B) \\ t &\equiv (p_D-p_B)^2 \equiv q^2 && \simeq -2\,(p_D\cdot p_B) \\ u &\equiv (p_A-p_D)^2 && \simeq -2\,(p_A\cdot p_D)\end{aligned} \tag{8.33}$$

where the approximation on the right again follows in the ultra-relativistic limit. From energy-momentum conservation ($p_A^\mu+p_B^\mu=p_C^\mu+p_D^\mu$) we have

$$\begin{aligned} (p_A+p_B)^2 &= (p_C+p_D)^2 & & & p_A\cdot p_B &= p_C\cdot p_D \\ (p_D-p_B)^2 &= (p_C-p_A)^2 & &\Longrightarrow & p_D\cdot p_B &= p_C\cdot p_A \\ (p_A-p_D)^2 &= (p_B-p_C)^2 & & & p_A\cdot p_D &= p_B\cdot p_C \end{aligned} \tag{8.34}$$

which gives

$$(p_A\cdot p_B)\ (p_C\cdot p_D) = \frac{1}{2}s\,\frac{1}{2}s=\frac{1}{4}s^2 \tag{8.35}$$

$$(p_A\cdot p_D)\ (p_C\cdot p_B) = \left(-\frac{1}{2}u\right)\left(-\frac{1}{2}u\right)=\frac{1}{4}u^2 \tag{8.36}$$

$$\tag{8.37}$$

Inserting this in the amplitude, we find

$$\boxed{\overline{|\mathcal{M}|^2} \simeq 2\,e^4\left(\frac{s^2+u^2}{t^2}\right)} \tag{8.38}$$

Finally, as we did for the spinless scattering in Lecture 5, consider again the scattering process in the centre-of-momentum system. The four-vectors can then be written as

$$\begin{aligned} p_A^\mu &= (|\vec{p}_A|,\vec{p}_A) & \qquad p_B^\mu &= (|\vec{p}_A|,-\vec{p}_A) \\ p_C^\mu &= (|\vec{p}_C|,\vec{p}_C) & \qquad p_D^\mu &= (|\vec{p}_C|,-\vec{p}_C)\end{aligned}$$

Define $p \equiv |\vec{p}_A|$ which, by four-vector conservation is also equal to $|\vec{p}_A|$. Define θ as the angle between $\vec{p}_A$ and $\vec{p}_C$ (see Fig. 8.3), which implies that

$$\vec{p}_A \cdot \vec{p}_C \;=\; \vec{p}_B \cdot \vec{p}_D \;=\; p^2 \, \cos\theta$$

We then find for the Mandelstam variables

$$\begin{aligned} s =& \; 4\,p^2 \\ t =& \; -2\,p^2 \;(1 - \cos\theta) \\ u =& \; -2\,p^2 \;(1 + \cos\theta) \end{aligned} \tag{8.39}$$

which gives for the amplitude squared

$$\overline{|\mathcal{M}|}^2 \;\simeq\; 8\,e^4\,\frac{4 + (1+\cos\theta)^2}{(1-\cos\theta)^2}. \tag{8.40}$$

Inserting this in the expression for the differential cross-section (which we obtained after integrating over the final state momenta in Lecture 5) we find

$$\left.\frac{d\sigma}{d\Omega}\right|_{\text{c.m.}} \;=\; \frac{1}{64\pi^2}\,\frac{1}{s}\,\overline{|\mathcal{M}|^2} \;\simeq\; \frac{\alpha^2}{2s}\,\frac{4 + (1+\cos\theta)^2}{(1-\cos\theta)^2} \tag{8.41}$$

with $\alpha \equiv e^2/4\pi$.

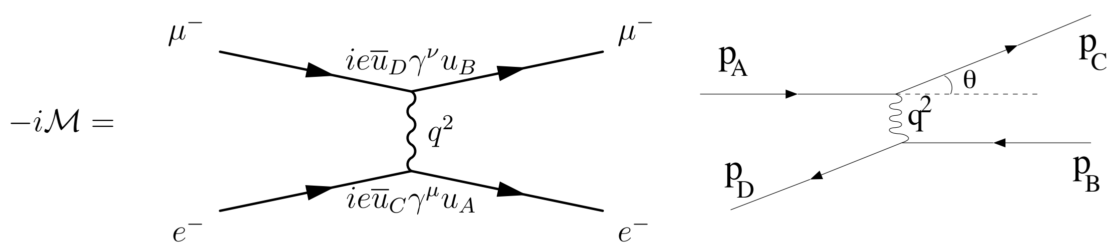


Figure 8.3: e⁻μ⁻ → e⁻μ⁻ scattering. Left: the Feynman diagram. Right: definition of scattering angle in C.M. frame.

8.3 Crossing: the process e⁻e⁺ → μ⁻μ⁺

We will now illustrate the method of "crossing" to obtain the amplitude for e⁻e⁺ → μ⁻μ⁺ scattering from the amplitude of e⁻μ⁻ → e⁻μ⁻ scattering. The method is illustrated in Fig. 8.4, and comes down to the following:

- assuming that you had computed the original amplitude in terms of particles, replace $p \to -p$ for every anti-particle in the diagram

- now relabel the momenta such that the ingoing and outgoing lines correspond to those in the original diagram
- 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 at the end of Lecture 6.)

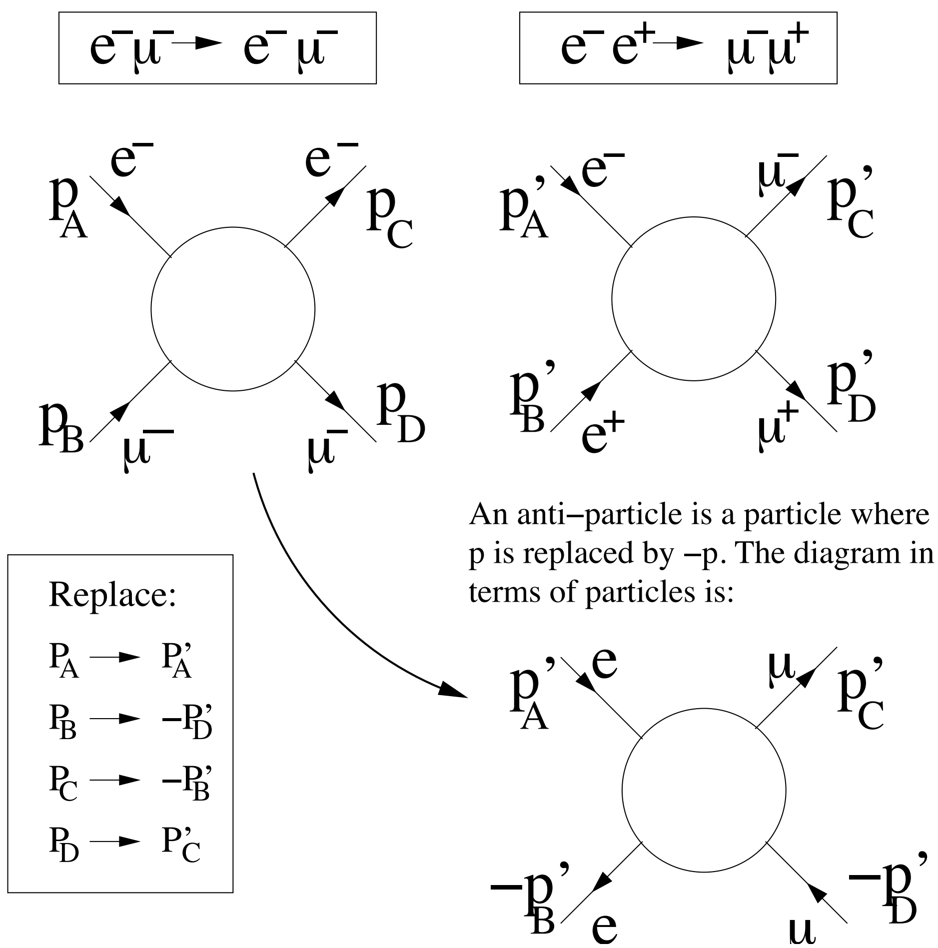


Figure 8.4: The principle of crossing. Use the anti-particle interpretation of a particle with the 4-momentum reversed in order to related the Matrix element of the "crossed" reaction to the original one.

Following these rules, we have

$$\overline{|\mathcal{M}|}^2_{\mathrm{e^-e^+\to\mu^-\mu^+}}(p_\mathrm{e}^-, p_\mathrm{e}^+, p_\mu^-, p_\mu^+) \;=\; \overline{|\mathcal{M}|}^2_{\mathrm{e^-\mu^-\to e^-\mu^-}}(p_\mathrm{e}^-, -p_\mu^+, -p_\mathrm{e}^+, p_\mu^-) \tag{8.42}$$

Labeling the outgoing momenta with a 'prime' such that $p'_A = p_\mathrm{e}^-$, $p'_B = p_\mathrm{e}^+$, $p'_C = p_\mu^-$, $p'_D = p_\mu^+$, we have

$$p_A = p'_A \qquad p_B = -p'_D \qquad p_C = -p'_B \qquad p_D = p'_C$$

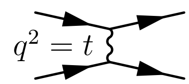
and find for the Mandelstam variables of the original 'particle' diagram

$$\begin{array}{rcccccl} s & \equiv & (p_A + p_B)^2 & = & (p'_A - p'_D)^2 & \equiv & u' \\ t & \equiv & (p_D - p_B)^2 & = & (p'_C + p'_D)^2 & = & s' \\ u & \equiv & (p_A - p_D)^2 & = & (p'_A - p'_C)^2 & = & t' \end{array} \tag{8.43}$$

(To express the results in 'primed' Mandelstam variables we have used that $p'_A + p'_B = p'_C + p'_D$.) Using the result in Eq. **??** the amplitude squared for the two processes are then

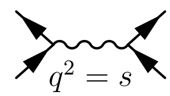
$$\overline{|\mathcal{M}|}^2_{e^-\mu^-\to e^-\mu^-} = 2\,e^4\,\frac{s^2+u^2}{t^2}$$

"t-channel":



$$\overline{|\mathcal{M}|}^2_{e^-e^+\to \mu^-\mu^+} = 2\,e^4\,\frac{u'^2+t'^2}{s'^2}$$

"s-channel":



It is customary to label these a *t-channel* and an *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 8.39.

We immediately get for the matrix element:

$$\overline{|\mathcal{M}|}^2_{\mathrm{c.m.}} = 2\,e^4\,\frac{t'^2+u'^2}{s'^2} \;=\; e^4\,\left(1+\cos^2\theta\right) \tag{8.44}$$

The differential cross-section becomes

$$\boxed{\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4s}\,\left(1+\cos^2\theta\right)} \tag{8.45}$$

Finally, to calculate the total cross section for the process we integrate over the azimuthal angle ϕ and the polar angle θ:

$$\boxed{\sigma\left(e^+e^- \to \mu^+\mu^-\right) = \frac{4\pi}{3}\,\frac{\alpha^2}{s}} \tag{8.46}$$

Note that the 'shape' of the angular distribution does not depend on the kinematics, but that the cross-section scales as $1/s$: it is inversely proportional to the square of the available energy.

Exercise 28.

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

$$\sigma^{\mu\nu} = \frac{i}{2} (\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu)$$

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 to exercise 27.)*

Exercise 29.

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 30. 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* $e^+e^- \to \pi^+\pi^-$

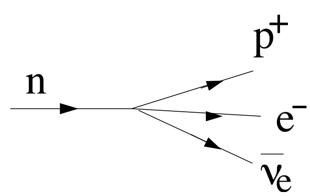
(d) *Determine the differential cross section* $d\sigma/d\Omega$ *for* $e^+e^- \to \pi^+\pi^-$ *in the center-of-momentum of the* e^+e^-*-system.*

Lecture 9

The Weak Interaction

In 1896 Henri Becquerel was studying 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.

We know now that the most nuclear decays of (not very) heavy nuclei in nature are the result of the transition of a neutron to an electron, a proton and an anti-neutrino,



or in a formula,

$$n \to p + e^- + \overline{\nu}_e \ . \tag{9.1}$$

A 'free' neutron has a lifetime of about 15 minutes. However, as you know, the lifetime of various elements spans a very wide range. (The reason is that neutrons are not free particles in the nucleus!)

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,

$$\tau = \frac{1}{\Gamma} . \tag{9.2}$$

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)

$$\Gamma(A \to B + C) = \int \frac{|\mathcal{M}|^2}{2E_A} \mathrm{d}\Phi = \frac{p_B}{8\pi m_A^2} |\mathcal{M}|^2 \tag{9.3}$$

For both spin-less and spin-$\frac{1}{2}$ scattering, 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 has net quark or lepton flavour and cannot decay to lighter states preserving flavour, then it can only decay through the weak interaction. Note that 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 20.

- 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. As it is charged, it cannot decay to photons. Its dominant decay is via the weak interaction to an electron and neutrinos.

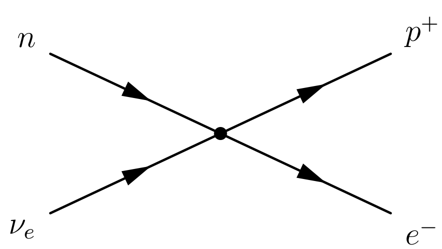
Considerations like these explain the gross features in the hierarchy of lifetimes. However, as you can also judge from the wide range in lifetimes of particles that decay weakly, kinematic effects must be important as well if we want to understand the lifetimes quantitatively.

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.

9.1 The 4-point interaction

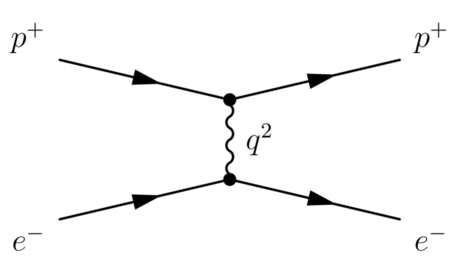
Let's turn back to nuclear decays. Based on the model of electromagnetic interactions Fermi proposed in 1932 the so-called *4-point interaction model*, introducing the Fermi constant as the strength of the interaction: $G_F \approx 1.166 \cdot 10^{-5}\text{GeV}^{-2}$.



The "Feynman diagram" of the 4-point interaction "neutrino scattering on a neutron" has the following matrix element:

$$\boxed{\mathcal{M} = G_F\ (\overline{u}_p \gamma^\mu u_n)\ (\overline{u}_e \gamma_\mu u_\nu)} \tag{9.4}$$

This is to be compared to the electromagnetic diagram for electron proton scattering:



Here the matrix element was:

$$\mathcal{M} = \frac{4\pi\alpha}{q^2}\ (\overline{u}_p \gamma^\mu u_p)\ (\overline{u}_e \gamma_\mu u_e) \tag{9.5}$$

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 rôle 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 in fact just a guess that turned out successful to describe many aspects of β-decay. There was no reason for this choice apart from similarity of quantum electrodynamics. In QED the reason that the interaction has a vector behaviour is the fact that the force mediator, the foton, is a spin-1 (or 'vector') particle.

In the most general case the matrix element of the 4-point interaction can be written as:

$$\mathcal{M} = G_F \left(\overline{\psi}_p \ (4 \times 4) \ \psi_n\right) \left(\overline{\psi}_e \ (4 \times 4) \ \psi_\nu\right) \tag{9.6}$$

where the (4×4) is a matrix. Lorentz invariance puts restrictions on the form of these matrices. In fact we have seen them already in lecture 6: 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 most general 4-point interaction now takes the following form:

$$\mathcal{M} = G_F \sum_{i,j}^{S,P,V,A,T} C_{ij} \left(\overline{u}_p \, O_i \, u_n\right) \ \left(\overline{u}_e \, O_j \, u_\nu\right) \tag{9.7}$$

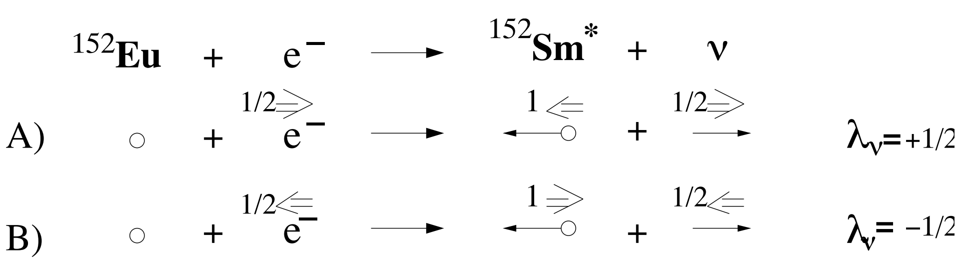
where O_i, O_j are operators of the form S, V, T, A, P. It will not surprise you that the kinematics of a decay depend on the type of operator involved. For example, it can be shown (see eg. Perkins: "Introduction to High Energy Physics", 3$^{\text{rd}}$ edition, appendix D) that for the decay $n \to pe^-\overline{\nu}_e$

- 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.

In 1958 Goldhaber et. al. measured experimentally that the weak interaction is of the type: V, A, (ie. it is <u>not</u> S, P, T). See Perkins ed 3, §7.5 for a full description of the experiment. The basic idea is the following.

Consider the following decay of a particular *Europium* isotope to *Samorium* via a so-called electron capture reaction, $^{152}\mathrm{Eu} + e^- \rightarrow ^{152}\mathrm{Sm}^*(\mathrm{J}=1) + \nu$.



The excited Samorium nucleus is in a $J = 1$ state. To conserve angular momentum, $\vec{J}$ must be parallel to the spin of the electron, but opposite to that of the electron-neutrino. Since in the rest frame of the decaying nucleus, the directions of the decay products are opposite, the polarisation (spin projection on momentum) of the nucleon and the neutrino are opposite as well.

The neutrino in this decay cannot be observed. However, the spin projection of the Samorium nucleus can be probed with the photon that is emitted in its decay to the ground state, $^{152}\mathrm{Sm}^* \rightarrow ^{152}\mathrm{Sm} + \gamma$. Measuring the photon spin is a work-of-art by itself, but assuming it can be done, this allows to distinghuish the two decay topologies above. The measurement by Goldhaber and collaborators showed that only case 'B' actually occurs. In other words: the neutrinos in this decay always have helicity $-\frac{1}{2}$. From this it was concluded that in the weak interaction only the V, A currents are involved and not S, P or T.

