1

An introduction to the Standard Model of Particle Physics

Michel H.G. Tytgat
Research Director FNRS
Professor ULB

May 10, 2021

[1]MA1 - February 2021 - Université Libre de Bruxelles

Contents

Chapter 1

Introduction

Key notions: basic vocabulary, dimensional arguments, natural units

This is a course on the Standard Model of particle physics. This is much more than a model. The SM it is the best theory of Nature at the microscopic level we have for the moment. Even if a more fundamental theory is elaborated in the future, the SM will remain, very much like the Maxwell equations have not disappeared from our text books. The SM describes, in a unified framework, three fundamental interactions: electromagnetism (EM, the theory of light, electricity and magnetism), weak interactions (which are responsible for β radioactivity) and strong interactions (that underlies the structure of proton, neutrons, which turned out to be made of quarks, and, also (in principle, the details are not fully understood), how protons and neutrons hold together in nuclei. There is a fourth interaction, gravity, but there is not yet a quantum theory of gravity, so this force is a bit on the side.

Like for gravity, you are familiar with several classical aspects of EM, which are valid at large distances $\sim$ macroscopic scales. The manifestations of weak and strong interactions are however intrinsically microscopic, so we need a theory that encompasses quantum mechanics to understand them. The Standard Model (SM) is that theory. It structure is amazingly simple, as you will see.

There are three similar families of quarks and leptons. This is the matter sector of the SM. Quarks and leptons interact through the exchange of spin-1 particles, the gauge bosons. Gauge interactions are determined by the *principle of gauge invariance*, or invariance of a theory under local symmetry transformations. There are theorems stating that the only consistent (*i.e.* unitary and renormalizable) theories of interacting spin-1 fields are gauge theories. This is advanced material, which we will not cover here. But these theorems are clearly supported by data. The Lagrangian of the SM is invariant under the *gauge group*

$$\mathrm{SU}(3)_c \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$$

Symmetry is central to the structure of the SM, so we will need revisit some aspects of group theory. Symmetries of space-time play also a critical role, so we will also look more closely at the representations of the Lorentz group. It turns out that left-handed leptons are charged only under $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$, while the right-handed leptons only interact with the $\mathrm{U}(1)_Y$ gauge bosons. (See table I.) The SM is thus a *chiral theory* and the electroweak interactions violate parity. This was a big surprise and remains one of the most intriguing aspects of the SM. This being said, leptons have weak interactions and the charged ones are also coupled to electromagnetism. So the SM allows to describe a process like

$$e^+ + e^- \to \mu^+ + \mu^-$$

or the decay of a muon,

$$\mu \to e + \bar{\nu}_e + \nu_\mu$$

There is a similar pattern for the quarks, which makes hadrons (baryons and mesons). Hence the SM allows (in principle, there are complications compared to muons) to explain the decay of Lambda baryon,

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

Now a bit of estimates. For this, we use natural units, with $\hbar = c = 1$, see Appendix A. The lifetime of the Λ is very long, about 10^{-10} s, if we compare it with another baryon of similar mass, say the Δ, with

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

for which the lifetime is of the order of 10^{-24} s. The ratio of decay rates is thus

$$\frac{\Gamma(\Delta)}{\Gamma(\Lambda)} \approx 10^{14}$$

On dimensional grounds, we expect these decay rate to be proportional to the particle mass, time some effective coupling parameter. Let us include for simplicity the phase space suppression factors, which are the same for both particles, into these couplings, so $\Gamma \sim (\alpha\, M)$. Then

$$\frac{\Gamma(\Delta)}{\Gamma(\Lambda)} \approx \frac{\alpha_\Delta}{\alpha_\Lambda} \sim 10^{14}$$

From Appendix A, we have that $\tau \sim 10^{-24}$ s $\sim (10\,\text{GeV})^{-1}$, so we estimate that $\alpha_\Delta \sim 1$. The smallness of $\alpha_\Lambda \sim 10^{-14}$ is what we mean by weak interactions, an interaction driven by a very small effective couplings. The decay of the Δ is driven instead by strong interactions (i.e. large effective coupling). Notice that we have used α to suggest a similarity with the fine structure constant,

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \approx \frac{1}{137} \sim 10^{-2}$$

On that basis, electromagnetic interactions are stronger than weak interactions but weaker than strong interactions.[1]

Looking forward, strong interaction are a property of quarks which, on top of weak and electromagnetic interactions, are also charged under the "color group" $SU(3)_c$ of quantum chromodynamics (QCD). The corresponding gauge bosons are the gluons. Unfortunately, we will say little about QCD here except for some general facts (as we did for the Δ decay) or emphasizing fundamental aspects that are shared by all gauge groups. Unlike the electroweak interactions,

[1]This is true a low energies. At very high energy, the strength of the weak, electromagnetic and strong interactions converge toward each others. This suggest that these three forces may be unified. See section 7.3.

the strong interactions are *vector-like*: both fermion chiralities interact with the same strength to gauge field. In technical term, they transform under the same representation of the gauge group. $SU(3)_c$ interactions exhibits the remarkable properties of asymptotic freedom in the UV (high energies) and, conversely, they confine in the IR (low energies). This lead to binding of quarks and gluons into colorless hadrons, and, also, spontaneous chiral symmetry breaking. Whether the two are correlated is not yet fully understood. This is because such effects are non-perturbative and also arise in a regime in which the theory is strongly interacting, meaning that calculation can not be done using perturbation theory. There are approaches to circumvent these problems (lattice calculations notably, but not only), but these are more advanced topics. So, unfortunately, these beautiful but difficult topics are beyond the scope of the present lectures. Instead, we will focus on electroweak interactions and the breaking of the $SU(2) \times U(1)$ symmetry through the Brout-Englert-Higgs mechanism.

According to Table I, the left-handed electron and its neutrino counterpart are components of a doublet of $SU(2)_L$. To us, however, these particles don't quite look the same. We know that the electron is massive and carries an electric charge, while ν_e is (almost) massless and electrically neutral. Moreover, if there are long range interactions mediated by the photon (a massless particle, as far as we know), weak interactions are short ranged (mediated by massive particles). All these apparent shortcomings are resolved if the gauge symmetry $SU(2) \otimes U(1)$ is spontaneously broken by the Brout-Englert-Higgs mechanism [2] The gauge bosons mediating weak interactions become massive, the photon arises as a mixture of SU(2) and U(1) gauge bosons (hence weak and electromagnetic interactions are related or "unified"), and the quarks and charged leptons naturally get a mass, something which a priori is excluded by the gauge symmetry. In the framework of the SM, the gauge symmetry is broken by the vacuum expectation of a scalar field transforming under the fundamental representation of $SU(2)$: the Higgs field. How this all works in practice will be the subject of the rest of these lectures.

Our plan is as follows. First we will revisit a simple process in electromagnetism, the annihilation of electron-positron into a muon and its antiparticle. Part of this process is described by electromagnetism. We will recap some basic facts about the Dirac equation, the use of Lagrangian, etc. and will look into more details into the differential cross section. That will tell us something about the Lorentz group. Next, motivated by this, we will do some group theory, emphasizing the relevance of Weyl spinors, the building block of matter in the SM. Next we will study gauge interactions, starting with QED. Extending to weak and strong interactions, we will have to review a few more aspects of Lie algebras. Symmetry breaking (both global and local) come next. Finally we will build the SM Lagrangian. The next-to-the-last chapter will consists of some simple applications of that theory. We will conclude with some shortcomings of the SM and their possible solutions. The experimental basis of the SM is peppered in the text.

[2]This topic has a fascinating history, with many different contributors, starting with Yoichiro Nambu (1961) and Philip W. Anderson (1963) with the phenomenology of superconductors (Meissner effect, 1933). The application to particle physics, is mostly associated to the names of Robert Brout and Francois Englert on one hand ("Broken symmetry and the mass of gauge vector boson", 1964) and Peter Higgs on the other hand ("Broken symmetries and the mass of gauge bosons", 1964).

	First family	Second family	Third family
Quarks	$Q_u = \begin{pmatrix} u_L \\ d_L \end{pmatrix}^{(1/3)}$ $U_u = u_R^{(4/3)}$ $D_d = d_R^{(-2/3)}$	$Q_c = \begin{pmatrix} c_L \\ s_L \end{pmatrix}^{(1/3)}$ $U_c = c_R^{(4/3)}$ $D_s = s_R^{(-2/3)}$	$Q_t = \begin{pmatrix} t_L \\ b_L \end{pmatrix}^{(1/3)}$ $U_t = t_R^{(4/3)}$ $D_b = b_R^{(-2/3)}$
Leptons	$L_e = \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}^{(-1)}$ $E_e = e_R^{(-2)}$	$L_\mu = \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}^{(-1)}$ $E_\mu = \mu_R^{(-2)}$	$L_\tau = \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}^{(-1)}$ $E_\tau = \tau_R^{(-2)}$
?	$N_e = \nu_{eR}^{(0)}$	$N_\mu = \nu_{\mu R}^{(0)}$	$N_\tau = \nu_{\tau_R}^{(0)}$

Table 1.1: **Matter content of the Standard Model: (almost) known.** The left-handed quarks and leptons are in doublets of SU(2) (or fundamental representation). The right-handed ones are singlets (they don't couple to the SU(2) gauge fields). The superscripts are their respective hypercharge Y. In this basis, the electric charge is given by $Q = T^3 + Y/2$. For instance, $Q = +2/3$ for a up-like quark, and $Q = -1/3$ for a down-like one. In the Lepton box, we have added right-handed neutrinos, *i.e.* singlets of SU(2) with zero hypercharge. We still don't know if neutrinos are of the Dirac or Majorana type.

Gauge group	Gauge bosons
$SU(3)_c$	eight spin-1 gluons
$SU(2)_L$	three spin-1 $W_3,\ W^\pm$
$U(1)_Y$	one spin-1 Y
Gravity	one spin-2 graviton

Table 1.2: **The gauge sector of the SM: known** (except for the graviton). All together 12 gauge bosons. By definition, the gauge bosons of non-abelian groups transform according to the adjoint representation.

Higgs field	$\Phi^{(1)} = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$

Table 1.3: **Scalar sector of the SM: little known, beyond the minimal scheme (one Higgs doublet field). A Higgs particle has been discovered in 2012.**

References

Books

Elementary

G. Kane, *Modern elementary particle physics*, Addison Wesley.
Donald Perkins, *Introduction to high energy particle physics*, Cambridge University Press.

Advanced

L. Okun, *Quarks and Leptons*, North Holland.
C. Quigg, *Gauge theories of the Strong, Weak and Electromagnetic Interactions*, Benjamin Cummings.
Michael E. Peskin and Daniel V. Schroeder, *An introduction to Quantum Field theory*, Addison-Wesley.
Ta-Pei Cheng and Ling-Fong Li, *Gauge theory of elementary particle physics*, Clarendon Press, Oxford.

More advanced

Steven Weinberg, *The Quantum Theory of Fields*, vol. I-III, Cambridge University Press.

Chapter 2

Matter: Weyl spinors as building blocks

2.1 An elementary process in QED

Key notions: angular momentum, helicity (informal), parity (informal)

To some extent, the purpose of a theory of fundamental interactions is to explain what happens in elementary processes, like

$$e^+ + e^- \to \mu^+ + \mu^-$$

This specific one, the annihilation of an electron with a positron into a pair of muon and its antiparticle (anti-muon), can in principle be described by QED (quantum electrodynamics). We work in the center-of-mass frame ($\vec{p}_{e^-} = -\vec{p}_{e^+}$). Let us assume that you know how to compute (or that a large set of data has been collected) the associated differential cross section (this is the fraction of, say, muons, emitted in the solid angle $d\Omega$, per unit of the incoming flux of electrons). We do it at very large energies E, so that we can neglect the mass of the fermions, $m_e, m_\mu \ll E$. As usual, we sum over the final helicities (more on that later) and average over the initial ones. The answer is (homework or Peskin-Schroeder)

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E_{\rm cm}^2}\left(1 + \cos\theta^2\right) \tag{2.1}$$

where $E_{\rm cm} = 2E$ is the total energy of the incoming electron-positron pair in center-of-mass (CM) frame. Integrating over angular directions gives

$$\sigma_{\rm tot} = \frac{4\pi\alpha^2}{3E_{\rm cm}^2}$$

We recognize here the fine structure constant α. The only dimensionful quantity is the energy, so the answer makes sense on dimensional grounds. How about the angular dependence? Here $\cos\theta = \hat{p}_{e^-} \cdot \hat{p}_{\mu^-}$ is the angle between the incoming momentum of the electron and that of the outgoing muon. We denote unit vectors with a "hat", eg $\hat{p} = \vec{p}/p$. The calculation or better, the data, tell us that there is a higher probability for the muon to go forward or backward than around $\theta = \pi/2$. This differential cross section is depicted in Fig.2.1.

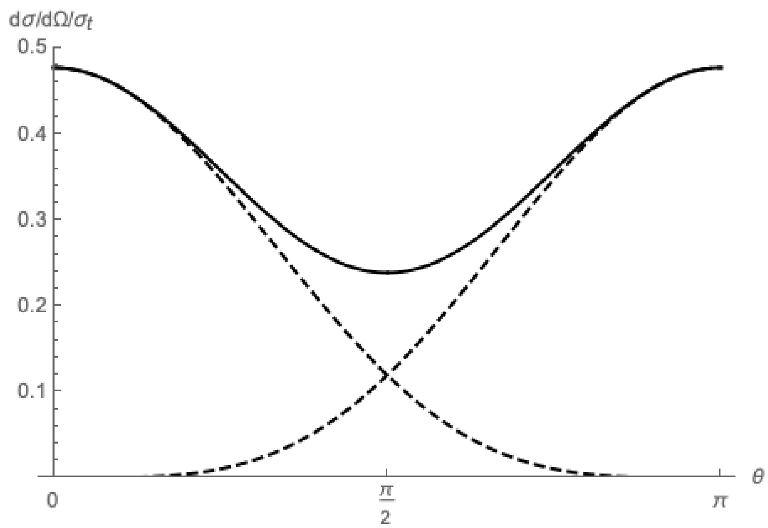


Figure 2.1: Differential cross section (normalized to total cross section) for unpolarized (solid) $e^+e^- \to \mu + \mu-$. The dashed curve with a maximum at $\theta = 0$ (forward) is the scattering cross section of a positive (or negative) helicity electron into a positive (or negative helicity) muon. The dashed curve with the maximum at $\theta = \pi$ (backward) the one of scattering of a positive (or negative) helicity electron into a negative (or positive helicity) muon. The latter two curves are normalized to 1/2.

Now, Eq.(2.1) is the averaged cross section. However, electrons (muons) and their antiparticles are also characterized by an extra quantum number, which is their spin or, more precisely for relativistic particles, their helicity. These are observable in principle. Hence, the average cross section is the average over the sum of, a priori, $2^4 = 16$ possible polarized cross sections. However, in QED, only 4 processes exist.

Let us first fix a notation problem. We note $f_\pm$ a positive (negative) helicity particle. The same for antiparticles, with a bar: $\bar{f}_\pm$ denotes a positive (negative) helicity antiparticle. As you know, for massless particles, helicity and chirality are related (more about this later). We will use (as in the tables in the chapter 1), L (left) and R (right) subscript for left and right particles respectively. For massless particles, helicity and chirality are the same, in the sense that a negative helicity particles is left-handed (ie turns to the left wrt to the momentum direction),

$$f_- \equiv f_L \quad \text{and} \quad f_+ \equiv f_R \quad \text{(massless particles!)}$$

while a positive helicity one is right-handed. It is the opposite for antiparticles,

$$\bar{f}_- \equiv \bar{f}_R \quad \text{and} \quad \bar{f}_+ \equiv \bar{f}_L \quad \text{(massless particles!)}$$

In the sequel, we use L and R subscripts to *label* the corresponding Weyl spinor fields. In plain words, in the present notes, a negative helicity antifermion is the antiparticle of a R-handed Weyl spinor. This is motivated by the fact that, in the SM, L-handed leptons are in doublets of SU(2), R-handed ones are singlets, but here, we are jumping ahead. [1]

[1] **Beware, this is not the convention of Peskin-Schroeder, which uses L and R for the helicity.** In that book,

Let's now go back to the cross sections for electron-positron annihilation. It can be shown (or again, our data show...) that the non-zero polarized cross sections are the following:

$$\begin{aligned}
\frac{d\sigma}{d\Omega}(e_+\bar{e}_- \to \mu_+\bar{\mu}_-) &= \frac{\alpha^2}{4E_{\rm cm}^2}(1+\cos\theta)^2 \\
\frac{d\sigma}{d\Omega}(e_+\bar{e}_- \to \mu_-\bar{\mu}_+) &= \frac{\alpha^2}{4E_{\rm cm}^2}(1-\cos\theta)^2
\end{aligned} \tag{2.2}$$

and, reversing the + and −,

$$\begin{aligned}
\frac{d\sigma}{d\Omega}(e_-\bar{e}_+ \to \mu_-\bar{\mu}_+) &= \frac{\alpha^2}{4E_{\rm cm}^2}(1+\cos\theta)^2 \\
\frac{d\sigma}{d\Omega}(e_-\bar{e}_+ \to \mu_+\bar{\mu}_-) &= \frac{\alpha^2}{4E_{\rm cm}^2}(1-\cos\theta)^2
\end{aligned} \tag{2.3}$$

All the other cross sections are zero (in the massless/high energy limit). For instance

$$\frac{d\sigma}{d\Omega}(e_+\bar{e}_+ \to \mu_+\bar{\mu}_-) \equiv 0$$

etc.

What is going on? First, it is clear that summing the four cross sections and dividing by 4 to average (again, we sum probabilities since each process is potentially observable/measurable) gives back (2.1). Second, what we see is consistent with angular momentum conservation, see figure 2.2. Take $e_+\bar{e}_- \to \mu_+\bar{\mu}_-$. The total angular momentum is 1 and it points along the momentum of the electron. Angular momentum conservation implies that a positive helicity muon should preferably go forward. In other words, in the language of quantum mechanics, we expect that the projection of the out-state (the polarized muon-antimuon pair) on the in-state, involve a factor of $\langle \mu_+\bar{\mu}_-|e_+\bar{e}_-\rangle \propto 1+\cos\theta$. Conversely, a negative helicity muon should go preferably backward, $\langle \mu_-\bar{\mu}_+|e_+\bar{e}_-\rangle \propto 1-\cos\theta$. Third, only states with total helicity one can transform into each others (ie interact with each others) in QED (again in the massless/high energy limit). Fourth, they do that with the same strength: helicity +1 and -1 lead to the same cross sections. This is QED. The mediator of QED is the photon, which is a particle of spin 1 (more precisely, helicity 1, since the photon is, we believe, massless).

There is a fifth lesson, which is more general. The relevant degrees of freedom at high energies are helicity states, or equivalently for massless particles, the chiral states. These are called Weyl spinors. QED treats Weyl spinors of both kind (L and R) on the same footing. We say that QED is a vector theory. The same holds of QCD. Weak interactions are however different. They treat differently L and R states. To take this point to the extreme, imagining that we were able to study experimentally the annihilation of a pair of electron neutrino and antineutrino followed by the creation of a pair of muonic neutrino and antineutrino. This is

the antiparticle of a R-handed particle/spinor/field is noted with a subscript L because it has negative (left) helicity. Sorry for the confusion.

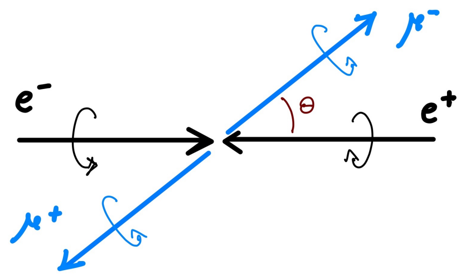


Figure 2.2: Electron-positron annihilation into a muon-antimuon pair. Only total helicity one processes are non-zero in QED, reflecting the fact that the photon, which is the QED mediator, is a spin one. This is illustrated for a positive helicity electron (muon) and negative helicity positron (antimuon). In this case, the muon is mostly emitted forward (as defined wrt the electron momentum) instead of backward.

probably quasi impossible to do experimentally, but never mind. The point is that if we were to test it, we would (probably) find that the only non-zero cross section (at very high energies, or technically, for $E_{\rm cm} \gg M_Z$, the Z boson mass) is

$$\frac{d\sigma}{d\Omega}(\nu_-^e \bar{\nu}_+^e \to \nu_-^\mu \bar{\nu}_+^\mu) = \frac{\alpha_W^2}{4E_{\rm cm}^2}\,(1+\cos\theta)^2$$

All the other cross section are zero (in the framework of the SM – perhaps some other theory could give a non-zero but, to be consistent with "data", a very small contribution). The factor α_W is similar (and actually is related) to α, but never mind. The angular dependence stems from angular momentum dependence, so there is no way around. But you need a theory to make explicit that only this cross section is non-zero. What about scattering positive helicity neutrinos? So far, we have not observed such a state. In the SM language, if they exist, they do not interact, or at least not with an interaction comparable to that of the negative helicity neutrinos.

Now, this means that parity is not a symmetry of weak interactions, for, under parity, a positive helicity state is transformed into a negative helicity one. Indeed, momentum is reversed by parity, but angular momentum is a pseudovector, so it does not change. By parity, we mean reversal of all spatial directions. This is equivalent to reversal of one axis, say along z (a mirror image), times a rotation of π around the same axis. If negative and positive (if they exist) neutrinos don't have the same cross section, the image in a mirror does not exist (as a physical process, not as an image). Very spooky. Conversely, electromagnetism, strong interactions, and for that matter, gravity, does not differentiate left from right. That only in the subtle effects of weak interactions we do see that parity is not a symmetry of Nature (at least at the energies at which this has been tested) explains why the discovery of "parity violation" came as a surprise when it was first theorized (T.D. Lee and C.N Yang, 1956 "Question of parity conservation in Weak Interactions") and then experimentally tested (Chien-Chiung Wu et al, 1956, "Experimental test of parity conservation in β decay'). This is a rare instance of an experimental test in particle physics that did not take years (or decades) to be realized.

2.2 Deconstructing QED

Key notions: Dirac equation, Weyl spinors, parity operator

The previous section emphasized the role of Weyl spinors. We quickly revisit this from a more formal point of view, again starting from QED. With start with the Dirac equation (1928). Dirac's aim, apparently, was to have a wave equation that is linear in time (instead of quadratic, like KG) to avoid negative energy solution (if so, that aspect failed). So something that looked as much as possible as the Schrödinger equation,

$$i\partial_t\psi = H\psi, \tag{2.4}$$

He wrote down

$$i\partial_t\psi_D = (\vec{\alpha}\cdot\vec{\nabla} + m\beta)\psi_D, \tag{2.5}$$

and showed that α and β must be 4×4 hermitian matrices. The 4 components column object ψ_D is called a Dirac spinor

$$\psi_D = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \tag{2.6}$$

Modern notations, which involve ∂_μ, emphasize the underlying Lorentz covariance of the Dirac equation (see Appendix B for a review of special relativity, including notations)

$$(i\gamma^\mu\partial_\mu - m)\psi_D = 0 \tag{2.7}$$

with $\alpha = \gamma^0\gamma^i$ and $\beta = \gamma^0$, so

$$H_D = i\gamma^0\vec{\gamma}\cdot\vec{\nabla} + m\gamma^0 \tag{2.8}$$

Note that $\gamma_i = -\gamma^i$, using the Minkowski metric to lower indices. Going from (2.5) to (2.7) is like going from the four Maxwell equations (differential form, using $\vec{E}$ and $\vec{B}$) to the form of Lorentz using $F_{\mu\nu}$ and its dual. Clearly γ^μ cannot be a mere 4-vector, as it would define a special direction in spacetime, in contradiction with the requirement of Lorentz invariance of the vacuum. We will use an even more compact notation, using the Feynman slash:

$$\partial_\mu\gamma^\mu = \not{\partial} \qquad k_\mu\gamma^\mu = \not{k} \quad \ldots$$

so

$$(i\gamma^\mu\partial_\mu - m)\psi_D = 0 \to (i\not{\partial} - m)\psi_D = 0$$

Now acting on (2.7) with $i\not{\partial} + m$, we get

$$(i\not{\partial} + m)(i\not{\partial} - m)\psi_D \equiv -(\gamma^\nu\gamma^\mu\partial_\mu\partial_\nu + m^2)\psi_D = 0 \tag{2.9}$$

If the γ^μ satisfy

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\mathbb{I} \tag{2.10}$$

where $\mathbb{I}$ is a unit operator, and $\eta^{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$, then ψ is, at the same time, solution of the KG equation,

$$(\Box + m^2)\psi_D = 0. \tag{2.11}$$

This is the equation of a free particle of mass m. In Fourier space (plane wave), $\psi \propto e^{-ip\cdot x}$, this gives an algebraic condition $(p^2 - m^2)\psi = 0$. However, there is more than $p^2 = m^2$ in the Dirac equation, since it is a collection of 4 functions.

The relation (2.10) defines what mathematicians recognize as a Clifford Algebra. Such structure arise naturally in representation theory of SO(n) groups. We deal with SO(3, 1), so no surprise. A concrete representation of the γ^μ is in terms of 4×4 matrices. We focus on the Weyl or chiral representation, as in Peskin-Schroeder,

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{I}_2 \\ \mathbb{I}_2 & 0 \end{pmatrix} \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \quad \text{(Weyl)} \tag{2.12}$$

where the σ_i, $i = 1, 2, 3$ are the Pauli matrices. Indeed,

$$(\gamma^0)^2 = \mathbb{I}_4, \quad \text{and} \quad \gamma^0\gamma^i = -\gamma^i\gamma^0 \tag{2.13}$$

Also

$$\gamma^i\gamma^j + \gamma^j\gamma^i = -2\delta_{ij}\mathbb{I}_4, \tag{2.14}$$

obtained using the relation

$$\sigma_i\sigma_j = \delta_{ij}\mathbb{I}_2 + i\epsilon_{ijk}\sigma_k. \tag{2.15}$$

2.2.1 Weyl spinors from Dirac spinors

If we define

$$\sigma^\mu = (\mathbb{I}_2, \vec{\sigma}) \tag{2.16}$$

and

$$\bar{\sigma}^\mu = (\mathbb{I}_2, -\vec{\sigma}) \tag{2.17}$$

we may rewrite the Dirac matrices as

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} \tag{2.18}$$

Then group the first and last two components of ψ_D together,

$$\psi_D = \begin{pmatrix} \chi_L \\ \chi_R \end{pmatrix} \tag{2.19}$$

The 2-components objects $\chi_{L,R}$ are called Weyl or chiral spinors. In this notation, the Dirac equation becomes

$$(i\partial\!\!\!/ - m)\psi_D = 0 \to \begin{pmatrix} -m & i\sigma^\mu\partial_\mu \\ i\bar{\sigma}^\mu\partial_\mu & -m \end{pmatrix} \begin{pmatrix} \chi_L \\ \chi_R \end{pmatrix} = 0 \tag{2.20}$$

or

$$\begin{aligned} i\bar{\sigma}^\mu \partial_\mu \chi_L &= m\chi_R \\ i\sigma^\mu \partial_\mu \chi_R &= m\chi_L \end{aligned} \tag{2.21}$$

a system of 2 coupled equations. However, in the massless limit (synonym, depending on the context, of massless particle, high energy limit, chiral limit), they decouple

$$\begin{aligned} i\bar{\sigma}^\mu \partial_\mu \chi_L &= 0 \\ i\sigma^\mu \partial_\mu \chi_R &= 0 \end{aligned} \tag{2.22}$$

These are called the Weyl equations for massless fermions.

This is interesting but what does that mean? Consider the extra 4×4 matrix,

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{2.23}$$

As all these matrices anti-commute with each others, we can also rewrite γ_5 using a Levi-Civita tensor

$$\gamma_5 = -\frac{i}{4!}\epsilon^{\mu\nu\rho\sigma}\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma \tag{2.24}$$

with the convention that $\epsilon^{0123} = +1$.[2]

You can check (using the explicit form of the γ^μ or their abstract properties given in the Appendix G) that γ_5 is hermitian and that

$$\gamma_5^2 = 1$$

and

$$\mathrm{Tr}(\gamma_5) = 0$$

In the Weyl basis,

$$\gamma_5 = \begin{pmatrix} -\mathbb{I}_2 & 0 \\ 0 & \mathbb{I}_2 \end{pmatrix} \tag{2.25}$$

which is diagonal with eigenvalues ± 1. A direct calculation shows that

$$\{\gamma_5, \gamma^\mu\} = 0$$

Now take the commutator of γ_5 with the Dirac hamiltonian, Eq.(2.8),

$$[\gamma_5, H_D] = 2m\gamma_5\gamma^0$$

In the massless limit, they commute, so we can label energy states as eigenstates of γ_5. The eigenvalue is called chirality. We write a negative chirality state with an L, for left-handed,

$$\gamma_5\psi_L = -\psi_L$$

[2] Never mind the fact that γ_5 is written with a lower index; this is just a notation.

and a positive chirality with an R for right-handed,

$$\gamma_5 \psi_R = +\psi_R$$

In the Weyl basis,

$$\psi_L = \begin{pmatrix} \chi_L \\ 0 \end{pmatrix} \tag{2.26}$$

and

$$\psi_R = \begin{pmatrix} 0 \\ \chi_R \end{pmatrix} \tag{2.27}$$

The projection on these states is obtained using the following operators,

$$L = \frac{\mathbb{I}_4 - \gamma_5}{2}$$

and

$$R = \frac{\mathbb{I}_4 + \gamma_5}{2}$$

There are indeed projectors,

$$L^2 = L \quad ; \quad R^2 = R \quad ; \quad RL = LR = 0 \quad ; \quad R + L = \mathbb{I}_4$$

so

$$\psi_L = L\psi$$

and

$$\psi_R = R\psi$$

and of course

$$\psi = \psi_L + \psi_R$$

Warning about notations: We will, in the sequel, work with $\psi_{L,R}$, hence projections of 4-dimensional objects. At the same time, and this may be confusing, we will convey the fact that Weyl spinors ($\chi_{L,R}$ hence 2-dimensional objects) are the truly fundamental building blocks. The reason for using $\psi_{L,R}$ is that the γ^μ matrices are easy to manipulate so it is often more convenient to work with projected Dirac spinors instead of Weyl spinors (although we will do both, depending on the context). There is a formalism (actually several versions of that formalism, but essentially this is all the van der Waerden notation of dotted spinors) that works directly on Weyl spinors. This is often very powerful (for instance in the framework of supersymmetry) but it requires more formalism and so more work. No books on the Standard Model, except those that focus on supersymmetry, use Weyl spinors.