9.2 Lorentz covariance and parity

Consider two coordinate frames and a corresponding Lorentz transformation,

$$x'^{\nu} = \Lambda^{\nu}_{\mu} x^{\mu} \; . \tag{9.8}$$

The Dirac equation the two frames is

$$\begin{aligned} i\gamma^\mu \frac{\partial \psi(x)}{\partial x^\mu} - m\psi(x) &= 0 \\ i\gamma^\nu \frac{\partial \psi'(x')}{\partial x'^\nu} - m\psi'(x') &= 0 \end{aligned} \tag{9.9}$$

For the wave function there must exist a relation with an operator S, such that:

$$\psi'(x') = S\psi(x) \tag{9.10}$$

Since the Dirac spinor is of the form $\psi(x) = u(p)\, e^{-ipx}$, where px is a Lorentz invariant, S must be independent of x and only acts on the spinor u. The Dirac equation after the Lorentz transformation becomes

$$i\gamma^\nu \frac{\partial S\,(\psi(x))}{\partial x'^\nu} - mS\,(\psi(x)) = 0 \tag{9.11}$$

and if we act on this equation by S^{-1} from the left:

$$iS^{-1}\gamma^\nu \frac{S\,(\partial\psi(x))}{\partial x'^\nu} - mS^{-1}S\psi(x) = 0 \tag{9.12}$$

Using the Lorentz transformation of the coordinate derivatives

$$\frac{\partial}{\partial x^\mu} = \Lambda^\nu_\mu \frac{\partial}{\partial x'^\nu}\,, \tag{9.13}$$

we find that this equation is consistent with the original Dirac equation if and only if

$$S^{-1}\gamma^\nu S = \Lambda^\nu_\mu \gamma^\mu\,. \tag{9.14}$$

Let us now take a look at the parity operator, which inverts the space coordinates:

$$P : f(t, \vec{x}) \longrightarrow f(t, -\vec{x}) \tag{9.15}$$

Written as a Lorentz transformation, the parity operator is

$$\Lambda^\mu_\nu = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \tag{9.16}$$

Which is the "Dirac" operator that gives: $\psi'(x') = S\psi(x)$? The easiest way to find it is to use the relation $S_p^{-1}\gamma^\mu S_p = \Lambda^\mu_\nu\gamma^\nu = (\gamma^0, -\gamma^1, -\gamma^2, -\gamma^3)$, or, more explicitly, to find the matrix S_p for which:

$$\begin{aligned} S_p^{-1}\gamma^0 S_p &= \gamma^0 \\ S_p^{-1}\gamma^k S_p &= -\gamma^k \end{aligned} \tag{9.17}$$

which has the solution $S_p = \gamma^0$.

Alternatively, we can get the parity operator from the Dirac equation. Assume that the wave function $\psi(\vec{r}, t)$ is a solution of the Dirac equation:

$$\left(\gamma_0 \frac{\partial}{\partial t} + \gamma_k \frac{\partial}{\partial x^k} - m\right) \psi(\vec{r}, t) = 0 \tag{9.18}$$

then, after a parity transformation we find:

$$\left(\gamma_0 \frac{\partial}{\partial t} - \gamma_k \frac{\partial}{\partial x^k} - m\right) \psi(-\vec{r}, t) = 0 \tag{9.19}$$

So, $\psi(-\vec{r}, t)$ is **not** a solution of the Dirac equation due to the additional minus sign! Multiply the Dirac equation of the parity transformed spinor from the left by γ^0, to find:

$$\begin{aligned}
& \gamma_0 \left(\gamma_0 \frac{\partial}{\partial t} - \gamma_k \frac{\partial}{\partial x^k} - m\right) \psi(-\vec{r}, t) = 0 \\
\Rightarrow \quad & \left(\gamma_0 \frac{\partial}{\partial t}\gamma_0 + \gamma_k \frac{\partial}{\partial x^k}\gamma_0 - m\gamma_0\right) \psi(-\vec{r}, t) = 0 \\
\Rightarrow \quad & \left(\gamma_0 \frac{\partial}{\partial t} + \gamma_k \frac{\partial}{\partial x^k} - m\right) \gamma_0 \psi(-\vec{r}, t) = 0
\end{aligned}$$

We conclude that if $\psi(\vec{r}, t)$ is a solution of the Dirac equation, then $\gamma_0 \psi(-\vec{r}, t)$ is also a solution (in the mirror world). In other words: under the parity operation ($S = \gamma^0$): $\psi(\vec{r}, t) \rightarrow \gamma_0 \psi(-\vec{r}, t)$.

An interesting consequence can be derived from the explicit representation of the γ^0 matrix:

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \tag{9.20}$$

The parity operator has opposite sign for the positive and negative solutions. Consequently, fermions and anti-fermions have opposite parity.

What does this mean for the currents in the interactions? Under the parity operation we find

$$\begin{array}{llllll}
S: & \overline{\psi}\psi & \rightarrow & \overline{\psi}\gamma^0\gamma^0\psi & = & \overline{\psi}\psi & \text{Scalar} \\
P: & \overline{\psi}\gamma^5\psi & \rightarrow & \overline{\psi}\gamma^0\gamma^5\gamma^0\psi & = & -\overline{\psi}\gamma^5\psi & \text{Pseudo Scalar} \\
V: & \overline{\psi}\gamma^\mu\psi & \rightarrow & \overline{\psi}\gamma^0\gamma^\mu\gamma^0\psi & = & \left\{ \begin{array}{c} \overline{\psi}\gamma^0\psi \\ -\overline{\psi}\gamma^k\psi \end{array} \right. & \text{Vector} \\
A: & \overline{\psi}\gamma^\mu\gamma^5\psi & \rightarrow & \overline{\psi}\gamma^0\gamma^\mu\gamma^5\gamma^0\psi & = & \left\{ \begin{array}{c} -\overline{\psi}\gamma^0\psi \\ \overline{\psi}\gamma^k\psi \end{array} \right. & \text{Axial Vector.}
\end{array}$$

We had concluded earlier that the weak matrix element in neutron decay is of the form

$$\mathcal{M} = G_F \sum_{i,j}^{V,A} C_{ij} \; (\overline{u}_p O_i u_p) \; (\overline{u}_e O_j u_\nu) \tag{9.21}$$

Note that the vector and axial vector currents have different behaviour under parity. This implies that if a process receives contributions from both currents, then it *must be violating parity.*

9.3 The $V - A$ interaction

The cumulative evidence from many experiments involving weak interactions leads 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

$$J^\mu \;=\; \overline{u}_e \, \gamma^\mu \, \tfrac{1}{2} \left(1 - \gamma^5\right) \; u_\nu \; . \tag{9.22}$$

The current for quarks looks identical, for example for a $u \to d$ transition

$$J^\mu \;=\; \overline{u}_u \, \gamma^\mu \, \tfrac{1}{2} \left(1 - \gamma^5\right) \; u_d \; . \tag{9.23}$$

We call this the "V-A" form. For massless particles (or in the ultrarelativistic limit), the projection operator

$$P_L \;\equiv\; \tfrac{1}{2} \left(1 - \gamma^5\right) \tag{9.24}$$

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

$$\mathcal{M} = \frac{G_F}{\sqrt{2}} \; \left(\overline{u}_p \gamma^\mu \left(C_V - C_A \gamma^5\right) u_n\right) \; \left(\overline{u}_e \gamma_\mu \left(1 - \gamma^5\right) u_\nu\right) \; . \tag{9.25}$$

For neutron decay, the vector and axial vector couplings are $C_V = 1.000 \pm 0.003$, $C_A = 1.260 \pm 0.002$

9.4 The propagator of the weak interaction

The Fermi theory has a 4-point interaction: there is no propagator involved to transmit the interaction from the lepton current to the hadron current. However, we know now that forces are carried by bosons:

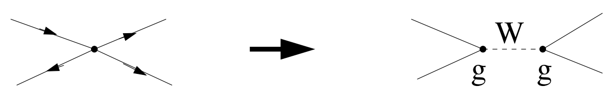
- the electromagnetic interaction is carried by the massless photon which gives rise to a propagator

$$\frac{1}{q^2}$$

- the weak interaction is carried by the massive W, Z bosons, for which we have the propagators:

$$\frac{1}{M_W^2 - q^2} \qquad \text{and} \qquad \frac{1}{M_Z^2 - q^2} \; .$$

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. The charged-current propagator reduces to $1/M_W^2$, which allows us to relate Fermi's coupling constant to the actual coupling constant 'g' of the weak interaction:



strength: $\sim \frac{G_F}{\sqrt{2}}$ $\sim \frac{g^2}{8M_W^2}$

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 electronmagnetic 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.

9.5 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.
Using the Feynman rules we can write for the matrix element:

$$\mathcal{M} = \frac{g}{\sqrt{2}} \left(\underbrace{\overline{u}(k)}_{\text{outgoing } \mu_\nu} \; \gamma^\mu \frac{1}{2}\left(1-\gamma^5\right) \; \underbrace{u(p)}_{\text{incoming } \mu} \right) \; \underbrace{\frac{1}{M_W^2}}_{\text{propagator}} \; \frac{g}{\sqrt{2}} \left(\underbrace{\overline{u}(p')}_{\text{outgoing } e} \; \gamma_\mu \frac{1}{2}\left(1-\gamma^5\right) \; \underbrace{v(k')}_{\text{outgoing } \overline{\nu}_e} \right) \tag{9.26}$$

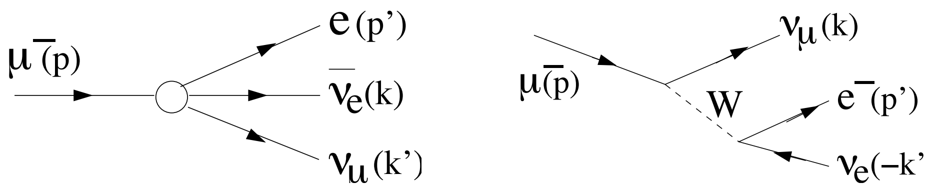


Figure 9.1: Muon decay: *left:* Labelling of the momenta, *right:* Feynman diagram. Note that for the spinor of the outgoing antiparticle we use: $u_{\nu_e}(-k') = v_{\nu_e}(k')$.

Next we square the matrix element and sum over the spin states, exactly similar to the case of $e^+e^- \to \mu^+\mu^-$. Then we use again the tric of Casimir as well as the completeness relations to convert the sum over spins into a trace. The result is:

$$\overline{|\mathcal{M}|}^2 = \frac{1}{2}\sum_{\text{Spin}} |\mathcal{M}|^2 = \frac{1}{2}\left(\frac{g^2}{8M_W^2}\right)^2 \cdot \text{Tr}\left\{\gamma^\mu \left(1-\gamma^5\right) \left(\not{p}' + m_e\right) \gamma^\nu \left(1-\gamma^5\right) \not{k}'\right\}$$
$$\cdot \text{Tr}\left\{\gamma_\mu \left(1-\gamma^5\right) \not{k}\, \gamma_\nu \left(1-\gamma^5\right) \left(\not{p} + m_\mu\right)\right\}$$

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:

$$\boxed{\overline{|\mathcal{M}|}^2 = 64\, G_F^2\ (k \cdot p')\ (k' \cdot p)} \tag{9.27}$$

Intermezzo: Trace theorems used (see also Halzen & Martin p 261):

$$\text{Tr}\left(\gamma^\mu \not{a}\, \gamma^\nu \not{b}\right) \cdot \text{Tr}\left(\gamma_\mu \not{c}\, \gamma_\nu \not{d}\right) = 32\ \left[(a\cdot c)\ (b\cdot d) + (a\cdot d)\ (b\cdot c)\right]$$
$$\text{Tr}\left(\gamma^\mu \not{a}\, \gamma^\nu \gamma^5 \not{b}\right) \cdot \text{Tr}\left(\gamma_\mu \not{c}\, \gamma_\nu \gamma^5 \not{d}\right) = 32\ \left[(a\cdot c)\ (b\cdot d) - (a\cdot d)\ (b\cdot c)\right]$$
$$\text{Tr}\left(\gamma^\mu \left(1-\gamma^5\right) \not{a}\, \gamma^\nu \left(1-\gamma^5\right) \not{b}\right) \cdot \text{Tr}\left(\gamma_\mu \left(1-\gamma^5\right) \not{c}\, \gamma_\nu \left(1-\gamma^5\right) \not{d}\right) = 256\ (a\cdot c)\ (b\cdot d)$$

The decay width we can find by applying Fermi's golden rule:

$$\begin{aligned}
d\Gamma &= \frac{1}{2E}\, \overline{|\mathcal{M}|}^2\, dQ \\
\text{where}:\ dQ &= \frac{d^3p'}{(2\pi)^3\ 2E} \cdot \frac{d^3k}{(2\pi)^3\ 2\omega} \cdot \frac{d^3k'}{(2\pi)^3\ 2\omega'} \cdot (2\pi)^4\, \delta^4\left(p - p' - k' - k\right) \\
\text{with}:E &= \text{muon energy} \\
E' &= \text{electron energy} \\
\omega' &= \text{electron neutrino energy} \\
\omega &= \text{muon neutrino energy}
\end{aligned}$$

First we evaluate the expression for the matrix element. We have the relation ($p = (m_\mu, 0, 0, 0)$):

$$p = p' + k + k' \qquad \text{so}: \qquad (k + p') = (p - k') \tag{9.28}$$

We can also see the following relations to hold:

$$\begin{aligned}(k+p')^2 &= \underbrace{k^2}_{=0} + \underbrace{p'^2}_{m_e^2\approx 0} + 2\,(k\cdot p') \\ (p-k')^2 &= \underbrace{p^2}_{m_\mu^2\equiv m^2} + \underbrace{k'^2}_{=0} - 2\,\underbrace{(p\cdot k')}_{m\omega'}\end{aligned}$$

Therefore we have the relation: $2\,(k\cdot p') = m^2 - 2m\omega'$, which we use to rewrite the matrix element as:

$$\overline{|\mathcal{M}|}^2 = 64\,G_F^2\,(k\cdot p')\,(k'\cdot p) = 32\,G_F^2\,\left(m^2 - 2m\omega'\right)\,m\omega' \tag{9.29}$$

We had the expression for the differential decay width

$$d\Gamma = \frac{1}{2E}\,\overline{|\mathcal{M}|}^2\,dQ = \frac{16G_F^2}{m}\,\left((m^2 - 2m\omega'\right)m\omega'\,dQ \tag{9.30}$$

(E is replaced by m since the decaying muon is in rest). For the total decay width we must integrate over the phase space:

$$\Gamma = \int \frac{1}{2E}\,\overline{|\mathcal{M}|}^2\,dQ = \frac{16G_F^2}{m}\int \left((m^2 - 2m\omega'\right)m\omega'\,dQ \tag{9.31}$$

We note that the integrand only depends on the neutrino energy ω'. So, let us first perform the integral in dQ over the <u>other</u> energies and momenta:

$$\begin{aligned}\int_{\text{other}} dQ &= \frac{1}{8\,(2\pi)^5}\int \delta\left(m - E' - \omega' - \omega\right)\,\delta^3\left(\vec{p}' + \vec{k}' + \vec{k}\right)\,\frac{d^3\vec{p}'}{E'}\,\frac{d^3\vec{k}'}{\omega'}\,\frac{d^3\vec{k}}{\omega} \\ &= \frac{1}{8\,(2\pi)^5}\int \delta\left(m - E' - \omega' - \omega\right)\,\frac{d^3\vec{p}'\,d^3\vec{k}'}{E'\omega'\omega}\end{aligned}$$

since the δ-function gives 1 for the integral over $\vec{k}$.

We also have the relation:

$$\omega = |k| = |\vec{p}' + \vec{k}'| = \sqrt{E'^2 + \omega'^2 + 2E'\omega'\cos\theta} \tag{9.32}$$

where θ is the angle between the electron and the electron neutrino. We choose the z-axis along $\vec{k}'$, the direction of the electron neutrino. From the equation for ω we derive:

$$d\omega = \frac{-2E'\omega'\sin\theta}{2\underbrace{\sqrt{E'^2 + \omega'^2 + 2E'\omega'\cos\theta}}_{\omega}}\,d\theta \qquad \Leftrightarrow \qquad d\theta = \frac{-\omega\,d\omega}{E'\omega'\sin\theta} \tag{9.33}$$

Next we integrate over $d^3\vec{p}' = E'^2 \sin\theta \, dE' \, d\theta \, d\phi$ with $d\theta$ as above:

$$\begin{aligned} dQ &= \frac{1}{8\,(2\pi)^5} \int \delta\,(m - E' - \omega' - \omega)\, \frac{E'^2 \sin\theta}{E'}\, dE' \, d\theta \, d\phi \, \frac{d^3\vec{k}'}{\omega'}\, \frac{1}{\omega} \\ &= \frac{1}{8\,(2\pi)^5}\, 2\pi \int \delta\,(m - E' - \omega' - \omega)\; dE' \, d\omega \, \frac{d^3\vec{k}'}{\omega'^2} \end{aligned}$$

(using the relation: $E' \sin\theta \, d\theta = -\frac{\omega}{\omega'}\, d\omega$).

Since we integrate over ω, the δ-function will cancel:

$$dQ = \frac{1}{8\,(2\pi)^4} \int dE' \, \frac{d^3\vec{k}'}{\omega'^2} \tag{9.34}$$

such that the full expression for Γ becomes:

$$\Gamma = \frac{2G_F^2}{(2\pi)^4} \int \left(m^2 - 2m\omega'\right) \omega' \, dE' \, \frac{d^3\vec{k}'}{\omega'^2} \tag{9.35}$$

Next we do the integral over k' as far as possible with:

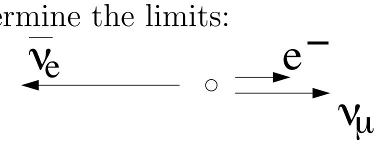
$$\int d^3\vec{k}' = \int \omega'^2 \, \sin\theta' \, d\omega' \, d\theta' \, d\phi' = 4\pi \int \omega'^2 \, d\omega \tag{9.36}$$

so that we get:

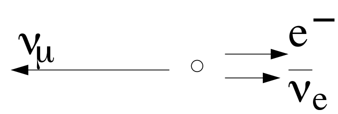
$$\Gamma = \frac{G_F^2 m}{(2\pi)^3} \int (m - 2\omega') \; \omega' \, d\omega' \, dE' \tag{9.37}$$

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

$$\frac{d\Gamma}{dE'} = \frac{G_F^2 m}{(2\pi)^3} \int_{\frac{1}{2}m - E'}^{\frac{1}{2}m} (m - 2\omega') \; \omega' \, d\omega' = \frac{G_F^2 m^2}{12\pi^3}\, E'^2 \left(3 - 4\frac{E'}{m}\right) \tag{9.38}$$

which can be measured experimentally. Finally, integrating the expression over the electron energy we find for the total decay width of the muon

$$\boxed{\Gamma \equiv \frac{1}{\tau} = \frac{G_F^2 \, m^5}{192\,\pi^3}} \tag{9.39}$$

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 $G_F = (1.16639 \pm 0.00002) \cdot 10^{-5}\mathrm{GeV}^{-2}$.