2.2.2 Plane wave solutions

This is quite abstract, so let us look directly at the explicit plane wave solutions that you derived in the QFT course. We won't repeat that. Imagine that we consider the implications of the Coulomb potential (which is a solution of the Maxwell equations) without bothering how we go it. Positive energy solutions

$$\psi(x) = u(p)e^{-ip\cdot x}$$

with $p^2 = m^2$ are given by

$$(\not p - m)u(p) = 0 \quad \rightarrow \quad u(p) = \frac{1}{\sqrt{2(E+m)}}\begin{pmatrix} (E+m-\vec\sigma\cdot\vec p)\,\phi_s \\ (E+m+\vec\sigma\cdot\vec p)\,\phi_s \end{pmatrix} \tag{2.28}$$

with ϕ_s two component objects, which we call 2-spinors, chosen to be normalized to $\phi_s^\dagger\phi_{s'} = \delta_{ss'}$ (don't worry about the overall normalization). There are thus 2 independent solutions, which we label as $s = \pm$. Also, a fermion at rest has an equal share of L and R components, since $\chi_L = \chi_R = \phi_s$. To clarify the meaning of the ϕ_s, consider a fermion solution at rest, $\vec p = 0$ and $E = m$,

$$u(m) = \sqrt{m}\begin{pmatrix} \phi_s \\ \phi_s \end{pmatrix} \tag{2.29}$$

so ϕ_s is (probably, we should check how rotations acts on them — see later) nothing but a spinor in standard non-relativistic quantum mechanics. Using σ^3 to label spin eigenstates

$$\sigma^3\phi_s = s\phi_s$$

gives

$$\phi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \phi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.30}$$

corresponding to spin up and down wrt to the direction $\hat z$. This suggests, and will be confirmed by a more refined analysis, that the spin operator (for a particle at rest) in the Dirac theory is simply

$$\vec S = \frac{1}{2}\begin{pmatrix} \vec\sigma & 0 \\ 0 & \vec\sigma \end{pmatrix} \tag{2.31}$$

Now take a spin up along $\hat z$ which move in the same direction, with momentum p_z. According to (2.28) the solution is

$$u_+(p_z) = \begin{pmatrix} \sqrt{E-p_z} \\ 0 \\ \sqrt{E+p_z} \\ 0 \end{pmatrix} \tag{2.32}$$

The + subscript now means that the spin points in the same direction as the momentum. This is clearly a very specific solution, but it is a useful one. It is an eigenstate of the so-called helicity operator, which is defined as the projection of the spin along the direction of propagation,

$$\hat h \mathrel{\hat=} \hat p\cdot\vec S = \frac{1}{2}\begin{pmatrix} \hat p\cdot\vec\sigma & 0 \\ 0 & \hat p\cdot\vec\sigma \end{pmatrix} = \frac{1}{2}\gamma_5\gamma^0\vec\gamma\cdot\vec p \tag{2.33}$$

$$\hat{h}\, u_+(p) = +\frac{1}{2}\, u_+(p)$$

Importantly, you can check that the helicity operator $\hat{h}$ commutes with the Dirac hamiltonian (2.8),

$$\left[\hat{h}, H_D\right] = 0 \tag{2.34}$$

so indeed eigenstates can be labelled using energy and helicity, and this for any fermion mass. Similarly, a negative helicity particle corresponds to a solution of the form

$$u_-(p_z) = \begin{pmatrix} 0 \\ \sqrt{E+p_z} \\ 0 \\ \sqrt{E-p_z} \end{pmatrix} \quad \text{with} \quad \hat{h}\, u_-(p) = -\frac{1}{2}\, u_+(p) \tag{2.35}$$

Up to this $1/2$ factor, the helicity of an electron is similar to the helicity states of a photon, also called circular polarization states. For a positive helicity photon, we picture a little arrow (an electric vector) that turns to the right along its direction of propagation. For spinor, we cannot quite draw a vector (this is precisely because they are spinors) but we can picture them as something that is spinning to the right (right-hand convention). Now is as important to appreciate that helicity is not preserved by Lorentz transformation (no contradiction; energy $\sim$ Hamiltonian is also frame dependent) as changing the direction of the momentum through a change of frame (not a mere rotation, a true Lorentz boost) can lead to a flip in the helicity. The exception are massless particles, for which this is impossible and so helicity of a particle is truly a Lorentz invariant.

This is further clarified if we study the link between helicity and chirality. To do so, consider the very relativistic limit $E \gg m$ of (2.32)

$$u_+(p) \approx \sqrt{2E}\begin{pmatrix} m/2E \\ 0 \\ 1 \\ 0 \end{pmatrix} \xrightarrow{m/EE\to 0} u_+(p) \equiv u_R(p) = \sqrt{2E}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \tag{2.36}$$

or (2.35)

$$u_-(p) \approx \sqrt{2E}\begin{pmatrix} 0 \\ 1 \\ 0 \\ m/2E \end{pmatrix} \xrightarrow{m/\to 0} u_-(p) \equiv u_L(p) = \sqrt{2E}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{2.37}$$

At high energies, the helicity eigenstates tend to chirality eigenstates: a spin aligned (anti-aligned) with its direction of propagation is essentially a R (L) state with a slight, $\mathcal{O}(m/E)$ component of the opposite chirality. Take for instance $E = 100$ GeV and $m = m_e$ (eg an electron at the former LEP accelerator), then the relative weight of the subdominant chiral component is indeed very small, $m/E \sim 5 \cdot 10^{-7}$. In the massless limit, their signs are strictly equal:

$$\boxed{\text{HELICITY} \;=\; \frac{1}{2}\;\text{CHIRALITY} \qquad \text{(massless fermionic particles)}}$$

How about antiparticle solutions?

A Dirac spinor has four degrees of freedom. The other two orthogonal plane wave solution are given by negative energy solutions, which correspond to antiparticles states. Taking $\psi(p) = \mathrm{v}(p)e^{ip\cdot x}$ (ie $E > 0$ and $\vec{p}$ reversed), the solutions are

$$(\not{p} + m)\mathrm{v}_s(p) = 0 \quad \rightarrow \quad \mathrm{v}_s(p) = \frac{1}{\sqrt{2(E+m)}}\begin{pmatrix} (E + m - \vec{\sigma}\cdot\vec{p})\,\xi_s \\ -(E + m + \vec{\sigma}\cdot\vec{p})\,\xi_s \end{pmatrix} \tag{2.38}$$

with ξ_s are 2-spinors, like the ϕ_s, but are distinct; so there indeed are four plane-wave solutions. The relative sign in (2.38) compared to (2.28) has a role to play, but a conclusion that the same as for particles is the L and R components are (up to a sign) the identical, $\chi_L = -\chi_R = \xi_s$. One thing that is distinct is that the solutions (2.38) have negative energy, while there is nothing like that in nature (ie negative mass). The interpretation, in the framework of QFT, is that the ψ are time-dependent operators which can be split into a destruction of particles part (positive energy solutions) and creation of antiparticles (negative energy solutions) when acting on incoming states (the reverse for the adjoint field). Another way to think about antiparticles is the old Dirac sea picture, in which an antiparticle is seen as the absence of a negative energy particle solution (a hole in the Dirac sea). Anyway, the conclusion is that the quantum numbers of antiparticles are opposite to that of the solution (2.38). In particular, an antiparticle at rest with a **spin up**, say along $\hat{z}$, corresponds to a 2-spinor with

$$\sigma^3\xi_+ = -\xi_+ \equiv -\begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{2.39}$$

So, an antiparticle at rest with spin up corresponds to the Dirac spinor (compare with (2.29)),

$$\mathrm{v}(m) = \sqrt{m}\begin{pmatrix} \xi_s \\ -\xi_s \end{pmatrix} \quad \rightarrow \quad \mathrm{v}(m, up) = \sqrt{m}\begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \tag{2.40}$$

Now, a priori, the momentum as the spin, should be reversed when going from the solution to the state. But this sign has already been reversed when we took the plane wave to be in the form $e^{ip\cdot x}$. Very confusing and I know no other way than to think about it so as to be convinced. Now, here is my advice. Consider specifically the very relativistic solutions

$$\mathrm{v}_s(p) = \sqrt{2E}\begin{pmatrix} (1 - \vec{\sigma}\cdot\hat{p})\,\xi_s \\ -(1 + \vec{\sigma}\cdot\hat{p})\,\xi_s \end{pmatrix} \tag{2.41}$$

In this expression, E is positive and is the antiparticle's energy. Similarly, $\vec{p}$ is its momentum. Take for instance an antiparticle whose spin is in the direction $\hat{z}$. The 2-spinor is the one of (2.39). Take it to move with momentum $\vec{p}$ in the opposite direction $-\hat{z}$. We are thus constructing a negative helicity state and it is described by the solution

$$\mathrm{v}_-(p) \equiv \mathrm{v}_R(p) = \sqrt{2E}\begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} \tag{2.42}$$

This corresponds to the antiparticle of a R-handed spinor. We see that for an antiparticle, helicity (here $-1/2$) and chirality (here positive) eigenvalues are opposite:

$$\boxed{\text{HELICITY} \;\; = -\frac{1}{2}\;\text{CHIRALITY} \qquad \text{(massless fermionic antiparticles)}}$$

This is actually general. Take the Dirac equation in the massless limit,

$$\not{p}\,\psi = 0 \rightarrow \vec{p}\cdot\vec{\gamma}\psi = p^0\gamma^0\psi$$

Multiplying both sides by γ_5 gives

$$\vec{S}\cdot\vec{p}\,\psi = \frac{1}{2}p^0\,\gamma_5\psi$$

So the sign of helicity and chirality eigenvalues are equal for particles eigenstates (positive energy solutions, $p^0 > 0$, $\vec{p}/p^0 = \hat{p}$) and opposite for the antiparticle ones (negative energy solutions, $p^0 < 0$, $\vec{p}/p^0 = -\hat{p}$),

$$\begin{aligned} \hat{h}\psi_R &= +\frac{1}{2}\,\psi_R \quad \text{and} \quad \hat{h}\psi_L = -\frac{1}{2}\,\psi_L \quad (p^0 > 0) \\ \hat{h}\psi_R &= -\frac{1}{2}\,\psi_R \quad \text{and} \quad \hat{h}\psi_L = +\frac{1}{2}\,\psi_L \quad (p^0 < 0) \end{aligned} \tag{2.43}$$

This is consistent with the explicit solutions we found above, as you can check. But this is also laborious. Fortunately, most of the time, we don't need to worry about all of these sign flips (or absence of) . The algebra of Dirac matrices, together with the projectors R and L allow to almost completely automatize amplitude calculations. Yet, it is important to keep in mind what we are talking about, and, at the end of a calculation, to understand whether it makes sense, if any. Anyway, the most important messages of this section are in the two boxes.

2.2.3 Parity and charge conjugation

From the previous section, we suspect that the Weyl spinors (or equivalently, the R and L projections of a Dirac spinor) are related by parity, noted P. Consider a particle, so as not to get confused. Indeed, a R-handed spinor corresponds to a particle of momentum say $\vec{p}$ with helicity $h = 1/2$, ie turning to the right wrt its direction of propagation. It's mirror image lead to a flip of the momentum but not of the spin; the spin is a pseudovector. So the helicity is reversed if you look in the mirror. Like the left and right hands, L and R spinors are transformed into each others.

To make this more formal, first define how parity acts on mundane objects. The reflection along any spatial dimension, say $\hat{x}$ does the job (mirror),

$$\vec{x} = (x, y, z) \quad \stackrel{P}{\rightarrow} \quad \vec{x}\,' = (-x, y, z)$$

and the same for any vectors (eg momentum, electric field, etc). As mentioned above, we also have pseudovectors in 3 dimensions. Three examples could come to mind: an oriented surface, a magnetic field and angular momentum. Consider the latter

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

It involves the cross product, with a conventional orientation, traditionally taken the right-hand rule (majority rules).[3] Now, it is easier to generalize a bit the notion of parity, so that momentum are fully reversed, instead of just one component. This is achieved by reflection along an axis (say $\hat{x}$) composed with a rotation of π around the same axis so that

$$\vec{x} = (x, y, z) \xrightarrow{P} \vec{x}\,' = -\vec{x}\,, \qquad \vec{p} \xrightarrow{P} -\vec{p}\,, \qquad \text{etc.}$$

Then

$$\vec{L} \quad \xrightarrow{P} \quad \vec{L}$$

This is the landmark of a pseudovector. Compared to a vector, a pseudovector has an extra sign flip (its is reversed and flipped). The reason for this peculiar behavior is related to the fact that in 3D, we can assign a collection of three numbers ("a vector", say B_i) to an antisymmetric matrix, say F_{ij}, using the Levi-Civita tensor, ϵ_{ijk}, through the identification

$$B_i = \frac{1}{2}\epsilon_{ijk}\, F_{jk}$$

Pseudovectors are really antisymmetric matrices. The latter can be built using two vectors, so this is consistent with the definition of an oriented surface, angular momentum or a magnetic field (think of the Biot-Savart law).

Let's now recall how parity is defined for spinors. This is our first explicit encounter with a problem of representation. Go back to the Dirac equation and ask how we could define an operator U_P whose action is to reverse all the spatial directions when acting on a Dirac spinor,

$$U_P\, \psi(t, \vec{x}) U_P^{-1} = \psi'(t, \vec{x}) \equiv P\psi(t, -\vec{x})$$

What does that mean? First the LHS. What we aim is to do QFT, so the Dirac spinor is actually a quantum field, ie an operator. So, what we mean on the LHS is that U_P is a representation of the parity operator that transforms $\psi(t, \vec{x})$ into a $\psi'(t, \vec{x})$ (a new operator in space-time). Now parity acting twice should give back the same thing modulo the sign ambiguity of fermions, which we do not consider here [4], so $U_P^2 = 1$. So U_P is also both unitary and hermitian. We expect its eigenvalues to be ± 1. How about the RHS? Dirac spinors are a collection of 4 operators which depend on time and position. Parity should reverse the spatial coordinates, but also, in general reshuffle the 4 operators[5]. So here P is a 4×4 matrix and $\psi'(t, \vec{x}) = P\psi(t, -\vec{x})$.

[3] This is a convention and it is important not to change convention when doing comparisons. Or better, to consider physical situations so as to consider something that is convention invariant.

[4] This sign ambiguity is specific to spinors. Coming back to itself is it like a rotation of zero angle or 2π? The latter gives a minus sign, hence the two possibilities. It is not clear whether there is physics in this choice, or if it is just a choice.

[5] This can be taken as an a priori assumption but notice that parity acts on the spatial dimension, not time, but the time coordinate is frame dependent, so parity cannot commute with Lorentz transformation, see next chapter. So we should expect that it mixes the components of a spinor.

To build it, we start with the Dirac equation and ask that $\psi'(t,\vec{x})$ satisfies the Dirac equation if $\psi(t,\vec{x})$ does (that we transform apples into apples)

$$(i\gamma^\mu\partial_\mu - m)\psi(t,\vec{x}) = 0 \rightarrow (i\gamma^\mu\partial_\mu - m)\psi'(t,\vec{x}) = 0$$

This is not magic. If we see the Dirac equation as a defining property of a Dirac spinor, then symmetry transformations must preserve that property. This can be rewritten as

$$(i\gamma^0\partial_0 + i\gamma^i\partial_i - m)\psi'(t,\vec{x}) = 0 \rightarrow (i\gamma^0\partial_0 + i\gamma^i\partial_i - m)P\psi(t,-\vec{x}) = 0$$

or

$$(i\gamma^0\partial_0 + i\gamma^i\partial_i - m)P\psi(t,-\vec{x}) = P(i\gamma^0\partial_0 + iP\gamma^iP\partial_i - m)\psi(t,-\vec{x})$$

We get the same equation as the original if we simultaneously flip $x^i \rightarrow -x^i$ and impose

$$P\gamma^iP = -\gamma^i$$

which is satisfied if we take

$$P = \eta\gamma^0 = \eta\begin{pmatrix} 0 & \mathbb{I}_2 \\ \mathbb{I}_2 & 0 \end{pmatrix} \tag{2.44}$$

Here η is a pure phase. Making twice P brings back a spinor to itself (up to the sign ambiguity we mentioned above) so we take $\eta^2 = \pm 1$, meaning $\eta = \pm 1$ (see Landau & Lifschitz, chapter 19 for a discussion on the other possible convention). We will come back to this factor when we will discuss charge conjugation. In the meantime, notice that acting on a Dirac spinor, Eq.(2.19), we have

$$\psi(t,\vec{x}) = \begin{pmatrix} \chi_L(t,\vec{x}) \\ \chi_R(t,\vec{x}) \end{pmatrix} \quad \stackrel{P}{\rightarrow} \quad P\psi(t,-\vec{x}) = \eta\begin{pmatrix} \chi_R(t,-\vec{x}) \\ \chi_L(t,-\vec{x}) \end{pmatrix}$$

or in term of Weyl spinors,

$$\chi_R(t,\vec{x}) \stackrel{P}{\rightarrow} \eta\chi_L(t,-\vec{x}) \quad \text{and} \quad \chi_L(t,\vec{x}) \stackrel{P}{\rightarrow} \eta\chi_R(t,-\vec{x})$$

or, equivalently,

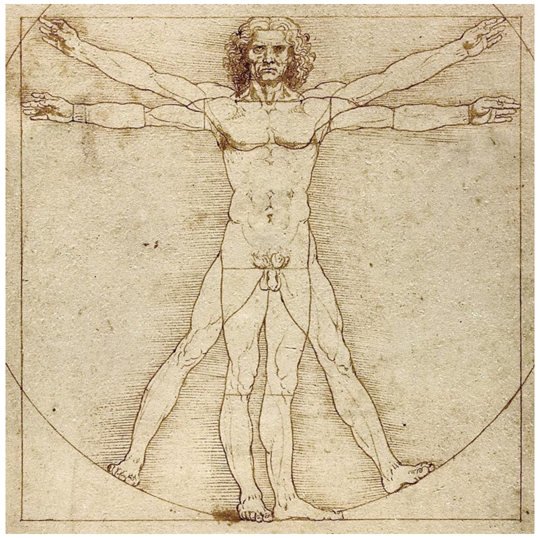
$$\psi_R(t,\vec{x}) \stackrel{P}{\rightarrow} \eta\psi_L(t,-\vec{x}) \quad \text{and} \quad \psi_L(t,\vec{x}) \stackrel{P}{\rightarrow} \eta\psi_R(t,-\vec{x})$$

So parity exchanges L and R spinors, very much like the reversal of a hand in a mirror.

Now another thing which perhaps will sound confusing. We have defined parity as the action on a Dirac spinor, an object with 4 components but the underlying degrees of freedom are really Weyl spinors. If we have a theory with an **unequal** number of L and R Weyl spinors, then parity is not be well-defined; we say that parity is broken. It is thus important that you realize that parity *exchanges* different objects, like a left hand and a right hand. I guess a mathematician would call parity an isomorphism. In some circumstance, the different objects can be grouped together, making something that is symmetric under parity, something like

$$\psi = \psi_L + \psi_R \stackrel{P}{\rightarrow} \psi = \psi_R + \psi_L$$

or that (Vitruvian man, L. da Vinci)



So a dull analogy, if it helps, is to think of a Weyl spinor as a one-armed person, say left-handed. Its image is a right-handed one-armed. She cannot be confused with the original. Under parity, she is transformed into something else.

Now a few words about charge conjugation, C. Naively, the idea of charge conjugation is to exchange particles and antiparticles, by mere complex conjugation. However, C can also reshuffle the components of spinor. Indeed the charge conjugate of a Dirac field is defined as

$$C\psi(x)C \equiv \psi^c = -i\gamma^2\psi^*$$

Why this qualifies as charge conjugation is because we have the complex conjugate of a field. The γ^2 can be understood by looking explicitly at the complex conjugate of the Dirac equation,

$$(i\partial\!\!\!/ - m)\psi = 0 \quad \rightarrow \quad (-i\gamma^{\mu *}\partial_\mu - m)\psi^* = 0 \quad \rightarrow \quad (i\gamma^2\gamma^{\mu *}\gamma^2\partial_\mu - m)\psi^c = 0$$

You can check that

$$\gamma^2\gamma^{\mu *}\gamma^2 = \gamma^\mu$$

directly in the Weyl basis, so that ψ^c satisfies the Dirac equation (in other words, C is a symmetry of this equation). Jumping ahead, we could have included the coupling to the electromagnetic field through the minimal substitution rule, ie

$$(i\gamma^\mu\partial_\mu - e\gamma^\mu A_\mu - m)\psi = 0$$

You can verify that ψ^c satisfies the same equation as ψ, but with $e \rightarrow -e$.

Notice that in the Weyl basis,

$$\psi(x)^c = \begin{pmatrix} -i\sigma^2\chi_R^* \\ i\sigma^2\chi_L^* \end{pmatrix} = \begin{pmatrix} \chi_R^c \\ \chi_L^c \end{pmatrix}$$

An equivalent way to write this is

$$\chi_R \stackrel{C}{\rightarrow} i\sigma^2\chi_L^* \quad \text{and} \quad \chi_L \stackrel{C}{\rightarrow} -i\sigma^2\chi_R^*$$

This emphasizes the fact that charge conjugation flips the chirality, ie transform Weyl spinor into another kind of spinors. So, C is not a symmetry of theory with an unequal number of chiral spinors of different chirality. On the other hand

$$\psi = \psi_L + \psi_R \quad \xrightarrow{C} \quad \psi^c = \psi^c_R + \psi^c_L$$

is invariant.

We can finally combine a parity and charge conjugation transformation, CP. We get

$$\boxed{\chi_R(t, \vec{x}) \quad \xrightarrow{CP} \quad i\eta\sigma^2\chi^*_R(t, -\vec{x}) \qquad \text{and} \qquad \chi_L(t, \vec{x}) \quad \xrightarrow{CP} \quad -i\eta\sigma^2\chi^*_L(t, -\vec{x})} \tag{2.45}$$

So, CP transforms a chiral spinor onto its antiparticle, and is thus *a priori* a good symmetry (an automorphism). We will see that in the SM, C and P are broken. The SM is a chiral theory. QED and QCD respect parity and charge conjugation; they are sometime called vector-like theories. We will see later on that CP is also violated within the SM (and QCD), but in very subtle ways.

There is a final discrete transformation, which reverses time flow, T. We will not discuss this here but to mention that a famous theorem states that CPT must be symmetry of Lorentz invariant local theories (ie with a real Lagrangian density). The consequences are that the mass and total decay rate (or lifetime) of a particle and its antiparticle must be strictly equal,

CPT THEOREM
mass particle = mass antiparticle
total lifetime particle = total lifetime antiparticle

2.2.4 Bilinear operators built from Weyl spinors - a first look

The Dirac equation derives from the following action

$$S = \int d^4x\, \mathcal{L} = \int d^4x\, (i\bar{\psi}\partial\!\!\!/\psi - m\bar{\psi}\psi) \tag{2.46}$$

where $\bar{\psi} = \psi^\dagger\gamma^0$ is called the adjoint of ψ. The reason for using $\bar{\psi}$ instead of $\psi^\dagger$ is because the spatial γ^i, $i = 1, 2, 3$ are anti-hermitian, $(\gamma^i)^\dagger = -\gamma^i$; the γ^0 in the definition of $\bar{\psi}$ makes all the terms in $\mathcal{L}$ hermitian. The Euler-Lagrange equation is simply

$$\frac{\delta\mathcal{L}}{\delta\bar{\psi}(x)} = 0 \to (i\partial\!\!\!/ - m)\,\psi = 0 \tag{2.47}$$

There is a similar equation for the adjoint. Integrate by part and take the derivative wrt to ψ,

$$\frac{\delta\mathcal{L}}{\delta\psi(x)} = 0 \to \bar{\psi}\left(i\overleftarrow{\partial\!\!\!/} + m\right) = 0 \tag{2.48}$$

The ugly reversed arrow means that the derivative acts on the left.

There are two terms in the Lagrangian, the kinetic term and the mass term. Consider first the mass term,

$$m\bar{\psi}\psi = m\bar{\psi}_R\psi_L + m\bar{\psi}_L\psi_R \tag{2.49}$$

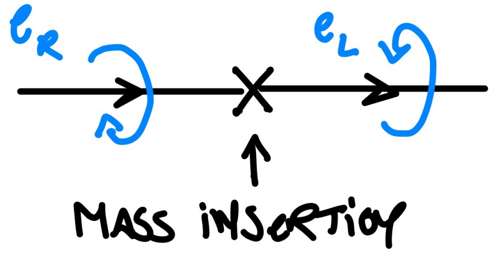


Figure 2.3: A mass insertion leads to a chirality flip.

where we have used the following properties of L and R projectors

$$R^\dagger = R \qquad R\gamma^0 L = R\gamma^0 = \gamma^0 L \qquad R\gamma^0 R = 0$$

and similarly exchanging R and L. We see that the mass term mixes ψ_L and ψ_R spinors. This reflects the fact that Weyl spinors of opposite chirality are coupled in presence of mass. It turns useful to consider a mass term as a sort of interaction, which induces transitions between opposite chiralities as a massive fermion propagates. I think this is related to something called the Zitterbewegung, but I am not sure, as I have never read a clear discussion about this. Regardless, we speak of a chirality flip, which diagrammatically, we will represent as a cross on a fermion propagator, see figure 2.3.(You may have the impression that this violates angular momentum conservation, but you have to imagine that temporarily the electron goes backward, see figure 2.4.

Now take the Dirac equation in the form (2.47) and multiply it on the left by the adjoint

$$\bar{\psi}\,(i\partial\!\!\!/ - m)\,\psi = 0$$

and the same with (2.48) multiplying on the right by ψ

$$\bar{\psi}\left(i\overleftarrow{\partial\!\!\!/} + m\right)\psi = 0$$

Adding we get a conserved current

$$\partial_\mu\,(\bar{\psi}\gamma^\mu\psi) = 0 \to \partial_\mu J^\mu = 0 \tag{2.50}$$

From the properties of the γ_5 and γ^μ we can write it as

$$J^\mu = \bar{\psi}\gamma^\mu\psi = \bar{\psi}_L\gamma^\mu\psi_L + \bar{\psi}_R\gamma^\mu\psi_R \tag{2.51}$$

For this, we used the fact that

$$(\gamma_5\psi)^\dagger\gamma^0 = -\psi^\dagger\gamma^0\gamma_5$$

Our notation is that

$$\overline{(\psi_R)} = \bar{\psi}_R$$

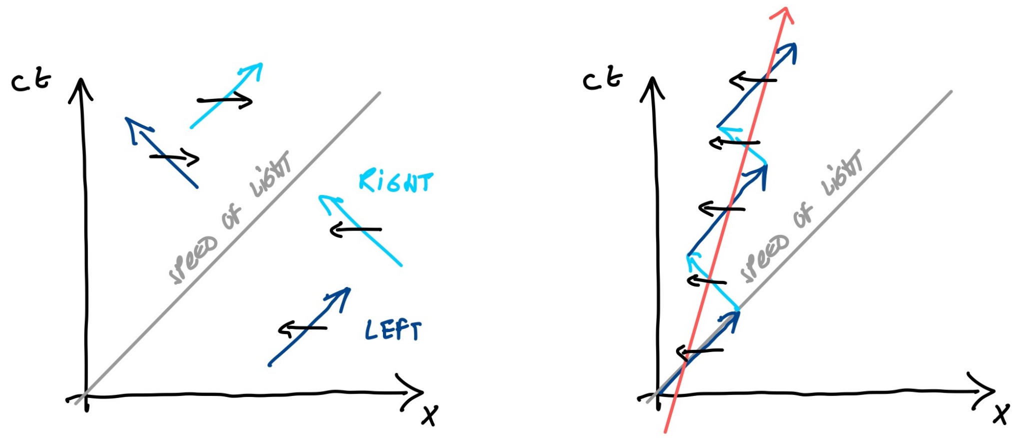


Figure 2.4: A picture of propagation of a massive fermion. Massless fermion (LHS) travel at the speed of light. A massive fermion (RHS) flips from one chirality to the other. Doing so, it travels at a speed less than 1. One may wonder how come that momentum is conserved in this process, but this is a quantum mechanical object, so the picture has its limits. The little black arrow refer to spin, so helicity states on the LHS, while on the RHS we have something suggestive of a spin pointing in the direction opposite to x.

However, beware that

$$\overline{(R\psi)} = \bar{\psi}L$$

where L is the L projector! So

$$\bar{\psi}_R\gamma^\mu\psi_R = \bar{\psi}R\psi \quad \text{and} \quad \bar{\psi}_L\gamma^\mu\psi_L = \bar{\psi}L\psi$$

The notation changes from one book to the other. For instance Peskin & Schroeder use R and L subscript for helicity... My notation tries to emphasize the fact that chirality is truly a label for Weyl spinors. For instance we will write

$$\bar{e}_R$$

for the adjoint of the spinor field of a R-handed electron. But this is not perfect world, and may lead to confusions.

These precautionary remarks being made, we see that the vector current preserves chirality as it couples a L (R) fermion with itself. It has also both L and R component, with the same weight. For that reason it is called the vector current.

As any conserved current, it is defined up to an overall factor. Using the normalization for the fields given in the Appendix G (same as Peskin- eder), we have, after normal ordering,

$$Q = \int d^3x\, j^0 = \int d^3x\, :\psi^\dagger(x)\psi(x): = \int \frac{d^3p}{(2\pi)^3} \sum_s \left(a^\dagger_{p,s}a_{p,s} - b^\dagger_{p,s}b_{p,s}\right) \tag{2.52}$$

summing over momentum and helicity creation/destruction operators. So Q counts, with weight one, the number of particles minus the number of anti-particles in a given incoming state. The coupling to the electromagnetic field A_μ is achieved by adding the following term to the Lagrangian (2.46)

$$\delta\mathcal{L} = -eA_\mu J^\mu \tag{2.53}$$

where $e = q_e/4\pi\epsilon_0$ is the charge of the electron (by convention, it is taken to be negative). In retrospect, the chiral structure of the current (2.51), the same for L and R components, explains a feature of the cross section $e^+e^- \to \mu^+\mu^-$ which we met in section 2.1. There we saw that the only non-zero cross-sections involve electrons and positrons (as well as muons and anti-muons) of opposite helicity. In the massless limit, this means spinors of the same chirality.

Going back to the Dirac equation and following the procedure of above to derive the vector current, but inserting this time the γ_5, leads to

$$\partial_\mu J_5^\mu = 2im\bar{\psi}\gamma_5\psi$$

with

$$J_5^\mu = \bar{\psi}\gamma^\mu\gamma_5\psi = \bar{\psi}_R\gamma^\mu\psi_R - \bar{\psi}_L\gamma^\mu\psi_L$$

which is conserved in the massless/chiral limit. We called it the chiral current because it changes sign when we exchange R and L-handed spinors. Another name that you will encounter is pseudovector. Indeed, the spatial components of the axial vector do not change sign under parity, like a pseudovector.

In the massless limit, both the chiral and vector current are conserved, and so are (a priori, there are subtle effects to which we will come back) linear combinations. In particular we can define purely R-handed and L-handed currents,

$$\boxed{J_R^\mu = \bar{\psi}_R\gamma^\mu\psi_R \equiv \bar{\psi}\gamma^\mu R\psi \quad \text{and} \quad J_L^\mu = \bar{\psi}_L\gamma^\mu\psi_L \equiv \bar{\psi}\gamma^\mu L\psi} \tag{2.54}$$

Again, the notation is ambiguous in the context of chiral theories, but hopefully you will get use to it.

A few words about early experiments. They have revealed 1) that parity is not conserved by weak interactions and 2) that the neutrinos that participates in weak processes are purely left-handed $\psi_{\nu L}$ (Golhaber et al, 1957). The vector and axial-vector couplings thus play a predominant role. For instance, the β decay process

$$n \to p + e^- + \bar{\nu}_e$$

involves the leptonic current

$$\bar{e}\gamma^\mu L\nu_e = \bar{e}\gamma^\mu\left(\frac{1-\gamma_5}{2}\right)\nu_e = \bar{e}\left(\frac{1+\gamma_5}{2}\right)\gamma^\mu\nu_e = \bar{e}_L\gamma^\mu\nu_e$$

which is of the (vector)-(axial vector) (or V-A) kind. An immediate consequence of this is that only the left-handed component of the electron (and generically of charged leptons) participate in charged weak processes.

The Wu experiment is very interesting to understand. It rests on the following β decay process

$${}^{60}_{27}\mathrm{Co} \to {}^{59}_{28}\mathrm{Ni} + e^- + \bar{\nu}_e \ (+2\gamma)$$

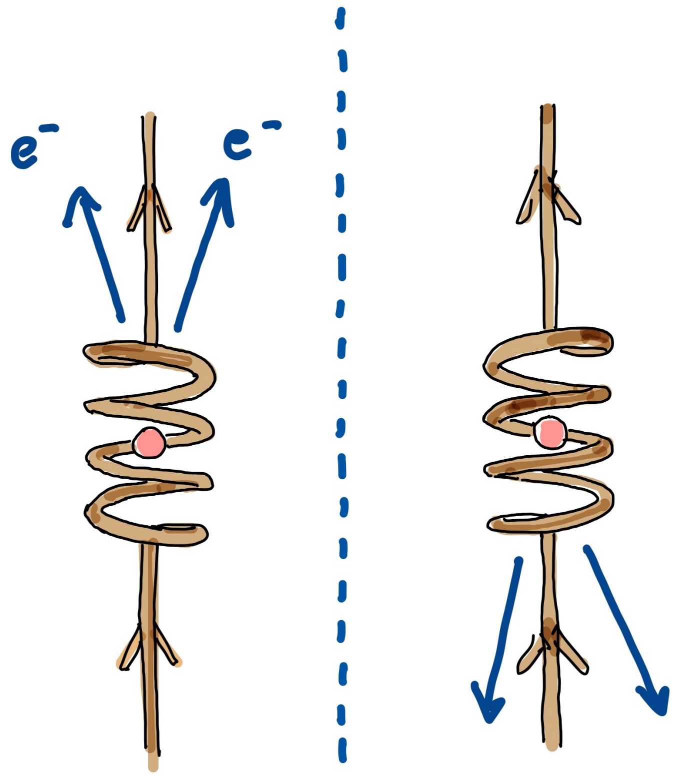


Figure 2.5: The Wu experiment and its mirror experiment. They differ, meaning that parity is broken. You should verify that the flux of electrons is opposite to the direction of the magnetic field.

Forget the gamma. What is important for us is that $J = 5$ for Cobalt and $J = 4$ for Nickel. The decay is observed in presence of a magnetic field $\vec{B}$ and imagine that the nuclear spin the initial and final states are aligned with the magnetic field (this is apparently very difficult experimentally). The main idea is that

$$O = \vec{p}_e \cdot \vec{B}$$

is a pseudoscalar as $\vec{p}_e$ is a vector and $\vec{B}$ a pseudovector. In other words, O changes sign under parity. If parity is conserved, if should be (on average) zero, $\langle O \rangle = 0$, meaning there should be no preferential direction for the electron to be emitted. Instead, the Wu experiment that electrons are preferably emitted in the direction opposite to the magnetic field, $\langle O \rangle < 0$. This, apparently, came as a shock. What is going on? Clearly the emitted electron and antineutrino must carry a total angular momentum equal to 1 (along the magnetic field direction) since $\Delta J = -1$. The Wu result is understood if the electron has a negative helicity (a left-handed electron) and the antineutrino is right-handed (the antiparticle of a L-handed neutrino). If there were both L and R handed neutrinos, coupled with the same strength, we could also have the emission along $\vec{B}$ of a positive helicity one electron and on average $\langle O \rangle = 0$, see Fig. 2.5.

As an another application of parity violation of weak interactions, consider the weak decay

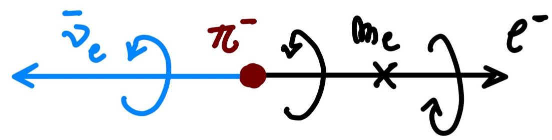


Figure 2.6: Pion decay, $\pi^- \to l^- + \bar{\nu}_l$. Conservation of angular momentum and chirality structure of weak decay requires a chirality flip (mass insertion = cross), so the amplitude for decay is $\propto m_l$, favoring decay into a muon rather than electron, despite the phase space suppression.

of a charged pion, say, π^- into a pair of leptons,

$$\pi^- \to l^- + \bar{\nu}_l$$

as shown in Figure 2.6 The leptons could be either an electron or a muon and their respective anti-neutrinos. According to the above discussion, the charged leptons have negative helicity and the anti-neutrinos have positive helicity. (Remember, anti-particles of negative chirality have positive helicity.) In the rest frame of the π^- the charged lepton and its antineutrino fly apart with opposite spatial momentum. Hence, this is a state of total angular momentum equal to one. Because a pion is a boson of spin zero, this process is forbidden to first order in the mass m of the charged lepton. To next order $O(m)$, it is allowed, as shown in the Figure 2.5. Hence, contrary to a naive expectation based on the available phase space, the pion decays preferably into a muon and its antineutrino, see section 6.3.2,

$$\frac{\Gamma(\pi \to e\bar{\nu}_e)}{\Gamma(\pi \to \mu\bar{\nu}_\mu)} \approx 1.2\,10^{-4}$$

Chapter 3

Representations of the Lorentz group

Key concepts: active transformation, generators, Lie algebra, structure constants, adjoint representation, defining representation, spinor.

After these introductory considerations, it is important to understand their link with group theory. Symmetry principles are indeed central to the construction of theories of fundamental interactions. The notations used in this chapter can be found in appendix B.

3.1 Lorentz transformations

Euclidian geometry is defined by limiting coordinate transformations to those that leave invariant the length of a vector,

$$\vec{x}\prime = R\vec{x} \rightarrow \vec{x}' \cdot \vec{x}' = \vec{x} \cdot \vec{x} = x^2$$

The 3x3 matrices R must be orthogonal, that is rotations (up to reversal of space directions, or improper rotations). All the implications of euclidian geometry can be derived from this condition.

Similarly, Minkowski geometry is built upon considering how mathematical objects (like a 4-vector) transform under Lorentz transformations. These are homogeneous transformations that preserve "lengths" in Minkowski space, $s^2 = \eta_{\mu\nu} x^\mu x^\nu$. They act linearly on the coordinate vectors

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu = \Lambda^\mu{}_0 x^0 + \Lambda^\mu{}_i x^i \tag{3.1}$$

They include, as we will see again, rotations and boosts, or changing of inertial frames,

More generally we write a coordinate transformation as

$$x'^\mu = x'^\mu(x^\nu) \tag{3.2}$$

with

$$\frac{\partial x'^\mu}{\partial x^\nu} \equiv \Lambda^\mu{}_\nu. \tag{3.3}$$

where the $\Lambda^\mu{}_\nu$ are constant (spacetime independent) for Lorentz transformations. The advantage of this notation is that it is to keep track of Lorentz indices.

To preserve $s^2 = \eta_{\mu\nu} x^\mu x^\nu$, we need that the 4×4 matrices Λ satisfy

$$s^2 = \eta_{\mu\nu} x'^\mu x'^\nu = \eta_{\mu\nu} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma x^\rho x^\sigma = \eta_{\rho\sigma} x^\rho x^\sigma \tag{3.4}$$

and thus

$$\boxed{\eta_{\mu\nu} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma = \eta_{\rho\sigma}} \tag{3.5}$$

or in matrix form

$$\Lambda^T \eta \Lambda = \eta. \tag{3.6}$$

This defines not only how Lorentz transformations must act on contravariant 4-vectors but actually that defines all of Minkowski geometry. Indeed, it leads to the invariance of the scalar product of two 4-vectors,

$$x' \cdot y' = \eta_{\mu\nu} x'^\mu y'^\nu = \eta_{\mu\nu} \Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma x^\rho y^\sigma = \eta_{\rho\sigma} x^\rho y^\sigma = x \cdot y \tag{3.7}$$

and also to the transformation law for covariant 4-vectors,

$$\begin{aligned} x' \cdot y' &= x'_\mu y'^\mu \hat{=} x_\alpha y^\alpha \\ &= x'_\mu \Lambda^\mu{}_\alpha y^\alpha \end{aligned} \tag{3.8}$$

which implies

$$x_\mu = x'_\nu \Lambda^\nu{}_\mu \tag{3.9}$$

As an application of the Lorentz as a coordinate transformation, we can write this as

$$x'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} x_\nu. \tag{3.10}$$

Indeed

$$\frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x'^\mu}{\partial x^\beta} = \frac{\partial x^\alpha}{\partial x^\beta} = \delta^\alpha_\beta. \tag{3.11}$$

and thus[1]

$$x'_\mu = x_\nu \, (\Lambda^{-1})^\nu{}_\mu \tag{3.12}$$

So far we have only considered four-vectors associated to coordinates, with contravariant or with covariant indices. Now the next step is to see that the very same transformations rules can be applied to any covariant and contravariant vectors. This is actually how mathematicians define a geometry. So,

$$V'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} V^\nu \tag{3.13}$$

[1]Here, to avoid cluttering of symbols, we refrain from writing matrix form for all these transformation laws. Indeed, x^μ and y_μ are quite distinct objects (*ie* they leave in different spaces, which are said to be dual to each others), which however may be related through the metric η, provided it is defined. Formally, x^μ are the components of a vector, let us write this as $\mathbf{x}$, while the y_μ are component of a so-called 1-form, that we may write as $\tilde{\mathbf{y}}$. A 1-form is defined to act on vectors so as to make a number $\tilde{\mathbf{y}}(\mathbf{x}) \hat{=} y_\mu x^\mu$. There is some beautiful geometry behind all this but which is however of limited practical interest at our level, so we don't elaborate further this point. General Relativity, which is deeply rooted in geometry, will take you to this next level of understanding.

which is is consistent with

$$\partial'_\mu = \frac{\partial x^\nu}{\partial x'^\mu}\partial_\nu \tag{3.14}$$

and the requirement that

$$\partial_\mu V^\mu \tag{3.15}$$

is a scalar quantity:

$$\partial'_\mu V'^\mu = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x'^\mu}{\partial x^\beta}\partial_\alpha V^\beta = \delta^\alpha_\mu \delta^\mu_\beta \partial_\alpha V^\beta = \partial_\mu V^\mu. \tag{3.16}$$

3.2 Elementary properties of Lorentz transformations

Consider the determinant of (3.5)

$$\det\eta = \det\Lambda^T\ \det\eta\ \det\Lambda \tag{3.17}$$

which gives

$$\det\Lambda = \pm 1 \tag{3.18}$$

The case $\det\Lambda = 1$ (-1) corresponds to so-called proper (resp. improper) Lorentz transformations. The following improper transformation $\det\Lambda = -1$ is $\Lambda = P = \mathrm{diag}(1,-1,-1,-1) \equiv \eta$ corresponds to space inversion, $t \to t, \mathbf{x} \to -\mathbf{x}$, hence the name P, with P for Parity, as defined in the previous section. Any improper transformation can be written as the product of P with a proper transformation.

Next consider the 00 entry of Eq.(3.5),

$$1 = \eta_{\mu\nu}\Lambda^\mu{}_0\Lambda^\nu{}_0 = (\Lambda^0{}_0)^2 - (\Lambda^i{}_0)^2 \tag{3.19}$$

A simple solution is

$$\Lambda^0{}_0 = \pm 1 \tag{3.20}$$

and

$$\Lambda^1{}_0 = 0 \tag{3.21}$$

A possibility is

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \tag{3.22}$$

Since the spatial components of the Minkowski metric is just (minus) the euclidian metric, R is a rotation matrix, which depends on three parameters (eg rotations along the 3 cartesian axis). There is another possibility,

$$\Lambda^0{}_0 = \cosh\eta \tag{3.23}$$

and (say for a boost in direction $x = 1$)

$$\Lambda^1{}_0 = \sinh\eta \tag{3.24}$$

and the same for the 2 other spatial directions. These are the three boosts (see below), or change of inertial frames.

Also, (3.17) implies that

$$|\Lambda^0{}_0| \geq 1. \tag{3.25}$$

Lorentz transformations with $\Lambda^0{}_0 \geq 1$ (≤ -1) are called orthochronous (resp. non-orthochronous) – they preserve the orientation of time. An element of the non-orthochronous transformation is T, the time reversion operator, $T = \mathrm{diag}(-1, 1, 1, 1)$. Any non-orthochronous transformation can be obtained from an orthochronous transformation followed by T. The set of Lorentz transformations form a group, which as we have seen has four disconnected components.[2] The subset of proper Lorentz transformations form a subgroup noted SO(3, 1) (sometime $\Lambda_+^\uparrow$). The subgroup of proper, orthochronous Lorentz transformations is called SO$^+$(3, 1), but is most of the time loosely referred to as SO(3, 1). This subgroup contains all the Lorentz transformations that may be continuously connected to the identity, the trivial Lorentz transformation. The others Lorentz transformations may be obtained from this subgroup and the discrete symmetry transformations P (parity) and T (time reversal).

In summary, the Lorentz group $\mathrm{SO}(3, 1) = \Lambda_+^\uparrow$ has dimension 6, in the sense that a general Lorentz transformation depends on six real parameters. Three of these parameters correspond to rotations in 3-dimensional euclidean space, $x'^0 = x^0$, $x'^i = R^{ij}x^j$, with R^{ij} is a 3×3 orthogonal matrix, with $\det R = 1$ (proper rotations),

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \tag{3.26}$$

For instance

$$\left\{ \begin{array}{lcc} t' & = & t \\ x' & = & x\cos\theta - y\sin\theta \\ y' & = & x\sin\theta + y\cos\theta \\ z' & = & z \end{array} \right. \tag{3.27}$$

for a rotation by an angle θ around the z direction. Proper rotations (*ie* rotations that preserve the orientation of space) form a subgroup of the Lorentz group, noted $\mathrm{SO}(3) \subset \mathrm{SO}(3, 1)$. The other three parameters corresponds to boosts. For instance

$$\left\{ \begin{array}{ccc} t' & = & t\,\cosh\eta + x\,\sinh\eta \\ x' & = & t\,\sinh\eta + x\,\cosh\eta \\ y', z' & = & y, z \end{array} \right. \tag{3.28}$$

or in matrix form

$$\Lambda = \begin{pmatrix} \cosh\eta & \sinh\eta & 0 & 0 \\ \sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.29}$$

[2] A group G is a set of elements g that satisfy the following properties:

Closure: For any two elements of the group $g_1, g_2 \in G$, then $g_1 \circ g_2 = g_3 \in G$ ($\circ$ is called the group law of G).

Associativity: For all g_1, g_2, g_3, $g_1 \circ g_2) \circ g_3 = g_1 \circ (g_2 \circ g_3)$.

Identity: There is an element e such that $e \circ g = g \circ e = g$ for all g.

Inverse: For all g, $g^{-1} \in G$, with $g^{-1} \circ g = g \circ g^{-1} = e$.

This transformation represents a boost (this is also call an active transformation) from a frame O to a frame O' moving in the positive direction x at velocity v. The parameter η (not be to confused with the Minkowski metric) is called rapidity. To see the relation with velocity, consider the wordline of an object at rest at $\mathbf{x} = 0$ in the unprimed frame. We boost it in the direction x. Then its coordinates (primed coordinates) become

$$t' = t\cosh\eta \quad \text{and} \quad x' = t\sinh\eta \tag{3.30}$$

and its velocity is given by

$$\frac{x'}{t'} = v = \frac{t\sinh\eta}{t\cosh\eta} = \tanh\eta. \tag{3.31}$$

Thus $\tanh\eta = v$ and

$$\cosh\eta = \frac{1}{\sqrt{1-v^2}} \quad , \quad \sinh\eta = \frac{v}{\sqrt{1-v^2}} \tag{3.32}$$

For non-relativistic velocities, $v \ll 1$ (meaning $v \ll c$), we should recover the Galilean transformation. There are hidden factors of c so we really mean $c \to \infty$ for fixed v, or

$$\begin{cases} t' &= t \\ x' &= vt \\ y', z' &= y, z \end{cases}$$

Boosts have nice properties. If we take two boosts of rapidity η_1 and η_2 along the same direction, then the total rapidity is $\eta_1 + \eta_2$. So boost along the same direction, like consecutive rotations in a given plane, form a an abelian subgroup, isomorphic to addition $\mathbb{R}_+$. However, boots do not form a subgroup of SO(3, 1), unlike the rotations. Actually, the product of two boosts (along distinct spatial directions) has a rotation part. This is most transparent in the language of generators, see below. Finally, we notice that since Lorentz transformations preserve $s^2 = t^2 - \vec{x}^2$, the orbits of Lorentz transformations are given by two upper and lower sheet of the hyperboloid $s^2 > 0$ for a time-like interval and the one-sheet hyperboloid of revolution $s^2 < 0$ for a space-like interval. The latter implies that simultaneity is relative to the observer.

3.3 Element of representation theory

Generators of rotations

Consider something that should be familiar, the group of rotations in 3-dimensions. Rotations may act on different mathematical objects: vectors, functions, etc. For each case, there is a specific representation of the group elements. The natural representation is the one acting on vectors. A word of caution first. We have been a bit vague so far, but we consider in these lectures transformations that act on physical objects themselves, not on the coordinate axis. This is sometime called an active transformation. A counter-clockwise rotation of a point at $\mathbf{x}$ lying in,

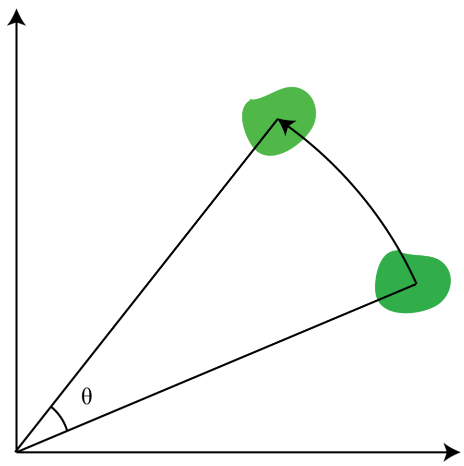


Figure 3.1: This is an active transformation on a coordinate point. All vectors are sent to other vector, $\vec{x} \to \vec{x}'$.

say the plane xy, by an angle θ around the axis z, brings it extremity to a point with coordinates[3]

$$\begin{cases} x' &= x\cos\theta - y\sin\theta \\ y' &= x\sin\theta + y\cos\theta \end{cases} \tag{3.33}$$

or

$$\vec{x}' = R(\theta)\vec{x} \tag{3.34}$$

A general rotation in $D = 3$ is represented by 3×3 matrice $R(\vec{\theta})$, with $\vec{\theta}$ is a 3-vector that defines the axis of the rotation and $|\vec{\theta}|$ the angle of the rotation. The 3×3 matrices $R(\vec{\theta})$ are characterized by three parameters (corresponding to the three components of $\vec{\theta}$), and they are orthogonal $R^T = R^{-1} \equiv R(-\vec{\theta})$ so as to preserve the scalar product $\mathbf{x} \cdot \mathbf{y}$. **Proper rotations** are such that $\det R(g) = 1$, and the subgroup of these transformations is noted SO(3) (the set of special orthogonal matrices of dimension 3).

The Lie algebra of SO(3) arises when one consider infinitesimal transformations, $\theta^a \ll 1$, and rewrite R as

$$R^{ij}x^j = x'^i = x^i + \delta x^i = (\delta^{ij} + \theta^{ij})x^j \tag{3.35}$$

Let us work out the requirement that $\mathbf{x}^2 = \mathbf{x}'^2$, which implies that $R^{-1} = R^T$. To leading order

$$\delta^{ij} - \theta^{ij} = \delta^{ji} + \theta^{ji} \tag{3.36}$$

or

$$\theta^{ij} + \theta^{ji} = 0 \tag{3.37}$$

which implies that θ^{ij} is antisymmetric. This means that there are three independent components, corresponding to three rotation angles.

[3] A passive transformation amounts to a clockwise rotation of the coordinate axes. There is no big deal, but we have to fix the convention about what we mean by a rotation by a positive angle.

An explicit representation of θ^{ij} in terms of the rotation angles $\vec{\theta}$ is given by

$$\theta^{ij} = -\epsilon^{ija}\theta^a \equiv \theta^a (T^a)^{ij} \tag{3.38}$$

where ϵ^{ijk} is the Levi-Civita in $D = 3$ and $a = 1, 2, 3$. Do not get confused by the fact that there are the same number of rotation angles as there are dimensions: this is specific to $D = 3$. The T^a are three antisymmetric matrices called the generators of rotations (in $D = 3$). Finite transformations may be formulated using the matrix exponent defined as an infinite series

$$R = \exp(\theta^a T^a) \equiv \sum_{n=0}^{\infty} \frac{(\theta^a T^a)^n}{n!}. \tag{3.39}$$

Note that $\det R = 1$ is automatic since $\mathrm{Tr} T^a = 0$.[4] Explicitely we have

$$T^1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad T^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad T^3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.40}$$

From these expressions, you may check that the matrices T^a obey the following commutation relations

$$[T^1, T^2] = T^3, \quad [T^2, T^3] = T^1, \quad [T^3, T^1] = T^2 \tag{3.41}$$

which we may rewrite as

$$[T^a, T^b] = \epsilon^{abc} T^c \tag{3.42}$$

These commutation relations fully characterize the so-called **Lie algebra** (sometime written so(3)) of the group SO(3). The ϵ^{abc} are called the structure constants of the Lie algebra. Incidentally, we have indirectly discovered that the structure constants themselves provide an explicit representation of the Lie algebra, called the **adjoint representation**,

$$\boxed{(T^a)^{ij} = -\epsilon^{aij}} \tag{3.43}$$

Now we are generally interested of the action of rotation on other objects than mere vectors. In particular in quantum mechanics we typically consider the action of a rotation on a wave function $\psi(\mathbf{x})$ of a particle. Up to a phase factor which we may choose to be unity, the wave function $\psi\prime$ describing the rotated state at $\mathbf{x}'$ must be equal to the original wave function ψ at $\mathbf{x}$ (this is again an active transformation, see Fig.3.3)

$$\psi'(\mathbf{x}') = \psi(\mathbf{x}) \tag{3.44}$$

This means that the wave function is a scalar quantity under a rotation (a transformation in general). Now we want to relate ψ and ψ' by a unitary operator such that

$$\psi'(\mathbf{x}) = U_R \psi(\mathbf{x}). \tag{3.45}$$

[4]For a generic square matrix A we may write $A = \exp(\log(A))$. Since $\det A$ is the product of the eigenvalues of A, $\det A = \exp(\mathrm{Tr} \log(A))$. If $A = 1 + B$, with B a matrix with small eigenvalues, we get that $\det(1 + B) = \exp(\mathrm{Tr} \log(1 + B)) \approx 1 + \mathrm{Tr} B$.

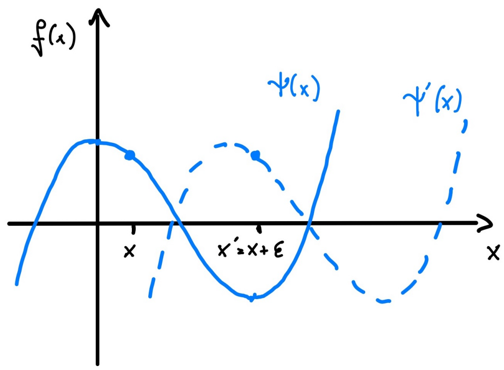


Figure 3.2: This is an active transformation (here a translation) on scalar function. A scalar function is translated to another function, $\psi(x) \to \psi'(x') = \psi(x)$ with $x' = x + \epsilon$. For ϵ infinitesimal, $\psi'(x) = \psi(x - \epsilon) \approx \psi(x) - \epsilon\, d\psi/dx|_x$

This we may find using the fact that

$$\psi'(\mathbf{x}) = \psi(R^{-1}\mathbf{x}). \tag{3.46}$$

To keep things simple let us work out this for the case of the rotation of (3.33) and an infinitesimal rotation angle θ:

$$\begin{aligned}
\psi'(x, y, z) &= \psi(x + y\theta, y - x\theta, z) \\
&= \psi(x, y, z) + \theta(y\partial_x - x\partial_y)\psi(x, y, z) \\
&= \psi(x, y, z) - i\theta\left(x(-i\partial_y) - y(-i\partial_x)\right)\psi(x, y, z) \\
&= (1 - i\theta L_z)\psi(\mathbf{x})
\end{aligned} \tag{3.47}$$

which shows that the angular momentum L_z is the generator of infinitesimal rotations around the axis z. For a general rotation

$$\psi'(\mathbf{x}) = e^{-i\theta^a L^a}\psi(\mathbf{x}). \tag{3.48}$$

As you know well, the angular momentum operators are hermitian and satisfy the commutation rotations

$$[L^a, L^b] = i\epsilon^{abc}L^c. \tag{3.49}$$

We have now two distinct representations of rotations: on vectors and on wave functions. Let us put this in a more unified perspective. In particular, let $T^a \to J^a = iT^a$, so that the J^a are Hermitian matrices ($(J^a)^\dagger = J^a$), which, from Eq.(3.42) satisfy

$$\boxed{[J^a, J^b] = i\epsilon^{abc}J^c} \tag{3.50}$$

We conclude that the commutation relations of the R^a and the L^a may be put in the same form, despite the fact that we have distinct representations. Of course this is in part why they are useful:

they completely define the Lie algebra of the rotation group. From now on, we write J^a for any representation of the rotation group. Furthermore we will write a finite rotation as

$$R(\theta) = \exp(-i\theta^a J^a) \tag{3.51}$$

One interest of these manipulations is that we all know three traceless Hermitian 2×2 matrices that satisfy the commutation relations (3.50). Take $J^a = \sigma^a/2$, where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.52}$$

are the Pauli matrices. In this representation, the J^a act on two-dimensional complex objects,

$$\chi = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} \tag{3.53}$$

called (surprise, surprise,...) spinors. The unitary matrices

$$U = e^{-\frac{i}{2}\theta^a\sigma^a} \tag{3.54}$$

tell us how a spinor transforms under a rotation

$$\chi' = e^{-\frac{i}{2}\theta^a\sigma^a}\chi \tag{3.55}$$

These are elements of the group of Special Unitary linear transformations of dimension 2, or SU(2).

We just saw that the group SO(3) and SU(2) shared the same Lie algebra. This hints at an isomorphism between the two groups. To see this, consider a generic 2 traceless Hermitian matrix

$$\mathbf{V} = V^a\sigma^a = \begin{pmatrix} V^3 & V^1 - iV^2 \\ V^1 + iV^2 & -V^3 \end{pmatrix} \tag{3.56}$$

where V^a are the components of a real vector in 3d.