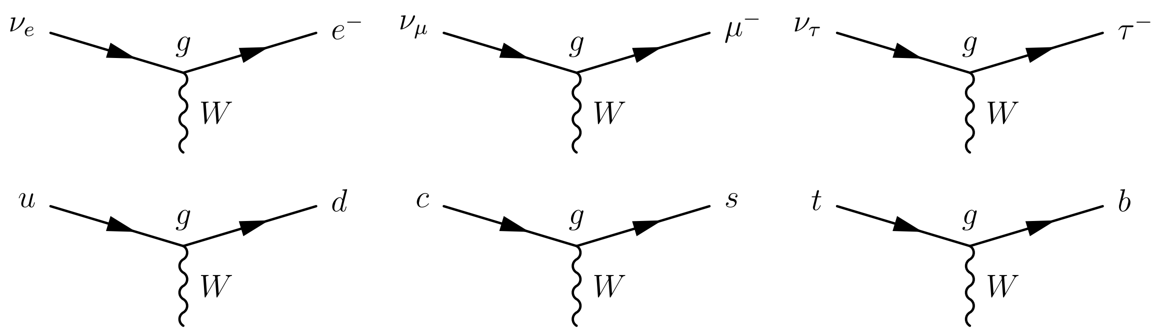
9.6 Quark mixing

For the decay of the muon we studied the weak interaction acting between leptons. We have seen in the process of neutron decay that the weak interaction also operates between quarks. All fundamental fermions are susceptible to the weak interaction. Both the leptons and quarks are usually ordered in a representation of three generations:

$$\underline{\text{Leptons :}} \quad \begin{pmatrix} \nu_e \\ e \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} \qquad \underline{\text{Quarks :}} \quad \begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \tag{9.40}$$

To leptons and quarks of each generation we assign a quantum number called 'flavour'. The strong and electromagnetic interaction conserve flavour. That is just another way of saying that they do not couple to currents that connect quarks or leptons from different generation doublets.

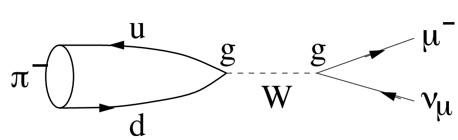
We could assume that the charged current weak interaction works inside the generation doublets as well, *e.g.* that we only have Feynman diagrams of the following types:



As far as we can tell experimentally, the weak interaction indeed conserves lepton flavour. However, consider the decays of charged pions and kaons to a muon and muon anti-neutrino,

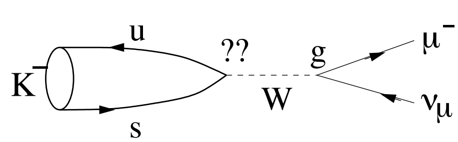
1. pion decay

   $\pi^- \to \mu^- \overline{\nu}_\mu$
   $\Gamma_{\pi^-} \propto \frac{g^4}{M_W^4} \propto G_F^2$



2. kaon decay

   $K^- \to \mu^- \overline{\nu}_\mu$
   This decay does occur!



Since the kaon decay occurs in nature, the weak coupling must violate quark flavour conservation.

9.6.1 Cabibbo - GIM mechanism

We have to modify the model by the replacements:

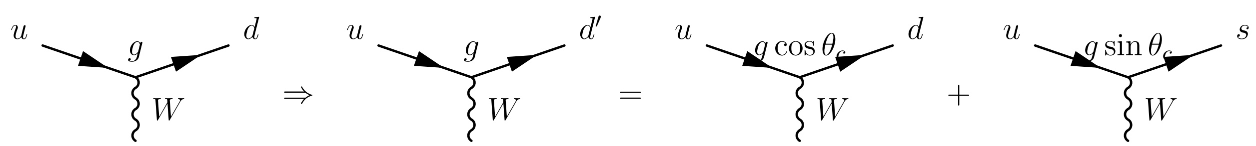
$$\begin{aligned} d &\rightarrow d' = d\ \cos\theta_c + s\ \sin\theta_c \\ s &\rightarrow s' = -d\ \sin\theta_c + s\ \cos\theta_c \end{aligned}$$

or, in matrix representation:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \tag{9.41}$$

where θ_c is the Cabibbo mixing angle.

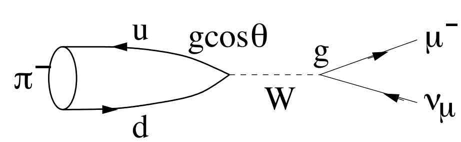
In terms of the diagrams the replacement implies:



Both the u, d coupling and the u, s coupling exist. In this case the diagrams of pion decay and kaon decay are modified:

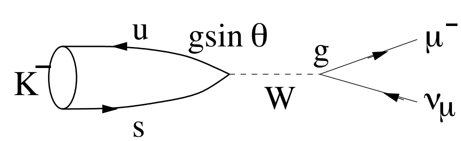
1. Pion decay

   $\pi^- \rightarrow \mu^- \overline{\nu}_\mu$
   $\Gamma_{\pi^-} \propto G_F^2 \cos^2\theta_c$



2. Kaon decay

   $K^- \rightarrow \mu^- \overline{\nu}_\mu$
   $\Gamma_{K^-} \propto G_F^2 \sin^2\theta_c$



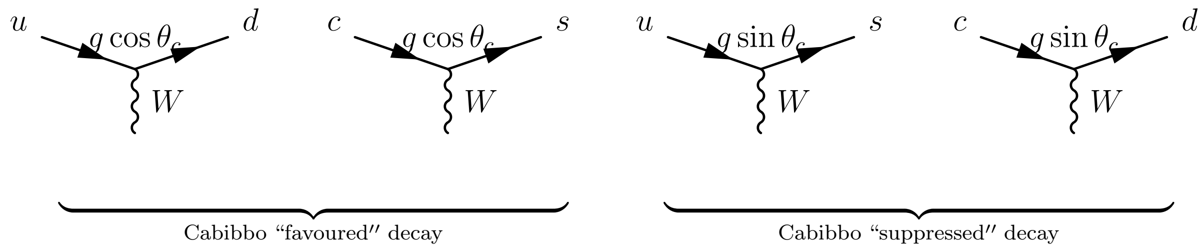
In order to check this we can compare the decay rate of the two reactions. A proper calculation gives:

$$\frac{\Gamma(K^-)}{\Gamma(\pi^-)} \approx \tan^2\theta_c \cdot \left(\frac{m_\pi}{m_K}\right)^3 \left(\frac{m_K^2 - m_\mu^2}{m_\pi^2 - m_\mu^2}\right)^2 \tag{9.42}$$

As a result the Cabibbo mixing angle is observed to be

$$\boxed{\theta_C = 12.8^\circ} \tag{9.43}$$

The couplings for the first two generations are:



Formulated in a different way:

- The flavour eigenstates u, d, s, c are the mass eigenstates. They are the solution of the total Hamiltonian describing quarks; ie. mainly strong interactions.
- 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.

The relation between the mass eigenstates and the interaction eigenstates is a rotation matrix:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} \tag{9.44}$$

with the Cabibbo angle as the mixing angle of the generations.

9.6.2 The Cabibbo - Kobayashi - Maskawa (CKM) matrix

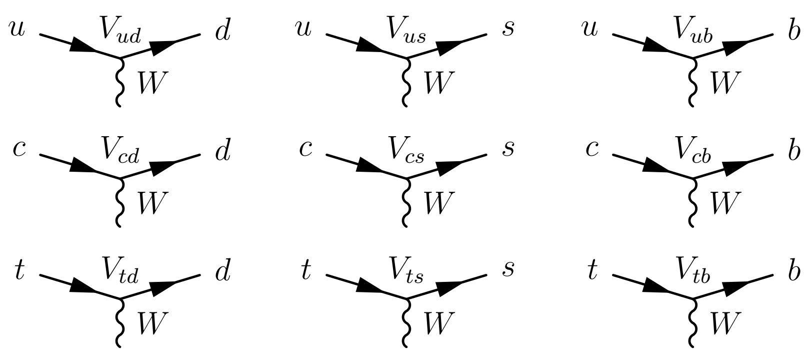
We extend the picture of the previous section to include all three generations. This means that we now make the replacement:

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} u \\ d' \end{pmatrix} \begin{pmatrix} c \\ s' \end{pmatrix} \begin{pmatrix} t \\ b' \end{pmatrix} \tag{9.45}$$

with in the most general way can be written as:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{\text{CKM}-\text{matrix}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{9.46}$$

The "g" couplings involved are:



It should be noted that the matrix is not uniquely defined since the phases of the quark wave functions are not fixed. The standard representation of this unitary 3×3 matrix contains three mixing angles between the quark generations θ_{12}, θ_{13}, θ_{23}, and one complex phase δ:

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}s_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{9.47}$$

where $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$.

In the Wolfenstein parametrization this matrix is:

$$V_{CKM} \approx \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} \tag{9.48}$$

It can be easily seen that it has 4 parameters:

3 real parameters : λ, A, ρ
1 imaginary parameter : $i\eta$

This imaginary parameter is the source of CP violation in the Standard Model. It means that it defines the difference between interactions involving matter and those that involve anti-matter.

We further note that, in case neutrino particles have mass, a similar mixing matrix also exists in the lepton sector. The Pontecorvo-Maki-Nakagawa-Sakata matrix U_{PMNS} is then defined as follows:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \underbrace{\begin{pmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \\ U_{31} & U_{32} & U_{33} \end{pmatrix}}_{PMNS-matrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \qquad (9.49)$$

In a completely similar way this matrix relates the mass eigenstates of the leptons (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. Alternatively, if neutrinos are Majorana particles, the U_{PMNS} matrix includes three complex phases.

It is currently not clear whether the explanation for a matter dominated universe lies in quark flavour physics (”baryogenesis”) or in lepton flavour physics (”leptogenesis”) and whether it requires physics beyond the Standard Model. It is however interesting to note that there exist 3 generations of particles!

Exercise 31.

Helicity vs Chirality

(a) *Write out the* **chirality operator** γ^5 *in the Dirac-Pauli representation.*

(b) *The helicity operator is defined as* $\lambda = \vec{\sigma} \cdot \hat{p}$. *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.

$$\gamma^5 \psi = \gamma^5 \begin{pmatrix} \chi^{(s)} \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi^{(s)} \end{pmatrix} \approx \lambda \begin{pmatrix} \chi^{(s)} \\ \frac{\vec{\sigma}\cdot\vec{p}}{E+m}\chi^{(s)} \end{pmatrix} = \lambda \psi$$

(c) *Explain why the weak interaction is called* **left-handed**.

Exercise 32. 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:* $\pi^- \to \mu^- \overline{\nu}_\mu$.

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.*

(b) *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:*

$$\frac{\Gamma(\pi^- \to e^- \overline{\nu}_e)}{\Gamma(\pi^- \to \mu^- \overline{\nu}_\mu)} = \left(\frac{m_e}{m_\mu} \right)^2 \left(\frac{m_\pi^2 - m_e^2}{m_\pi^2 - m_\mu^2} \right)^2 \approx 1.2 \cdot 10^{-4} \quad !!$$

(c) *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.*

Lecture 10

Local Gauge Invariance

In the last three lectures of this course we will introduce you to the theory of the electroweak interaction, the so-called "Glashow-Salam-Weinberg model". This theory can be formulated by using the principle of local gauge invariance. As we did before, we try to focus on the concepts rather than on formal derivations.

A good book on this topic is "Gauge Theories of the Strong, Weak, and Electromagnetic Interactions", by Chris, Quigg, in the series of "Frontiers in Physics", Benjamin Cummings. You can find the material for Lecture 10 also in Griffiths chapter 10 and Halzen and Martin chapter 14.

10.1 Introduction

In the introduction to Lecture 5 we briefly discussed the formulation of the laws of physics in terms of a Lagrangian and the principle of least action (Hamilton's principle). The reason that this approach is popular is that it is particularly suitable to understand symmetries and the conservation laws that follow from them.

Symmetries play a fundamental role in particle physics. In general one can distinguish[1] four types of symmetries. There is a theorem stating that a symmetry is always related to a quantity that is fundamentally unobservable. Some of these unobservables are mentioned below:

- permutation symmetries: Bose Einstein statistics for integer spin particles and Fermi Dirac statistics for half integer spin particles. The unobservable is the identity of a particle;

[1] T.D. Lee: "Particle Physics and Introduction to Field Theory"

- 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. A famous example in this respect is to try and make an absolute definition of matter and anti-matter. Is this possible? This question will be addressed in the particle physics II course;
- unitary symmetries or internal symmetries:, also called 'gauge invariance'. These are the symmetries discussed in this lecture. As an example of an unobservable quantity we can mention the absolute phase of a quantum mechanical wave function.

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, and invariance under rotation to conservation of angular momentum. Invariances under *internal* operations, like the rotation of the complex phase of wave functions lead to conserved currents, or more specific, conservation of charge.

We believe that the fundamental elementary interactions of the quarks and leptons can be understood as being the result of gauge symmetries. Starting from a Lagrangian that describes free quarks and leptons, the interactions can be constructed simply by requiring the Lagrangian to be symmetric under particular transformations. The idea of local gauge invariance will be discussed in this first lecture and will be further applied in the unified electroweak theory in the second lecture. In the third lecture we will calculate the electroweak process $e^+e^- \to \gamma, Z \to \mu^+\mu^-$, using the techniques we developed before.

10.2 The Lagrangian

In Lecture 5 we briefly introduced the concept of a Lagrangian and generalized coordinates. We describe the dynamics of a system by a finite set of coordinates q_i with time-derivatives (or velocities) $\dot{q}_i$. For the classical Lagrangian

$$L \;=\; T(\dot{q}) \;-\; V(q) \tag{10.1}$$

(where the potential energy only depends on q and the kinetic energy only on $\dot{q}$), we defined the action (or 'action integral') of a trajectory (or 'path') that starts at t_1 and ends at t_2 with

$$S(q) \;=\; \int_{t_0}^{t_1} L(q,\dot{q})\mathrm{d}t. \tag{10.2}$$

The principle of least action, or Hamilton's principle, not states that the actual trajectory $q(t)$ followed by the system is the one for which the action is stationary, that is

$$\delta S(q, \delta q, \dot{q}, \delta \dot{q}) \;=\; 0 \tag{10.3}$$

You will show in exercise 33 that from Hamilton's principle we can derive the Euler-Lagrange equation of motion for each of the coordinates q_i:

$$\frac{\partial L}{\partial q_i} \;=\; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) \tag{10.4}$$

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. A solution is found by going 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), such that the action becomes

$$S \;=\; \int_{x_1}^{x_2} \mathrm{d}^4x \; \mathcal{L}\left(\phi(x, \partial_\mu \phi(x)\right) \tag{10.5}$$

Following the principle of least action we can also obtain an Euler-Lagrange equation for the fields:

$$\frac{\partial \mathcal{L}}{\partial \phi\left(x\right)} = \partial_\mu \frac{\partial \mathcal{L}}{\partial\left(\partial_\mu \phi\left(x\right)\right)} \tag{10.6}$$

To create a Lorentz covariant theory, we will require the Lagrangian to be a Lorentz scalar. As we have seen before this imposes certain conditions on the Lorentz transformation properties of the fields. Furthermore, although we consider complex fields, we will require the Lagrangian to be real. As we shall see below, this effectively leads to a global $U(1)$ symmetry.

In *quantum* field theory, the coordinates ϕ become operators that obey the standard quantum mechanical commutation relation with their associated generalized momenta. The *wave functions* that we have considered before can be viewed as single particle excitations that occur when the field operators 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:

$$\mathcal{L} \;=\; (\partial_\mu \phi^*)(\partial^\mu \phi) - m^2 \phi^* \phi \tag{10.7}$$

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 Lagrangian of the Dirac field.

Exercise 33.

Use Hamilton's principle in Eq. 10.3 to derive the Euler-Lagrange equation of motion (Eq. 10.4).

Exercise 34. Lagrangians versus equations of motion

(a) Show that the Euler-Lagrange equations of the Lagrangian

$$\mathcal{L} = \mathcal{L}_{KG}^{free} = \frac{1}{2}\left(\partial_\mu \phi\right)\left(\partial^\mu \phi\right) - \frac{1}{2} m^2 \phi^2 \tag{10.8}$$

of a real scalar field ϕ leads to the Klein-Gordon equation.

For a complex scalar field one can show that the Lagrangian becomes:

$$\mathcal{L} = \left|\partial^\mu \phi\right|^2 - m^2 \left|\phi\right|^2 \tag{10.9}$$

(b) Show that the Euler-Lagrange equations for the Lagrangian

$$\mathcal{L} = \mathcal{L}_{Dirac}^{free} = i\overline{\psi}\gamma_\mu \partial^\mu \psi - m\overline{\psi}\psi \tag{10.10}$$

leads to the Dirac equation:

$$\left(i\gamma^\mu \partial_\mu - m\right)\psi(x) = 0 \tag{10.11}$$

and its adjoint. Note: *To do this, you need to consider ψ and $\overline{\psi}$ as independent fields.*

(c) Show that the Lagrangian

$$\mathcal{L} = \mathcal{L}_{EM} = -\frac{1}{4}\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right)\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) - j^\mu A_\mu = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - j^\mu A_\mu \tag{10.12}$$

leads to the Maxwell equations:

$$\partial_\mu \left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) = j^\nu \tag{10.13}$$

Hence the current is conserved ($\partial_\nu j^\nu = 0$), since $F^{\mu\nu}$ is antisymmetric.

10.3 Where does the name "gauge theory" come from?

The idea of gauge invariance as a dynamical principle is due to Hermann Weyl. He called it *"eichinvarianz"* ("gauge" = "calibration"). Hermann Weyl[2] was trying to find a geometrical basis for both gravitation and electromagnetism. 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:

$$f(x + \delta x) = f(x) + \partial^\mu f(x)\, \delta x_\mu \tag{10.14}$$

But if in addition the scale, or the unit of measure, for f changes by a factor $(1 + S^\mu dx_\mu)$ between x and $x + \delta x$, then the value of f becomes:

$$\begin{aligned} f(x+\delta x) &= (f(x) + \partial^\mu f(x)\delta x_\mu)\ (1 + S^\nu \delta x_\nu) \\ &= f(x) + (\partial^\mu f(x) + f(x) S^\mu)\, \delta x_\mu + O(\delta x)^2 \end{aligned} \tag{10.15}$$

So, to first order, the increment is:

$$\delta f = (\partial^\mu + S^\mu)\, f\, \delta x_\mu \tag{10.16}$$

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:

$$(1 + S^\mu\, \delta x_\mu) \to (1 + iq\, A^\mu\, \delta x_\mu) \approx \exp\left(iq\, A^\mu\, \delta x_\mu\right) \tag{10.17}$$

then we see that the change of scale gets the form of a change of a phase. When he later on studied the invariance under phase transformations, he kept using the terminology of "gauge invariance".