The norm of a vector may be defined as

$$|\det \mathbf{V}| = (V^1)^2 + (V^2)^2 + (V^3)^2 \tag{3.57}$$

Now a natural action of SU(2) transformations U on $\mathbf{V}$ is

$$\mathbf{V} \ \to \mathbf{V}' = U\mathbf{V}U^\dagger \tag{3.58}$$

This transformation mixes the components V^a while preserving the norm,

$$\det \mathbf{V}' = \det(U\mathbf{V}U^\dagger) = \det \mathbf{V}. \tag{3.59}$$

Hence to any U element of SU(2) we may associate a rotation of a 3-vector, hence an element of SO(3). However U and $-U$ give the same rotation, hence the isomorphism is actually

$$\mathrm{SO}(3) \simeq \mathrm{SU}(2)/\mathbb{Z}_2 \tag{3.60}$$

This means that transformations of SU(2) that differ by a sign are identified (this is the meaning of the quotient). Notice that, since

$$\exp(-\frac{i}{2}\theta^a\sigma^a) = \cos\frac{\theta}{2} - i\hat{\theta}^a\sigma^a \sin\frac{\theta}{2} \tag{3.61}$$

where $\hat{\theta}$ is the unit vector corresponding to θ, a rotation by $\theta = 2\pi$ corresponds to $U = -\mathbb{I}_2$. Obviously these are identified for rotations on vector, but a full rotation of a spinor corresponds to $\theta = 4\pi$. This may be strange, but they are many things, not as weird as spinors, that require a 4π rotation to come back to the identity (see course).

As a preparation for Lorentz transformations, let us now take a slightly different road. We have just seen that the angular momentum operators are generators of rotations,

$$\mathbf{J} = \mathbf{x} \times \mathbf{p} \tag{3.62}$$

or

$$\mathbf{J} = \mathbf{x} \times (-i\nabla) \tag{3.63}$$

in the $\mathbf{x}$ representation.

Now the cross-product is specific to 3 dimensions, so we would like to have something more general, possibly valid in any dimensions. So consider the antisymmetric operator

$$\boxed{J^{ij} = -i(x^i\nabla^j - x^j\nabla^i)}. \tag{3.64}$$

We may use the Levi-Civita tensor with three indices

$$J^i = \frac{1}{2}\epsilon^{ijk}J^{jk} \tag{3.65}$$

to relate the J^i and the J^{ij}, but this is only valid in three dimensions. Conversely

$$J^{ij} = \epsilon^{ijk}J^k. \tag{3.66}$$

Now, although this is a bit cumbersome, you should check that Eq.(3.64) implies that the J^{ij} satisfy

$$\boxed{\left[J^{ij}, J^{kl}\right] = i\,\delta^{ik}J^{jl} - i\,\delta^{jk}J^{il} + i\,\delta^{jl}J^{ik} - i\,\delta^{il}J^{jk}} \tag{3.67}$$

Now, these commutation relations may be used to define the Lie algebra of the group of rotation in any dimensions n, SO(n) and should be valid for any representation. The form given in Eq.(3.50) is specific to three dimensions, accordingly an important case. Notice finally that in D dimensions, there are $D(D-1)/2$ independent rotation angles and thus generators.

Generators of Lorentz transformations

For Lorentz transformation, we basically mimick the approach followed for rotations. After all, rotations form a subgroup of Lorentz transformations. We pose

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu} \tag{3.68}$$

with $\omega^{\mu}{}_{\nu}$ a set of infinitesimal parameters, so that the transformations is close to the identity. Then

$$x'^{\mu} = x^{\mu} + \delta x^{\mu} \equiv x^{\mu} + \omega^{\mu}{}_{\nu} x^{\nu}. \tag{3.69}$$

with δx^{μ} infinitesimal. The transformation rule of Eq.(3.5) becomes

$$0 = \omega^{\alpha}{}_{\mu}\eta_{\alpha\nu} + \eta_{\mu\beta}\omega^{\beta}{}_{\nu} \quad \text{or} \quad \boxed{0 = \omega_{\nu\mu} + \omega_{\mu\nu}} \tag{3.70}$$

which implies that $\omega_{\mu\nu}$ is a 4×4 antisymmetric matrix, that altogether may have $4 \cdot 3/2 = 6$ independent components, as expected: three rotations and three boosts.

Let us rewrite this infinitesimal transformation using generators. We know that three of the Lorentz transformations are rotations in 3-dimensions (rotations are a subgroup of the Lorentz group). We have considered these transformations in the previous section. Since the group of rotations is a subgroup of the Lorentz group, from Eq.(3.64) it is very natural to introduce the antisymmetric operator

$$J^{\mu\nu} = i(x^{\mu}\partial^{\nu} - x^{\nu}\partial^{\mu}) \tag{3.71}$$

as the generator of Lorentz transformation acting on a scalar function of x. When writing this remember that $\partial^{\mu} = (\partial_t, -\nabla)$. You may verify that the $J_{\mu\nu}$ satisfy

$$\boxed{\left[J_{\mu\nu}, J_{\rho\sigma}\right] = i\,\eta_{\nu\rho}J_{\mu\sigma} - i\,\eta_{\mu\rho}J_{\nu\sigma} - i\,\eta_{\nu\sigma}J_{\mu\rho} + i\,\eta_{\mu\sigma}J_{\nu\rho}} \tag{3.72}$$

From our experience with rotations, we expect that these relations completely define the Lie algebra so(3, 1) of the group SO(3, 1). Again, they do not depend on the specific representation chosen for the $J^{\mu\nu}$.

Notice that we worked in a reversed way compared the case of rotations. Can we get the representation of Lorentz generators on 4-vectors? Sure, let us act with our $J_{\mu\nu}$ on x^{α}. We have

$$\begin{aligned} J_{\mu\nu}x^{\alpha} &\equiv i(x_{\mu}\partial_{\nu} - x_{\nu}\partial_{\mu})x^{\alpha} \\ &= i(x_{\mu}\delta^{\alpha}_{\nu} - x_{\nu}\delta^{\alpha}_{\mu}) \qquad (3.73) \\ &= i(\eta_{\mu\beta}\delta^{\alpha}_{\nu} - \eta_{\nu\beta}\delta^{\alpha}_{\mu})x^{\beta} \qquad (3.74) \end{aligned}$$

If we define

$$\left(\mathcal{J}_{\mu\nu}\right)^{\alpha}_{\ \beta} = i\left(\delta^{\alpha}_{\mu}\eta_{\nu\beta} - \delta^{\alpha}_{\nu}\eta_{\mu\beta}\right) \tag{3.75}$$

then you may check

$$\delta x^{\alpha} = -\frac{i}{2}\omega^{\mu\nu}(\mathcal{J}_{\mu\nu})^{\alpha}{}_{\beta}x^{\beta} \tag{3.76}$$

that we have the right transformation law of Eq.(3.69),

$$\delta x^{\alpha} = \omega^{\alpha}_{\ \beta}x^{\beta}. \tag{3.77}$$

For the sake of clarity, we consider two specific instances. First the generator for a rotation by an angle θ around the third axis. This is given by $\mathcal{J}^{12}$ and $\omega_{12} = \theta$,

$$\Lambda(\theta) = \mathbb{I}_4 - i\theta\mathcal{J}^{12} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -\theta & 0 \\ 0 & \theta & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.78}$$

For a boost in the x direction we need $\mathcal{J}^{01}$ and $\omega_{01} = \eta$, the rapidity, and

$$\Lambda(\eta) = \mathbb{I}_4 - i\eta\mathcal{J}^{01} = \begin{pmatrix} 1 & \eta & 0 & 0 \\ \eta & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.79}$$

Finally, from all this we may write a finite Lorentz transformation acting on a (here covariant) vector V^μ as

$$V'^\alpha = (e^{-\frac{i}{2}\omega^{\mu\nu}\mathcal{J}_{\mu\nu}})^\alpha{}_\beta V^\beta \equiv \Lambda^\alpha{}_\beta V^\beta \tag{3.80}$$

where here the components of $\omega_{\mu\nu}$ are arbitrary (not specially small).

3.4 Building Lorentz invariants

You might think that we have done little but rewriting known things in a more complicated way but we have actually learned a great deal. To start with we now know how to transform an arbitrary (contra- or covariant-) vector. We have already seen how to transform the 4-momentum p^μ but in this case it was trivially related to the transformation of the x^μ. Another instance of a 4-vector is the electromagnetic 4-potential field

$$A^\mu = (\phi, \mathbf{A}) \tag{3.81}$$

where ϕ is the electrostatic potential and $\mathbf{A}$ the 3-vector potential. According to the previous discussion, this object must transform as

$$A'^\alpha = \frac{\partial x'^\alpha}{\partial x^\beta} A^\beta \tag{3.82}$$

That's it, we don't have to work more. We can lower the index of A^μ using the metric, $A_\mu = \eta_{\mu\nu}A^\nu$. Since

$$A'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} A_\nu \tag{3.83}$$

we may write with A^μ various Lorentz invariant quantities. For instance

$$A'_\mu A'^\mu = \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x'^\mu}{\partial x^\sigma} A_\rho A^\sigma = \delta^\rho_\sigma A_\rho A^\sigma = A_\rho A^\rho \tag{3.84}$$

or

$$\partial'_\mu A'^\mu = \partial_\nu A^\nu \tag{3.85}$$

Why do we focus on Lorentz invariants? Well, we want the physics to be the same in all inertial frame, related by Lorentz transformations (and translations - see later). If we start with invariants, we do not have worry anymore; this will be automatic.

How about more complicated objects? Consider for instance the electromagnetic field tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{3.86}$$

This is a complicated object, with six physical quantities, the components of the electric and magnetic fields

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix} \tag{3.87}$$

However the transformation law is easy

$$F'_{\mu\nu} = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}F_{\alpha\beta}. \tag{3.88}$$

For instance the following quantity

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} = \frac{1}{2}(\mathbf{E}^2 - \mathbf{B}^2) \tag{3.89}$$

is scalar under Lorentz (beware, $F_{\mu\nu} = -F_{\nu\mu}$).

Another important object is the 4d Levi-Civita tensor, $\epsilon^{\alpha\beta\rho\sigma}$, the completely antisymmetric tensor in 4d, with $\epsilon^{0123} = +1$ (so that $\epsilon_{\alpha\beta\rho\sigma} = -1$), which is invariant under Lorentz transformations,

$$\frac{\partial x'^\alpha}{\partial x^\mu}\frac{\partial x'^\beta}{\partial x^\nu}\frac{\partial x'^\gamma}{\partial x^\rho}\frac{\partial x'^\delta}{\partial x^\sigma}\epsilon^{\mu\nu\rho\sigma} \equiv J(\partial x'/\partial x)\epsilon^{\alpha\beta\gamma\delta} = \det(\Lambda)\epsilon^{\alpha\beta\gamma\delta} = \epsilon^{\alpha\beta\gamma\delta} \tag{3.90}$$

where J is the Jacobian of the transformation $x' = x'(x)$.

From the Levi-Civita and the electromagnetic field tensors we may build another antisymmetric tensor (with two indices), called the dual,

$$\tilde{F}_{\alpha\beta} = \frac{1}{2}\epsilon_{\alpha\beta\mu\nu}F^{\mu\nu}. \tag{3.91}$$

Then the following quantity is invariant:

$$\frac{1}{4}F_{\mu\nu}\tilde{F}^{\mu\nu} = -\mathbf{E}\cdot\mathbf{B} \tag{3.92}$$

Note that these two Lorentz invariants have distinct properties under parity transformations. Indeed, you know that $\mathbf{E}$ is a vector under parity (an improper transformation), $\mathbf{E} \to -\mathbf{E}$, while $\mathbf{B}$ is a pseudopseudovectorvector, $\mathbf{B} \to \mathbf{B}$. Hence the quantity of (3.89) is a scalar, while (3.92) is a pseudo-scalar.

Note that in practice Lorentz covariance[5] is as much useful as Lorentz invariance, in particular when we need to consider the possible relations between some physical quantities – *ie* want

[5] By this is meant the quality of transforming in a *a bona fide* way under Lorentz transformations

to build equations. By covariance we mean a physical quantity that transforms under a given representation of the Lorentz transformations. For instance

$$\partial_\mu F^{\mu\nu} = J^\nu_E \tag{3.93}$$

transforms as a contravariant vector and so may be equal to a contravariant vector J^ν. Also, by antisymmetry of $F^{\mu\nu}$,

$$\partial_\nu \partial_\mu F^{\mu\nu} = 0 \rightarrow \partial_\nu J^\nu_E = 0 \tag{3.94}$$

In components, this gives $J^\nu_E = (\rho_E, \mathbf{j_E})$,

$$\partial_t \rho_E + \vec{\nabla} \cdot \vec{j}_E = 0 \tag{3.95}$$

which is the conservation of a current density. Of course this is nothing but the inhomogeneous Maxwell equations (note that the magnetic permeability of vacuum $\mu_0 = 1$ in our system of units), with J^ν the electric current (E is for electric). As for the homogeneous Maxwell equations they are given by

$$\partial_\alpha \tilde{F}^{\alpha\beta} = 0 \tag{3.96}$$

where $\tilde{F}^{\mu\nu}$ is given by Eq.(3.91).[6] This equation can be rewritten in the form $\partial_\gamma F_{\alpha\beta} + \partial_\beta F_{\gamma\alpha} + \partial_\alpha F_{\beta\gamma} = 0$.

3.5 Lorentz symmetry and spinors

How about Dirac spinors? Consider the following 6 matrices,

$$S^{\mu\nu} = \frac{i}{4}\left[\gamma^\mu, \gamma^\nu\right] \tag{3.97}$$

A standard but laborious exercise is to verify that these 6 matrices obey the Lie algebra of the Lorentz group, Eq.(3.72). Alternatively, you can verify that

$$\psi \rightarrow e^{-i\frac{1}{2}\omega_{\mu\nu}S^{\mu\nu}}\psi$$

leaves the Dirac equation invariant. Either way, a massive Dirac spinor can be defined as a geometric object that transforms according to an the above representation of the Lorentz group SO(3, 1),

In the limit of zero mass, however, this representation is reducible. Indeed, in the Weyl representation the boost and rotation generators are block diagonal,

$$S^{0k} = \frac{i}{4}\left[\gamma^0, \gamma^k\right] = \frac{-i}{2}\begin{pmatrix} \sigma^k & 0 \\ 0 & -\sigma^k \end{pmatrix} = \begin{pmatrix} K^i & 0 \\ 0 & -K^i \end{pmatrix} \tag{3.98}$$

[6] There is an intriguing symmetry between Eq.(3.93) and (3.96), at least in vacuum (no electric charge). It is interesting to consider the possibility of introducing a magnetic current J^ν_M (and thus of magnetic charge or magnetic monopole), but this is not observed in Nature (as of today) and is anyway beyond the scope of this course.

and

$$S^{ij} = \frac{i}{4}\left[\gamma^i, \gamma^j\right] = \frac{1}{2}\epsilon_{ijk}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix} = \epsilon^{ijk}\begin{pmatrix} J^k & 0 \\ 0 & J^k \end{pmatrix} = \epsilon^{ijk}S^k \tag{3.99}$$

The latter is nothing but the spin operator introduced in Eq.(2.31). We wrote J^i and $\pm K^i$ the generators of respectively rotations and boosts. In the Weyl basis, these generators act on Weyl spinors χ_L and χ_R. The sign difference between the boost generators is the reason why χ_L and χ_R are distinct geometrical objects!

It is easy to check that the generators satisfy the following commutation relations

$$[\, J_i, J_j] \; = i\epsilon_{ijk}J_k \tag{3.100}$$

$$[\, K_i, K_j] \; = -i\epsilon_{ijk}J_k \tag{3.101}$$

and

$$[\, J_i, K_j] \; = i\epsilon_{ijk}K_k \tag{3.102}$$

The first one shows that rotations form a subgroup of Lorentz transformations; the second means that doing two boost along different directions and then the reverse is equivalent to a rotation. This can also be checked directly using the representation of boosts on 4-vectors, but is clear in terms of the Lie algebra. Thus boosts do not form a subgroup of SO(3, 1). Now consider the following linear combinations,

$$\vec{N}_L = \frac{1}{2}(\vec{J} + i\vec{K}) \qquad \text{and} \qquad \vec{N}_R = \frac{1}{2}(\vec{J} - i\vec{K}) \tag{3.103}$$

There form 2 sets of 3 hermitian operators. At this level, it is important to consider that the $\vec{J}$ and $\vec{K}$ are independent objects which satisfy the above commutation relations. You can verify that

$$[\, N_L^i, N_L^j] \; = i\epsilon_{ijk}N_L^k \tag{3.104}$$

$$[\, N_R^i, N_R^j] \; = i\epsilon_{ijk}N_R^k \tag{3.105}$$

and

$$[\, N_L^i, N_R^j] \; = 0 \tag{3.106}$$

where are isomorphic to 2 independent su(2) Lie algebras. These show that so(3, 1) $\sim$ su(2) $\times$ su(2) (isomorphism of the algebras). Furthermore, the su(2) can be mapped on each other through

$$P: \qquad J_i \to J_i \quad , \quad K_i \to -K_i$$

Analogous to the square of the angular momentum for rotations, also have two independent hermitian Casimir operators, which can be used to label the states, N_L^2 and N_R^2 with

$$N_L^2 \quad \text{with eigenvalues} \quad n_L(n_L + 1) \tag{3.107}$$
$$N_R^2 \quad \text{with eigenvalues} \quad n_R(n_R + 1) \tag{3.108}$$
$$\tag{3.109}$$

so that Lorentz representations are labelled by $(n_L/2, n_R/2)$ (so books use (n_L, n_R) with no factor $1/2$). Since $\vec{J} = \vec{N}_L + \vec{N}_R$, the spin of a representation is given by $(n_L + n_R)/2$. For Weyl spinors,

$$\psi_L \equiv \chi_L \sim (1/2, 0)$$

and

$$\psi_R \equiv \chi_R \sim (0, 1/2)$$

What is a Dirac spinor? Parity exchanges the n_L and n_R and a Dirac spinor is parity symmetric, this is simply

$$\psi_D \sim (1/2, 0) \oplus (0, 1/2)$$

The same as combining spins, we can combine Weyl spinors to obtain higher spins states or lower spins. But what is, say, a photon? Is it $(1, 0)$, $(0, 1)$ or $(1/2, 1/2)$? While this can be worked out, I always find it more confusing than useful for my practice. It is much important to understand thee the explicit action of the Lorentz transformations on Weyl spinors. To be a bit explicit, consider a boost along the $x = x^1$ direction, so that $\omega_{01} = \eta^1$, with η^1 the rapidity in direction 1. Then

$$\chi_L \to \Lambda_L \chi_L = e^{-\frac{\eta^1}{2}\,\sigma^1} \chi_L = \left(\cosh\frac{\eta^1}{2} - \sigma^1 \sinh\frac{\eta^1}{2}\right)\chi_L$$

while

$$\chi_R \to \Lambda_R \chi_R = e^{+\frac{\eta^1}{2}\,\sigma^1} \chi_R = \left(\cosh\frac{\eta^1}{2} + \sigma^1 \sinh\frac{\eta^1}{2}\right)\chi_R$$

Also

$$\chi_L^\dagger \to \chi_L^\dagger e^{-\frac{\eta^1}{2}\sigma^1}$$

and

$$\chi_R^\dagger \to \chi_R^\dagger e^{+\frac{\eta^1}{2}\sigma^1}$$

We see that $\Lambda_L^{-1} = \Lambda_R^\dagger$, so boost transformations are not unitary. [7] However, we have

$$\sigma^2 \chi_L^* \to \sigma^2 \Lambda_L^* \chi_L^* = \Lambda_R(\sigma^2 \chi_L^*)$$

as can be checked using $\sigma^2 \sigma^i \sigma^2 = -\sigma^{i*}$, and so

$$\sigma^2 \Lambda_L\, \sigma^2 = \Lambda_R \quad \text{and} \quad \sigma^2 \Lambda_L\, \sigma^2 = \Lambda_L$$

Thus

$$\sigma^2 \chi_L^* \sim (0, 1/2)$$

[7] This is unlike the rotations,This is no problem. There is a theorem which states that there are no finite dimensional unitary representations of the Lorentz group. One has to consider instead the Poincaré group, of which representations are infinite dimensional, labelled by momentum and classified by two invariant operators (Casimir), whose eigenvalues are the mass and the spin of the particle, see course on representation theory. If I had time, I would speak about that, but the supersymmetry course should cover this, in relation with the Coleman-Mandula theorem.

and

$$\sigma^2\chi_R^* \sim (1/2, 0)$$

From the above considerations, we can easily combine Weyl spinors to get other representations. We had a Dirac spinors as a direct sum of a χ_L and χ_R,

$$\psi_D = \begin{pmatrix} \chi_L \\ \chi_R \end{pmatrix} \tag{3.110}$$

but notice that we can also combine

$$\chi_L \oplus \sigma^2\chi_L^* \sim (1/2, 0) \oplus (0, 1/2) \sim (1/2, 1/2)$$

Explicitly,

$$\psi_M^L = \begin{pmatrix} \chi_L \\ -i\sigma^2\chi_L^* \end{pmatrix} \tag{3.111}$$

Another possibility, using a χ_R instead, is

$$\psi_M^R = \begin{pmatrix} i\sigma^2\chi_R \\ \chi_R^* \end{pmatrix} \tag{3.112}$$

These are called Majorana spinors. To see why there are funny factor of $\pm i$, you should remember that we introduced charge conjugation in the previous chapter, under which

$$\psi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \quad \xrightarrow{C} \quad \psi^c = -i\gamma^0\gamma^2\gamma^0\psi^* = \begin{pmatrix} i\sigma^2\phi_2^* \\ -i\sigma^2\phi_1^* \end{pmatrix} \tag{3.113}$$

Thus

$$(\psi_M^{L,R})^c = \psi_M^{L,R}$$

are self-conjugate states, meaning they are their own antiparticles.

From the above, we can also take tensorial product (ie multiply) Weyl spinors to get objets that transform as scalar (spin 0) or vector (spin 1) object. If you want, this is yet another perspective on the bilinear operators we considered in section 2.2.4. We start with scalars.

Suppose you have two Weyl spinors which transform as $\phi_1 \sim (1/2, 0)$ and $\phi_2 \sim (0, 1/2)$. The two most obvious possibilities are

$$\phi_1^\dagger\phi_2$$

and

$$\phi_2^\dagger\psi_1$$

They are hermitian of each others, so we we can consider the two linear combinations which are hermitian,

$$\phi_1^\dagger\phi_2 + \phi_2^\dagger\psi_1 \sim \bar{\psi}\,\psi$$

or

$$i(\phi_1^\dagger\phi_2 - \phi_2^\dagger\psi_1) \sim i\bar{\psi}\gamma_5\psi$$

Notice that we have specified the ϕ_1 and ϕ_2. Now take, eg, $\phi_2 \sim -i\sigma^2\phi_1^*$ so that ψ is Majorana instead of Dirac. Then, in components,

$$i\phi_2^\dagger\psi_1 = i(\phi_{1,1}, \phi_{1,2})\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}\begin{pmatrix} \phi_{1,1} \\ \phi_{1,2} \end{pmatrix} = \phi_{1,1}\phi_{1,2} \; - \phi_{1,2}\phi_{1,1}$$

The latter would be zero if the components were mere number. This is one of the situations for which it is imperative to take into account that spinors are fermions and so anticommute. With a scalar out of a Dirac of Majorana spinor one can write down a mass term in a Lagrangian, so neutral massive fermions could be of a Dirac or Majorana type. It is not yet known if the neutrinos are the latter (likely) or the former (less likely, but who knows).

How about spin 1? This is described by 4-vector A^μ. It is easy to check that

$$\chi_L^\dagger\bar{\sigma}^\mu\chi_L \tag{3.114}$$

(notice the bar on the $\bar{\sigma}^\mu$)) or

$$\chi_R^\dagger\sigma^\mu\chi_R \tag{3.115}$$

(no bar) transform as a 4-vectors under boosts. Indeed,

$$\chi_L^\dagger\chi_L \to \chi_L^\dagger e^{-\eta^1\sigma^1}\chi_L = \cosh\eta^1\chi_L^\dagger\chi_L + \sinh\eta^1(-\chi_L^\dagger\sigma^1\chi_L)$$

and

$$-\chi_L^\dagger\sigma^1\chi_L \to -\chi_L^\dagger e^{-\eta^1\sigma^1/2}\sigma^1 e^{-\eta^1\sigma^1/2}\chi_L = \cosh\eta^1(-\chi_L^\dagger\sigma^1\chi_L) + \sinh\eta^1\chi_L^\dagger\chi_L$$

while

$$-\chi_L^\dagger\sigma^{2,3}\chi_L \to -\chi_L^\dagger e^{-\eta^1\sigma^1/2}\sigma^{2,3}e^{-\eta^1\sigma^1/2}\chi_L = -\chi_L^\dagger\sigma^{2,3}\chi_L$$

since σ^1 anticommute with $\sigma^{2,3}$. These are indeed the transformation laws of a 4-vector under a boost. Idem for $\chi_L \to \chi_R$. These are purely axial current bilinear operators. If we have both χ_L and χ_R we can make linear combinations, as in the previous chapter, building vector, axial, axial-vector, etc. currents. And we can contract with a 4-vector to make a scalar, like the coupling of a current to a gauge field.

Chapter 4

Forces : gauge symmetries

4.1 Abelian symmetries

Consider a free massive Dirac field, $\psi(x)$ and its lagrangian,

$$\mathcal{L} = i\bar{\psi}\partial\!\!\!/\psi - m\bar{\psi}\psi$$

This theory is invariant under the redefinition of the field ψ by phase multiplication

$$\psi \to \psi' = e^{i\alpha}\psi$$

This is an example of a global symmetry transformation.

What happens if we take the parameter α to be space-time dependent,

$$\psi(x) \to e^{i\alpha(x)}\psi(x) \tag{4.1}$$

and ask the Lagrangian to be still invariant? The transformation (4.1) is an example of an *abelian* or $U(1)$ *gauge symmetry*.[1] Abelian means that if we perform two transformations, it doesn't matter in which order we do so,

$$e^{i\alpha_1(x)}e^{i\alpha_2(x)} = e^{i\alpha_2(x)}e^{i\alpha_1(x)}.$$

Obviously, the mass term is invariant under (4.1),

$$m\bar{\psi}\psi \to m\bar{\psi}'\psi'$$

The derivative, or kinetic term poses a problem however, because

$$\partial_\mu\psi \to e^{i\alpha(x)}(\partial_\mu\psi + i(\partial_\mu\alpha)\psi) \neq e^{i\alpha(x)}\partial_\mu\psi$$

To restore the invariance of the Lagrangian under the gauge symmetry, we introduce a *covariant derivative* D_μ, defined by the requirement that $D_\mu\psi$ transforms like ψ,

$$D_\mu\psi(x) \to e^{i\alpha(x)}D_\mu\psi(x) \tag{4.2}$$

[1] One possible motivation to do so is that of locality and Lorentz invariance: it seems sound to ask that, as we transform a field at one point, it doesn't have to know that it is transformed the same at the other side of the universe.

To construct D_μ, we introduce a vector field A_μ,

$$D_\mu = \partial_\mu - ieA_\mu \tag{4.3}$$

We have also introduced an arbitrary parameter e, whose meaning will become clear in a minute. The field $A_\mu(x)$ must transform in a specific way under (4.1) to guarantee that (4.2) holds:

$$(\partial_\mu - ieA'_\mu(x))\psi'(x) = e^{i\alpha(x)}(\partial_\mu - ieA_\mu(x))\psi(x) \tag{4.4}$$

Developing the LHS of (4.4) gives

$$e^{i\alpha(x)}(\partial_\mu + ieA'_\mu(x) + i\partial_\mu\alpha(x))\psi(x)$$

Comparing with the RHS of (4.4) gives

$$A'_\mu(x) = A_\mu(x) + \frac{1}{e}\partial_\mu\alpha(x) \tag{4.5}$$

The vector field A_μ is an example of a *gauge field*: its *raison d'être* is to restore the invariance of our theory under local symmetry transformations.

Our modified, gauge invariant, Lagrangian becomes

$$\mathcal{L} = i\bar{\psi}\not{D}\psi - m\bar{\psi}\psi$$

Developing, we get

$$\mathcal{L} = i\bar{\psi}\not{\partial}\psi + eA_\mu\bar{\psi}\gamma^\mu\psi - m\bar{\psi}\psi$$

If ψ is identified with the electron field, we recognize in the second term the coupling of the electromagnetic current to the photon. Hence, e is to be identified with the electromagnetic coupling constant, and $-e$ with the charge of the electron. In natural units, $e^2/4\pi = \alpha \approx 1/137$, the fine structure constant

Our work is not done yet, as we have introduced a vector field which can also be dynamical (well, at least the photon does propagate, doesn't it?). What we need is a kinetic term which is gauge invariant (this is what we ask for) and two derivatives of the gauge field (to give a *bona fide* kinetic term). Although you probably already know the answer or can easily guess it, let me take a longer road. This will be useful when we will get to discuss more complicated gauge transformations.

Consider the commutator of covariant derivatives

$$[D_\mu, D_\nu]\psi(x) = -ie(\partial_\mu A_\nu - \partial_\nu A_\mu)\psi(x)$$

As $D_\mu\psi(x)$ and $\psi(x)$ transform the same under (4.1), so does the commutator:

$$[D_\mu, D_\nu]\psi(x) \to e^{i\alpha(x)}[D_\mu, D_\nu]\psi(x)$$

Thus the combination

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{4.6}$$

is gauge invariant. Of course, this is nothing but the familiar electromagnetic field strength, whose components are the electric and magnetic fields. Again, what is important is how its expression has emerged from the requirement of gauge invariance.

In its full glory, our gauge invariant Lagrangian is finally

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\not{D}\psi - m\bar{\psi}\psi \tag{4.7}$$

Again, if ψ is identified with the electron and A_μ the photon field, this is just the Lagrangian of electromagnetism. The equation of motion are

$$\partial_\mu F^{\mu\nu} = -e\bar{\psi}\gamma^\nu\psi$$

and

$$\not{\partial}\psi = ieA_\mu\gamma^\mu\psi - m\psi$$

Furthermore,

$$\partial_\mu\partial_\nu F^{\mu\nu} = 0$$

(because $F_{\mu\nu}$ is antisymmetric) implies

$$\partial_\mu(e\bar{\psi}\gamma^\mu\psi) = 0$$

which is the conservation of the electromagnetic current.

Note that a mass term for the gauge field, $m^2A_\mu A^\mu$ is forbidden by gauge invariance. Hence, the photon must be massless.

exercise 1 *We could have add to the Lagrangian (4.7) the dimension 4, gauge invariant operator,*

$$\delta\mathcal{L} = c\epsilon_{\alpha\beta\mu\nu}F^{\alpha\beta}F^{\mu\nu}$$

where $\epsilon_{\alpha\beta\mu\nu}$ is the totally antisymmetric Levi-Civita tensor. Show that such a term breaks P and CP. Show also that it can be rewritten as a total derivative.

To conclude this rather lengthly derivation of the Lagrangian of electromagnetism, let us consider how things are modified if there are other matter fields. To be specific, consider a complex scalar field, $\phi(x)$, with free Lagrangian

$$\mathcal{L} = \partial_\mu\phi^\dagger\partial^\mu\phi - m^2\phi^\dagger\phi$$

Again, this theory is invariant under a global phase redefinition of the field ϕ. The mass term is also invariant under local transformations, as would be any monomial of $(\phi^\dagger\phi)$, as in

$$V[\phi(x)] = m^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$$

for instance. You can verify that substituting the partial derivative for the covariant derivative

$$D_\mu = \partial_\mu - ieQA_\mu$$

makes the full theory invariant under (4.4) and

$$\phi(x) \to e^{iQ\alpha(x)} \tag{4.8}$$

Note that we have introduced an arbitrary factor Q. As you might guess, or check by writing down the equations of motion, eQ is to be interpreted as the electric charge of the field ϕ. Thus, such a theory gives us no clue as why all charges are rational multiple of the charge of the electron e in Nature.

4.2 Non-abelian symmetries

We want to generalize the preceding discussion to more complicated transformation laws. As mentioned in the introduction, a motivation for doing so is that years of experiments have indicated that the (left-handed component of the) electron and its neutrino are two facets of the same object. A convenient choice of basis is:[2]

$$L = \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix} \tag{4.9}$$

This structure is familiar. It looks very much like a *spinor*, an object extensively studied in elementary courses on quantum mechanics. Incidentally, an SU(2) group of transformations acts naturally on L. The main difference is that spinors transform under *space-time* symmetries, while the SU(2) symmetry considered here is an example of an *internal symmetry*, sometime called the *weak isospin*. Otherwise, the mathematics is just the same. For instance, using isospin doublets as building blocks, we could construct higher isospin (1, 3/2, etc) objects, by applying the same rules as in the composition of spinors. For this reason, an object like L is said to transform in the *fundamental representation* of SU(2).

Consider now the Lagrangian

$$\mathcal{L} = i\bar{L}\partial\!\!\!/ L \tag{4.10}$$

It is invariant under global SU(2) transformations

$$U = e^{i\alpha^a\sigma^a/2}, \quad L \to L' = UL$$

where $a = 1, 2, 3$ and the σ^a are the hermitian (*i.e.* $\sigma^{a\dagger} = \sigma^a$) Pauli matrices,

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

These are the simplest example of *non-abelian unimodular unitary* transformations. Non-abelian because two different SU(2) transformations do not commute in general,

$$U_1U_2 \neq U_2U_1,$$

unimodular because

$$\det(U) = 1$$

and unitary because the complex-valued matrices U satisfy

$$U^\dagger U = UU^\dagger = 1.$$

The latter condition implies that $U^{-1} = U^\dagger$, hence the invariance of (4.10).

[2]Note that we could also have chosen another basis, for instance.

$$L = \begin{pmatrix} e_L \\ \nu_L \end{pmatrix}$$

As long as we don't break the SU(2) symmetry the distinction between e_L and ν_L is meaningless.

exercise 2 *Consider an infinitesimal transformation,* $U \approx 1 + i\alpha^a\sigma^a/2$. *Show that* $\det(U) = 1$ *implies* $Tr\sigma^a = 0$*, while unitarity implies the hermiticiy of the Pauli matrices.*

Now, the least we can say is that (4.10) is not a very interesting theory: it merely describes the propagation of free, massless electrons and neutrinos. In order to introduce interactions between these particles, we follow Yang and Mills who had the brilliant idea to proceed in analogy with quantum electrodynamics by imposing invariance under local symmetry transformations.

Local SU(2) transformations are parameterized by three space-time dependent functions $\alpha^a(x)$.

$$L(x) \to U(x)L(x) \tag{4.11}$$

This is our first example of a *non-abelian gauge symmetry*. From the previous section we know that we must introduce a covariant derivative D_μ, such that $D_\mu L$ transforms the same as L,

$$D_\mu L(x) \to U(x)D_\mu L(x)$$

In analogy with quantum electrodynamics, we introduce three gauge fields $W^a_\mu(x)$ (one for each α^a) and define

$$D_\mu = \partial_\mu + igW^a_\mu \frac{\sigma^a}{2} \tag{4.12}$$

We have also introduced the coupling g, analogous to e. We must now determine how the W^a_μ transform in order to restore gauge invariance. To simplify notations, we define $W_\mu \equiv W^a_\mu\sigma^a/2$.

Then, by definition,

$$(\partial_\mu + igW'_\mu)L' = U(x)(\partial_\mu + igW_\mu)L$$

Developing the LHS gives

$$(U\partial_\mu + igW'_\mu U + \partial_\mu U)L = U(\partial_\mu + igU^\dagger W'_\mu U + U^\dagger \partial_\mu U)L$$

Identifying with the RHS, gives

$$U^\dagger W'_\mu U = W_\mu + \frac{i}{g}U^\dagger \partial_\mu U$$

or

$$W'_\mu = UW_\mu U^\dagger - \frac{i}{g}U\partial_\mu U^\dagger$$

where, in the last term, we have used $(\partial_\mu U)U^\dagger = -U(\partial_\mu U^\dagger)$. In components, taking α^a as infinitesimal parameters, this gives to $O(\alpha^2)$

$$W'^a_\mu \frac{\sigma^a}{2} = W^a_\mu \frac{\sigma^a}{2} - \frac{1}{g}\partial_\mu\alpha^a\frac{\sigma^a}{2} + i\alpha^b W^c_\mu\left[\frac{\sigma^b}{2}, \frac{\sigma^c}{2}\right] \tag{4.13}$$

exercise 3 *Check this!*

We can go one step further by using the following properties of the Pauli matrices:

$$\left[\frac{\sigma^a}{2}, \frac{\sigma^b}{2}\right] = i\epsilon_{abc}\frac{\sigma^c}{2} \tag{4.14}$$

and

$$\mathrm{Tr}(\sigma^a\sigma^b) = 2\delta_{ab}$$

where δ_{ab} is Kronecker's delta. The relation (4.14) defines a structure known as the *Lie algebra* of the group SU(2). The ϵ_{abc} is the completely antisymmetric tensor with three indices. In technical jargon, its components are called the *structure constants* of the group SU(2). Using (4.14), we finally get (compare with (4.4))

$$W'^a_\mu = W^a_\mu - \frac{1}{g}\partial_\mu\alpha^a - \epsilon_{abc}\alpha^b W^c_\mu \tag{4.15}$$

The last term, which arises because SU(2) gauge transformations do not commute in general is new compared to the abelian case. In particular, it implies that the W^a_μ transform non-trivially even under constant α^a gauge transformations. (See next section.)

Another important consequence of this result is that gauge invariance under SU(2) *transformations implies that all fields, be it the electron or muon doublet and the field W_μ itself, couple to W_μ with the same coupling g.*

Substituting the normal derivative by the covariant derivative in (4.10) gives

$$\mathcal{L} = i\bar{L}\not{D}L = i\bar{L}\not{\partial}L - gW^a_\mu\bar{L}\frac{\sigma^a}{2}\gamma^\mu L$$

The second term describes the coupling of the current

$$j^a_\mu = g\bar{L}\gamma_\mu\frac{\sigma^a}{2}L$$

to the gauge field W^a_μ. In components, the interaction Lagrangian is

$$\mathcal{L}_{int} = -\frac{g}{2}W^3_\mu(\bar{\nu}_L\gamma^\mu\nu_L - \bar{e}_L\gamma^\mu e_L) - \frac{g}{\sqrt{2}}W^-_\mu\bar{e}_L\gamma^\mu\nu_L - \frac{g}{\sqrt{2}}W^+_\mu\bar{\nu}_L\gamma^\mu e_L \tag{4.16}$$

where $W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \mp iW^2_\mu)$.

exercise 4 *Consider a matter field with covariant derivative*

$$\tilde{D}_\mu = \partial_\mu + i\tilde{g}W_\mu$$

Show that gauge invariance imposes $\tilde{g} = g$.

Our last task is to find a gauge invariant kinetic term for the non-abeliane gauge bosons. As in the abelian case, let us consider the commutator of two covariant derivatives. As $D_\mu L$ transforms like L, so does the commutator,

$$[D_\mu, D_\nu]L(x) \rightarrow U(x)[D_\mu, D_\nu]L(x)$$

Hence,

$$[D_\mu, D_\nu] \rightarrow U(x)[D_\mu, D_\nu]U^\dagger(x) \tag{4.17}$$

The explicit expression of the commutator gives

$$\begin{aligned}
[D_\mu, D_\nu] &= ig(D_\mu W_\nu - D_\nu W_\mu) \\
&= ig(\partial_\mu W_\nu - \partial_\nu W_\nu + ig[W_\mu, W_\nu]) &(4.18)\\
&= ig(\partial_\mu W_\nu^a - \partial_\nu W_\mu^a - g\epsilon_{abc}W_\mu^b W_\nu^c)\frac{\sigma^a}{2} &(4.19)\\
&\hat{=} igW_{\mu\nu} \equiv igW_{\mu\nu}^a\frac{\sigma^a}{2}
\end{aligned}$$

where on the rhs of the first line, the derivatives only act on the $W_{\mu,\nu}$ (i.e. not on the right). The matrix-valued combination $W_{\mu\nu}$ (or, equivalently, in components $W_{\mu\nu}^a$) is the non-abelian analog of the electromagnetic field strength $F_{\mu\nu}$. The new feature is the second term in (4.18) or (4.19), which involves the structure constants of SU(2). Unlike $F_{\mu\nu}$, the non-abelian field strength $W_{\mu\nu}$ is not gauge invariant. Rather, from (4.17), we deduce that it transforms *covariantly* under gauge transformations,

$$W_{\mu\nu} \rightarrow W'_{\mu\nu} = U(x)W_{\mu\nu}U^\dagger(x) \tag{4.20}$$

This suggests the following generalization of the a gauge invariant kinetic term for non-abelian gauge fields:

$$\mathcal{L}_W = -\frac{1}{2}\mathrm{Tr}(W_{\mu\nu}W^{\mu\nu}) = -\frac{1}{4}W_{\mu\nu}^a W^{a\mu\nu}$$

More explicitely,

$$\begin{aligned}
\mathcal{L}_W &= -\frac{1}{4}(\partial_\mu W_\nu^a - \partial_\nu W_\mu^a)(\partial^\mu W^{a\nu} - \partial^\nu W^{a\mu}) \\
&\quad + g\epsilon_{abc}\partial_\mu W_\nu^a W^{b\mu}W^{c\nu} &(4.21)\\
&\quad - \frac{g^2}{4}(W_\mu^a W_\nu^b W^{a\mu}W^{b\nu} - W_\mu^a W_\nu^b W^{a\nu}W^{b\mu}) &(4.22)
\end{aligned}$$

where we have used the relation $\epsilon_{abc}\epsilon_{ab'c'} = \delta_{bb'}\delta_{cc'} - \delta_{bc'}\delta_{cb'}$ for the last term.

This is quite remarkable. We have discovered that, on top of a now familiar kinetic term, the requirement of gauge invariance introduces, in a unique way[3] trilinear (4.21) and quadratic (4.22) couplings between the non-abelian gauge fields W_μ^a.

[3]Well, almost. For one thing, we have limited ourself to an operator of dimension 4. This is related to our desire of having a renormalizable gauge theory. Furthermore, we could have added a term of the form $\epsilon_{\mu\nu\alpha\beta}W_{a\mu\nu}W^{a\mu\nu}$ which is gauge invariant. This term however violates P and CP and furthermore, is a total derivative. As such, it doesn't contribute to the equation of motion. It however plays an important role in topological aspects of non-abelian gauge theories.

exercise 5 *From (4.20), derive how the components $W^a_{\mu\nu}$ of the non-abelian field strength transforms under infinitesimal gauge transformations, $U(x) \approx 1 + i\alpha^a(x)\sigma^a/2$.*

We will not need here the classical equations of motions which can be derived from this Lagrangian. These equations are non-linear in the gauge fields and, thus, difficult to solve in general. Some very specific classical solutions are know however, which play an important role, for instance in relation with the topological structure of non-abelian gauge theories. Unfortunately, we won't have time to discuss here these beautiful topics. As you could expect, these non-linearities also gives rise to extremely important physical effects at the quantum level. We will perhaps say a bit more about some of these phenomena later on.

4.3 The Standard Model: gauge and fermion sectors

We have learned how to couple spin-1 bosons to fermions, using the principle of gauge invariance. Considering our L_e doublet of left-handed fermions, the gauge invariant Lagrangian under SU(2) gauge symmetry is:

$$\mathcal{L} = i\bar{L}_e \not{D} L_e - \frac{1}{4} W^a_{\mu\nu} W^{a\mu\nu} \tag{4.23}$$

Decomposing or from (4.16), we see that the gauge boson W^3_μ couples diagonally to the left-handed electron and its neutrino and this with opposite charges. This is an instance of a *neutral current* coupling. The exchange of $W^\pm_\mu$ gauge bosons on the other hand change the charge of the fermions, since they transform an electron into its neutrino and *vice versa*. Obviously, W^3_μ cannot be directly identified with the photon. First because it couples to neutrinos. Then because the photon must couple to Dirac fermions. If we want to include electromagnetism in this theory, we are forced to introduce e_R and a new chiral gauge sector, with gauge field $U(1)_Y$, and covariant derivative for e_R

$$D_\mu = \partial_\mu + ig' \frac{Y_{e_R}}{2} B_\mu$$

Here, g' is again an *a priori* arbitrary coupling constant, and Y is a number called the *hypercharge* of e_R. Furthermore, nothing prevents the doublet field L_e to couple to B_μ, so we take $Y_{L_e} \neq 0$ or

$$D_\mu = \partial_\mu + igW^a_\mu \frac{\sigma^a}{2} + ig' \frac{Y_{L_e}}{2} B_\mu$$

Hence, we are now considering a theory with gauge group $\mathrm{SU}(2) \times \mathrm{U}(1)_Y$. Adding these extra fields and couplings to (4.23) and concentrating on W^3_μ and B_μ we get

$$\mathcal{L}_{int} = -\frac{1}{2}(gW^3_\mu + g'Y_{L_e}B_\mu)\bar{\nu}_L\gamma^\mu\nu_L + \frac{1}{2}(gW^3_\mu - g'Y_{L_e}B_\mu)\bar{e}_L\gamma^\mu e_L - \frac{1}{2}g'Y_{e_R}B_\mu\bar{e}_R\gamma^\mu e_R \tag{4.24}$$

This pretty messy expression can be a bit clarified by working in a rotated basis for the gauge fields:

$$\begin{pmatrix} W^3_\mu \\ B_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} \tag{4.25}$$

We want to identify A_μ with the photon field, so let us work this out. Substituting (4.25) in (4.24), the couplings to A_μ read

$$
\begin{aligned}
\mathcal{L}_{int} \;=\; & -\frac{1}{2}(g\sin\theta_W + g' Y_{L_e}\cos\theta_W)\, A_\mu\, \bar{\nu}_L\gamma^\mu\nu_L \\
& +\frac{1}{2}(g\sin\theta_W - g' Y_{L_e}\cos\theta_W)\, A_\mu\, \bar{e}_L\gamma^\mu e_L \\
& -\frac{1}{2}g' Y_{e_R}\cos\theta_W A_\mu \bar{e}_R\gamma^\mu e_R
\end{aligned}
$$

Because the neutrino is electrically neutral, the first term must vanish identically. Hence,

$$g\sin\theta_W \equiv -g' Y_{L_e}\cos\theta_W$$

Furthermore, identifying the couplings to e_L and e_R with e the electric charge of the electron, implies

$$Y_{L_e} = \frac{Y_{e_R}}{2}$$

and

$$e = -gY_{L_e}\sin\theta_W$$

If we choose g' so that $Y_{L_e} = -1$ (and thus $Y_{e_R} = -2$), we finally get

$$e = g\sin\theta_W = g'\cos\theta_W \tag{4.26}$$

while the mixing angle can be expressed in function of g and g',

$$\sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}}, \quad \cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}} \tag{4.27}$$

The mixing parameter θ_W is called the *Weinberg angle*.

Furthermore, we see that the electric charge Q can be identified with the a linear combination of the (third component of the) isospin and the hypercharges,

$$Q = T_f^3 + \frac{Y_f}{2} \tag{4.28}$$

Indeed, the right-handed electron e_R transforms as *singlet* under SU(2), which means that it carries no isospin charge and $T_3|e_R\rangle = 0$. On the other hand, ν_L and e_L are eigenstates of T_3 with eigenvalues $1/2$ and $-1/2$ respectively. You can easily verify (4.28) using the hypercharge assignments $Y_{L_e} = -1$ and $Y_{e_R} = -2$.

exercise 6 *What would have happened if we had chosen*

$$L = \begin{pmatrix} e_L \\ \nu_L \end{pmatrix}$$

instead? Is (4.28) still valid?

The neutral currents interaction Lagrangian then becomes,

$$
\begin{aligned}
\mathcal{L}_{int} &= eA_\mu\, \bar{e}_L\gamma^\mu e_L \\
&- \frac{e}{\cos\theta_W \sin\theta_W}\,\frac{1}{2}\, Z_\mu \bar{\nu}_L\gamma^\mu \nu_L \\
&- \frac{e}{\sin\theta_W \cos\theta_W}(-\frac{1}{2} + \sin^2\theta_W)\, Z_\mu \bar{e}_L\gamma^\mu e_L \\
&- \frac{e}{\sin\theta_W \cos\theta_W}(-\sin^2\theta_W)\, Z_\mu\, \bar{e}_R\gamma^\mu e_R
\end{aligned}
$$

We have used the following useful identities:

$$
\sqrt{g^2 + g'^2} = \left[\frac{e^2}{\cos^2\theta_W} + \frac{e^2}{\sin^2\theta_W}\right] = \frac{e}{\cos\theta_W \sin\theta_W},
$$

for the second term,

$$
\frac{g^2 - g'^2}{\sqrt{g^2 + g'^2}} = \frac{e}{\sin\theta_W \cos\theta_W}(1 - 2\sin^2\theta_W)
$$

for the third term, and finally

$$
\frac{g'^2}{\sqrt{g^2 + g'^2}} = \frac{e}{\sin\theta_W \cos\theta_W}\sin^2\theta_W
$$

for the last one. We have written the coupling to Z_μ to help one notice that they can all be expressed as

$$
\frac{e}{\sin\theta_W \cos\theta_W}(T^3_f - Q_f \sin^2\theta_W) \tag{4.29}
$$

We can add the other leptons to our Lagrangian, since the charge assignment must be the same as for the electron and its neutrino. We have no explicitely verified this here, but adding the quarks is no more complicated, using the hypercharge and weak isospin assignments of Table I and the relations (4.28) and (4.29).

exercise 7 *Verify that the couplings of A_μ and Z_μ to quarks satisfy (4.28) and (4.29).*

Altogether, and using a pretty condensed, manifestly $\mathrm{SU}(2) \times U(1)_Y$ invariant notation, the SM part of the Lagrangian describing the coupling of matter to gauge fields is:

$$
\begin{aligned}
\mathcal{L} &= -\frac{1}{4}W^a_{\mu\nu}W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \\
&+ \sum_{i=e,\mu,\sigma} (\bar{L}_i \not{D} L_i + i\bar{E}_i \not{D} E_i) \\
&+ \sum_{i=u,c,t} (i\bar{Q}_i \not{D} Q_i + i\bar{U}_i \not{D} U_i) + \sum_{i=d,s,b} i\bar{D}_i \not{D} D_i
\end{aligned}
$$

This Lagrangian cannot be the end of the story however. For one thing, to define the electromagnetic field A_μ we had to break apart the manifest $\mathrm{SU}(2) \times U(1)_Y$ invariance of the underlying theory. Furthermore, the Z_μ and $W^\pm_\mu$ fields are strictly massless at this level of our discussion. Adding a mass term to the Lagrangian would break *explicitely* the gauge invariance of the theory. As we have stated, without proof, here and there, gauge invariance is mandatory to have a well-defined, unitary and renormalizable *quantum theory*. Finally, the quarks and charged leptons are massless. Again, adding a mass term would break the gauge invariance of the theory. This is because the SM is a chiral theory, thus preventing fermions masses to appear in the Lagrangian. All this issues can be solved at once, through a mechanism of *spontaneous breaking of the gauge symmetry.* Because our theory naturally embeds electromagnetism, we will look for a pattern of symmetry breaking $\mathrm{SU}(2) \times U(1)_Y \to U(1)_Q$. This topic will be the subject of the next chapter.

4.4 Basics of Lie algebras

All the things we have said about SU(2) generalize to complex objects with N components. The group of symmetries is then SU(N). The SU(N) family of groups are example of continuous groups or *Lie groups*. These are group which contain elements arbitrarily close to the identity. Then, any infinitesimal group element U can be written as

$$U(\alpha) \approx 1 + i\alpha^a T^a$$

The coefficients T^a of the infinitesimal parameters are called the *generators* of the gauge group. Unitarity of U implies that the generators are hermitian matrices. If the group elements of SU(N) act on complex vector with N components, the generators, which we denote by t^a, are $N \times N$ hermitian matrices. This is the *fundamental representation* of the group SU(N). (Other representations T^a exist, as you know from SU(2) and as we will see in a moment.) They span a vector space of dimension $a = 1, \ldots, N^2 - 1$.

To derive the commutation relation of the T^a, consider for instance the product

$$U(\alpha)U(\beta)U^{-1}(\alpha)U^{-1}(\beta) = U(\gamma)$$

Developing this product for infinitesimal representations to second order in α and β gives

$$-\alpha^a \beta^b [T^a, T^b] = i\gamma^c T^c$$

As the T^a give a basis of a given representation R, one can write

$$[T^a, T^b] = i f^{ab}_c T^c \tag{4.30}$$

where the coefficients are manifestly antisymmetric in the first two indices. Imposing the normalization condition

$$\mathrm{Tr} T^a T^b = \frac{1}{K_R} \delta^{ab}$$

for the representation R, we can extract the f_{abc},

$$f_{abc} = -K_R i \mathrm{Tr} T_a [T_b, T_c]$$

which shows that they are completely antisymmetric in all three indices. These coefficients are called the *structure constants* of the group and the vector space spanned by the generators, with the additional commutation relation, is called the *Lie algebra* of the group.

The commutation relation (4.30) together with the identity

$$[T^a, [T^b, T^c]] + [T^b, [T^b, T^a]] + [T^c, [T^a, T^b]] = 0$$

implies that the structure constants obey the *Jacobi identity*

$$f^{ade} f^{bcd} + f^{bde} f^{bad} + f^{cde} f^{abd} = 0$$

What interests us here is that, defining

$$(T^a)_{ij} = i f_{iaj}$$

the Jacobi identity can be rewritten as the commutator (4.30). Hence, the structure constants defines a new representation called the *adjoint representation* of the group.

To gain some hindsight, consider ψ and χ transforming in the fundamental representation of SU(N). Then

$$\psi^\dagger \chi$$

is invariant. Moreover, the $N^2 - 1$ objects

$$V^a = \psi^\dagger t^a \chi$$

transform, for α^a infinitesimal, as

$$\begin{aligned} V'^a = (\psi^\dagger T^a \chi)' &= \psi^\dagger T^a \chi + i\alpha^b \psi^\dagger [T^a, T^b] \chi \\ &= V^a - f_{abc} \alpha^b V^c \end{aligned} \tag{4.31}$$

Specializing to SU(2), in which case $f_{abc} = \epsilon_{abc}$, we recognize in the transformation law of the gauge field W_μ under constant gauge transformations. Gauge fields thus transform in the adjoint representation of the gauge group.

To conclude this short section on Lie algebras, consider the generators of SU(N) in the fundamental representation. These are $N \times N$ hermitian, traceless matrices. We can work in a basis in which a subset of these $N^2 - 1$ matrices can be diagonalized. Obviously, there a $N - 1$ linearly independent of these matrices, which we denote by H^i. They form a commutative sub-algebra of the Lie algebra called the *Cartan algebra*, *i.e.*

$$[H^i, H^j] = 0$$

The dimension of the Cartan algebra (which is $N - 1$ for SU(N) groups) is called the *rank* of the group. For instance SU(2) is of rank 1. The Cartan algebra is one-dimensional, with generator $\sigma^3/2$. To this generator, corresponds the abelian subgroup of $U(1) \subset$ SU(2), parameterized by

$$U(\alpha^3) = e^{i\alpha^3 \frac{\sigma^3}{2}}$$

The left-handed electron couples which charge $-1/2$ to the corresponding gauge field W_μ^3 and the neutrino with the opposite charge.

As another instance, consider SU(3), which is of rank 2. It has eight generators. $\mathrm{SU}(3)_c$ is the gauge group of quantum chromodynamics, the theory of strong interactions. Quarks (of both chiralities) transform in the fundamental representation of SU(3). They thus carry extra three indices, called color for historical reasons. To the eight generators correspond eight gauge bosons, the gluons.

The gauge group of the SM of strong and electroweak interactions is $\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. This group is of rank 4, and so his SU(5). Incidentally SU(5) is the smallest group which contains the gauge groups of the SM as a subgroup. It provides the simplest example of a unified theory of the strong, weak and electromagnetic interactions.

Chapter 5

Broken symmetries

5.1 Introduction

There are many instance of manifest symmetries of *physical laws*. For instance, Newton's equations are invariant under spatial and temporal translations and rotations. To these symmetries correspond conservation laws, like the conservation of energy for temporal translations, or angular momentum for rotations. This is the content of Noether's theorem. Other examples includes the Lorentz invariance of Maxwell's equations. Also, we have met example of more subtle symmetries, like the abelian gauge symmetry of electromagnetism. We said *physical laws* to emphasize that these are symmetries of the equation of motions and not of specific solutions. For instance, planetary orbits around the sun are generically ellipses and are not rotationally invariant. Likewise, a stair is not a rotationally invariant object (contrary to, say, a sphere), although the equations describing its motion are symmetric. This is an important distinction. When we talk of symmetries, we mean of the underlying laws describing the system, not of a specific solution of these.

Likewise, in quantum mechanics, symmetry principles are utterly important. The existence of symmetries often help to solve otherwise difficult problems. For instance, the hydrogen atom problem is easy because it has many symmetries. The orbitals can be labeled according to quantum numbers why corresponds to the eigenvalues of, for instance, the angular momentum operator. At the level of the spectrum, the existence of symmetries implies the degeneracy of some energy levels. This is an instance of what is called an *explicit symmetry*. In the technical jargon, we say that the symmetry of the Hamiltonian is realized in the Wigner phase.

In Nature, however, we have also many instances of other ways in which symmetries can be realized. A pedagogical example is that of a vertical rigid rod, hold fixed at its bottom. This object is obviously invariant for rotation around its vertical axe. Now, if we apply a vertical force on the stick, the system is still invariant under such rotations. However, above some threshold value force, the value of which depends of the resistance of the stick to deformations, chances are that the stick will bend in one direction or another. The end configuration will have lost its rotational invariance. This is an instance of a *broken symmetry*. Of course, you could argue that this conclusion must depends of how the stick has been build, whether it is truly homogeneous, or maybe the forces applied was not perfectly symmetric. So let us consider another, more instructive example.

You have all plays with ferro-magnets. Ferro-magnets can be think of as materials in which there are strong attractive interactions between the spins of its atoms, which tend to align them. To make things more concrete and neglecting irrelevant details, the interaction Hamiltonian of a spin system can be modeled as

$$\mathcal{H} = -J \sum_{\{i,j\}} \vec{S}_i \cdot \vec{S}_j$$

where i, j labeled the different spins of the system and the sum is over pairs of $\{i, j\}$. Obviously, if the coupling $J > 0$, the interaction energy is lowered if the spins are aligned. Hence, the Hamiltonian is minimal if they are all aligned (ordered phase). Note, however, that this Hamiltonian is at the same time *invariant* under *global* spatial rotations of the spins, $R \cdot \vec{S} = \vec{S}'$, with $R \in O(3)$, or $R^T \cdot R = 1$. Obviously, a configuration with all the spin misaligned (disordered phase) would respect the symmetries of the Hamiltonian. On the contrary, the configuration with all the spins aligned would break the symmetry under spatial rotations. Which phase is realized depends on temperature. Consider then the corresponding partition function

$$Z = \sum_{\text{all configurations}} e^{-\mathcal{H}/T}$$

where T is the temperature.

At high temperatures, there are very many configurations that can equally contribute to the partition function. Then, although it is energically favorable for the spins to aligned, the entropy is maximized if they are disordered. The symmetry of the underlying Hamiltonian is manifest. At low temperatures, however, the Boltzmann exponential is highly peaked around a configuration which minimizes the Hamiltonian and a configuration with all the spin aligned will be preferable. We call this *spontaneous magnetization.* and this is an instance of *spontaneous symmetry breaking* (SSB).[1] This behavior is an example of a *Phase Transition.* Phase transition are ubiquitous in Nature and understanding them is a very important topic in Statistical Mechanics. Note the connection with the notion of broken or explicit symmetries and the dichotomy order/disorder. This circle of ideas has also had deep influence in our understanding of Fundamental Physics. We also think that phase transition may have played an important role in the early Universe.