10.4 Global phase invariance and Noether's theorem

As we have seen above Lagrangians for complex fields are constructed such that the Lagrangian is real (and the Hamiltonian is hermitian). As a result, the Lagrangian is

[2] H. Weyl, *Z. Phys.* **56**, 330 (1929)

not sensitive to a *global* shift in the complex phase of the field. Such a global phase change is called an $U(1)$ 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 'intrinsic' symmetries of the Lagrangian this works as follows. Consider a transformation of the field components with a small variation ϵ,

$$\phi_i \to \phi_i + \epsilon_i(x) \tag{10.18}$$

The resulting change in the Lagrangian is

$$\begin{aligned} \delta\mathcal{L} = & \sum_i \left(\frac{\partial \mathcal{L}}{\partial \phi_i}\epsilon_i + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)}\partial_\mu \epsilon_i \right) \\ = & \, \partial_\mu \left(\sum_i \epsilon_i \frac{\partial \mathcal{L}}{\partial_\mu \phi_i} \right) \end{aligned} \tag{10.19}$$

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

$$j^\mu \; = \; \sum_i \epsilon_i \frac{\partial \mathcal{L}}{\partial_\mu \phi_i} \tag{10.20}$$

is a conserved current.

Applying this to the complex scalar field for a (small) $U(1)$ phase translation

$$\begin{aligned} \phi \to \phi e^{i\alpha} \; &\approx \; \phi(1 + i\alpha) \\ \phi^* \to \phi^* e^{-i\alpha} \; &\approx \; \phi^*(1 - i\alpha) \end{aligned} \tag{10.21}$$

leads to the current

$$j^\mu \; = \; \alpha \, i \, \left(\phi(\partial^\mu \phi^*) - \phi^*(\partial^\mu \phi) \right) \tag{10.22}$$

Since α is an arbitrary constant, we can omit it from the current. Note that we have obtained exactly the current that we constructed for the Klein-Gordon wave in Lecture 2.

10.5 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:

$$\langle O \rangle = \int \psi^* O \psi \tag{10.23}$$

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.)

But this 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,

$$\psi(x) \to \psi'(x) = e^{i\alpha(x)}\psi(x)? \tag{10.24}$$

In general, we cannot do this without breaking the symmetry. The problem is that the Lagrangian density $\mathcal{L}(\psi(x), \partial_\mu\psi(x))$ depends on both on the fields $\psi(x)$ and on the derivatives $\partial_\mu\psi(x)$. The derivative term yields:

$$\partial_\mu\psi(x) \to \partial_\mu\psi'(x) = e^{i\alpha(x)}\ (\partial_\mu\psi(x) + i\partial_\mu\alpha(x)\psi(x)) \tag{10.25}$$

and therefore the Lagrangian is not invariant. However, suppose that we now introduce 'local $U(1)$ symmetry' as a hypothesis. Is it possible to modify the Lagrangian such that it obeys this symmetry?

The answers is 'yes'! We can do this by introducing a new field, the so-called *gauge field*. First, we replace the derivative ∂_μ by the *gauge-covariant derivative*:

$$\partial_\mu \to D_\mu \equiv \partial_\mu + iqA_\mu \tag{10.26}$$

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

$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) - \frac{1}{q}\partial_\mu\alpha(x) \tag{10.27}$$

By inserting the expression for A in the covariant derivative, we find that it just rotates with the *local* phase $\alpha(x)$:

$$\begin{aligned} D_\mu\psi(x)\ \to\ D'_\mu\psi'(x) &= e^{i\alpha(x)}\left(\partial_\mu\psi(x) + i\partial_\mu\alpha(x)\psi(x) + iqA_\mu(x)\psi(x) - iq\frac{1}{q}\partial_\mu\alpha(x)\psi(x)\right) \\ &= e^{i\alpha(x)}D_\mu\psi(x) \end{aligned}$$

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.

10.6 Application to the Lagrangian for a Dirac field

Now consider the effect of the replacement of the derivative

$$\partial_\mu \to D_\mu \equiv \partial_\mu + iqA_\mu(x) \tag{10.28}$$

in the Dirac Lagrangian,

$$\begin{aligned}\mathcal{L} &= \overline{\psi}\left(i\gamma^\mu D_\mu - m\right)\ \psi \\ &= \overline{\psi}\left(i\gamma^\mu \partial_\mu - m\right)\psi - qA_\mu\overline{\psi}\gamma^\mu\psi \\ &\equiv \mathcal{L}_{\text{free}} - \mathcal{L}_{\text{int}}\end{aligned} \tag{10.29}$$

We defined the interaction term

$$\mathcal{L}_{\text{int}} = J^\mu A_\mu \tag{10.30}$$

with the familiar Dirac current

$$J^\mu = q\overline{\psi}\gamma^\mu\psi \tag{10.31}$$

Note that this is exactly the form of the electromagnetic interaction that we discussed in the previous lectures! 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. 10.27 corresponds exactly to 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. It is given by

$$\mathcal{L}_A^{\text{free}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{10.32}$$

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

$$\mathcal{L}_{\text{QED}} = \overline{\psi}\left(i\gamma^\mu \partial_\mu - m\right)\psi - qA_\mu\overline{\psi}\gamma^\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{10.33}$$

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:

$$\mathcal{L}_\gamma = \frac{1}{2}m^2 A^\mu A_\mu \tag{10.34}$$

However, this term violates local gauge invariance, since:

$$A^\mu A_\mu \to \left(A^\mu - \partial^\mu\alpha\right)\left(A_\mu - \partial_\mu\alpha\right) \neq A^\mu A_\mu \tag{10.35}$$

Therefore, the requirement of $U(1)$ invariance, automatically implies that the photon is massless. This actually holds for other gauge symmetries as well. Later on, in the PPII course, it will be discussed how masses of vector bosons can be generated in the Higgs mechanism, by 'breaking' the symmetry.

Exercise 35. Gauge invariance

(a) (i) *Consider the Lagrangian for a complex scalar field:*

$$\mathcal{L} = |\partial^\mu \phi|^2 - m^2 |\phi|^2 \ . \tag{10.36}$$

Make a transformation of these fields:

$$\phi(x) \to e^{iq\alpha}\phi(x) \qquad ; \qquad \phi^*(x) \to e^{-iq\alpha}\phi^*(x) \ . \tag{10.37}$$

Show that the Lagrangian does not change.

(ii) *Do the same for the Dirac Lagrangian while considering the simultaneous transformations:*

$$\psi(x) \to e^{iq\alpha}\psi(x) \qquad ; \qquad \overline{\psi}(x) \to e^{-iq\alpha}\overline{\psi}(x) \tag{10.38}$$

(iii) *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 conservation of charge:* $\partial_\mu j^\mu = 0$, *with:*

$$j^\mu = \frac{ie}{2}\left(\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi\right)}\psi - \overline{\psi}\frac{\mathcal{L}}{\partial\left(\partial_\mu\overline{\psi}\right)}\right) = -e\overline{\psi}\gamma^\mu\psi \tag{10.39}$$

(b) (i) *Start with the Lagrange density for a complex Klein-Gordon field*

$$\mathcal{L} = \left(\partial_\mu\phi\right)^*\left(\partial^\mu\phi\right) - m^2\phi^*\phi \tag{10.40}$$

and show that a **local** *field transformation:*

$$\phi(x) \to e^{iq\alpha(x)}\phi(x) \qquad ; \qquad \phi^*(x) \to e^{-iq\alpha(x)}\phi^*(x) \tag{10.41}$$

does **not** *leave the Lagrangian invariant.*

(ii) *Replace now in the Lagrangian:* $\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu$ *and show that the Lagrangian now* **does** *remain invariant, provided that the additional field transforms with the gauge transformation as:*

$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) - \partial_\mu\alpha(x) \ . \tag{10.42}$$

(c) (i) *Start with the Lagrange density for a Dirac field*

$$\mathcal{L} = i\overline{\psi}\gamma^\mu\partial_\mu\psi - m\overline{\psi}\psi \tag{10.43}$$

and show that a **local** *field transformation:*

$$\psi(x) \to e^{iq\alpha(x)}\psi(x) \qquad ; \qquad \overline{\psi}(x) \to e^{-iq\alpha(x)}\overline{\psi}(x) \tag{10.44}$$

also does **not** *leave the Lagrangian invariant.*

(ii) *Again make the replacement:* $\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu$ *where again the gauge field transforms as:*

$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) - \partial_\mu\alpha(x) \ . \tag{10.45}$$

and show that the physics now **does** *remain invariant.*

10.7 Yang-Mills theory

The concept of *non-abelian* gauge theories is introduced here in a somewhat historical context as this helps to also understand the origin of the term weak isospin and the relation to (strong-) isospin.

In the 1950s Yand and Mills tried to understand the strong interaction in the proton-neutron system in terms of a gauge symmetry. Ignoring the electric charge, we can write the free Lagrangian for the nucleons as

$$\mathcal{L} = \overline{p}\ (i\gamma^\mu \partial_\mu - m)\ p + \overline{n}\ (i\gamma^\mu \partial_\mu - m)\ n \tag{10.46}$$

or, in terms of a composite spinor $\psi = \begin{pmatrix} p \\ n \end{pmatrix}$,

$$\mathcal{L} = \overline{\psi}\ (i\gamma^\mu\, I\, \partial_\mu - I\, m)\ \psi \qquad \text{with} \qquad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{10.47}$$

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 which effectively '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

$$U \ = \ e^{i\theta} \exp\left(\frac{i}{2}\vec{\tau}\cdot\vec{\alpha}\right) \ , \tag{10.48}$$

where $\vec{\alpha}$ and θ are real and $\vec{\tau} = (\tau_1, \tau_2, \tau_3)$ are the Pauli Matrices [3]. 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 a 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. For an infinitesimal transformation

$$\psi \to \psi' = (1 + \frac{i}{2}\vec{\tau}\cdot\vec{\alpha})\psi \tag{10.49}$$

[3] Our default representation is: $\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$

(and similar for $\overline{\psi}$) we find for the Lagrangian

$$\begin{aligned}
\delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\psi}\delta\psi + \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi\right)}\delta\left(\partial_\mu\psi\right) + \frac{\partial\mathcal{L}}{\partial\overline{\psi}}\delta\overline{\psi} + \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\overline{\psi}\right)}\delta\left(\partial_\mu\overline{\psi}\right) \\
&= \frac{\delta\mathcal{L}}{\partial\psi}\,\frac{i}{2}\vec{\tau}\cdot\vec{\alpha}\psi + \frac{\delta\mathcal{L}}{\partial\left(\partial_\mu\psi\right)}\,\frac{i}{2}\vec{\tau}\cdot\vec{\alpha}\;\left(\partial_\mu\psi\right) + \ldots \\
&= \left(\partial_\mu\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi\right)}\right)\,\frac{i}{2}\vec{\tau}\cdot\vec{\alpha}\;\psi + \frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi\right)}\,\frac{i}{2}\vec{\tau}\cdot\vec{\alpha}\left(\partial_\mu\psi\right) \\
&= \partial_\mu\,\vec{\alpha}\cdot\left(\frac{i}{2}\frac{\partial\mathcal{L}}{\partial\left(\partial_\mu\psi\right)}\vec{\tau}\psi\right)
\end{aligned}$$

where the Euler-Lagrange relation has been used to eliminate $\partial\mathcal{L}/\partial\psi$. Consequently, from the required invariance of the Lagrangian we obtain three continuity equations $\partial_\mu\vec{J}^\mu = 0$ for the conserved currents

$$\vec{J}^\mu \;=\; \overline{\psi}\,\gamma^\mu\,\frac{\vec{\tau}}{2}\,\psi\;. \tag{10.50}$$

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 subtle. The first step is to make the parameters α depend on space time. To simplify the notation we define the gauge transformation as follows,

$$\begin{aligned}
\psi(x)\rightarrow\psi'(x) &= \;G(x)\psi(x) \\
\text{with}\quad G(x) &= \;\exp\left(\frac{i}{2}\,\vec{\tau}\cdot\vec{\alpha}(x)\right)
\end{aligned} \tag{10.51}$$

We have again, as in the case of QED, the problem with the transformation of the derivative:

$$\partial_\mu\psi(x)\rightarrow G\left(\partial_\mu\psi\right) + \left(\partial_\mu\,G\right)\;\psi \tag{10.52}$$

So, also here, we must introduce a new gauge field to keep the Lagrangian invariant:

$$\mathcal{L} = \overline{\psi}\;\left(i\gamma^\mu\,D_\mu - Im\right)\;\psi \quad\text{with}\quad \psi = \begin{pmatrix} p \\ n \end{pmatrix} \quad\text{and}\quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{10.53}$$

where we introduce the new covariant derivative:

$$I\partial_\mu \rightarrow D_\mu = I\partial_\mu + igB_\mu \tag{10.54}$$

with g is a new coupling constant that replaces the charge e in electromagnetism. The object B_μ is now a (2×2) matrix:

$$B_\mu \;=\; \frac{1}{2}\,\vec{\tau}\cdot\vec{b}_\mu \;=\; \frac{1}{2}\tau^a b^a_\mu \;=\; \frac{1}{2}\begin{pmatrix} b_3 & b_1 - ib_2 \\ b_1 + ib_2 & -b_3 \end{pmatrix} \tag{10.55}$$

where $\vec{b}_\mu = (b_1, b_2, b_3)$ are three new real fields. We call them the gauge fields of the $SU(2)$ symmetry. We need three fields rather than one, because $SU(2)$ has three generators.

In order to obtain a Lagrangian that is invariant, we again want the gauge transformation to take the form

$$D_\mu \psi \to D'_\mu \psi' = G\,(D_\mu \psi) \tag{10.56}$$

If we write out the covariant derivative term we find

$$\begin{aligned} D'_\mu \psi' &= \left(\partial_\mu + igB'_\mu\right)\, \psi' \\ &= G\,(\partial_\mu \psi) + (\partial_\mu G)\,\psi + igB'_\mu\,(G\psi) \end{aligned}$$

If we compare this to the desired result:

$$\begin{aligned} D'_\mu \psi' &= G\,(\partial_\mu \psi + igB_\mu)\,\psi \\ &= G\,(\partial_\mu \psi) + igG\,(B_\mu \psi) \end{aligned}$$

then we see that the desired behaviour is obtained if the gauge field transforms simultaneously as:

$$igB'_\mu\,(G\psi) = igG\,(B_\mu \psi) - (\partial_\mu G)\,\psi \tag{10.57}$$

which must then be true for all values of the nucleon field ψ. Multiplying this operator equation from the right by G^{-1} we find

$$B'_\mu \;=\; GB_\mu G^{-1} \;+\; \frac{i}{g}\,(\partial_\mu G)\,G^{-1}\;. \tag{10.58}$$

Although this looks rather complicated we can again try to interpret this by comparing to the case of electromagnetism, where $G_{em} = e^{iq\alpha(x)}$.
Then:

$$\begin{aligned} A'_\mu &= G_{em} A_\mu G_{em}^{-1} + \frac{i}{q}\,(\partial_\mu G_{em})\,G_{em}^{-1} \\ &= A_\mu - \partial_\mu \alpha \end{aligned}$$

which is exactly what we had before.

Exercise E36. (Extra exercise, not obligatory!)

Consider an infinitesimal gauge transformation:

$$G = 1 + \frac{i}{2}\,\vec{\tau}\cdot\vec{\alpha} \qquad |\vec{\alpha}| \ll 1 \tag{10.59}$$

Use the general transformation rule for B'_μ and use $B_\mu = \frac{1}{2}\vec{\tau}\cdot\vec{b}_\mu$ to demonstrate that the fields transform as:

$$\vec{b}'_\mu = \vec{b}_\mu - \vec{\alpha}\times\vec{b}_\mu - \frac{1}{g}\partial_\mu \vec{\alpha} \tag{10.60}$$

(use: the Pauli-matrix identity: $(\vec{\tau}\cdot\vec{a})(\vec{\tau}\cdot\vec{b}) = \vec{a}\cdot\vec{b} + i\vec{\tau}\cdot(\vec{a}\times\vec{b})$*).*

So for isospin symmetry the b^a_μ fields transform as an isospin rotation and a gradient term. The gradient term was already present in QED. The rotation term is new. It arises due to the non-commutativity of the 2x2 isospin rotations. If we write out the gauge field transformation formula in components:

$$b'^l_\mu = b^l_\mu - \epsilon_{jkl}\,\alpha^j\,b^k - \frac{1}{g}\partial_\mu\alpha^l \tag{10.61}$$

we can see that there is a coupling *between* the different components of the field. We call this the self-coupling.

The effect of this becomes clear if one also considers the kinetic term of the isospin gauge field. Analogous to the QED case, the three new fields require their own free Lagrangian, which we write as

$$\mathcal{L}_b^{\text{free}} = -\frac{1}{4}\sum_l F_l^{\mu\nu}F_{\mu\nu,l} \;=\; -\frac{1}{4}\vec{F}^{\mu\nu}\cdot\vec{F}_{\mu\nu}\;. \tag{10.62}$$

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

$$F_l^{\mu\nu} = \partial^\nu b_l^\mu - \partial^\mu b_l^\nu + g\,\epsilon_{jkl}\,b_j^\mu\,b_k^\nu \tag{10.63}$$

or in vector notation

$$\vec{F}^{\mu\nu} \;=\; \partial^\mu\vec{b}^\nu - \partial^\nu\vec{b}^\mu \;-\; g\,\vec{b}^\mu\times\vec{b}^\nu \tag{10.64}$$

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 os only the bilinear photon propagator term. In the isospin 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 $B_\mu = \frac{1}{2}\vec{\tau}\cdot\vec{b}_\mu$. The full Lagrangian of the theory (which is called the

Yang-Mills theory) is then given by

$$\begin{aligned}\mathcal{L}_{SU(2)} =& \ \overline{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}\vec{F}^{\mu\nu}\cdot\vec{F}_{\mu\nu} \\ &= \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi - g\vec{J}^\mu\vec{b}_\mu - \frac{1}{4}\vec{F}^{\mu\nu}\cdot\vec{F}_{\mu\nu} \\ &\equiv \ \mathcal{L}_\psi^{\text{free}} \ + \ \mathcal{L}^{\text{interaction}} \ + \ \mathcal{L}_b^{\text{free}}\end{aligned} \tag{10.65}$$

where we now absorbed the coupling constant g in the definition of the conserved current,

$$\vec{J}^\mu = \frac{g}{2}\overline{\psi}\gamma^\mu\vec{\tau}\psi \tag{10.66}$$

Comparing this to the QED Lagrangian

$$\mathcal{L}_{U(1)} = \mathcal{L}_{U(1)}^{free} - A_\mu\cdot J^\mu - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{10.67}$$

(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.

Comments

We see a symmetry in the $\begin{pmatrix} p \\ n \end{pmatrix}$ system: the isospin rotations.