Some more remarks are of interest here. Note first that the symmetry is not totally broken as rotations around the spontaneous magnetization axe, say $\vec{1}_z$, are still allowed. The pattern of symmetry breaking in the present case is $O(3) \to O(2)_z$. Moreover, there is a remnant of the underlying $O(3)$ symmetry because, in the absence of an external disturbance, the direction in which the spin will point is *a priori* undetermined. In particular, it would cost only little energy to let all the spins globally rotate in directions orthogonal to the $\vec{1}_z$ axe. One physical consequence of this is that there are *massless excitations* (*i.e.* with dispersion relation $\omega \to 0$ as $k \to 0$) in the spectrum of a ferromagnet. This feature is generic. That there are massless excitations in system with spontaneously broken global continuous symmetries is a consequence of *Goldstone's theorem*. Incidentally, these excitations are called *Goldstone modes.*

[1] To is to be contrasted with *explicit* symmetry breaking. For instance, a non-zero external magnetic field couples to the spin through an extra term $\delta\mathcal{H} \propto \sum_i \vec{B} \cdot S_i$ which breaks explicitely the $O(3)$ symmetry. The result is that the spins try to align along $\vec{B}$, thus lifting the degeneracy of the vacuum.

5.2 SSB: global abelian symmetry

Let us be a bit more systematic. Consider a complex scalar field ϕ with Lagrangian

$$\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - \mu^2 \phi^\dagger \phi - \lambda(\phi^\dagger \phi)^2 \tag{5.1}$$

This theory has a global $U(1)$ symmetry, $\Phi \to e^{i\alpha}\Phi$. It is the simplest example of a theory with a continuous global symmetry. How is the symmetry realized in the vacuum or ground state? The potential for Φ is

$$V[\Phi] = \mu^2 \Phi^\dagger \Phi + \lambda(\Phi^\dagger \Phi)^2$$

To get some insight, it is convenient to parameterize the complex scalar field as

$$\Phi = \frac{1}{\sqrt{2}}(\varphi_1 + i\varphi_2)$$

In this basis, the potential reads

$$V[\varphi_1, \varphi_2] = \frac{1}{2}\mu^2(\varphi_1^2 + \varphi_2^2) + \frac{\lambda}{4}(\varphi_1^2 + \varphi_2^2)^2$$

The potential is invariant under SO(2) $\sim U(1)$ rotations

$$\begin{pmatrix} \varphi_1' \\ \varphi_2' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}$$

which preserve the scalar product (norm) $\varphi_1^2 + \varphi_2^2$.

If the parameter $\mu^2 > 0$, the potential is the two-dimensional parabola showed in Figure 5.2 (for $\lambda > 0$). A quick inspection of this figure reveals that the potential energy is minimized for $\langle\varphi_1\rangle = \langle\varphi_2\rangle = 0$. (We use the bra and kets to note the vacuum expectation value (*vev*) of the fields (*vev*). Alternatively, minimizing the potential with respect to φ_1 and φ_2 gives,

$$\frac{\delta V}{\delta \varphi_{1,2}} = \mu^2 \varphi_{1,2} + \varphi_{1,2}(\varphi_1^2 + \varphi_2^2) = 0$$

gives $\langle\varphi_1\rangle = \langle\varphi_2\rangle = 0$ as the unique solution. The second variation at this extremum gives

$$\frac{\delta^2 V}{\delta \varphi_1^2} = \frac{\delta^2 V}{\delta \varphi_1^2} = \mu^2$$

and

$$\frac{\delta^2 V}{\delta \varphi_1 \delta \varphi_2} = 0$$

Hence, the two real fields have the same mass $M_1^2 = M_2^2 = \mu^2$ and the SO(2) symmetry of the theory is manifest at the level of the spectrum.

Consider now the case $\mu^2 < 0$. The potential has now the shape of a Mexican hat (See Figure 5.2). Obviously, the origin $\varphi_1 = \varphi_2 = 0$ is no more the minimum of the potential. Rather, there is a whole family of vacua in the hollow of the potential. Obviously, any of these is a good vacuum,

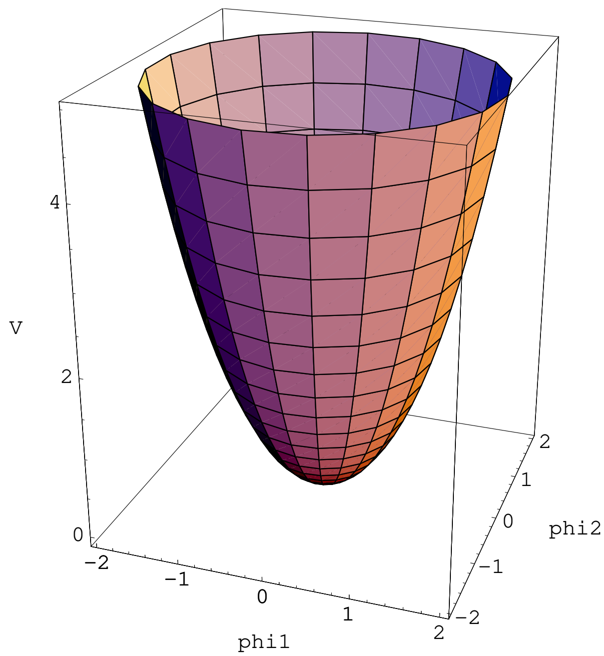


Figure 5.1: Potential for $\mu^2 > 0$.

but picking one breaks the invariance of the theory under rotations. Among these, let us choose for convenience the one along the $\varphi_2 = 0$ axis, and pose $\langle\varphi_1\rangle = v$. The potential for v is then

$$V[v] = \frac{1}{2}\mu^2 v^2 + \frac{\lambda}{4}v^4$$

and minimizing it gives

$$\frac{\delta V}{\delta v} = \mu^2 v + \lambda v^3 = 0$$

If $\mu^2 < 0$, there are two possible extrema,

$$v_0 = 0$$

and

$$v_1 = \pm\sqrt{\frac{-\mu^2}{\lambda}}$$

The first solution v_0 is a local maximum, the true vacuum corresponds to v_1.

What is the spectrum of excitations around this vacuum? Let us first redefine our field in such a way that $\varphi_{1,2}$ represents infinitesimal fluctuations around a vacuum expectation value v, $\varphi_1 \to v + \varphi_1$:

$$\begin{aligned}
V &= \frac{1}{2}\mu^2(v+\varphi_1)^2 + \frac{1}{2}\mu^2\varphi_2^2 + \frac{\lambda}{4}\left((v+\varphi_1)^2 + \varphi_2^2\right)^2 \\
&= V[v] + (\mu^2 v + \lambda v^3)\varphi_1 + \frac{1}{2}(\mu^2 + 3\lambda v^2)\varphi_1^2 + \frac{1}{2}(\mu^2 + \lambda v^2)\varphi_2^2 \\
&+ \lambda v\varphi_1(\varphi_1^2 + \varphi_2^2) + \frac{\lambda}{4}(\varphi_1^2 + \varphi_2^2)^2
\end{aligned}$$

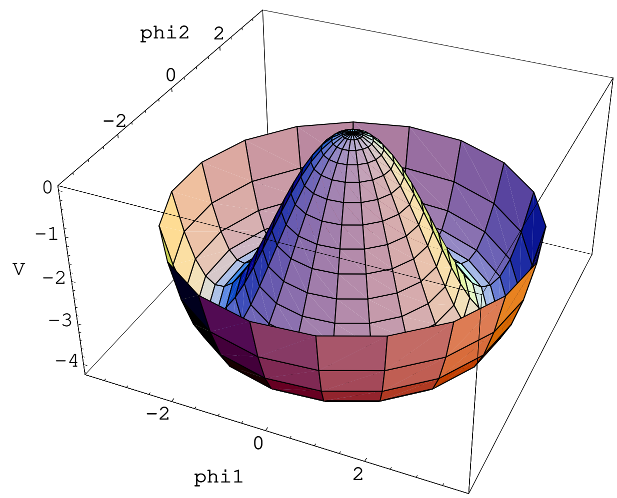


Figure 5.2: Potential for $\mu^2 < 0$.

The first term is just the potential for v. Also, we recognize in the second term the first variation of V. Also, note the appearance of a new cubic interaction term. Substituting v_1 for v, gives finally

$$V[\varphi_1, \varphi_2] = V[v_1, 0] - \mu^2\varphi_1^2 + \lambda v_1\varphi_1(\varphi_1^2 + \varphi_2^2) + \frac{\lambda}{4}(\varphi_1^2 + \varphi_2^2)^2$$

Something remarkable has happened. First we see that φ_1 excitations now have a mass $M_1^2 = -2\mu^2 > 0$. On the other hand, the mode in the direction φ_2 is strictly massless, $M_2^2 = 0$. The latter is called a Goldstone boson. Physically, it is clear why there is a massless excitation. It arises because the potential is flat along the direction φ_2 near the vacuum $\varphi_1 = v_1$, a remnant of the underlying SO(2) symmetry.

exercise 8 *In the broken phase, it would have been more natural to use polar rather than cartesian coordinates,*

$$\Phi(x) = \frac{1}{\sqrt{2}}(v + \sigma(x))\exp(i\pi(x)/v)$$

What is the spectrum in term of the fields σ and π? Hint: consider infinitesimal fluctuations of σ and π.

As a final remark, let us note that we could add a constant term to the potential V, so as to ensure that $V[v] = 0$, *i.e.* At this level of discussion, such a term is innocuous as it can only couple to gravity.

$$V[\phi] = \lambda(\phi^\dagger\phi - v^2/2)^2$$

The advantage is that we see right-away that the potential is minimized for $\langle\phi^\dagger\phi\rangle = v^2/2$.

5.3 SSB: global non-abelian example

We now study an example of spontaneously broken non-abelian symmetry, which is relevant for the breaking of the gauge symmetries of the SM. Consider a real scalar field with four components,

$$\vec{\Phi} = (\varphi_1, \varphi_2, \varphi_3, \varphi_4)$$

and the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu \vec{\Phi} \cdot \partial^\mu \vec{\Phi} - \frac{1}{2}\mu^2 \vec{\Phi} \cdot \vec{\Phi} - \frac{\lambda}{4}(\vec{\Phi} \cdot \vec{\Phi})^2$$

This theory is invariant under the group of rotations SO(4) which leave the norm of the vector $\vec{\Phi}$ invariant. If $\mu^2 > 0$, the symmetry is manifest and the spectrum consists of four real degrees of freedom with mass $M^2 = \mu^2$. If $\mu^2 < 0$, the symmetry is spontaneously broken, with

$$\langle|\vec{\Phi}|\rangle = \sqrt{\frac{-\mu^2}{\lambda}}$$

The vacua in the spontaneously broken phase is a surface of constant $|\vec{\Phi}|$, or three-sphere S^3. Using SO(4) transformations, we can always choose coordinates so that $\langle|\vec{\Phi}|\rangle$ point to the "North Pole" of S^3,

$$\langle|\vec{\Phi}|\rangle = (v, 0, 0, 0)$$

Clearly, the symmetry is not totally broken, as we are still free to make rotations which leave the vacuum vector invariant, or "North Pole" axis invariant. Obviously, these form an SO(3) subgroup. The pattern of symmetry breaking is thus

$$\mathrm{SO}(4) \to \mathrm{SO}(3)$$

How many Goldstone modes are there? We have learned that Goldstone modes correspond to flat directions at the minimum of the potential. In the present case, the minimum is a sphere S^3, which is described by three parameters. Hence, there should be three Goldstone modes. Is this counting correct? To each Goldstone mode corresponds a generator of a spontaneously broken symmetry. The SO(4) group has 6 generators, while SO(3), which is isomorphic to SU(2), has 3. Hence, there are $6 - 3 = 3$ broken symmetries, which matches our counting of 3 Goldstone modes.

exercise 9 *It can be shown that the group* SO(4) *is isomorphic to* SU(2) × SU(2) *by looking directly at their respective Lie algebras. Here, we consider a pretty indirect evidence. Add the* 2×2 *unit matrix to the set of three Pauli matrices, so that* $\vec{\sigma} = (1, \sigma^1, \sigma^2, \sigma^3)$*, and consider the following* 2×2 *complex matrix (actually hermitian),*

$$U = \frac{1}{2}\vec{\Phi} \cdot \vec{\sigma} = \frac{1}{2}\begin{pmatrix} \varphi_4 + \varphi_3 & \varphi_1 + i\varphi_2 \\ \varphi_1 - i\varphi_2 & \varphi_4 - \varphi_3 \end{pmatrix}$$

Show that the Lagrangian for $\vec{\Phi}$ *can be rewritten in this basis as*

$$\mathcal{L} = \mathrm{Tr}(\partial_\mu U^\dagger \partial^\mu U) - \mu^2 \mathrm{Tr}(U^\dagger U) - \lambda(\mathrm{Tr}(U^\dagger U))^2$$

The field U transform naturally under the following SU(2) × SU(2) *transformations,*

$$U \to g^\dagger U \tilde{g}.$$

with $g \neq \tilde{g}$. Show that the Lagrangian is invariant under these transformations. Show that the vacuum expectation value

$$\langle U \rangle = \frac{1}{2}\begin{pmatrix} v & 0 \\ 0 & v \end{pmatrix}$$

breaks the $SU(2)\times SU(2)$ symmetry to the diagonal subgroup SU(2) *(*i.e. *is left invariant by $g = \tilde{g}$ transformations). This is the equivalent in this basis of the SSB of* SO(4) → SO(3) ∼ SU(2).

Yet another possibility is to rewrite $\vec{\Phi}$ as

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} \varphi_1 + i\varphi_2 \\ \varphi_4 + i\varphi_3 \end{pmatrix} \tag{5.2}$$

This object has the same number of components as $\vec{\Phi}$ and U. However, it naturally transforms as a doublet, or fundamental representation of only one SU(2) group. In this parameterization, one of the SU(2) symmetries is thus hidden (The field U is said to transform as bi-doublet, or bi-fundamental representation.) These observations will have some interest when we will come to discuss the breaking of the SU(2) × U(1) symmetries of the SM.

In the latter basis, the Lagrangian is

$$\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - \mu^2 \phi^\dagger \phi - \frac{\lambda}{4}(\phi^\dagger \phi)^2$$

If $\mu^2 < 0$, the SU(2) is totally broken, with three Goldstone bosons. Using $SU(2)$ transformations, the *vev* of ϕ can be chosen to be in the φ_4 direction,

$$\langle \phi \rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix} \tag{5.3}$$

with

$$v^2 = \frac{-\mu^2}{\lambda}$$

For infinitesimal fluctuations around this vacuum, the fields $(\varphi_1, \varphi_2, \varphi_3)$ are Goldstone modes, with zero mass, while excitation along φ_4 are massive, with mass

$$M^2 = -\mu^2 > 0$$

Finally note that, as in the abelian case, in the broken phase, it is often more sensible to work using polar variable,

$$\phi(x) = \frac{1}{\sqrt{2}}\,(v + H(x))\, e^{i\pi^a \sigma^a/2v}$$

where $a = 1, 2, 3$. The π^a fields are the Goldstone modes. The field H corresponds to massive radial excitations. In the framework of the SM, it is called the Higgs field.

5.4 SSB: local abelian symmetry

What happens if these scalar field theories with SSB are gauged? Again, it is perhaps better to start with an abelian example. Let us thus gauge this theory by substituting the normal derivative in the Lagrangian (5.1) by a covariant derivative, $D_\mu = \partial_\mu + ieA_\mu$. The Lagrangian then reads

$$\mathcal{L} = (D_\mu\phi)^\dagger D^\mu\phi - \lambda(\phi^\dagger\phi - v^2/2)^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

where we have redefined the potential so that it vanishes in the broken phase. In this parametrization, $\mu^2 = -\lambda v^2$ in the broken phase. Here we directly consider the case $\mu^2 < 0$ so that

$$\phi = \frac{1}{\sqrt{2}}(v + \varphi_1 + i\varphi_2)$$

We know already what happens to the potential term. The φ_1 field is massive, $M_1^2 = -2\mu^2$, while, for infinitesimal fluctuations, the field φ_2 is massless, $M_2^2 = 0$. So let us concentrate on the kinetic term. Decomposing it gives

$$\begin{aligned} D_\mu\phi^\dagger D^\mu\phi &= \frac{1}{2}(\partial_\mu\varphi_1\partial^\mu\varphi_1 + \partial_\mu\varphi_2\partial^\mu\varphi_2) + eA_\mu(\varphi_1\partial^\mu\varphi_2 - \varphi_2\partial^\mu\varphi_1) + \frac{e^2}{2}\,(\varphi_1^2 + \varphi_2^2)A_\mu A^\mu \\ &\quad + ev\,A_\mu\partial^\mu\varphi_2 + \frac{e^2v^2}{2}A_\mu A^\mu + e^2v\varphi_1\,A_\mu A^\mu \end{aligned} \tag{5.4}$$

The terms in the first line on the RHS are standard features. They are the kinetic terms for the field φ_1 and φ_2 and the coupling of A_μ to the electric current, $j_\mu = e(\phi^\dagger\partial_\mu\phi - \partial_\mu\phi^\dagger\phi)$, including a quadratic term. The terms of the second line are new. The first remarkable feature is that, because of SSB, the gauge field A_μ *has become massive*,

$$M_A^2 = e^2v^2.$$

However, the presence of

$$-evA_\mu\partial^\mu\varphi_2.$$

makes this interpretation less clear, as it introduces a mixing between the the fields A_μ and φ_2. To remove this annoying term, we first parameterize the field ϕ using polar variables,

$$\begin{aligned} \phi(x) &= \frac{1}{\sqrt{2}}(v + \sigma(x))\exp(i\pi(x)/v) \\ &= \frac{1}{\sqrt{2}}\,[v + \sigma(x) + i\pi(x) + \ldots] \end{aligned}$$

For small fluctuations, the fields σ and π are just φ_1 and φ_2. You can verify that the free part of the Lagrangian also keeps the same form in these variables and these fields have the same particle interpretation.

We can now remove the unwanted term by using the gauge freedom of our theory. To do this, we make the gauge transformation

$$\begin{aligned} \phi(x) &\rightarrow \phi^U(x) = e^{-i\pi(x)/v}\phi = \frac{1}{\sqrt{2}}(v+\sigma) \\ A_\mu(x) &\rightarrow A_\mu^U(x) = A_\mu(x) - \frac{1}{ev}\partial_\mu\pi(x) \end{aligned} \tag{5.5}$$

In this basis, also called the *unitary gauge* (although this is a bit misleading), the Lagrangian becomes

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma - \frac{|\mu^2|}{2}\sigma^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}e^2v^2A_\mu A^\mu \\ &+ \frac{1}{2}e^2A_\mu A^\mu\sigma(2v+\sigma) - \lambda v^2\sigma^3 - \frac{1}{4}\sigma^4 + \text{const} \end{aligned} \tag{5.6}$$

We have dropped the U(nitary) upper-script to avoid cluttering of symbols. Clearly, the free Lagrangian describes the propagation of a massive gauge boson A_μ^u, with mass $M_A = ev$, and a real scalar particle σ, with mass $M_\sigma = 2|\mu|$. All traces of the Goldstone boson $\pi(x)$ have disappeared from the Lagrangian. Let us check that the counting of degrees of freedom is preserved. Initially, we had one massless gauge field, with two degrees of freedom, and a complex scalar field, thus altogether 4 degrees of freedom. In the new basis, we a one scalar field, and a massive gauge boson, with three degrees of freedom (corresponding to the three possible polarization states of a massive spin-1 object), thus altogether also 4 free degrees of freedom, as required. We see that the massless gauge field A_μ has combined with the scalar field $\pi(x)$ to become a massive vector field, A_μ^u. This is the Brout-Englert-Higgs mechanism. The field π is called a *would-be Goldstone boson.*

exercise 10 *Derive the propagator of a massive abelian gauge boson A_μ^u, Hint: The propagator of a massive vector particle is obtained from the Proca equation*

$$\partial^\mu F_{\mu\nu} + M_A^2 A_\nu = j_\nu$$

Taking the derivative of the LHS and RHS sides gives

$$\partial^\mu\partial^\nu F_{\mu\nu} + M_A^2\partial^\nu A_\nu = \partial^\nu j_\nu$$

Since $F_{\mu\nu}$ is antisymmetric, this equation yields,

$$\partial^\nu A_\nu = \frac{\partial^\nu j_\nu}{M_A^2}$$

Use the last equation to show that

$$D_{\mu\nu} = \frac{-1}{k^2 - M_A^2}\left(\eta_{\mu\nu} - \frac{k_\mu k_\nu}{M_A^2}\right).$$

with

$$A_\mu(k) = D_{\mu\nu}(k)j^\nu(k)$$

The unitary gauge has the advantage that it exhibits the correct number of propagating degrees of freedom. Nevertheless, it is instructive to understand how things work if the Golstone mode is kept explicit. At weak coupling e, the term quadratic in the gauge field can be treated as an interaction:

$$ie^2 v^2 \eta_{\mu\nu}$$

The mixing term on the other hand gives rises to a contribution which is of the form,

$$(-evk_\mu) \times \left(\frac{i}{k^2}\right)(evk_\nu) = -ie^2 v^2 \frac{k_\mu k_\nu}{k^2}$$

(See Figure 5.4.)

Putting these to contribution together gives

$$i\Pi_{\mu\nu} = ie^2 v^2 \left(\eta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2}\right) \tag{5.7}$$

$\Pi_{\mu\nu}$ is the polarization tensor of the gauge field, *i.e.* the correlator

$$i\Pi_{\mu\nu}(k^2) = e^2 \int d^4x\, e^{ik\cdot x} \langle 0|J_\mu(x) J_\nu(0)|0\rangle$$

where, in the broken phase, the bosonic current is $J_\mu = v\partial_\mu \phi_2 + O(\phi^2)$. Note that the polarization tensor is transverse,

$$k^\mu \Pi_{\mu\nu}(k^2) = k^\nu \Pi_{\mu\nu}(k^2) = 0$$

as it should be because of current conservation $\partial_\mu J^\mu = 0$. This is an example of a Ward identity. Ward identities are the expression at the quantum level of the symmetries of the Lagrangian. (For Yang-Mills theories, Ward identities are called Slavnov-Taylor identities.) We see that the contribution of the Goldstone mode is critical in insuring that the polarization tensor is transverse, and thus that the gauge invariance of the theory is preserved even in the broken phase.

In terms of Feynman diagrams, the polarization tensor gives a correction to the propagator of the gauge field,

$$\begin{aligned}
D_{\mu\nu} &= \frac{-\eta_{\mu\nu}}{k^2} + \frac{-\eta_{\mu\alpha}}{k^2}\Pi^{\alpha\beta}(k^2)\frac{-\eta_{\beta\nu}}{k^2} + \dots \\
&= \frac{-\eta_{\mu\nu}}{k^2} \times \left(1 + \frac{e^2 v^2}{k^2}\right) + O\left(\frac{k_\mu k_\nu}{k^2}\right) \\
&\approx \frac{-\eta_{\mu\nu}}{k^2 - e^2 v^2} + \dots
\end{aligned}$$

We see that the gauge field gets a mass $M_A = ev$.

5.5 Breaking the gauge symmetry of the Standard Model

The coupling of the MS fermions to gauge bosons has required the introduction of the gauge group $SU(2) \times U(1)_Y$. We have also seen that the massless photon field can be naturally identified as a mixture of B_μ and W^3_μ. This leaves us with three gauge fields (the field Z_μ, that couples to neutral currents and the charge fields $W^\pm_\mu$) which must be massive to match the fact that 1/ the gauge symmetry of the SM is not manifest at low energies and 2/ to make the weak interactions short ranged. The simplest way to break the $SU(2) \times U(1)_Y$ gauge symmetry is to introduce a scalar field in the doublet representation of $SU(2)$. Indeed, we have seen that a vacuum expectation for such a field completely break $SU(2)$ and gives rise to three Golstone modes, which is precisely what we need to give a mass to three gauge bosons. So we introduce an extra field to our Lagrangian,

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \varphi_1(x) + i\varphi_2(x) \\ v + H(x) + i\varphi_3(x) \end{pmatrix} \tag{5.8}$$

This field is called the *Higgs field* and $H(x)$ the *Higgs boson*. The field $\phi(x)$ can also couple to the $U(1)_Y$ gauge field, with an hypercharge Y_ϕ which we will determined in a second. The covariant derivative of ϕ is thus

$$D_\mu\phi = (\partial_\mu + igW^a_\mu \frac{\sigma^a}{2} + ig' \frac{Y_\phi}{2} B_\mu)\phi$$

We consider right-away the unitary gauge, so that the fields φ_i, $i = 1, 2, 3$ have been removed from the Lagrangian. Substituting ϕ by its vacuum expectation value gives a term

$$\mathcal{L}_{W,B} = \frac{1}{4} \left| \begin{pmatrix} gW^3_\mu + g'Y_\phi B_\mu & g(W^1_\mu + iW^2_\mu) \\ g(W^1_\mu - iW^2_\mu) & -gW^3_\mu + g'Y_\phi B_\mu \end{pmatrix} \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \right|^2$$

in the Lagrangian. Expanding this term gives

$$\mathcal{L}_{WB} = \begin{pmatrix} W^+_\mu & W^3_\mu & B_\mu \end{pmatrix} \begin{pmatrix} \frac{g^2v^2}{4} & 0 & 0 \\ 0 & \frac{g^2v^2}{8} & \frac{-gg'Y_\phi v^2}{8} \\ 0 & \frac{-gg'Y_\phi v^2}{8} & \frac{g'^2 Y_\phi^2 v^2}{8} \end{pmatrix} \begin{pmatrix} W^-_\mu \\ W^3_\mu \\ B_\mu \end{pmatrix} \tag{5.9}$$

where we have used $W^\pm_\mu = 1/\sqrt{2}(W^1_\mu \mp iW^2_\mu)$. We see that the charged W bosons get a mass

$$M_W = \frac{gv}{2} \tag{5.10}$$

If $Y_\phi \neq 0$, W^3_μ and B_μ are mixed. Note that the mass-mixing matrix has a vanishing determinant, hence one of its eigenvalues vanishes: we want to identify the corresponding eigenvector with the photon. This is

$$A_\mu(x) = \frac{g'Y_\phi}{\sqrt{g^2 + g'^2 Y_\phi^2}} W^3_\mu(x) + \frac{g}{\sqrt{g^2 + g'^2 Y_\phi^2}} B_\mu(x) \tag{5.11}$$

Comparing with (4.25) and (4.27), we see that we must set

$$Y_\phi = +1$$

to identify A_μ with the photon field

$$A_\mu = \sin\theta_W W^3_\mu + \cos\theta_W B_\mu.$$

Finally, the orthogonal combination

$$Z_\mu = \cos\theta_W W^3_\mu - \sin\theta_W B_\mu$$

is massive, with

$$M_Z = \frac{\sqrt{g^2 + g'^2}v}{2} \tag{5.12}$$

or

$$M_Z = \frac{M_W}{\cos\theta_W} \tag{5.13}$$

Let us recapitulate. We have successfully broken the symmetry of the SM, $SU(2)\times U(1)_Y \to U(1)_Q$. We have had to introduce four parameters,

$$g,\ g',\ \mu^2 \text{ and } \lambda$$

which can be re-expressed in term of

$$e,\ M_Z, \tan\theta \text{ and } M_H$$

for instance. Until a few years ago, only the first three were known and the Higgs was missing. This particle was finally discovered in 2012 by the CMS and ATLAS experiments at the LHC. The numerical values are

$$M_W = 80.4\,\text{GeV} \qquad M_Z = 91.2\,\text{GeV} \qquad M_H = 125\,\text{GeV}$$

From this $\sin^2\theta_W = 1 - M_W^2/M_Z^2 = 0.223$. Note however that the Weinberg angle can also be defined from, e.g. the Z boson decay or of the muon decay rates. These processes occur at different energies, and so the value of the Weinberg angle, which is energy dependent (as the gauge couplings do) is different. Z boson decay gives $\sin^2\theta_W|_Z = 0.2312$ (see e.g. Peskin).

The last thing we need to understand is how the Higgs doublet couples to the fermions of the SM. To simplify, consider first its coupling to the leptons of the first family. Consider first its coupling to lepton fields. A minimal coupling is

$$\mathcal{L}_{Yukawa} = y_e \bar{L}_e \phi E_e + h.c.$$

This is called a *Yukawa term* and y_e a Yukawa coupling. Because the Higgs doublet is a scalar under Lorentz transformation, this coupling must flip chiralities. Thus it couples the left and right-handed leptons. It is invariant under $SU(2)$ transformations and, looking at the hypercharge assignments,

$$-Y_L + Y_\phi + Y_E = -(-1) + 1 - 2 = 0 \tag{5.14}$$

we see that it is also invariant under the group $U(1)_Y$. Substituting the vev of the Higgs doublet in, we get

$$\mathcal{L}_{Yukawa} = y_e \frac{v}{\sqrt{2}} \bar{e}_L e_R + y_e^* \frac{v}{\sqrt{2}} \bar{e}_R e_L \tag{5.15}$$

Assuming for a while that the Yukawa coupling is real, we see that the SSB also gives a mass to the electron, with

$$M_e = y_e \frac{v}{\sqrt{2}}$$

What about the quarks? Again to simplify, we consider only one family of quarks. The term

$$\mathcal{L}_{Yukawa} = y_d \bar{Q}_i \phi D_j \tag{5.16}$$

is manifestly $SU(2)$ invariant. The only issue is the hypercharge assignment, but

$$-Y_Q + Y_\phi + Y_D = -\frac{1}{3} + 1 - \frac{2}{3} = 0$$

so this term is also acceptable. After SSB, it leads to a mass for the down quark

$$\mathcal{L}_{Yukawa} = y_d \frac{v}{\sqrt{2}} \bar{d}_L d_R + y_d^* \frac{v}{\sqrt{2}} \bar{d}_R d_L \tag{5.17}$$

Taking y_d real, gives

$$M_d = y_d \frac{v}{\sqrt{2}}$$

Now, it seems we have a problem if we want to give a mass to the up quark. But consider

$$\tilde{\phi} = i\sigma^2 \phi^*$$

Using properties of the Pauli matrices, it is easy to verify that this object transforms just like ϕ under $SU(2)$ transformation,

$$\begin{aligned}
\tilde{\phi} &\rightarrow i\sigma^2 (e^{i\alpha^a \sigma^a/2} \phi)^* \\
&= i\sigma^2 e^{-i\alpha^a \sigma^{a*}/2} \phi^* \\
&= e^{i\alpha^a \sigma^a/2} \tilde{\phi}
\end{aligned}$$

exercise 11 *Show that the Pauli matrices satisfy*

$$\sigma^2 \sigma^a \sigma^2 = -\sigma^{a*}.$$

Use this to verify that $\tilde{\phi}$ transforms like ϕ. What is $\tilde{\phi}$ in components?

Hence, we can also consider the SU(2)×U(1)$_Y$ invariant Yukawa term

$$\mathcal{L}_{Yukawa} = y_u \bar{Q} \tilde{\phi} U$$

Indeed,

$$-Y_Q - Y_L + Y_U = -\frac{1}{3} - 1 + \frac{4}{3} = 0$$

which, after SSB, gives a mass to the up quark,

$$M_u = y_u \frac{v}{\sqrt{2}}$$

This construction can be generalized to the other families of quarks and leptons. Note however that nothing prevents the Yukawa couplings to mix different families and even to be complex. In particular, mass eigenstates and flavor eigenstates will generically be different. A complex Yukawa matrix may also lead to CP violation in the SM (provided there are at least three generations of quarks.) These subtleties, and other phenomenological consequences of the SM, will be discussed in the next chapter.

Chapter 6

Aspects of Electroweak interactions

6.1 Low energy limit of the SM and Fermi's theory

Consider for instance the decay of the muon,

$$\mu^- \to \nu_\mu + e^- + \nu_e$$

In the SM, this decay proceeds through the exchange of a W_μ^- boson. This is a low energy process compared to the mass of the W, so we can take the limit $k \to 0$ in the propagator,

$$-\frac{i\left(\eta_{\mu\nu} - \frac{k_\mu k_\nu}{M_W^2}\right)}{k^2 - M_W^2} \to +\frac{i\eta_{\mu\nu}}{M_W^2}$$

In this limit, the effective interaction reduces to

$$\mathcal{A} = \left(\frac{-ig}{\sqrt{2}}\right)^2 \frac{i\eta_{\mu\nu}}{M_W^2}(\bar{e}_L\gamma_\mu\nu_e)(\bar{\nu}_\mu\gamma^\nu\mu_L)$$

This amplitude can be derived from an effective interaction Lagrangian

$$\mathcal{L}_{eff} = \frac{-g^2}{2M_W^2}(\bar{\nu}_\mu\gamma_\mu\mu_L)(\bar{e}_L\gamma^\mu\nu_e) + h.c. \tag{6.1}$$

Such an interaction term was introduced by Fermi in 1934, soon after Pauli had proposed his neutrino hypothesis. He was concerned with the β-decay of the neutron (the muon was still to be discovered) but the spirit is the same. Also, it did not have the correct $V - A$ structure of weak couplings. This also had to wait a little bit. Note that, because the operator in this effective Lagrangian is of dimension 6, the "coupling constant" is dimension-full, $\propto g^2/M_W^2$. Fermi introduced what is now called Fermi's coupling constant G_F. Matching with the SM gives

$$\frac{g^2}{8M_W^2} = \frac{G_F}{\sqrt{2}}$$

Since $g = e/\sin\theta_W$, one needed the value of $\sin\theta_W$ to allowed to predict the mass of the W and Z bosons. At low energies, one can extract $\sin\theta_W$ from the scattering of neutrinos through neutral currents.

exercise 12 *Check that the ratio*

$$\frac{\sigma(\nu_\mu e \to \nu_\mu e)}{\sigma(\bar{\nu}_\mu e \to \bar{\nu}_\mu e)}$$

only depends on $\sin\theta_W$.

Furthermore, using $G_F \approx 10^{-5} GeV^{-2}$ as measured from the muon lifetime and $M_W^2 = g^2 v^2/4$, gives

$$v = 2^{-1/4} G_F^{-1/2} \approx 246\,\mathrm{GeV}$$

(I don't give full justice to these numbers. They are actually known with much better precision, but these expressions are also affected by quantum corrections, which makes the extraction of numbers not totally straightforward.)

Although Fermi's theory is successful in describing low energy processes, it has many shortcomings. Note that the effective Lagrangian (6.1) is a dimension 6 operator, with a dimension-full coupling $[G_F] = E^{-2}$. The jargon is that such a coupling is irrelevant in the infrared, $E \to 0$. For instance, just on dimensional grounds, we expect the neutrino scattering cross-section on a nucleon at low energies $E \ll m_W$ to scale like

$$\sigma \sim G_F^2 s \sim G_F^2 E m_N$$

with $s = (p_\nu + p_n)^2$ and E the neutrino energy in the rest frame (so-called lab frame) of the nucleon (mass m_N). Conversely, it has a bad UV behavior, as, taken at face value, this cross-section growth without limit with E. This kind of behavior is in contradiction with unitarity bounds on high energies scattering cross-sections. Technically, it is said that such an effective theory is non-renormalizable. This means that quantum corrections require the introduction of counter-terms of arbitrary order. For instance, the first one-loop diagram of Figure 6.1 for instance superficially behaves like $O(G_F^2 \Lambda^2)$ whose divergence can be absorbed in a redefinition of the Fermi constant. The other one-loop diagram is also divergent, $\mathcal{M} \sim G_F^3 \Lambda$ but, since it involves 6 particles, its divergence requires to introduce another effective coupling. The new coupling can induce higher order interactions that are also divergent and so the process repeats, requiring an infinite number of parameters (and thus experiments) to fully describe the theory. The SM resolves all these problems at once. The same diagram in the SM is indeed much better behaved in the UV.[1]

6.2 Aspects of flavor physics

When we discussed the coupling of the Higgs field to the quarks and leptons, we limited ourself to one family. We now generalize this construction to the three known generations of the SM.

[1]We are a bit sketchy here. In the unitary gauge, the propagator of the W boson has a part which behaves like g^2/M_W^2 at high energies, thus spoiling renormalizability; at least naively. The way to deal with this has been shown by 't Hooft and consists in keeping the Goldstone mode explicit and to add a gauge fixing term to the Lagrangian, like $1/2\xi(\partial_\mu W^\mu)^2$, to invert the gauge boson propagator,

$$-\frac{\eta_{\mu\nu} - \frac{(\xi-1)k_\mu k_\nu}{M^2\xi - k^2}}{k^2 - M^2}$$

In these gauges, the propagator falls like $1/k^2$ at high energies and has a much better behavior from the point of view of renormalization.

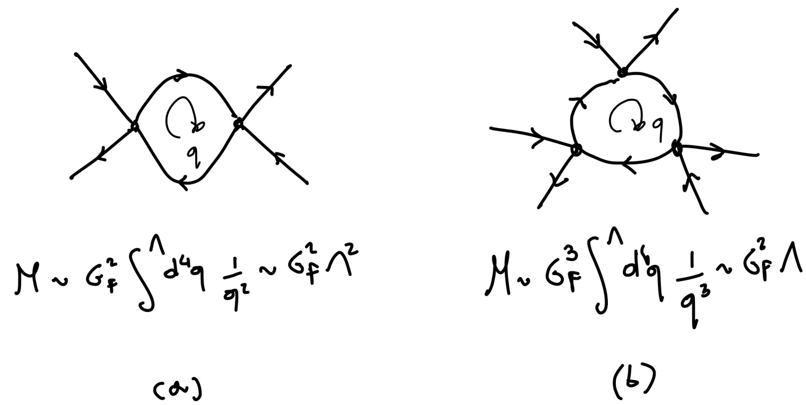


Figure 6.1: Divergent amplitudes in the Fermi effective theory. The amplitude with 6 fermions (b) diverges, which requires to introduce a dim 9 effective operator to absorb its divergence, thus limiting the predictive power of the Fermi theory.

The Yukawa couplings to three generations of quarks and leptons reads (assuming the neutrinos are strictly massless to begin with),

$$\mathcal{L}_{yukawa} = \sum_{i,j=1\ldots3} y^D_{ij} \bar{Q}_i \phi D_j + \sum_{k,l=1\ldots3} y^U_{kl} \bar{Q}_i \tilde{\phi} U_l + \sum_{i,j=e,\mu,\tau} y^E_{ij} \bar{L}_i \phi E_j + h.c. \tag{6.2}$$

where we have introduced three 3×3 complex matrices, y^D, y^U,y^E with labels corresponding to whether the down-like and up-like quarks or charged leptons get a mass after SSB. Note that in the absence of any specific global symmetry among the three families, there is no restriction on the structure of these matrices. In particular, they don't have to be diagonal, nor real. That there can be non-diagonal will imply that propagating (or mass) eigenstates and gauge interactions eigenstates will in general be different. That they can be complex will lead to the possibility of CP violation in the SM.

exercise 13 *Why are the gauge couplings g and g' real? (hint: go back to the definition of gauge transformations).*

6.2.1 Gauge *vs* mass eigenstates

Because the gauge and propagating eigenstates can be different, we have to make a decision about nomenclature. The most natural thing to do is to keep the name "electron" or "up-quark" for mass eigenstates. To distinguish them from the gauge interaction eigenstates, we put a prime on the latter,

$$\begin{array}{rcl}
(e,\mu,\tau) & \rightarrow & (e',\mu',\tau') \\
(\nu_e,\nu_\mu,\nu_\tau) & \rightarrow & (\nu'_e,\nu'_\mu,\nu'_\tau) \\
(u,c,t) & \rightarrow & (u',c',t') \\
(d,s,b) & \rightarrow & (d',s',b')
\end{array}$$

in such a way that SU(2) doublets (gauge eigenstates) become

$$L_i = \begin{pmatrix} \nu'_{iL} \\ e'_{iL} \end{pmatrix} \quad \text{and} \quad Q_i = \begin{pmatrix} u'_{iL} \\ d'_{iL} \end{pmatrix}$$

After SSB, we have (with summation over indices implicit)

$$\mathcal{L}_{Yukawa} \;=\; \frac{v+H(x)}{\sqrt{2}}\left[y^E_{ij}\,\bar{e}'_{iL}e'_{jR} + y^U_{ij}\,\bar{u}'_{iL}u'_{jR} + y^D_{ij}\,\bar{d}'_{iL}d'_{jR}\right] + h.c.$$

Thus in the gauge-eigenstate basis, the fermion mass matrices are of the form

$$M^{E,D,U} = -\frac{v}{\sqrt{2}}y^{E,D,U}$$

There is no *a priori* reason for these matrices to be diagonal. Actually, in general, there are neither symmetric nor hermitian. Now, it turns out that an arbitrary square matrix can always be diagonalized using a bi-unitary transformation, *i.e.* given M_{ij}, there exist two unitary matrices S and T such that

$$S^\dagger M T = M_{diag}$$

exercise 14 *Complete the steps in the following demonstration. As $MM^\dagger$ is hermitian and positive definite, there is a unitary transformation S such that*

$$S^\dagger M M^\dagger S = M^2_{diag}$$

with

$$M^2_{diag} = \begin{pmatrix} m_1^2 & 0 & 0 \\ 0 & m_2^2 & 0 \\ 0 & 0 & m_3^2 \end{pmatrix}$$

Show that the matrix S is unique up to a phase transformation,

$$S \rightarrow SF$$

with

$$F = \begin{pmatrix} e^{i\phi_1} & 0 & 0 \\ 0 & e^{i\phi_2} & 0 \\ 0 & 0 & e^{i\phi_3} \end{pmatrix}$$

This freedom can be used to ensure that the eigenvalues of $M_{\rm diag}$ are real and positive,

$$M_{diag} = \begin{pmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{pmatrix}$$

with $m_i \geq 0$. Now define the hermitian matrix

$$H = S M_{\rm diag} S^\dagger$$

and show that $V = H^{-1}M$ is a unitary matrix. Then, by definition,

$$S^\dagger H S = S^\dagger M V^\dagger S = M_{\rm diag}$$

Hence,

$$S^\dagger M T = M_{diag}$$

with S and $T = V^\dagger S$ both unitary.

Using this result, the relation between gauge and mass eigenstates is

$$\begin{aligned} \bar{\psi}'_L M \psi'_R &= \bar{\psi}' S M_{\rm diag} T^\dagger \psi'_R \\ &= \bar{\psi}_L M_{diag} \psi_R \end{aligned}$$

with

$$\psi'_L = S\psi_L \tag{6.3}$$
$$\psi'_R = T\psi_R \tag{6.4}$$

6.2.2 Mixing in quark charged current couplings

Consider the coupling of quarks to the $W^\pm_\mu$ gauge bosons,

$$\begin{aligned} \mathcal{L} &= -\frac{g}{\sqrt{2}} W^+_\mu \bar{u}'_{i_L} \gamma^\mu d'_{iL} + h.c. \\ &= -\frac{g}{\sqrt{2}} W^+_\mu \bar{u}_{i_L} \gamma^\mu \left[S^\dagger_u S_d \right]_{ij} d_{jL} + h.c. \\ &= -\frac{g}{\sqrt{2}} W^+_\mu \bar{u}_{i_L} \gamma^\mu U_{ij} d_{jL} + h.c. \end{aligned} \tag{6.5}$$

where

$$u'_{iL} = (S_u)_{ij} u_{jL} \quad \text{and} \quad d'_{iL} = (S_d)_{ij} d_{jL}$$

The mixing matrix $U_{CKM} = S^\dagger_u S_d$ is clearly unitary. It is called the Cabibbo-Kobayashi-Maskawa matrix (CKM).

6.2.3 CP violating phases

If the Yukawa couplings to quarks where real, the CKM matrix would be orthogonal and all the quark-gauge couplings would be real. In general, an $n \times n$ complex matrix depends of n^2 complex or $2n^2$ real parameters. These are reduced to n^2 by the unitarity condition. (Prove this.) These can be split into $n(n-1)/2$ real angles, corresponding to the parametrization of a $n \times n$ orthogonal matrix and $n^2 - n(n-1)/2 = n(n+1)/2$ complex phases. Not all these phases are physical because $(2n-1)$ of these can be re-absorbed in a redefinition of the quark fields. (At first sight, we could think that we can remove $2n$ phases. But the CKM matrix is invariant under a redefinition of all the quarks by the same phase. Hence, we can only eliminate $2n-1$ phases.) Altogether, the CKM matrix can be parameterized by

$$\frac{N_f(N_f-1)}{2} \text{ real mixing angles}$$

and

$$\frac{(N_f-1)(N_f-2)}{2} \text{ CP violating phases}$$

where N_f is the number of families. For $N_f = 2$, there is one mixing angle and zero phase, hence no room for CP violation. In this case, which dates back to the time when there was no hint of a third family, but also no charm quark, the mixing matrix was introduced by Cabibbo, with interactions eigenstates defined as

$$\begin{pmatrix} d_\theta \\ s_\theta \end{pmatrix} = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}$$

in term of the mass eigenstates. $\theta_c \approx 15°$ is called the Cabibbo angle. For $N_f = 3$, there are three real mixing angle and one CP violating phase. The mixing matrix is called the CKM matrix.

CP violation is an important, still not fully understood problem of the SM. We don't have time to discuss this here. For a review, one can consult the Particle Data Book.

6.2.4 Flavor conservation in neutral currents

The CKM matrix implies that hadronic charged currents are non-diagonal in flavor space. For instance, it implies that strangeness is not conserved in weak interactions, such as in

$$K^0 \to \pi^- + e + \bar{\nu}_e$$

which has $\Delta S = +1$. On the other hand, no flavor changing neutral currents have been observed so far. In the framework of the SM, this is guaranteed by the unitarity of the matrix U_{CKM}.

exercise 15 *Prove this. Check that the coupling of the Z boson to quarks is the same in all basis.*

6.2.5 No mixing in the leptonic sector of the SM?

In this section, we first assume that neutrino are strictly massless. We can do the same reasoning as for the quarks. The couplings to charged gauge bosons is

$$\begin{aligned}\mathcal{L} &= -\frac{g}{\sqrt{2}}W_\mu^- \bar{e}'_{iL}\gamma^\mu \nu'_{iL} + h.c. \\ &= -\frac{g}{\sqrt{2}}W_\mu^- \bar{e}_{iL}\gamma^\mu \left[S_e^\dagger S_\nu\right]_{ij} e_{jL} + h.c. \\ &= -\frac{g}{\sqrt{2}}W_\mu^- \bar{e}_{iL}\gamma^\mu E_{ij} e_{jL} + h.c.\end{aligned}$$

with E unitary. However, in the limit of massless (actually, degenerate mass is enough) neutrinos, any neutrino basis is as good as another. In particular, we can take the gauge interaction and mass eigenstates to be the same. In this respect, note that the left-handed neutrinos are precisely identified as mixing particles in decays like

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

or in charged lepton production in scatterings.

In Nature, it seems well established that the neutrinos are massive (see the Particle Data Book for a review). Mixing effects are proportional to the difference in masses of the neutrinos, which in any case, must be very tiny. Hence, lepton flavor violating processes in the SM, if any, should be very suppressed. Related to this issue is the existence of global lepton number conservation laws. In the limit of massless quarks, the separate electronic L_e, muonic L_μ and tauonic L_σ lepton number are strictly conserved because there is no mixing between the different lepton flavors. If the neutrinos have a Dirac mass, only the total lepton number would be conserved, $L = L_e + L_\mu + L_\sigma$. Finally, if the neutrinos are of the Majorana kind, the lepton number is not conserved. We will go back to these questions in the last chapter. Note that there are rather strong experimental limits on the existence of lepton flavor and/or lepton number violation. See the PDB for details.

6.3 Some explicit calculations

In this section, we compute in some details two decay processes of the SM. For lack of time, we limit ourself to processes with only two particles in the final state. Otherwise, we would waste a lot of time doing phase-space integrals. In case you are interested, more involved processes are described in the textbooks given in the references.

6.3.1 W-boson decay into leptons

We want to evaluate the amplitude for W decay in its rest frame. The polarization vector can be chosen to lie along the $\vec{1}_z$ axis. Starting either from the Lagrangian or using the Feynman rules, we can show that the amplitude is

$$\mathcal{M} = \frac{g}{\sqrt{2}}\epsilon_\mu \bar{u}_{\nu_e}(q)\gamma^\mu \frac{1-\gamma_5}{2} v_e(p)$$

Summing over the outgoing lepton spins, and neglecting the electron mass, we get

$$\begin{aligned}
\sum_{spins} |\mathcal{M}|^2 &= \frac{g^2}{2}\mathrm{Tr}\left[\not\epsilon \not p \not\epsilon \frac{1-\gamma_5}{2} \not q\right] \\
&= g^2 \left(\epsilon \cdot p\, \epsilon \cdot q - \epsilon^2\, p \cdot q + \epsilon \cdot q\, \epsilon \cdot p\right)
\end{aligned}$$

To evaluate this take, for instance,

$$\begin{aligned}
\epsilon^\mu &= (0, 0, 0, 1) \\
p^\mu &= (p, 0, p \sin\theta, p\cos\theta) \\
q^\mu &= (p, 0, -p \sin\theta, -p\cos\theta) \\
p &= \frac{M_W}{2}
\end{aligned}$$

to get

$$\sum_{spins} |\mathcal{M}|^2 = \frac{g}{\sqrt{2}} M_W^2 (1 - \cos^2\theta)$$

The partial decay width for this process is

$$d\Gamma(W \to \bar{e}\nu_e) = (2\pi)^4 \delta^4(p_W - p - q) \frac{|\mathcal{M}|^2}{2\omega_W} \frac{d^3p}{2\omega_e (2\pi)^3} \frac{d^3q}{2\omega_\nu (2\pi)^3}$$

which, all (trivial) integrations done, gives finally

$$\Gamma(W \to \bar{e}\nu_e) = \frac{g^2}{48\pi} M_W$$

exercise 16 *Do the integrals. It's easy thanks to the δ-functions.*

exercise 17 *Compute the decay width of the Z boson into neutrinos. This is of interest because the total decay width of the Z affects the shape of the total cross-section for e^+e^- annihilation at the Z-pole. This measurement was done at LEPI and SLAC with the conclusion that there are three families of light (active) neutrinos in Nature.*

6.3.2 The pion decay rate and f_π

The decay $\pi^+ \to \mu^+ + \nu_\mu$ is described by the effective Lagrangian for the four-fermion interaction,

$$\mathcal{L}_{eff} = -\frac{4G_F}{\sqrt{2}} \cos\theta_c \left[\bar{u}_L \gamma^\mu d_L\right] \left[\bar{\mu}_L \gamma_\mu \nu_\mu\right]$$

To perform the calculation of the decay rate, we need some informations on matrix elements involving the pion. First, it can be shown that the following matrix element must vanish,

$$\langle 0 | \bar{u} \gamma^\mu d | \pi^+ \rangle = 0$$

because strong interactions conserve parity.

The pseudovector current on the other has a non-vanishing matrix element, which it is convenient to parameterize as

$$\langle 0|\bar{u}\gamma^\mu\gamma_5 d|\pi^+(p)\rangle = i\sqrt{2}f_\pi p^\mu$$

This structure arises because the matrix element transforms as a vector, and the only one available is p^μ. The form-factor f_π is called the pion decay constant. Its value can be determined by computing the pion decay rate and matching it to the experimental value.

The matrix element for the decay is of the form,

$$\begin{aligned}
\mathcal{M} &= -i\frac{G_F}{\sqrt{2}}\cos\theta_c\langle 0|\bar{u}\gamma_\mu\gamma_5 d|\pi^+(p)\rangle\ \bar{v}_{\mu^+}(k_2)\gamma^\mu(1-\gamma_5)u_{\nu_\mu}(k_1) \\
&= G_F f_\pi\cos\theta_c p_\mu\bar{v}_{\mu^+}(k_2)\gamma^\mu(1-\gamma_5)u_{\nu_\mu}(k_1) \\
&= G_F f_\pi\cos\theta_c m_\mu\bar{v}_{\mu^+}(k_2)(1-\gamma_5)u_{\nu_\mu}(k_1)
\end{aligned}$$

where $p = k_1 + k_2$, with k_2 the momentum of the pion, and we took the neutrino to be massless (a very good approximation indeed).

The decay rate in the rest frame of the pion is then,

$$\Gamma(\pi^+ \to \mu^+ + \nu_\mu) = \frac{1}{2m_\pi}\int(2\pi)^4\delta^4(p-k_1-k_2)\sum_{spin}|\mathcal{M}|^2\frac{d^3k_1}{2\omega_1(2\pi)^3}\frac{d^3k_2}{2\omega_2(2\pi)^3}$$

The square of the amplitude is easily evaluated. Just for fun, let us be a bit more pedestrian than usual:

$$\begin{aligned}
\sum_{spins}|\mathcal{M}|^2 &= G_F^2 f_\pi^2\cos^2\theta_c m_\mu^2\mathrm{Tr}\left[\bar{v}(k_2)(1-\gamma_5)u(k_1)u^\dagger(k_1)(1-\gamma_5)\gamma^0 v(k_2)\right] \\
&= G_F^2 f_\pi^2\cos^2\theta_c m_\mu^2\mathrm{Tr}\left[\bar{v}(k_2)(1-\gamma_5)u(k_1)\bar{u}(k_1)(1+\gamma_5)v(k_2)\right] \\
&= G_F^2 f_\pi^2\cos^2\theta_c m_\mu^2\mathrm{Tr}\left[(\not{k}_2 - m_\mu)(1-\gamma_5)\not{k}_1(1+\gamma_5)\right] \\
&= G_F^2 f_\pi^2\cos^2\theta_c m_\mu^2\mathrm{Tr}\left[(\not{k}_2 - m_\mu)\not{k}_1\right] \\
&= G_F^2 f_\pi^2\cos^2\theta_c m_\mu^2 8k_1\cdot k_2 \\
&= G_F^2 f_\pi^2\cos^2\theta_c m_\mu^2 4(m_\pi^2 - m_\mu^2)
\end{aligned}$$

where we have used

$$\sum_{spin} v(k)\bar{v}(k) = \not{k} - m$$

and

$$\sum_{spin} u(k)\bar{u}(k) = \not{k} + m$$

and, finally, $2k_1\cdot k_2 = (k_1+k_2)^2 - k_1^2 - k_2^2$.

Let us now compute the phase-space integral:

$$\begin{aligned}
\rho &= \int (2\pi)^4 \delta^4(p - k_1 - k_2) \frac{d^3k_1}{(2\pi)^3 2\omega_1} \frac{d^3k_2}{(2\pi)^3 2\omega_2} \\
&= \frac{1}{(2\pi)^2} \int \delta(m_\pi - \omega_1 - \omega_2) \frac{d^3k_1}{4\omega_2\omega_1} \\
&= \frac{1}{4\pi} \int \delta\left(m_\pi - \omega_1 - (\omega_1^2 + m_\mu^2)^{1/2}\right) \frac{\omega_1 d\omega_1}{(m_\mu^2 + \omega_1^2)^{1/2}}
\end{aligned}$$

where we have used $\mathbf{k_1}^2 = \omega_1^2$ and $\omega_2 = \sqrt{m_\mu^2 + \mathbf{k_2}^2} = \sqrt{m_\mu^2 + \omega_1^2}$.

Using the well-known formula

$$\delta(f(x)) = \frac{1}{|f'(x^*)|} \delta(x - x^*)$$

we get,

$$\begin{aligned}
\delta(m_\pi - \omega_1 - \omega_2) &= \frac{m_\pi^2 + m_\mu^2}{2m_\pi^2} \delta\left(\omega_1 - \frac{m_\pi^2 - m_\mu^2}{2m_\pi}\right) \\
&= \frac{\omega_2}{m_\pi} \delta\left(\omega_1 - \frac{m_\pi^2 - m_\mu^2}{2m_\pi}\right)
\end{aligned}$$

or

$$\rho = \frac{1}{4\pi} \frac{m_\pi^2 - m_\mu^2}{2m_\pi^2}$$

The final expression for the decay rate is then

$$\Gamma(\pi^+ \to \mu^+ + \nu_\mu) = \frac{G_F^2}{4\pi} f_\pi^2 m_\mu^2 m_\pi \left(1 - \frac{m_\mu^2}{m_\pi^2}\right)^2 \cos\theta_c^2 = \frac{1}{\sigma_\pi}$$

Using the life-time of the pion, the value of Fermi's constant and of the Cabibbo mixing angle, one can deduce that $f_\pi \sim 90 MeV$.

6.3.3 Decay of the top quark into W^+b

On dimensional grounds, we would think that the decay process

$$t \to W^+ + b$$

is

$$\Gamma \sim \frac{g^2}{4\pi} m_t$$

Actually, for $m_t \gg m_W$, the dominant decay channel is into the longitudinal polarization of the W^+, which brings in a factor of m_t/M_W in the decay amplitude. This is an instance of a general phenomenon, known as the “Golstone boson equivalence theorem”, which is a the heart of the Brout-Englert-Higgs mechanism.

The amplitude for this process is

$$\mathcal{M} = \frac{ig}{\sqrt{2}}\bar{u}(q)_L\gamma^\mu u(p)_L\epsilon^*_\mu(k)$$

(we set the CKM factor to 1 and neglect the mass of the bottom), with $p = q + k$. Squaring the amplitude gives, [2]

$$\begin{aligned}
\frac{1}{2}\sum_{spins}|\mathcal{M}|^2 = \frac{1}{2}\frac{g^2}{2}\mathrm{Tr}\left[\bar{u}(q)\gamma^\mu\left(\frac{1-\gamma_5}{2}\right)u(p)\bar{u}(p)\gamma^\nu\left(\frac{1-\gamma_5}{2}\right)u(q)\right]\sum_{pol.}\epsilon^*_\mu(k)\epsilon_\nu(k)\\
&= \frac{1}{2}\frac{g^2}{2}\mathrm{Tr}\left[\not{q}\gamma^\mu\left(\frac{1-\gamma_5}{2}\right)(\not{p}+m_t)\gamma^\nu\left(\frac{1-\gamma_5}{2}\right)\right]\left[-\eta_{\mu\nu}+\frac{k_\mu k_\nu}{m_W^2}\right]\\
&= \frac{g^2}{8}\mathrm{Tr}\left[\not{q}\gamma^\mu\not{p}\gamma^\nu\right]\left[-\eta_{\mu\nu}+\frac{k_\mu k_\nu}{m_W^2}\right]\\
&= \frac{g^2}{2}\left[q^\mu p^\nu + q^\nu p^\mu - \eta^{\mu\nu}q\cdot p\right]\left[-\eta_{\mu\nu}+\frac{k_\mu k_\nu}{m_W^2}\right]\\
&= \frac{g^2}{2}\left[q\cdot p + 2\frac{k\cdot q\; k\cdot p}{m_W^2}\right]
\end{aligned}$$

In the first line, we have sum over the final polarization states of the W^+ and bottom quark and averaged over the initial polarizations of the top. In the second line, we have neglected the mass of the bottom, since $m_b \ll m_t$. In going from the second to the third line, we used the fact that $P_L^2 = P_L$ with $P_L = (1-\gamma_5)/2$. Also, although this is not obvious, the γ_5 term drops (check this) in doing the trace, so I don’t write iit down. Finally I have used the usual expression for the trace over 4 gamma matrices. Cf appendix and QFT course.