- If we require local gauge invariance of such transformations we need to introduce $\vec{b}_\mu$ gauge fields.
- But what are they? $\vec{b}_\mu$ must be three massless vector bosons that couple to the proton and neutron. It cannot be the π^-, π^0, π^+ since they are pseudo-scalar particles rather then vector bosons. It turns out this theory does not describe the strong interactions. We know now that the strong force is mediated by massless gluons. In fact gluons have 3 colour degrees of freedom, such that they can be described by 3x3 unitary gauge transformations (SU(3)), for which there are 8 generators, listed here:

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

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

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

The strong interaction will be discussed later on in the particle physics course. Next lecture we will instead look at the weak interaction and introduce the concept of weak isospin.

- Also, we have started to say that the symmetry in the p, n system is only present if we neglect electromagnetic interactions, since obviously from the charge we can absolutely define the proton and the neutron state in the doublet. In such a case where the symmetry is only approximate, we speak of a *broken symmetry* rather then of an *exact symmetry*.

Lecture 11

Electroweak Theory

In the previous lecture we have seen how imposing a local gauge symmetry requires a modification of the free Lagrangian such that a theory with interactions is obtained. We studied:

- local $U(1)$ gauge invariance:

$$\overline{\psi}\ (i\gamma^\mu D_\mu - m)\ \psi = \overline{\psi}\ (i\gamma^\mu \partial_\mu - m)\ - \underbrace{q\overline{\psi}\gamma^\mu\psi}_{J^\mu}\ A_\mu \tag{11.1}$$

- local $SU(2)$ gauge invariance:

$$\overline{\psi}\ (i\gamma^\mu D_\mu - m)\ \psi = \overline{\psi}\ (i\gamma^\mu \partial_\mu - m)\ - \underbrace{\frac{g}{2}\overline{\psi}\gamma^\mu\vec{\tau}\psi}_{\vec{J}^\mu}\ \vec{b}_\mu \tag{11.2}$$

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. For the $SU(2)$ case we hoped that we could describe the strong nuclear interactions, but this failed. We will now show that the model with an $SU(2)$ local gauge symmetry is still useful, but then to explain the *weak* interaction. (The strong interaction is derived from $SU(3)$ gauge invariance.)

We define for any Dirac field ψ the 'left' and 'right' handed projections

$$\psi_L \equiv \frac{1}{2}\,(1-\gamma_5)\ \psi \qquad \psi_R \equiv \frac{1}{2}\,(1+\gamma_5)\ \psi \tag{11.3}$$

These are called the chiral projections. As we have seen in chapter 9, for particles with $E \gg m$ these correspond to the negative and positive helicity states, respectively. Note that, unless the particle is massless, the chiral projections are *not* solutions of the Dirac equation even if ψ is a solution. However, using the fact that (see exercise 38)

$$\overline{\psi}\,\gamma^\mu\,\psi\ =\ \overline{\psi}_L\,\gamma^\mu\,\psi_L\ +\ \overline{\psi}_R\,\gamma^\mu\,\psi_R \tag{11.4}$$

where the chiral projections of the adjoint spinors are given by

$$\overline{\psi}_L \;=\; \frac{1}{2}\overline{\psi}\;(1+\gamma_5) \qquad \text{and} \qquad \overline{\psi}_R \;=\; \frac{1}{2}\overline{\psi}\;(1-\gamma_5) \tag{11.5}$$

we can rewrite the Dirac Lagrangian for ψ as

$$\mathcal{L} \;=\; \overline{\psi}_R(i\gamma^\mu\partial_\mu - m)\psi_R \;+\; \overline{\psi}_L(i\gamma^\mu\partial_\mu - m)\psi_L \;. \tag{11.6}$$

Let us now introduce the following doublets for the left-handed chirality states of the leptons and quarks in the first family:

$$\Psi_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \qquad \text{and} \qquad \Psi_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \tag{11.7}$$

We call these "weak isospin" doublets. Again, note that Ψ is not a Dirac spinor, but a doublet of Dirac spinors. As an example, consider the Lagrangian for the electron and neutrino and verify that it can be written as

$$\begin{aligned}\mathcal{L} =\; & \overline{e}_R(i\gamma^\mu\partial_\mu - m_e)\overline{e}_R \;+\; \overline{\nu}_R(i\gamma^\mu\partial_\mu - m_\nu)\overline{\nu}_R \;+ \\ & + (\overline{\nu}_L\;\overline{e}_L)\begin{pmatrix} i\gamma^\mu\partial_\mu - m_\nu & 0 \\ 0 & i\gamma^\mu\partial_\mu - m_e \end{pmatrix}\begin{pmatrix} \nu_L \\ e_L \end{pmatrix}\end{aligned} \tag{11.8}$$

Now it comes: We impose a $SU(2)$ gauge symmetry on the left-handed doublets. 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 or mass, problems that we will clearly need to deal with later. In fact, the mass problem is big enough, that we need to choose all fermions massless.

The fact that we only impose the gauge symmetry on left-handed states will lead to a weak interaction that is completely left-right asymmetric. (Intuitively this is very difficult to accept: why would there be a symmetry for the left-handed states only?!). This is called *maximal violation of parity*.

As we shall soon see, the three vector fields (b_1, b_2, b_3 from the previous lecture) can be associated with the carriers of the weak interaction, the W^+, W^-, Z bosons. However, these bosons are now massless as well. An explicit mass term ($\mathcal{L}_M = Kb_\mu b^\mu$) would in fact 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 where we start perturbation theory, is not symmetric. Future lectures will discuss this aspect in more detail.

To construct the weak $SU(2)_L$ theory[1] we start again with the free Dirac Lagrangian and we impose $SU(2)$ symmetry (but now on the weak isospin doublets):

$$\mathcal{L}_{free} = \overline{\Psi}_L \, i\gamma^\mu \partial_\mu \, \Psi_L \tag{11.9}$$

After introducing the covariant derivative

$$\partial_\mu \to D_\mu = \partial_\mu + igB_\mu \qquad \text{with} \qquad B_\mu = \frac{1}{2}\,\vec{\tau}\cdot\vec{b}_\mu \tag{11.10}$$

the Dirac equation obtains an interaction term,

$$\mathcal{L}_{free} \to \mathcal{L}_{free} - \vec{b}_\mu \cdot J^\mu_{weak} \tag{11.11}$$

where the weak current is

$$J^\mu_{weak} \;=\; \frac{g}{2}\,\overline{\Psi}_L\,\gamma^\mu\,\vec{\tau}\Psi_L \tag{11.12}$$

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 will now show how these gauge fields can be recast into the 'physical' fields of the 3 (massive) bosons W^+, W^-, Z^0 in order to have them interact with currents of the physical electrons and neutrinos.

11.1 The Charged Current

We choose for the representation of the $SU(2)$ generators the Pauli spin matrices,

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{11.13}$$

Note that τ_1 and τ_2 mix the components of a doublet, while τ_3 does not. We now define the fields $W^\pm$ as follows,

$$W^\pm_\mu \;\equiv\; \frac{1}{\sqrt{2}}\left(b^1_\mu \mp i\, b^2_\mu\right) \tag{11.14}$$

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. We will not look at that coupling in

[1]The subscript L is used to indicate that we only consider $SU(2)$ transformations of the left-handed doublet.

[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!

this lecture. As an alternative, we will show now that these W fields couple to charge-lowering and charge-raising currents. 'Charge conservation' at each Feynman diagram vertex then implies a fixed charge of the virtual fields.

We define the charged current term of the interaction Lagrangian as

$$\mathcal{L}_{CC} \;=\; -\,b^1_\mu J^{1\mu} \;-\; b^2_\mu J^{2\mu} \tag{11.15}$$

with

$$J^{1\mu} = \frac{g}{2}\,\overline{\Psi}_L\,\gamma^\mu\,\tau_1\,\Psi_L \qquad J^{2\mu} = \frac{g}{2}\,\overline{\Psi}_L\,\gamma^\mu\,\tau_2\,\Psi_L \tag{11.16}$$

As you will show in exercise 37 we can rewrite the charged current Lagrangian as

$$\mathcal{L}_{CC} \;=\; -W^+_\mu J^{+\mu} - W^-_\mu J^{-\mu} \tag{11.17}$$

with

$$J^{\mu,\pm} \;=\; \frac{g}{\sqrt{2}}\,\overline{\Psi}_L\,\gamma^\mu\,\tau^\pm\,\Psi_L \tag{11.18}$$

and $\tau^\pm \;=\; \frac{1}{2}(\tau_1 \pm i\tau_2)$, or in our representation

$$\tau^+ \;=\; \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \tau^- \;=\; \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{11.19}$$

The leptonic currents can then be written as

$$J^{+\mu} = \frac{g}{\sqrt{2}}\,\overline{\nu}_L\,\gamma^\mu\,e_L \qquad \text{and} \qquad J^{-\mu} = \frac{g}{\sqrt{2}}\,\overline{e}_L\,\gamma^\mu\,\nu_L \tag{11.20}$$

or written out with the left-handed projection operators:

$$J^{+\mu} = \frac{g}{\sqrt{2}}\,\overline{\nu}\,\frac{1}{2}\left(1+\gamma^5\right)\,\gamma^\mu\,\frac{1}{2}\left(1-\gamma^5\right)\,e \tag{11.21}$$

and similar for $J^{-\mu}$. Verify for yourself that

$$\left(1+\gamma^5\right)\gamma^\mu\left(1-\gamma^5\right) \;=\; 2\gamma^\mu\left(1-\gamma^5\right) \tag{11.22}$$

such that we can rewrite the leptonic charge raising current (W^+)

$$\boxed{J^{+\mu} = \frac{g}{2\sqrt{2}}\,\overline{\nu}\,\gamma^\mu\left(1-\gamma^5\right)e} \tag{11.23}$$

and for the leptonic charge lowering current (W^-)

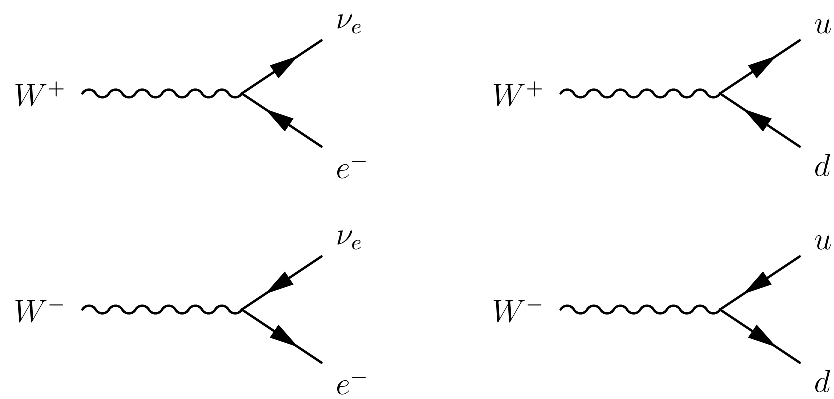
$$\boxed{J^{-\mu} = \frac{g}{2\sqrt{2}}\,\overline{e}\,\gamma^\mu\left(1-\gamma^5\right)\nu}\;. \tag{11.24}$$

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:



11.2 The Neutral Current

11.2.1 Empirical Approach

In the discussion of the weak interaction in Lecture 9 we have been rather brief about the neutral current, so before we discuss the it in detail we catch up with some terminology. The Lagrangian for weak and electromagnetic interactions that arrises as the result of $SU(2)_L$ and $U(1)$ gauge symmetries can be written as

$$\begin{aligned}
\mathcal{L}_{EW} &= \mathcal{L}_{free} - \mathcal{L}_{weak} - \mathcal{L}_{EM} \\
\mathcal{L}_{weak} &= W^+_\mu J^{+\mu} + W^-_\mu J^{-\mu} + b^3_\mu J^\mu_3 \\
\mathcal{L}_{EM} &= a_\mu J^\mu_{EM}
\end{aligned}$$

We have on purpose not written the electromagnetic field as the photon field A^μ. The reason is that the two neutral vector bosons corresponding b_3 and the $U(1)$ symmetry (called 'hypercharge') mix. In the Higgs mechanism this occurs in such a way that one gauge boson becomes the massive Z^0, which is responsible for the neutral weak current, while the other one, the photon, remains massless.

Experiments have shown that in contrast to the charged weak interaction, the neutral weak current associated to the Z^0-boson is *not* purely left-handed. We suggestively parameterize it as

$$J^{\mu f}_{NC} = \frac{g}{2} \overline{\psi}^f \gamma^\mu \left(C^f_V - C^f_A \gamma^5 \right) \psi^f \tag{11.25}$$

where the constants C^f_V and C^f_A express the relative strength of the vector and axial vector components of the interaction. Their value depends on the type of fermion f in the interaction. You can verify yourself that we can rewrite this current as

$$J^{\mu f}_{NC} = \frac{g}{2} \left(C^f_L \overline{\psi}^f_L \gamma^\mu \psi^f_L + C^f_R \overline{\psi}^f_R \gamma^\mu \psi^f_R \right) \tag{11.26}$$

where the left- and right-handed couplings have been defined as

$$C_R \;\equiv\; C_V - C_A \qquad \text{and} \qquad C_L \;\equiv\; C_V + C_A \tag{11.27}$$

Now, although the neutral current couples to both left-handed and right-handed charge lepton and quark field, experimental evidence suggests that it does not couple to right-handed neutrinos. Therefore, we set $C_L^\nu = 1$ and $C_R^\nu = 0$ (or $C_V^\nu = C_A^\nu = 1/2$). The observed neutral current for the lepton doublet is therefore written as

$$J^\mu_{NC} = \frac{g}{2}\;(\overline{\nu}_L\,\gamma^\mu\,\nu_L) \;+\; \frac{g}{2}\;(C_L^e\,\overline{e}_L\,\gamma^\mu\,e_L) \;+\; \frac{g}{2}\;(C_R^e\,\overline{e}_R\,\gamma^\mu\,e_R) \tag{11.28}$$

For the electron, we have in addition the known electromagnetic current, which we can write as

$$J^\mu_{EM} \;=\; q\;(\overline{e}_L\,\gamma^\mu\,e_L) \;+\; q\;(\overline{e}_R\,\gamma^\mu\,e_R) \tag{11.29}$$

with (obviously) equal couplings to the left- and right-handed projection.

Written explicitly in terms of the lepton currents (using Eq. 11.12 and the definition of τ_3), the current that couples to the b_3^μ gauge field is

$$J_3^\mu \;=\; \frac{g}{2}\overline{\nu}_L\gamma^\mu\nu_L - \frac{g}{2}\overline{e}_L\gamma^\mu e_L\;. \tag{11.30}$$

If the physical currents J^μ_{NC} and J^μ_{EM} are the result of mixing of the b_3^μ gauge field and the a^μ gauge field, then they are a linear combination of the associated currents. In that case, the inverse is also true: the gauge currents are linear combinations of J^μ_{NC} and J^μ_{EM}. Therefore, we can write

$$J_3^\mu = a\cdot J^\mu_{NC} + b\cdot J^\mu_{EM}\;. \tag{11.31}$$

By comparing these expressions we find

- look at the ν_L terms: $a = 1$
- look at the e_R terms: $\frac{g}{2}C_R^e + q\cdot b = 0 \Rightarrow C_R^e = -\frac{2qb}{g}$
- look at the e_L: $\frac{g}{2}C_L^e + q\cdot b = -\frac{g}{2} \qquad \Rightarrow C_L^e = -1 - \frac{2qb}{g}$

Summarizing this, we obtain for the vector and axial vector couplings of the electron

$$C_V = \frac{1}{2}\;(C_R + C_L) \qquad \Rightarrow \qquad C_V^e = -\frac{1}{2} - \frac{2q}{g}b$$
$$C_A = \frac{1}{2}\;(C_L - C_R) \qquad \Rightarrow \qquad C_A^e = -\frac{1}{2}$$

We find that the vector coupling contains a new constant b that expresses the relative strength of the $SU(2)$ current and the electromagnetic current. As we shall see below, the constant can be expressed as $b = \sin^2\theta_w$, where θ_w is called the weak mixing angle. We will now look at this more carefully.

11.2.2 $SU(2)_L \otimes U(1)_Y$

In the previous lecture we explained that we study the $SU(2)$ subgroup of the $U(2)$ isospin rotations because we had already separately looked at the $U(1)$ symmetry. The latter generated the photon field. When we now combine the two symmetries we need to be careful since the gauge fields mix.

To solve this problem it is customary to start from a new $U(1)$ gauge symmetry with a charge operator Y. This charge is called 'hypercharge'. Under the combination of $SU(2)_L$ and $U(1)_Y$ a left handed doublet transform as

$$\Psi_L \;\to\; \Psi'_L \;=\; \exp\left[i\vec{\alpha}(x)\vec{T} + i\beta(x)Y\right]\Psi_L \tag{11.32}$$

where for $\vec{T} = \vec{\tau}/2$ 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 in the doublet transform only under hypercharge,

$$\Psi_R \;\to\; \Psi'_R \;=\; e^{i\beta Y}\Psi_R \tag{11.33}$$

The electroweak Lagrangian following from local $SU(2)_L \otimes U(1)_Y$ symmetry now takes the form (see H&M[3], Chapter 13):

$$\boxed{\mathcal{L}_{EW} = \mathcal{L}_{free} - g\,\vec{J}^{\mu}_{SU(2)}\cdot\vec{b}_\mu - \frac{g'}{2}J^\mu_Y a_\mu} \tag{11.34}$$

where a_μ is the gauge field corresponding to $U(1)_Y$ and $g'/2$ is its coupling strength. The charged currents follow as usual from

$$W^\pm_\mu = \frac{b^1_\mu \mp i b^2_\mu}{\sqrt{2}}\,.$$

while the physical neutral current fields are orthogonal linear combinations of the gauge fields,

$$\begin{aligned} A_\mu &= \quad a_\mu\cos\theta_w + b^3_\mu\sin\theta_w \\ Z_\mu &= -a_\mu\sin\theta_w + b^3_\mu\cos\theta_w \end{aligned} \tag{11.35}$$

We can now write the interaction terms for b^3_μ and a_μ in the Lagrangian as

$$\begin{aligned} -gJ^\mu_3 b^3_\mu - \frac{g'}{2}J^\mu_Y a_\mu \;&=\; -\left(g\sin\theta_w J^\mu_3 + g'\cos\theta_w\frac{J^\mu_Y}{2}\right)A_\mu \\ &\quad\; -\left(g\cos\theta_w J^\mu_3 - g'\sin\theta_w\frac{J^\mu_Y}{2}\right)Z_\mu \\ &\equiv\; -eJ^\mu_{EM}A_\mu - g_Z J^\mu_{NC}Z_\mu \end{aligned} \tag{11.36}$$

[3]Halzen and Martin, Quarks & Leptons: "An Introductory Course in Modern Particle Physics"

where in the last line we defined the currents and coupling constants associated to the physical fields.

Up to now all these fields are massless. We have stated before that in the standard model the gauge fields and fermions can be given mass via something that is called the Higgs mechanism. We cannot discuss this in more detail in these lectures, but we will use some predictions (or conjectures). The Higgs mechanism allows that three of the four vector fields ($\vec{b}_\mu$ and a_μ) acquire a mass. In order that the massless field becomes the photon A^μ, we must impose the following relation between the generators (or also 'charge operators') of the $SU(2)_L$, $U(1)_Y$ and $U(1)_{em}$ symmetry groups:

$$Q \;=\; T_3 \;+\; \frac{Y}{2} \tag{11.37}$$

(This relation is also called the Gellmann-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 $SU(3)$, namely $\tau_3/2$. 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 in fact be equal, since the $SU(2)_L$ doublet is a singlet under $U(1)_Y$.)

A direct consequence of Eq. 11.37 is that also the currents are related, namely by

$$\boxed{J^\mu_{EM} = J^\mu_3 + \frac{1}{2} J^\mu_Y} \;. \tag{11.38}$$

Comparing this to Eq. 11.36 we find that

$$e = g \sin\theta_w = g' \cos\theta_w \tag{11.39}$$

Given a value for the mixing angle, this equation fixes the ratio between the $SU(2)_L$ and $U(1)_Y$ coupling constants,

$$g'/g = \tan\theta_w \tag{11.40}$$

Analogously, we find for the Z current,

$$\begin{aligned} & -\left(g\cos\theta_w J_3^\mu - \frac{g'}{2}\sin\theta_w \cdot 2\left(J_{EM}^\mu - J_3^\mu\right)\right) Z_\mu \\ = & \ \ldots \\ = & -\frac{e}{\cos\theta_w \sin\theta_w}\left(J_3^\mu - \sin^2\theta_w J_{EM}^\mu\right) Z_\mu \end{aligned}$$

We can rewrite this as

$$\boxed{J_{NC}^\mu = J_3^\mu - \sin^2\theta_w \, J_{EM}^\mu} \qquad (11.41)$$

which is in accordance with the expression in Eq. 11.31 provided that we choose $a = 1$ and $b = \sin^2\theta_w$.

Summary

We have introduced a local gauge symmetry $SU(2)_L \otimes U(1)_Y$ to obtain a Lagrangian for electroweak interactions,

$$-\left(g\,\vec{J}_L^\mu \cdot \vec{b}_\mu + \frac{g'}{2}J_Y^\mu \cdot a_\mu\right) \qquad (11.42)$$

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:

$$\begin{aligned} J_{EM}^\mu &= J_3^\mu + \frac{1}{2}J_Y^\mu \\ J_{NC}^\mu &= J_3^\mu - \sin^2\theta_w J_{EM}^\mu = \cos^2\theta_w J_3^\mu - \sin^2\theta_w \frac{J_Y^\mu}{2} \end{aligned}$$

and the interaction term in the Lagrangian becomes:

$$-\left(eJ_{EM}^\mu \cdot A_\mu + \frac{e}{\cos\theta_w \sin\theta_w}J_{NC}^\mu \cdot Z_\mu\right) \qquad (11.43)$$

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. Rather, we have 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, which we will discuss next.

11.3 The Mass of the W and Z bosons

In Lecture 9 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

$$\frac{G}{\sqrt{2}} = \frac{g^2}{8M_W^2} \tag{11.44}$$

For neutral current processes we can also compute the coupling-constant of the four-point interaction. It is given by

$$\rho\frac{G}{\sqrt{2}} = \frac{g^2}{8M_Z^2\cos^2\theta_w} \tag{11.45}$$

The parameter ρ specifies the relative strength between the charged and neutral current weak interactions. Comparing the two expressions, we have

$$\rho = \frac{M_W^2}{M_z^2\cos^2\theta_w} \tag{11.46}$$

Experimentally, this number is 1 within small uncertainties. In fact, this is exactly the prediction of the Higgs mechanism. In the Higgs mechanism the mass generated for the W and Z are respectively

$$M_W \;=\; \frac{1}{2}\,v\,g \qquad \text{and} \qquad M_Z \;=\; \frac{1}{2}\,v\,\sqrt{g^2+g'^2}\,, \tag{11.47}$$

where v is the so-called *vacuum expectation value* of the Higgs field. With $g'/g = \tan\theta_w$ we find that $\rho = 1$.

We can also turn the argument around, from a measurement of the Fermi coupling constant in charged current and neutral current processes, we can predict the masses of the W and the Z,

$$M_W \;=\; \sqrt{\frac{\sqrt{2}}{8G_F}}\,\frac{e}{\sin\theta_w} = 81\text{ GeV} \tag{11.48}$$

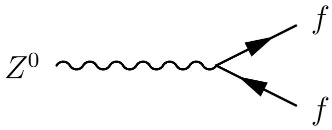
$$M_Z \;=\; M_W\;(g_z/g) = M_W/\cos\theta = 91\text{ GeV} \tag{11.49}$$

11.4 The Coupling Constants for $Z \to f\overline{f}$

For the neutral Z-current interaction we have for the interaction in general:

$$\begin{aligned} -ig_Z\,J^\mu_{NC}\,Z_\mu &= -i\frac{g}{\cos\theta_w}\,\left(J^\mu_3 - \sin^2\theta_w J^\mu_{EM}\right)\,Z_\mu \\ &= -i\frac{g}{\cos\theta_w}\,\overline{\psi}_f\gamma^\mu\underbrace{\left[\frac{1}{2}\left(1-\gamma^5\right)T_3 - \sin^2\theta_w Q\right]}_{\frac{1}{2}\left(C_V^f - C_A^f\gamma^5\right)}\psi_f\cdot Z_\mu \end{aligned}$$

which we can represent with the following vertex and Feynman rule



$$-i\frac{g}{\cos\theta_w}\gamma^\mu\frac{1}{2}\left(C_V^f - C_A^f\gamma^5\right)$$

For the coefficients of the left- and right-handed couplings we can readily find

$$\begin{aligned} C_L^f &= T_3^f - Q^f\sin^2\theta_w \\ C_R^f &= -Q^f\sin^2\theta_w \end{aligned} \tag{11.50}$$

which implies for the vector and axial vector couplings

$$\boxed{\begin{aligned} C_V^f &= T_3^f - 2Q^f\sin^2\theta_w \\ C_A^f &= T_3^f \end{aligned}} \tag{11.51}$$

Table 11.1 lists the quantum numbers and resulting couplings for all fermions in the standard model. The model can be experimentally scrutinized by measuring all these couplings.

fermion	T_3	Q	Y	C_A^f	C_V^f
$\nu_e\ \nu_\mu\ \nu_\tau$	$+\frac{1}{2}$	0	-1	$\frac{1}{2}$	$\frac{1}{2}$
$e\ \ \mu\ \ \tau$	$-\frac{1}{2}$	-1	-1	$-\frac{1}{2}$	$-\frac{1}{2}+2\sin^2\theta_w$
$u\ \ c\ \ t$	$+\frac{1}{2}$	$+\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}-\frac{4}{3}\sin^2\theta_w$
$d\ \ s\ \ b$	$-\frac{1}{2}$	$-\frac{1}{3}$	$\frac{1}{3}$	$-\frac{1}{2}$	$-\frac{1}{2}+\frac{2}{3}\sin^2\theta_w$

Table 11.1: The neutral current vector and axial vector couplings for each of the left-handed fermions in the Standard Model.

Exercise 37.

Show how we get from Eq. 11.15 to Eq. 11.17.

Exercise 38.

Show explicitly that:

$$\overline{\psi}\,\gamma^\mu\,\psi = \overline{\psi}_L\,\gamma^\mu\,\psi_L + \overline{\psi}_R\,\gamma^\mu\,\psi_R \tag{11.52}$$

making use of $\psi = \psi_L + \psi_R$ *and the projection operators* $\frac{1}{2}\,(1-\gamma_5)$ *and* $\frac{1}{2}\,(1+\gamma_5)$

Exercise 39.

What do you think is the difference between an exact and a broken symmetry?
Can you make a (wild) guess what spontaneous symmetry breaking means?
Which symmetry is involved in the gauge theories below? Which of these gauge symmetries are exact? Why/Why not?

(a) U1(Q) symmetry

(b) SU2(u-d-flavour) symmetry

(c) SU3(u-d-s-flavour) symmetry

(d) SU6(u-d-s-c-b-t) symmetry

(e) SU3(colour) symmetry

(f) SU2(weak-isospin) symmetry

(f) SU5(Grand unified) symmetry

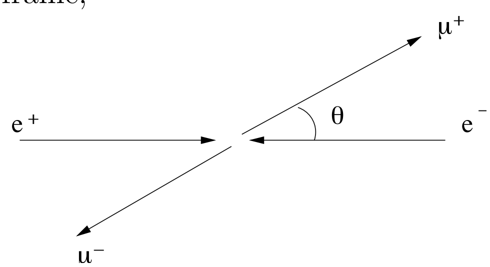
(g) SUSY

Lecture 12

The Process $e^-e^+ \to \mu^-\mu^+$

12.1 The cross section of $e^-e^+ \to \mu^-\mu^+$

Equipped with the Feynman rules of the electroweak theory we proceed to calculate 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^+.

There are two Feynman diagrams that contribute to the process:

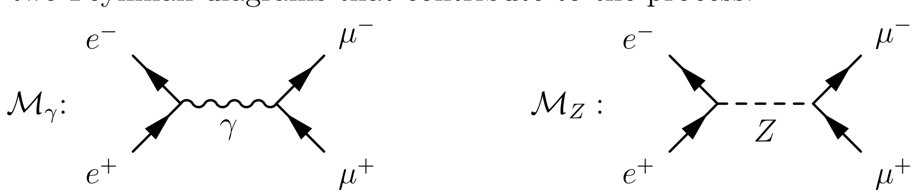


Figure 12.1: Feynman diagrams contributing to $e^-e^+ \to \mu^-\mu^+$

In complete analogy with the calculation of the QED process $e^+e^- \to e^+e^-$ we obtain the cross section using Fermi's Golden rule:

$$d\sigma = \frac{\overline{|\mathcal{M}|}^2}{F}\, dQ \tag{12.1}$$

with the phase factor dQ flux factor F given by

$$\begin{aligned} \mathrm{d}Q &= \frac{1}{4\pi^2}\,\frac{p_f}{4\sqrt{s}}\,\mathrm{d}\Omega \\ F &= 4p_i\sqrt{s} \end{aligned}$$

where p_i and p_f are the magnitudes of the electron and muon momentum vectors, respectively. Consequently, we have for the differential cross-section the usual formula

$$\frac{\mathrm{d}\sigma(e^-e^+ \to \mu^-\mu^+)}{\mathrm{d}\Omega} = \frac{1}{64\pi^2}\,\frac{1}{s}\,\frac{p_f}{p_i}\,\overline{|\mathcal{M}|}^2 \tag{12.2}$$

12.1.1 Propagators

Before we write down the matrix element we need to discuss the Z and W propagator, which we have omitted in the last lecture. The recipe for finding a propagator is to take a free-particle wave equation $P\psi = 0$ and then find the inverse of the operator P. We will spare you the details and just list the results:

- spinless complex field

$$(\Box + m^2)\phi = 0 \qquad \Longrightarrow \qquad \frac{i}{p^2 - m^2} \tag{12.3}$$

- spin-$\frac{1}{2}$ field

$$(\not{p} - m)\psi = 0 \qquad \Longrightarrow \qquad \frac{i}{\not{p} - m} = \frac{i(\not{p} + m)}{p^2 - m} = \frac{i\sum_{\text{spin}} u\overline{u}}{p^2 - m^2} \tag{12.4}$$

- massless spin-1 boson (in the Lorentz gauge)

$$(\Box - \partial^\mu\partial^\nu)A_\nu = 0 \qquad \Longrightarrow \qquad \frac{-\,i\,g^{\mu\nu}}{p^2} \tag{12.5}$$

- massive spin-1 boson

$$(\Box + m^2 - \partial^\mu\partial^\nu)A_\nu = 0 \qquad \Longrightarrow \qquad \frac{-\,i\;(g^{\mu\nu} - p^\mu p^\nu/m^2)}{p^2 - m^2} \tag{12.6}$$

Note that 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 we shall see later, this leads to an infinite cross-section when we tune the beam energies such that $\sqrt{s} = M_Z$, which is clearly non-physical.

In reality this does not happen due to the fact that the Z-particle itself decays and has an intrinsic decay width Γ_Z. We can account for the decay width by replacing the mass in the propagator with

$$M_Z \;\to\; M_Z - \frac{i}{2}\Gamma \tag{12.7}$$

where Γ is the total decay width of *real* on-shell Z-bosons. A heuristic explanation (Halzen and Martin, §2.10) goes as follows: An unstable particle decays following the exponential law

$$|\psi\,(t)|^2 \;=\; |\psi\,(0)|^2\; e^{-\Gamma t} \tag{12.8}$$

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

$$\psi(t) \;=\; \psi(0)e^{-imt}e^{-\Gamma t/2} \tag{12.9}$$

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.

12.1.2 Matrix elements

Using the Feynman rules (see *e.g.* appendix A) we find for the matrix element of the photon exchange

$$\mathcal{M}_\gamma \;=\; -e^2\left(\overline{\psi}_m\gamma^\mu\psi_m\right)\cdot\frac{g_{\mu\nu}}{q^2}\cdot\left(\overline{\psi}_e\gamma^\nu\psi_e\right) \tag{12.10}$$

while we have for the Z-boson exchange

$$\mathcal{M}_Z \;=\; -\frac{g^2}{4\cos^2\theta_w}\left[\overline{\psi}_m\gamma^\mu\left(C_V^m - C_A^m\gamma^5\right)\psi_m\right]\cdot\frac{g_{\mu\nu}-q_\mu q_\nu/M_Z^2}{q^2-M_Z^2}\cdot\left[\overline{\psi}_e\gamma^\nu\left(C_V^e - C_A^e\gamma^5\right)\psi_e\right] \tag{12.11}$$

We can simplify the Z propagator if we ignore the lepton masses ($m_\ell \ll \sqrt{s}$). In that case the Dirac equation becomes:

$$\overline{\psi}_e\;(i\partial_\mu\gamma^\mu - m) = 0 \quad\Rightarrow\quad \overline{\psi}_e\;(\gamma^\mu p_{\mu,e}) = 0 \tag{12.12}$$

Since $p_e = \frac{1}{2}q$ we also have:

$$\frac{1}{2}\,\overline{\psi}_e\;(\gamma^\mu q_\mu) = 0 \quad\Rightarrow\quad q_\mu q_\nu/M_z^2 = 0 \tag{12.13}$$

As a result the $q_\mu q_\nu$ in the propagator vanishes and we get for the matrix element

$$\mathcal{M}_Z = \frac{-g^2}{4\cos^2\theta_w}\,\frac{1}{q^2-M_Z^2}\cdot\left[\overline{\psi}_m\,\gamma^\mu\left(C_V^m - C_A^m\gamma^5\right)\;\psi_m\right]\left[\overline{\psi}_e\,\gamma_\mu\left(C_V^e - C_A^e\gamma^5\right)\;\psi_e\right]\quad . \tag{12.14}$$

To calculate the cross section by summing over $\mathcal{M}_\gamma$ and $\mathcal{M}_Z$ is now straightforward but a rather lengthy procedure: applying Casimir's trick, trace theorems, etc. Let us here try to follow a different approach.

12.1.3 Polarized and unpolarized cross-sections

We rewrite the $\mathcal{M}_Z$ matrix element in terms of right-handed and left-handed couplings, using the definitions: $C_R = C_V - C_A$; $C_L = C_V + C_A$. As before we have:

$$\left(C_V - C_A\gamma^5\right) = \left(C_V - C_A\right) \cdot \frac{1}{2}\left(1+\gamma^5\right) + \left(C_V + C_A\right) \cdot \frac{1}{2}\left(1-\gamma^5\right) \quad . \tag{12.15}$$

Thus:

$$\left(C_V - C_A\gamma^5\right)\psi = C_R\psi_R + C_L\psi_L \quad . \tag{12.16}$$

Let us now look back at the QED process:

$$\mathcal{M}_\gamma = \frac{-e^2}{s}\,\left(\overline{\psi}_m\gamma^\mu\psi_m\right)\,\left(\overline{\psi}_e\gamma_\mu\psi_e\right) \tag{12.17}$$

with (see previous lecture):

$$\begin{aligned} \left(\overline{\psi}_m\gamma^\mu\psi_m\right) &= \left(\overline{\psi}_{Lm}\,\gamma^\mu\,\psi_{Lm}\right) + \left(\overline{\psi}_{Rm}\,\gamma^\mu\,\psi_{Rm}\right) \\ \left(\overline{\psi}_e\gamma_\mu\psi_e\right) &= \left(\overline{\psi}_{Le}\,\gamma_\mu\,\psi_{Le}\right) + \left(\overline{\psi}_{Re}\,\gamma_\mu\,\psi_{Re}\right) \end{aligned}$$

The fact that there are no terms connecting L-handed to R-handed $(\overline{\psi_{Rm}}\ \gamma^\mu\ \psi_{Lm})$ actually implies that we have helicity conservation for high energies (i.e. neglecting $\sim m/E$ terms) at the vertices:

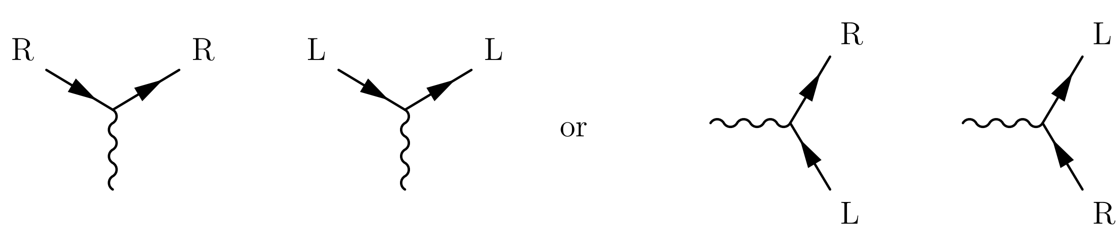


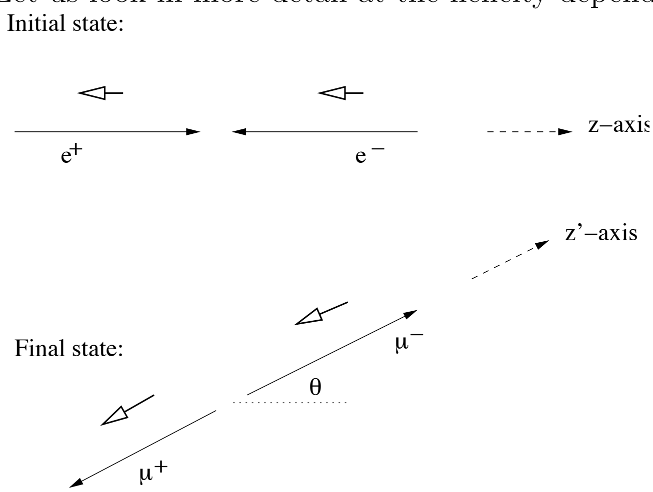
Figure 12.2: Helicity conservation. *left:* A right-handed incoming electron scatters into a right-handed outgoing electron and vice versa in a vector or axial vector interaction . *right:* In the crossed reaction the energy and momentum of one electron is reversed: i.e. in the e^+e^- pair production 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.)

A different way of saying this is that at high energy spin is a 'good' quantum number. Therefore, if we choose a basis with spin eigen functions, then different spin configurations do not interfere. As a consequence we can decompose the unpolarized QED scattering process as a sum of 4 cross section contributions (*Note:* $e^+_R \equiv \overline{\psi}_{Le}$ *etc.(!)*)

$$\begin{aligned} \frac{d\sigma}{d\Omega}^{\text{unpolarized}} = \frac{1}{4}\{ \; & \frac{d\sigma}{d\Omega}\left(e^-_L e^+_R \to \mu^-_L \mu^+_R\right) + \frac{d\sigma}{d\Omega}\left(e^-_L e^+_R \to \mu^-_R \mu^+_L\right) \\ & \frac{d\sigma}{d\Omega}\left(e^-_R e^+_L \to \mu^-_L \mu^+_R\right) + \frac{d\sigma}{d\Omega}\left(e^-_R e^+_L \to \mu^-_R \mu^+_L\right) \} \end{aligned}$$

where we *average* over the incoming spins and *sum* over the final state spins.

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^+.

So, 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 then given by the rotation matrices[1]

$$d^j_{m'm}(\theta) \equiv \left\langle jm' | e^{-i\theta J_y} | jm \right\rangle \tag{12.19}$$

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 also called 'Wigner d-matrices'. Computing them is not entirely trivial, so we take the elements that we need from a table (see *e.g.* H&M exercise 2.6),

$$\begin{aligned} d^1_{1\,1}(\theta) &= d^1_{-1-1}(\theta) = \frac{1}{2}\left(1 + \cos\theta\right) \\ d^1_{1-1}(\theta) &= d^1_{-1\,1}(\theta) = \frac{1}{2}\left(1 - \cos\theta\right) \end{aligned}$$

[1]See H&M§2.2:

$$e^{-i\theta J_2}\,|j\;m\rangle = \sum_{m'} d^j_{m\,m'}(\theta)\,|j\;m'\rangle \tag{12.18}$$

and also appendix H in Burcham & Jobes

Using this result, we find considering only the electromagnetic contribution

$$\begin{aligned}\frac{d\sigma}{d\Omega}\left(e^-_L e^+_R \to \mu^-_L \mu^+_R\right) &= \frac{\alpha^2}{4s}\,(1+\cos\theta)^2 = \frac{d\sigma}{d\Omega}\left(e^-_R e^+_L \to \mu^-_R \mu^+_L\right)\\ \frac{d\sigma}{d\Omega}\left(e^-_L e^+_R \to \mu^-_R \mu^+_L\right) &= \frac{\alpha^2}{4s}\,(1-\cos\theta)^2 = \frac{d\sigma}{d\Omega}\left(e^-_R e^+_L \to \mu^-_L \mu^+_R\right)\end{aligned}$$

Indeed the unpolarised cross section is obtained as the spin-averaged sum over the allowed helicity combinations (*see lecture 8*): $\frac{1}{4}\cdot[(1)+(2)+(3)+(4)] =$

$$\frac{d\sigma}{d\Omega}^{\text{unpol}} = \frac{1}{4}\,\frac{\alpha^2}{4s}\,2\,\left[(1+\cos\theta)^2 + (1-\cos\theta)^2\right] = \frac{\alpha^2}{4s}\left(1+\cos^2\theta\right) \quad . \tag{12.20}$$

Now we go back to the γ, Z scattering. We have the individual contributions of the helicity states, so let us compare the expressions for the matrix-elements $\mathcal{M}_\gamma$ and $\mathcal{M}_Z$:

$$\mathcal{M}_\gamma = -\frac{e^2}{s}\,\left[\,\left(\overline{\psi}_{Lm}\,\gamma^\mu\,\psi_{Lm}\right) + \left(\overline{\psi}_{Rm}\,\gamma^\mu\,\psi_{rm}\right)\,\right]\cdot\left[\,\left(\overline{\psi}_{Le}\,\gamma_\mu\,\psi_{Le}\right) + \left(\overline{\psi}_{Re}\,\gamma_\mu\,\psi_{Re}\right)\,\right] \tag{12.21}$$

$$\begin{aligned}\mathcal{M}_Z = -\frac{g^2}{4\cos^2\theta_w}\,\frac{1}{s-M_Z^2}\,&\left[\,C_L^m\left(\overline{\psi}_{Lm}\,\gamma^\mu\,\psi_{Lm}\right) + C_R^m\left(\overline{\psi}_{Rm}\,\gamma^\mu\,\psi_{rm}\right)\,\right]\\ &\cdot\left[\,C_L^e\left(\overline{\psi}_{Le}\,\gamma_\mu\,\psi_{Le}\right) + C_R^e\left(\overline{\psi}_{Re}\,\gamma_\mu\,\psi_{Re}\right)\,\right]\end{aligned} \tag{12.22}$$

We denote the ingoing and outgoing particles as $e^-_{L,R}(p) \equiv \psi_{L,R_e}(p)$, $e^+_{L,R}(p) \equiv \psi_{R,L_e}(-p)$, $\mu^-_{L,R}(p) \equiv \psi_{L,R_m}(p)$, $\mu^+_{L,R}(p) \equiv \psi_{R,L_m}(-p)$. Since the helicity processes do not interfere, we can see (Exercise 40 (a)) that

$$\frac{d\sigma}{d\Omega}_{\gamma,Z}\left(e^-_L e^+_R \to \mu^-_L \mu^+_R\right) = \frac{\alpha^2}{4s}\,(1+\cos\theta)^2\cdot|1 + r\,C_L^m C_L^e|^2 \tag{12.23}$$

$$\frac{d\sigma}{d\Omega}_{\gamma,Z}\left(e^-_L e^+_R \to \mu^-_R \mu^+_L\right) = \frac{\alpha^2}{4s}\,(1-\cos\theta)^2\cdot|1 + r\,C_R^m C_L^e|^2 \tag{12.24}$$

with the relative contribution of the Z and γ parameterized as

$$r = \frac{g^2}{e^2}\,\frac{1}{4\cos^2\theta_w}\,\frac{s}{s-M_z^2} = \frac{\sqrt{2}G_F M_Z^2}{e^2}\,\frac{s}{s-M_Z^2} \quad . \tag{12.25}$$

where we used that

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2} = \frac{g^2}{8M_Z^2\cos^2\theta_w} \quad . \tag{12.26}$$

Similar expressions hold for the other two helicity configurations. Note that, as discussed before, the result is not finite because of the pole at the Z mass. With the replacement $M_Z \to M_Z - i\Gamma_Z/2$ in the propagator, the value of r becomes

$$r = \frac{\sqrt{2}G_F M_Z^2}{e^2}\,\frac{s}{s-\left(M_Z - i\frac{\Gamma_Z}{2}\right)^2} \tag{12.27}$$

The total unpolarized cross section finally is the average over the four L, R helicity combinations. Inserting "lepton universality" $C_L^e = C_L^\mu$; $C_R^e = C_R^\mu$ and therefore also: $C_V^e = C_V^\mu$; $C_A^e = C_A^\mu$, the expression becomes (by writing it out):

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\alpha^2}{4s} \left[A_0 \left(1+\cos^2\theta\right) + A_1 \left(\cos\theta\right)\right] \tag{12.28}$$

with

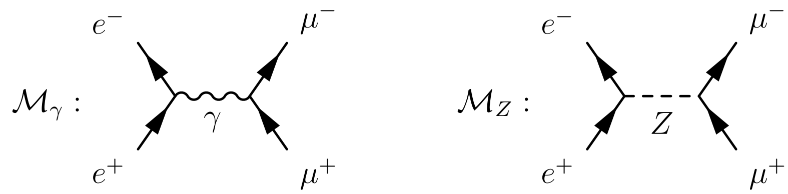
$$\begin{aligned} A_0 &= 1 + 2\,Re(r)\,C_V^2 + |r|^2 \left(C_V^2 + C_A^2\right)^2 \\ A_1 &= 4\,Re(r)\,C_A^2 + 8|r|^2 C_V^2 C_A^2 \end{aligned}$$

In the Standard Model we have: $C_A = -\frac{1}{2}$ and $C_V = -\frac{1}{2} + 2\sin^2\theta_w$. If we allow for different couplings in the final state, the general expression for $e^-e^+ \to \gamma, Z \to f\bar{f}$ can be obtained with

$$A_0 = 1 + 2Re(r)\,C_V^e C_V^f + |r|^2 \left({C_V^e}^2 + {C_A^e}^2\right)\left({C_V^f}^2 + {C_A^f}^2\right) \tag{12.29}$$

$$A_1 = 4Re(r)\,C_A^e C_A^f + 8|r|^2\,C_V^e C_V^f C_A^e C_A^f \tag{12.30}$$

To summarize, on the *amplitude level* there are two diagrams that contribute:



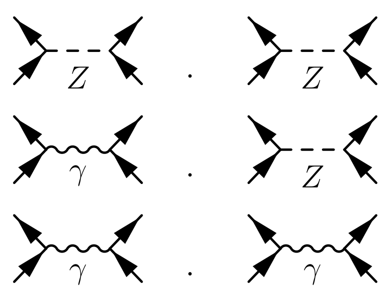
(12.31)

Using the following notation:

$\frac{d\sigma}{d\Omega}[Z,Z] =$ ... $\propto |r|^2$

$\frac{d\sigma}{d\Omega}[\gamma Z] =$ ... $\propto Re(r)$

$\frac{d\sigma}{d\Omega}[\gamma,\gamma] =$ ... $\propto 1$



the expression for the differential cross-section becomes

$$\frac{d\sigma}{d\Omega} = \frac{d\sigma}{d\Omega}[\gamma,\gamma] + \frac{d\sigma}{d\Omega}[Z,Z] + \frac{d\sigma}{d\Omega}[\gamma,Z]$$

with

$$\frac{d\sigma}{d\Omega}[\gamma,\gamma] = \frac{\alpha^2}{4s}\left(1+\cos^2\theta\right)$$

$$\frac{d\sigma}{d\Omega}[Z,Z] = \frac{\alpha^2}{4s}|r|^2 \left[\left({C_V^e}^2 + {C_A^e}^2\right)\left({C_V^f}^2 + {C_A^f}^2\right)\left(1+\cos^2\theta\right) + 8C_V^e C_V^f C_A^e C_A^f \cos\theta\right]$$

$$\frac{d\sigma}{d\Omega}[\gamma,Z] = \frac{\alpha^2}{4s} Re|r| \left[C_V^e C_V^f \left(1+\cos^2\theta\right) + 2C_A^e C_A^f \cos\theta\right]$$

12.1.4 Near the resonance

Let us take a look at the cross section close to the peak of the distribution:

$$r \propto \frac{s}{s - \left(M_z - i\frac{\Gamma_Z}{2}\right)^2} = \frac{s}{s - \left(M_z^2 - \frac{\Gamma_Z^2}{4}\right) + i\mathcal{M}_Z\Gamma_Z} \tag{12.32}$$

The peak is located at $s_0 = M_Z^2 - \frac{\Gamma_Z^2}{4}$.

In Exercise 40 (b) we show that:

$$Re(r) = \left(1 - \frac{s_0}{s}\right)|r|^2 \quad \text{with} \quad |r|^2 = \frac{s^2}{\left(s - \left(M_Z^2 - \frac{\Gamma_Z^2}{4}\right)\right)^2 + M_Z^2\Gamma_Z^2} \tag{12.33}$$

This shows that the interference term is 0 at the peak.

In that case (i.e. at the peak) we have for the cross section terms:

$$\begin{aligned} A_0 &= 1 + |r|^2 \left(C_V^{e\,2} + C_A^{e\,2}\right)\left(C_V^{f\,2} + C_A^{f\,2}\right) \\ A_1 &= 8|r|^2 \left(C_V^e\, C_A^e\, C_V^f\, C_A^f\right) \end{aligned}$$

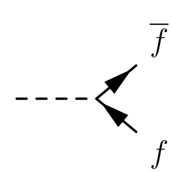
The total cross section (integrated over $d\Omega$) is then:

$$\sigma(s) = \frac{G_F^2 M_Z^4}{\left(s - \left(M_Z^2 - \frac{\Gamma_Z^2}{4}\right)\right)^2 + M_Z^2\Gamma_Z^2} \cdot \frac{s}{6\pi}\left(C_V^{e\,2} + C_A^{e\,2}\right)\left(C_V^{f\,2} + C_A^{f\,2}\right) \quad . \tag{12.34}$$

12.2 The Z^0 decay widths

We can also calculate the decay width:

$$\Gamma\left(Z \to f\overline{f}\right) \tag{12.35}$$



which is according Fermi's golden rule:

$$\begin{aligned} \Gamma\left(Z \to f\overline{f}\right) &= \frac{1}{16\pi}\frac{1}{M_Z}\left|\overline{\mathcal{M}}\right|^2 \\ &= \frac{g^2}{48\pi}\frac{M_z}{\cos^2\theta_w}\left(C_V^{f\,2} + C_A^{f\,2}\right) \\ &= \frac{G_F}{6\sqrt{2}}\frac{M_Z^3}{\pi}\left(C_V^{f\,2} + C_A^{f\,2}\right) \end{aligned}$$

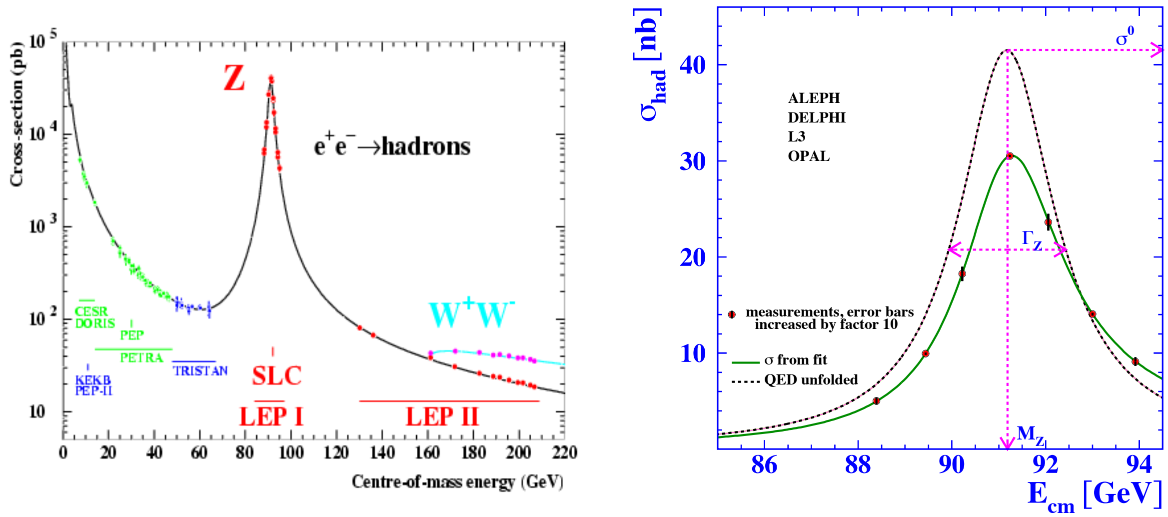


Figure 12.3: *left:* The Z-lineshape: the cross-section for $e^+e^- \rightarrow$ hadrons as a function of $\sqrt{s}$. *right:* Same but now near the resonance. The Lineshape parameters for the lowest order calculations and including higher order corrections.

Using this expression for $\Gamma_e \equiv \Gamma(Z \rightarrow e^+e^-)$ and $\Gamma_f \equiv \Gamma(Z \rightarrow f\overline{f})$ we can re-write:

$$\sigma(s) = \frac{12\pi}{M_Z^2} \cdot \frac{s}{\left(s - \left(M_Z^2 - \frac{\Gamma_Z^2}{4}\right)\right)^2 + M_Z^2\Gamma_Z^2} \cdot \Gamma_e\Gamma_f \quad . \tag{12.36}$$

Close to the peak we then find:

$$\sigma_{peak} \approx \frac{12\pi}{M_Z^2}\frac{\Gamma_e\Gamma_f}{\Gamma_Z^2} = \frac{12\pi}{M_Z^2} BR(Z \rightarrow ee) \cdot BR(Z \rightarrow ff) \tag{12.37}$$

Let us now finally consider the case when $f = q$ (a quark). Due to the fact that quarks can be produced in 3 colour states the decay width is:

$$\Gamma(Z \rightarrow \overline{q}q) = \frac{G_F}{6\sqrt{2}}\frac{M_Z^3}{\pi}\left({C_V^f}^2 + {C_A^f}^2\right) \cdot N_C \tag{12.38}$$

with the colour factor $N_C = 3$. The ratio between the hadronic and leptonic width: $R_l = \Gamma_{had}/\Gamma_{lep}$ can be defined. This ratio can be used to test the consistency of the standard model by comparing the calculated value with the observed one.

12.3 Forward-backward asymmetry

A direct consequence of the photon-Z interference is that the angular distribution is not symmetric. Figure 12.4 shows the $\cos\theta$ distribution observed at the Jade experiment,

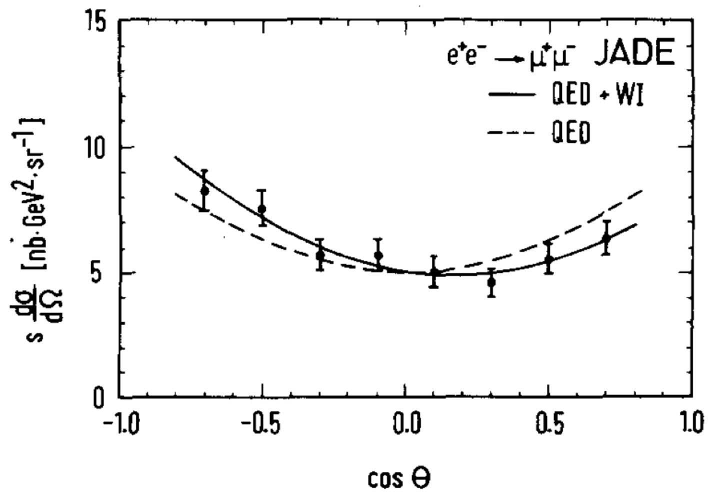


Figure 12.4: Angular distribution for $e^+e^- \to \mu^+\mu^-$ for $\sqrt{s} > 25$ GeV at the JADE experiment. θ is the angle between the outgoing μ^+ and the incoming e^+. The curves show fits to the data $p(1+\cos^2\theta) + q\cos\theta$ (full curve) and $p(1+\cos^2\theta)$ (dashed curve). (Source: JADE collaboration, PLB, Vol108B, p108, 1981.)

operating at the PETRA collider in Hamburg. The beam energy in this experiment was not yet sufficient to directly produce Z bosons. Still, the effect of the interference was clearly visible long before the direct discovery of the Z resonance.

The forward-backward asymmetry can be defined using the polar angle distribution. At the peak and ignoring the pure photon exchange (because it is negligibly small)

$$\frac{d\sigma}{d\cos\theta} \propto 1 + \cos^2\theta + \frac{8}{3}A_{FB}\cos\theta \qquad (12.39)$$

where

$$A_{FB}^{0,f} = \frac{3}{4}A_e A_f \qquad \text{with} \qquad A_f = \frac{2C_V^f C_A^f}{C_V^2 + C_A^2}\ . \qquad (12.40)$$

The precise measurements of the forward-backward asymmetry can be used to determine the couplings C_V and C_A.

12.4 The number of light neutrinos

Since the total decay width of the Z must be equal to the sum of all partial widths the following relation holds:

$$\Gamma_Z = \Gamma_{ee} + \Gamma_{\mu\mu} + \Gamma_{\tau\tau} + 3\Gamma_{uu} + 3\Gamma_{dd} + 3\Gamma_{ss} + 3\Gamma_{cc} + 3\Gamma_{bb} + N_\nu \cdot \Gamma_{\nu\nu} \qquad (12.41)$$

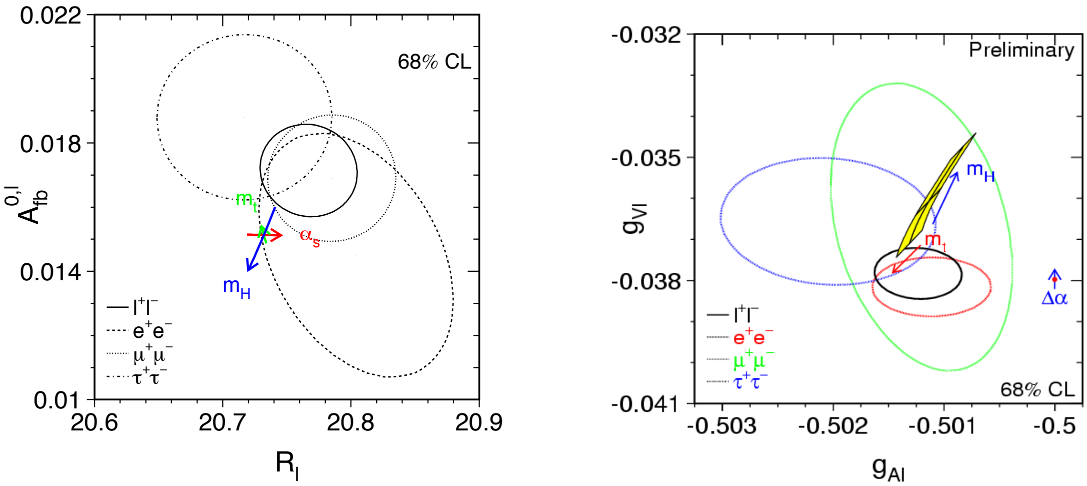


Figure 12.5: *left:* Test of lepton-universality. The leptonic A_{fb} vs. R_l. The contours show the measurements while the arrows show the dependency on Standard Model parameters. *right:* Determination of the vector and axial vector couplings.

Using all available data to extract information on the couplings we can now determine the decay widths to all final states within the standard model,

$$\begin{aligned}
\Gamma_{ee} &\approx \Gamma_{\mu\mu} \approx \Gamma_{\tau\tau} = 84 \text{ MeV} & C_V \approx 0 \qquad & C_A = -\frac{1}{2} \\
\Gamma_{\nu\nu} &= 167 \text{ MeV} & C_V = \frac{1}{2} \qquad & C_A = \frac{1}{2} \\
\Gamma_{uu} &\approx \Gamma_{cc} = 276 \text{ MeV} & C_V \approx 0.19 \qquad & C_A = \frac{1}{2} \\
\Gamma_{dd} &\approx \Gamma_{ss} \approx \Gamma_{bb} = 360 \text{ MeV} & C_V \approx -0.35 \qquad & C_A = -\frac{1}{2}
\end{aligned}$$

(Of course $\Gamma_{tt} = 0$ since the top quark is heavier than the Z.) A measurement of the lineshape (the cross-section as function of $\sqrt{s}$) gives for the total decay width of the Z,

$$\Gamma_Z \approx 2490 \text{ MeV}$$

So, even though we cannot see the neutrino contribution, we can estimate the number of neutrinos from the total width of the Z. The result is (see figure 12.6),

$$N_\nu = \frac{\Gamma_Z - 3\Gamma_l - \Gamma_{had}}{\Gamma_{\nu\nu}} = 2.984 \pm 0.008 \, . \tag{12.42}$$

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 weakly. In either case, this generation would be very different from the known generations of quarks and leptons.

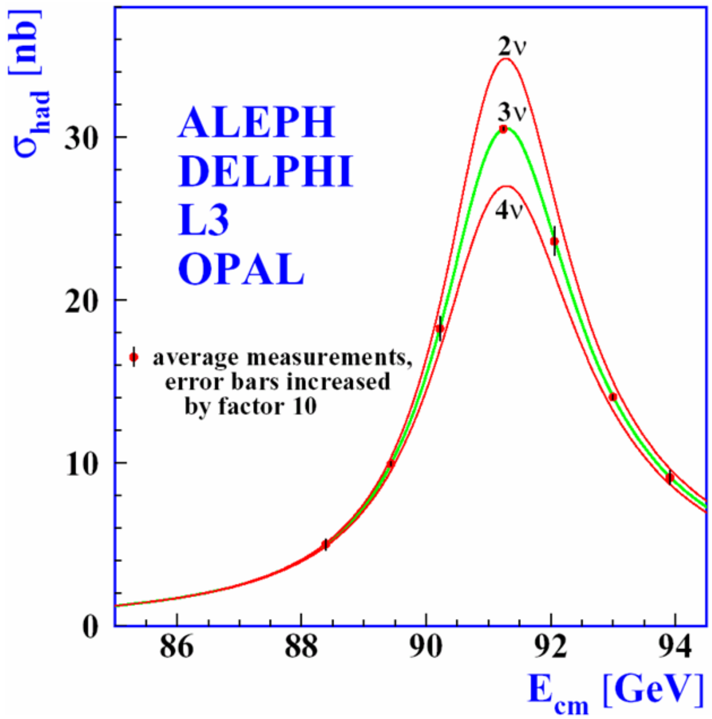


Figure 12.6: The Z-lineshape for resp. $N_\nu = 2,\ 3,\ 4$.

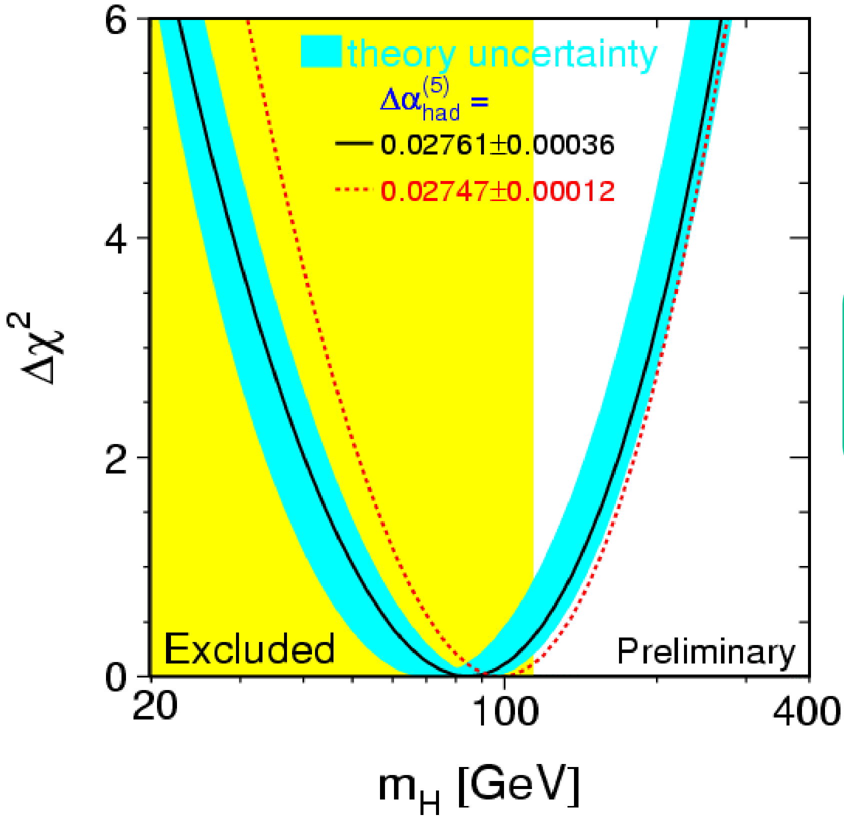


Figure 12.7: Standard Model fit of the predicted value of the Higgs boson.

Exercise 40.

(a) *Show how the unpolarised cross section formula for the process $e^+e^- \to Z, \gamma \to \mu^+\mu^-$ can be obtained from the expression of the helicity cross sections in the lecture:*

$$\frac{d\sigma}{d\Omega}\left(e^-_{L/R}e^+_{R/L} \to \mu^-_{L/R}\mu^+_{R/L}\right) = \frac{\alpha^2}{4s}\,(1 \pm \cos\theta)^2\,\left|1 + r\,C^e_{L/R}C^\mu_{L/R}\right|^2$$

(b) *Show, using the expression of r from the lecture, that close to the peak of the Z-lineshape the expression*

$$Re(r) = \left(1 - \frac{s_0}{s}\right)\,|r|^2$$

with $s_0 = M_z^2 - \Gamma_z^2/4$ holds.

(c) *Show also that at the peak:*

$$\sigma_{peak} \approx \frac{12\pi}{M_z^2}\frac{\Gamma_e\Gamma_\mu}{\Gamma_Z^2}$$

(d) *Calculate the relative contribution of the Z-exchange and the γ exchange to the cross section at the Z peak.*
Use $\sin^2\theta_W = 0.23$, $M_z = 91\,GeV$ and $\Gamma_Z = 2.5\,GeV$.

(e) *The actual line shape of the Z-boson is not a pure Breit Wigner, but it is asymmetrical: 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?

(f) *The number of light neutrino generations is determined from the "invisible width" of the Z-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 $Z \to \nu\bar{\nu}$ directly?
Do you think this method is more precise or less precise?

Appendix A

Summary of electroweak theory

Take the Lagrangian of free fermions (leptons and quarks)

$$\mathcal{L} \;=\; \sum_f \overline{\psi_f}(i\gamma^\mu\partial_\mu - m)\psi_f \tag{A.1}$$

Arrange the left-handed projections of the lepton and quark fields in doublets

$$\Psi_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \qquad \text{or} \qquad \Psi_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \tag{A.2}$$

Ignore their masses (or choose them equal within the doublet). Now consider that the Lagrangian remains invariant under

$U(1)_Y$:

$$\psi \to \psi' = e^{iY\beta(x)}\psi \tag{A.3}$$

$SU(2)_L$:

$$\Psi_L \to \Psi'_L \;=\; \exp\left[iY\vec{\alpha}(x)\cdot\vec{\tau}\right]\,\Psi_L \tag{A.4}$$

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

$$\partial_\mu \to D_\mu = \partial_\mu + ig'\frac{Y}{2}a_\mu \tag{A.5}$$

$SU(2)_L$: weak isospin fields b^1_μ, b^2_μ, b^3_μ (only couple to left-handed doublet):

$$\partial_\mu \to D_\mu = \partial_\mu + ig\vec{\tau}\cdot\vec{b}_\mu \tag{A.6}$$

Ignoring the kinetic and self-coupling terms of the gauge fields, the Lagrangian becomes

$$\mathcal{L} \;=\; \mathcal{L}^{\text{free}} - i\frac{g'}{2}J_Y^{\mu}a_{\mu} - ig\vec{J}_L^{\mu}\cdot\vec{b}_{\mu} \tag{A.7}$$

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"

$$W_{\mu}^{\pm} \;=\; \frac{b_{\mu}^1 \mp ib_{\mu}^2}{\sqrt{2}} \tag{A.8}$$

"neutral currents"

$$\begin{aligned} Z_{\mu} &= -a_{\mu}\sin\theta_w + b_{\mu}^3\cos\theta_w \\ A_{\mu} &= a_{\mu}\cos\theta_w + b_{\mu}^3\sin\theta_w \end{aligned} \tag{A.9}$$

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

$$\tan\theta_w \;=\; \frac{g'}{g} \tag{A.10}$$

The coupling of the massless field becomes proportional to a charge

$$Q = T_3 + \tfrac{1}{2}Y \tag{A.11}$$

Furthermore, the W and Z masses obey the relation

$$M_W \;=\; M_Z\cos\theta \tag{A.12}$$

The interaction Lagrangian for the doublet can now be written as

$$\begin{aligned} \mathcal{L}^{\text{int}} \;=& -\frac{g}{\sqrt{2}}\overline{\psi}_u\gamma^{\mu}\tfrac{1}{2}\left(1-\gamma^5\right)\psi_d\,W_{\mu}^{+} \\ & -\frac{g}{\sqrt{2}}\overline{\psi}_d\gamma^{\mu}\tfrac{1}{2}\left(1-\gamma^5\right)\psi_u\,W_{\mu}^{-} \\ & -e\left[\sum_{f=u,d}Q_f\overline{\psi}_f\gamma^{\mu}\psi_f\right]A_{\mu} \\ & -g_z\left[\sum_{f=u,d}\overline{\psi}_f\gamma^{\mu}\tfrac{1}{2}\left(C_V^f - C_A^f\gamma^5\right)\psi_f\right]Z_{\mu} \end{aligned} \tag{A.13}$$

with

$$e = g\sin\theta_w \qquad g_z = g/\cos\theta_w \qquad \tan\theta_w = g'/g \tag{A.14}$$

and

$$C_V^f = T_3^f - 2Q^f\sin^2\theta_w \qquad C_A^f = T_3^f \tag{A.15}$$

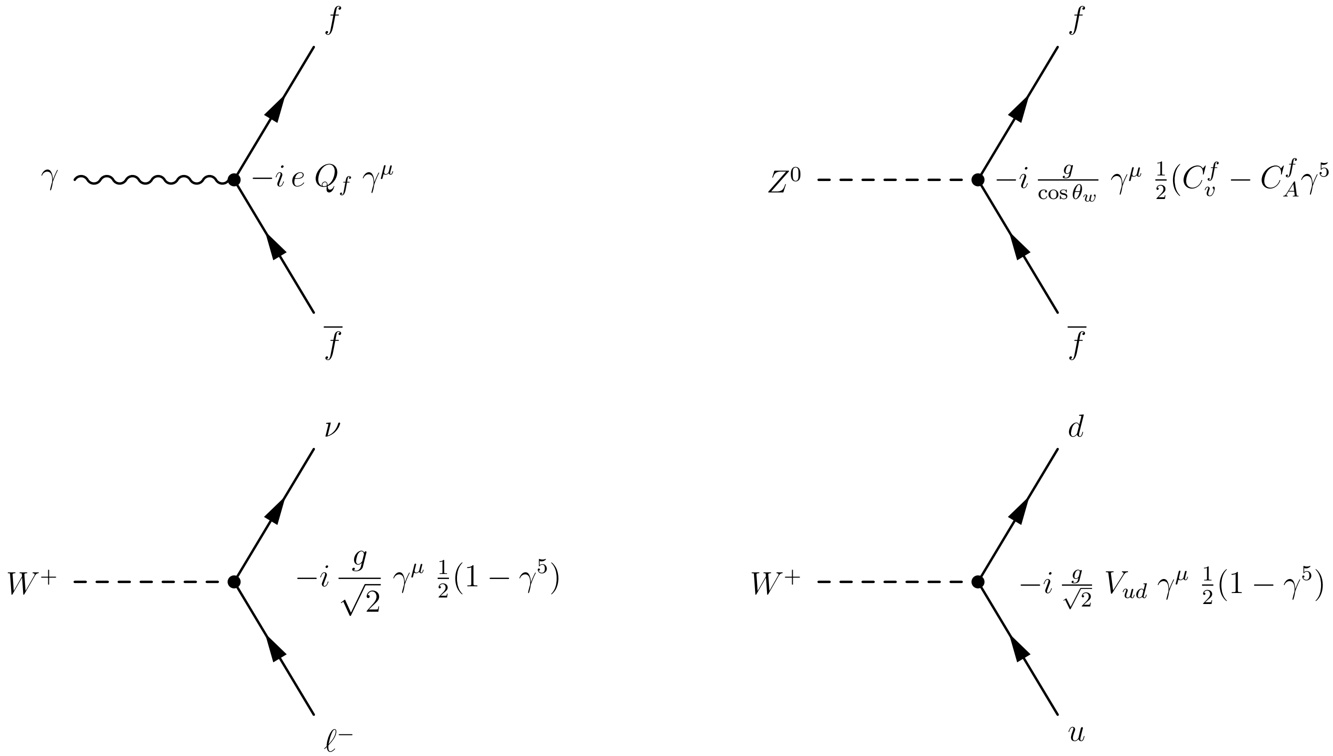
The relevant quantum numbers for our fields are

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

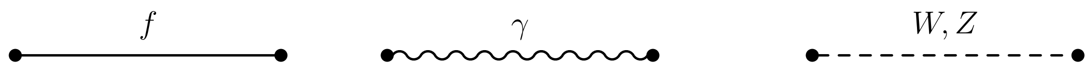
Till now we have ignored that the weak interaction mixes the quark fields. Inserting the CKM matrix we get for the charged currents,

$$\begin{aligned}\mathcal{L}^{\text{C.C.}} &= -\,i\frac{g}{\sqrt{2}}\,V_{ud}\,\overline{\psi}_u\gamma^\mu\tfrac{1}{2}\left(1-\gamma^5\right)\psi_d\,W_\mu^+ \\ &\quad -\,i\frac{g}{\sqrt{2}}\,V_{ud}^*\,\overline{\psi}_d\gamma^\mu\tfrac{1}{2}\left(1-\gamma^5\right)\psi_u\,W_\mu^-\end{aligned} \tag{A.16}$$

The Feynman rules for the vertex factors are then as follows



while those for the propagators are



$$\frac{i}{\not{p} - m} \qquad \frac{-ig^{\mu\nu}}{p^2} \qquad \frac{-i(g^{\mu\nu} - p^\mu p^\nu/m^2)}{p^2 - m^2}$$

The photon propagator is not unique: the form above holds in the Lorentz gauge.

Appendix B

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.