Setting the mass of the bottom, $m_b = 0$,

$$2q\cdot p \equiv (p-q)^2 + q^2 + p^2 = 2q\cdot k = m_t^2 - m_W^2, \qquad 2k\cdot p = m_t^2 + m_W^2$$

[2]The sum over the polarizations of the gauge boson can be obtained in various ways. Note first that, on-shell $k^2 = m_W^2$, $k_\mu W^\mu = k_\mu\epsilon^\mu = 0$, as can be seen by taking the divergence of the free equation of motion,

$$\partial_\nu\partial^\nu W^\mu + \partial^\mu(\partial_\nu W^\nu) + m_W^2 W^\mu = 0$$

This constraint is satisfied by the projector

$$\sum_{pol}\epsilon^\mu\epsilon^{*\nu} = -\left(\eta_{\mu\nu} - \frac{k_\mu k_\nu}{m_W^2}\right) \tag{6.6}$$

Incidentally, in the rest frame of the W, it reduces to a projector over the three spatial indices. Alternatively, one can build the sum over polarizations directly: the two transverse polarizations, which satisfy $\epsilon\cdot\mathbf{k}$ can be taken to be $\epsilon_\mu^{(1,2)} = 1/\sqrt{2}(0,1,\pm i,0)$ for left and right circular polarizations (or any linear combinations); the longitudinally polarized state can be obtained by $\epsilon_\mu^{(3)} = (0,0,0,1)$ (which satisfies $\epsilon_\mu k^\mu = 0$ in the W rest frame) along the z axis.

Then

$$\frac{1}{2}\sum_{spins}|\mathcal{M}|^2 = \frac{g^2}{4}m_t^2\left(1-\frac{m_W^2}{m_t^2}\right)\left(2+\frac{m_t^2}{m_W^2}\right)$$

The phase-space factor is the same as in the decay $\pi \to e\bar{\nu}$, with proper renaming. Altogether,

$$\Gamma(t \to Wb) \quad = \quad \frac{g^2}{64\pi}m_t\left(1-\frac{m_W^2}{m_t^2}\right)^2\left(2+\frac{m_t^2}{m_W 2}\right)$$

The term 2 in the last factor is from the emission of the spin ±1 components of the W boson. The enhancement factor of $\frac{m_t^2}{m_W^2}$ comes in because the top can also decay into the longitudinal component of the W-boson. Because this is a spin-0 state, the top quark must flip chirality[3], hence the factor of m_t in the amplitude. Incidentally, if we rewrite the decay rate as

$$\Gamma(t \to Wb) = \frac{g^2}{64\pi}m_t\left(1-\frac{m_W^2}{m_t^2}\right)^2\left(\frac{m_t^2}{m_W^2}+2\right)$$

the last factor is the sum of the longitudinal ($\propto m_t^2/m_W^2$) and of the 2 transverse polarizations of the W^+. This can also be checked by computing the partial decay rate into transverse or longitudinal modes.

Assuming, for the sake of the argument, that $m_t \gg m_W$, the leading term is

$$\Gamma = \frac{g^2}{64\pi}\frac{m_t^3}{m_W^2}$$

Using $m_t = \frac{1}{\sqrt{2}}\lambda_t v$ and $m_W = gv/2$, we get that

$$\Gamma = \frac{g^2}{64\pi}m_t\frac{2\lambda_t^2}{g^2} = \frac{\lambda_t^2}{32\pi}m_t \tag{6.7}$$

This is precisely the decay rate for the top to decay into a bottom and the would-be Goldstone mode ϕ^+ (in the limit in which the top mass is much larger than the other two masses). This is a generic feature of theories with SSB of a gauge symmetry: in the limit of high energy (*i.e.* here $m_t \gg m_W$), the emission or absorption of a longitudinally polarized W^+ can be equivalently seen as the emission or absorption of the would-be Goldstone ϕ^+.

exercise 18 *Compute the decay rate of a top quark into a scalar ϕ (the would-be Goldstone in the discussion above) and a bottom, using*

$$\Delta\mathcal{L} = \lambda_t \bar{b}_L \phi^+ t_R$$

where $\phi^+ = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$. Show that, for $m_t \gg m_\phi$, it is equal to (6.7).

[3]This is educated understanding, as it comes, at the end of the day, from the equivalence theorem. A way is to consider the reverse process of decay of the W into two quarks, one much more massive than the others (that excludes the top, but never mind). Then, it is easy to convince oneself that the decay of the spin zero component of the W clearly requires a spin flip.

6.3.4 Vacuum polarization: effect of m_t

The structure of the SM Lagrangian gives rise to a number of relation between the low-energy parameters of the theory, like

$$e = \sin\theta_W g$$

or

$$m_W^2 = \cos\theta_W^2 m_Z^2$$

These relations hold at three level, but can be corrected by quantum effects. An important feature is that the SM Lagrangian is renormalizable: infinities, which generically arise when computing quantum corrections, can be absorbed by redefining a finite number of parameters (couplings, vev). Once this is done, quantum effects lead to finite corrections, which are thus predictions of the theory. We will now discuss one such correction. Although you don't have practice in computing loops effects in QFT, the present calculation is easy enough that you don't have to master all the subtleties of the renormalization program. In short, we will use a number of tricks to get as fast as possible to the answer.

Consider the relation

$$m_W^2 = \cos\theta_W m_Z^2$$

This relation arises solely because of the symmetry breaking pattern of the SM. In particular, it is protected by the existence of a global SU(2) symmetry in the Higgs sector. This symmetry is however explicitely broken by the Yukawa couplings. As could be expected, the largest correction to this tree level relation will be due to quantum loops with a top quark. These are corrections which can shift the mass of the W and Z boson. Each of these diagram is logarithmically divergent. The building blocks we have to compute are the vacuum polarization tensors of the W and Z bosons,

$$i\Pi_{\mu\nu}^{(W,Z)}$$

Intermezzo

In QED the vacuum polarization tensor is transverse,

$$i\Pi_{\mu\nu}^{(\gamma)} = i(\eta_{\mu\nu} - \frac{q_\mu q_\nu}{q^2})\Pi(q^2)$$

with $\Pi(q^2) \propto q^2$. This is a consequence of gauge invariance and vector current conservation, which imposes that $q^\mu \Pi_{\mu\nu}^{(\gamma)} = q_\nu \Pi_{\mu\nu}^{(\gamma)} = 0$ and prevents the appearance of a mass term for the photon. Indeed, the vacuum polarization tensor enters in the quantum corrected photon propagator as

$$iD_{\mu\nu} = \frac{-i\eta_{\mu\nu}}{q^2} + \frac{-ig_{\mu\alpha}}{q^2} i\Pi^{\alpha\beta} \frac{-ig_{\beta\nu}}{q^2} + \ldots$$

or

$$iD_{\mu\nu} = \frac{-i\eta_{\mu\nu}}{q^2 - \Pi(q^2)} \equiv \frac{-i\eta_{\mu\nu}}{q^2(1 - \Pi'(q^2))}$$

with $\Pi'(q^2) = d\Pi(q^2)/dq^2$, in such a way that the photon pole is at $q^2 = 0$. The effect of the vacuum polarization tensor is to give rise to a logarithmic running of e,

$$e^2(Q^2) = \frac{e^2}{1 - \Pi'(Q^2)} \propto e^2 \log \frac{Q^2}{\mu^2}$$

for $Q^2 \gg \mu^2$, with $Q^2 = -q^2 > 0$. Π' is logarithmically divergent, but this divergence can be absorbed by imposing that $e^2(\mu)/4\pi \approx 1/137$ at some scale μ, for instance $\mu = m_e$. That e is running with energy was spectacularly demonstrated at LEP, which gave $e^2(M_Z^2)/4\pi \approx 1/128$.

Back to W and Z

In the broken phase, the vacuum polarization tensors of the W and Z boson however don't have to vanish at zero momentum and can give rise to a shift of their masses. Defining

$$i\Pi_{\mu\nu} = i\eta_{\mu\nu}\Pi(q^2) - iq_\mu q_\nu \Delta(q^2)$$

We will consider an experimental setting in which the gauge bosons couple only to light external fermions, such as at LEP. Then, the last term gives corrections which are of order $\frac{m_{light}^2}{m_W^2}$ and is thus negligible. To simplify things a bit more, we will compute the first term only in the zero-momentum limit, $i\Pi_{\mu\nu} \to i\eta_{\mu\nu}\Pi(0)$, so that

$$\begin{aligned} M_Z^2 &= (g^2 + g')\frac{v^2}{4} + \Pi_Z(0) \\ M_W^2 &= g^2\frac{v^2}{4} + \Pi_W(0) \end{aligned}$$

Taken independently, these quantities are logarithmically divergent, but the relative mass shift

$$\frac{\Pi_W(0)}{m_W^2} - \frac{\Pi_Z(0)}{m_Z^2}$$

is finite. This quantity appears as a quantum correction to mass relation

$$\rho = \frac{m_W^2}{\cos\theta_W m_Z^2} - 1$$

which is predicted to satisfy $\rho = 0$ at tree-level in the SM.

exercise 19 *Show that*

$$\rho \approx \frac{\Pi_{WW}}{m_W^2} - \frac{\Pi_{ZZ}}{m_Z^2}$$

where $m_W^2 = \cos\theta_W m_Z^2$ *are the tree-level boson masses.*

Calculation of $\Pi_{\mu\nu}$

The Feynman rules for the quark-gauge boson vertices are:

$$-i\frac{e}{\sin\theta_W\sqrt{2}}\gamma^\mu\left(\frac{1-\gamma_5}{2}\right)$$

for the $W^+b\bar{t}$ vertex,

$$i\frac{e}{\cos\theta_W\sin\theta_W}\left(\frac{1}{2}\left(\frac{1-\gamma_5}{2}\right)-\frac{2}{3}\sin\theta_W\right)$$

for the $Z\bar{t}t$ one and, finally,

$$i\frac{e}{\cos\theta_W\sin\theta_W}\left(\frac{-1}{2}\left(\frac{1-\gamma_5}{2}\right)+\frac{1}{3}\sin\theta_W\right)$$

for the $Z\bar{b}b$ vertex.

We compute all the diagrams at once, using the (accordingly ugly) notation

$$i(g_L L+g_V)$$

for the left and vectorial couplings, with $L=(1-\gamma_5)/2$. We also introduce two quark masses, m_1 and m_2 and set $g_{L,V}$ and $m_{1,2}$ to their values only at the end.

Applying the Feynman rules, the structure of the polarization tensor of our gauge bosons is

$$i\Pi_{\mu\nu}=(-)(i)^2N_c\int\frac{d^4}{(2\pi)^4}\,\mathrm{Tr}\left[\frac{i}{\not{k}-m_1}\gamma^\nu(g_L L+g_V)\frac{i}{(\not{k}+\not{q})-m_2}\gamma^\mu(g_L L+g_V)\right]$$

(The $(-)$ sign comes in because of Fermi statistic. $N_c=3$ is the number of colors.)

Consider first the trace. Using

$$\frac{i}{\not{p}-m}=i\frac{\not{p}+m}{p^2-m^2}$$

the properties of the γ_5 matrix, we get three distinct contributions,

$$\Pi_{\mu\nu}=g_L^2\Pi_{\mu\nu}^{LL}+2g_Lg_V\Pi_{\mu\nu}^{LV}+g_V^2\Pi_{\mu\nu}^{VV}$$

The first comes purely from left couplings:

$$\Pi_{\mu\nu}^{LL}=\mathrm{Tr}\left[\not{k}\gamma^\nu(\not{k}+\not{q})\left(\frac{1-\gamma_5}{2}\right)\right] \tag{6.8}$$

The other two arise, respectively, from the mixing between the couplings g_L and g_V, and purely from g_V. They both are of the form

$$\mathrm{Tr}\left[\not{k}\gamma^\nu(\not{k}+\not{q})\gamma^\mu\right]-m_1m_2\mathrm{Tr}\left[\gamma^\mu\gamma^\nu\right]$$

where we have used the fact that the γ_5 part drops out, since no antisymmetric tensor can be build out of only q^μ. Instead of computing these two contribution $\propto g_Lg_V$ or g_V^2, we use what we now from QED. First, these contribution only arise for the boson Z, as the boson W has only couplings to left-handed quarks. Hence $m_1 = m_2$. Then, these contributions to the polarization tensor have the same structure as the polarization tensor of a photon field, which has only vector-like couplings. Hence, we expect these part to vanish at zero momentum, as discussed in the *intermezzo*. If you do things correctly, for instance using a regularization of divergent integrals which preserves gauge invariance, like *dimensional regularization*, this is what you would get. We leave this as an exercise for those who master this technique. At any rate, we only have to compute the part (6.8) of the polarization tensor in the limit of zero external momentum.

We can compute the trace in (6.8) using the formula of Appendix B:

$$\begin{aligned} i\Pi_{\mu\nu}(q) &= -4N_c\frac{g_L^2}{2}\int\frac{d^4k}{(2\pi)^4}\left((k^\nu(k+q)^\mu + k^\mu(k+q)^\nu - k\cdot(k+q)g^{\mu\nu})\right) \\ &\quad\times\frac{1}{(k^2-m_1)^2((k+q)^2-m_2)^2}+\ldots \end{aligned}$$

where the dots represent the terms which are proportional to q^2. In the low momentum limit $q \to 0$, we are thus left with

$$i\Pi_{\mu\nu}(0) = -2N_c g_L^2\int\frac{d^4k}{(2\pi)^4}\left((-1/2)\,\eta_{\mu\nu}k^2\right)\frac{1}{(k^2-m_1^2)(k^2-m_2^2)}$$

where we have replaced the terms in $k_\mu k_\nu$ by $1/4\eta_{\mu\nu}k^2$, since this is the only structure which may arise from the integration over k in the $q^\mu \to 0$ limit.

This integral seems to be badly divergent, but let us go on and put the contributions of the Z and W bosons together. Using the values of g_L from the Z and W couplings, the mass of the quarks top and bottom, and $m_W^2 = \cos^2\theta_W m_Z^2$ and $g = e/\sin\theta_W$, we get

$$\begin{aligned} \frac{\Pi_W(0)}{m_W^2} - \frac{\Pi_Z(0)}{m_Z^2} &= i2N_c\frac{e^2}{2\sin^2\theta_W\cos^2\theta_W m_Z^2}\int\frac{d^4k}{(2\pi)^4}(-1/2)k^2 \\ &\quad\times\left\{\frac{1}{(k^2-m_t^2)(k^2-m_b^2)} - \frac{1}{2}\frac{1}{(k^2-m_t^2)^2} - \frac{1}{2}\frac{1}{(k^2-m_b^2)^2}\right\} \end{aligned}$$

To simplify things a little bit, suppose that $m_b = 0$. Taken independently, all these integrals seems to be quadratically divergent in the ultraviolet. However, expanding the propagators at high k^2,

$$\frac{1}{(k^2-M^2)^2} \approx \frac{1}{k^4}\left(1+\frac{2M^2}{k^2}\right)$$

you can readily check that not only the quadratic divergence, but also a logarithmic divergence cancel! Hence, as announced, the relative shift in the mass of the W and Z bosons is finite in the SM.

To finish our calculation, we just need to compute the integral. After some rearrangement of the terms, we get

$$\begin{aligned} \rho &= A\int\frac{d^4k}{(2\pi)^4}\frac{-1}{2}\left\{\frac{1}{(k^2-m_t^2)} - \frac{1}{2}\frac{k^2}{(k^2-m_t^2)^2} - \frac{1}{2}\frac{1}{k^2}\right\} \\ &= A\int\frac{d^4k}{(2\pi)^4}\left(-\frac{1}{2}\right)^2\frac{m_t^4}{(k^2-m_t^2)^2k^2} \end{aligned}$$

where

$$A = iN_c \frac{e^2}{\sin^2\theta_W \cos^2\theta_W m_Z^2}$$

Thus we are left with the integral

$$I = \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2(k^2 - m_t^2)^2}$$

One last trick is to perform a Wick rotation from Minkowski to euclidian momentum, $k^0 = ik_E^4$, $k^2 = -k_E^2$, so that[4]

$$\begin{aligned} I &= -i \int \frac{d^4k_E}{(2\pi)^4} \frac{1}{k_E^2(k_E^2 + m_t^2)^2} \\ &= -i \int_{S^3} d\Omega \int_0^\infty \frac{dk_E k_E^3}{(2\pi)^4} \frac{1}{k_E^2(k_E^2 + m_t^2)^2} \\ &= -i \frac{2\pi^2}{(2\pi)^4} \int_0^\infty dk_E \frac{k_E}{(k_E^2 + m_t^2)^2} \end{aligned}$$

If you know that $2\pi^2$ is the volume of a 3-sphere S^3 of unit radius, you can finish the work. Your result should be

$$\rho = \frac{e^2}{\sin\theta_W^2 \cos\theta^2 M_Z^2} \frac{3}{16\pi^2} \frac{1}{4} m_t^2$$

From the precision measurement at LEP of m_Z, and from the values of e and $\sin\theta_W$, people could infer that $m_t = 169 \pm 24 GeV$. When top quark was eventually found at Fermilab, we learned that $m_t = 180 \pm 13 GeV$. Now you understand why 't Hooft and Veltman got the big Prize.

6.4 The Higgs particle

The Higgs was discovered in 2012. Its mass is about 125 GeV. Here we just summarize some salient features.

$\mathbf{M_W} = \cos\theta_{\mathbf{W}} \mathbf{M_Z}$

This relation follows from the doublet structure of the Higgs field and is well established experimentally. Most Higgs structures violate this relation. There is an interesting connection between this symmetry and the hidden SU(2)′ symmetry of the Higgs sector. This so called *custodial symmetry* is broken by the Yukawa couplings.In particular, it receives an important correction from virtual loops with top quarks. (See the tutorials.)

[4] You might not be familiar with this. Remember that the propagator are actually defined as $i/k^2 - m^2 + i\epsilon$. These have poles in the complex k_0 plane, situated at $\omega = \sqrt{\mathbf{k^2} + m^2} - i\epsilon'$ and $\omega = -\sqrt{\mathbf{k^2} + m^2} + i\epsilon'$. Hence, the integration over k_0 which runs from $-\infty$ to $+\infty$ can be rotated to run from $-i\infty$ to $+i\infty$ without encountering these poles. In turn, the redefinition of the integration variable $k_0 = ik_E^4$, gives the analytical continuation from Minkowski space-time to Euclidian space, integrals which are easier to evaluate.

exercise 20 *How is the relation*

$$\rho = \frac{M_W^2}{\cos^2\theta_W M_Z^2} - 1$$

modified if, on top of the SM doublet, a scalar field in the adjoint representation of SU(2) develops a vacuum expectation value?

Higgs couplings to fermions

The couplings of the Higgs to fermions is proportional to the fermion mass. In the case of many families, you can check that the diagonalization of the fermion mass matrices also diagonalize the Higgs couplings. This is because the Higgs belongs to a single irreducible representation.

A direct consequence, is that the Higgs couples more easily to heavy fermions. This is of experimental interest, both for Higgs decays and Higgs production.

Higgs couplings to gauge bosons

These couplings are quite large, being proportional to the mass of the gauge bosons,

$$\mathcal{L}_{HVV} = gH(x)\left(M_W W_\mu^+ W^{\mu-} + \frac{1}{2\cos\theta_W} M_Z Z_\mu Z^\mu\right)$$

Higgs mass

The Higgs mass is

$$M_H^2 = -2\mu^2 = 2\lambda v$$

A theoretical prejudice is that $M_H < 350GeV$ from the requirement that $\lambda < 1$. Otherwise the Higgs sector is strongly coupled and our theory of spontaneous symmetry breaking is not under control. More precise estimates, based on unitarity, predicts that the mass should below the TeV scale. The particle was discovered in 2012, and has a mass of 125 GeV, hence substantially lighter than these theoretical constraints.

Chapter 7

Shortcomings and beyond the Standard Model

7.1 Adding neutrinos

In the strict framework of the SM, neutrinos are massless if they only come as left-handed Weyl fermions. However, there is mounting evidence that neutrinos must be massive particles, although very light ones. In particular, the solar and atmospheric neutrinos problems could be neatly solved if neutrinos are massive. The recent result of the SNO experiment have demonstrated that the total flux of neutrinos coming from the sun is consistent with the predictions of the the standard model of the sun. It has also proven that the flux contains a non-vanishing fraction of ν_μ or ν_τ neutrinos, although only ν_e and their anti-particles can be created in the core of the sun, where nuclear reactions take place. If neutrinos are massive, chances are that their interaction and propagation eigenstates are different, pretty much like what happens in the quark sector. Then a phenomenon of neutrino oscillations is possible and could explain the suppression of the fluxes of ν_e from the sun and of ν_μ from the atmosphere.

At first, it seems like a trivial matter to have massive neutrinos in the SM. All we have to do is to add right-handed neutrinos, possibly three of them. Note however that a right-handed neutrino is not charged under SU(2)×U(1), for it is a right-handed particle and furthermore, because

$$Q = \frac{T_L^3}{2} + \frac{Y}{2}$$

requires $Y_{\nu_R} = 0$. This is quite strange, or at least very singular compared to all the known form of matter, which are known to interact according to the gauge principle. Right-handed neutrinos would only interact through Yukawa couplings of the kind

$$\mathcal{L}_{Yukawa} = y_\nu \bar{L}\phi\nu_R + h.c.$$

A further problem is then to understand why the Yukawa couplings of the neutrinos are so small, giving to the neutrinos a mass much much smaller than to the charged counterparts. This is another instance of an *hierarchy problem.*

Since the neutrino carry no electric charge, there is however another way to give a mass to a neutrino. For this, we must give up lepton number conservation. To see this, consider a Dirac spinor

$$\psi = \begin{pmatrix} \chi \\ \omega \end{pmatrix}$$

where χ and ω are respectively Left and Right-handed Weyl spinors. A Lorentz invariant mass term is

$$\Delta\mathcal{L} = m\bar{\psi}\psi = m(\chi^\dagger\omega + \omega^\dagger\chi)$$

as you saw in your course on QED. This is called a Dirac mass term if χ and ω are distinct Weyl spinors. Then ψ is a Dirac spinor, with four complex components. An important feature is that a Dirac mass term is invariant under a *global* symmetry transformation,

$$\psi \to e^{i\alpha}\psi$$

To this symmetry is associated a conserved current,

$$j_\mu = \bar{\psi}\gamma_\mu\psi$$

and the conserved charge

$$Q = \int d^3x\, j^0(t, x)$$

is the fermion number.

There is another possibility. The following term is also Lorentz invariant:

$$\Delta\mathcal{L} = \frac{m}{2}\bar{\psi}^c\psi + h.c. = \frac{m}{2}\left((\chi^c)^\dagger\chi + (\omega^c)^\dagger\omega\right) + h.c.$$

where ψ^c is the charge conjugate of ψ. This is because (exercise) ψ and ψ^c transform the same under Lorentz transformations,

$$\psi^{(c)} \to e^{-\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}}\psi^{(c)}$$

Note that such a mass term couples χ with χ^c and ω with ω^c. If we have only, say, Left-handed particles, we could introduce an object ψ_M, which has twice less degrees of freedom than a Dirac spinor,

$$\psi_M = \begin{pmatrix} \chi \\ \chi^c \end{pmatrix}$$

and which satisfies

$$\psi_M^c = \psi_M$$

This is a Majorana spinor and

$$\Delta\mathcal{L} = m\psi_M^\dagger\psi_m$$

is called a Majorana mass term.[1] A very important feature of a Majorana mass term is that it violates fermion number conservation. Another, equivalent, way to state this, is that a Majorana particle is its own antiparticle.

[1]Note that there is a basis of the Clifford algebra in which a Majorana particle can be written as an object with 4 real components, $\psi_M^* = \psi_M$.

Certainly the electron cannot be a Majorana particle, because it carries an electric charge. However the neutrino is a neutral particle, and a neutrino mass term can be either of the Dirac or Majorana kind. In the former case, there must be right-handed neutrinos and lepton number is conserved. In the latter, lepton number is violated.

Can we write a Yukawa coupling which, after SSB, gives a Majorana mass term to the LH neutrinos of the SM? Because the neutrinos a part of an $SU(2)$ doublet, the simplest Yukawa coupling is

$$\Delta\mathcal{L} = y_\nu \bar{L}^c \Delta L + h.c.$$

where

$$L^c = -i\sigma^2 \begin{pmatrix} \nu_L^c \\ e_L^c \end{pmatrix}$$

to insure proper transformation under both Lorentz and $SU(2)$. Although $\bar{L}^c L$ is an $SU(2)$ singlet, is breaks $U(1)_Y$, hence we must introduce a field Δ which transforms as

$$\Delta \to g\Delta g^\dagger$$

under $SU(2)$ transformations and has hypercharge $Y_\Delta = +2$. In components,

$$\Delta = \begin{pmatrix} \delta^+ & \delta^{++} \\ \delta^0 & -\delta^- \end{pmatrix}$$

Such an object is said to belong to a triplet representation of $SU(2)$. One would have to extend the Higgs sector to give a Majorana mass to the neutrinos, through a *vev* of δ^0. The trouble with this approach is that such a modification of the Higgs structure spoils the mass relation

$$M_W = \cos\theta_W M_Z$$

which is well-established experimentally. Hence $\langle\delta^0\rangle$ cannot be large and we have not made much progress.

There is however an alternative. Consider the following product of Higgs fields:

$$(\tilde{\phi}\phi^\dagger)$$

It transforms precisely as Δ under $SU(2) \times U(1)$. Hence, we can write a gauge and Lorentz coupling as

$$\Delta\mathcal{L} = \frac{1}{\Lambda}\bar{L}^c\tilde{\phi}\phi^\dagger L + h.c.$$

Note that we had to introduce a dimension-full coupling $\frac{1}{\Lambda}$, because $\tilde{\phi}\phi^\dagger$ is an operator of dimension 2. Altogether this is an operator of dimension 5. Hence it spoils the renormalizability of the SM. On the other hand, it goes to zero as $\Lambda \to \infty$. The modern point of view on this question is that, such an effective coupling will or may arise at a scale at which the SM as a theory breaks down. At low energies, these higher dimensional operators are said to be *irrelevant*, *i.e.* they don't affect much the low energy predictions. As such, the new scale Λ is to be seen as a cut-off beyond which the SM has to be modified. The beauty of this argument is that we can understand right-away why the neutrinos are so light. This is simply because, by assumption, $v \ll \Lambda$, so that

$m_\nu \sim v^2/\lambda$ can be very small. In practice, Λ is expected to be of the order $10^{10} GeV$, to account for the smallness of the neutrino masses. Furthermore, this argument also sets the scale for lepton number violating processes.[2] It is so beautiful that it must be true. However, at the time of this writing, we still don't know if neutrinos are of the Majorana or Dirac kind.

7.2 Grand Unification

In the SM, there is a partial unification of the weak and electromagnetic interactions. The unification is partial because there are two independent coupling g and g'. The idea of Grand Unification is to embed the gauge group and matter fields in representation of a larger gauge group.The simplest solution is $SU(5)$. It has rank 4, like the SM, and the matter fields can be all fitted into a $\bar{\mathbf{5}}$ fundamental representation and a 5×5 antisymmetric matrix (hence with 10 components) the **10**. Altogether, these are 15 degrees of freedom, just like in one generation of the SM (3×4 quarks and 3 leptons).

The most striking consequences of this scheme are:

- There is just one coupling constant:

$$g_5 = g_c = g = \sqrt{\frac{5}{3}} g'$$

  This relation is not quite satisfied in the SM, but couplings run logarithmically with energy, and this relation is only supposed to hold at energies where the $SU(5)$ symmetry is manifest. In the minimal $SU(5)$ model, they almost unify at $10^{16} GeV$. In a supersymmetric version, the unification is (too ?) perfect.

- The electric charge is quantized. This is because the electric (or hypercharge) is now a generator (in the Cartan sub-algebra) of $SU(5)$. Then, as for any generator,

$$TrY = 0$$

  in some given representation, and the charge of the quarks and leptons are all quantized in unit of e (or g').[3]

- Baryon number is not conserved: This is because quarks and leptons are parts of the same multiplets. Then transitions which transform quarks into leptons are possible. A dramatic consequence is that the proton becomes unstable. Experimental limits on the proton lifetime impose $M_{GUT} \gtrsim 10^{16} GeV$ and rule out the minimal $SU(5)$ model.

There are other predictions, like the existence of magnetic monopoles, and, also, many shortcomings but we don't have time to discuss these aspects here.

[2] Lepton number violation might have played a crucial role in the generation of the asymmetry between matter and antimatter in the Universe.

[3] At the level of the SM, consistency of the theory at the quantum level requires the cancellation of *chiral gauge anomalies*. The condition is precisely $TrQ = 0$. This can be seen as a strong hint for Grand Unification.

An even nicer scheme is $SO(10)$. The spinorial representation of $SO(10)$ has 16 chiral degrees of freedom. This permits to accommodate all the fermions of one generation plus, remarkably, a right-handed neutrino. After SSB, these GUT's give rise to lepton number violation and Majorana masses.

7.3 Yet another hierarchy problem and how to solve it

Consider the Higgs boson. Its mass is suppose to be below $1TeV$, for the theory to be perturbative, $\lambda \ll 1$. Like any other particles in the SM, the Higgs mass may suffer quantum corrections. One particular contribution comes from the Higgs sector itself with the Higgs particle in a loop. This amplitude gives a correction to the Higgs mass which is

$$\delta m_H^2 \sim \lambda \int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2}$$

The problem with this diagram is that it is (at least superficially) quadratically divergent

$$\delta m_H^2 \propto \lambda \Lambda^2$$

Remember that the photon is protected from such correction because of gauge invariance,

$$\delta m_\gamma \propto e^2 q^2 \log \Lambda^2$$

Likewise, a fermion is protected from getting large correction because of chiral symmetry. For instance, corrections to the electron mass in QED are of the form

$$\delta m_e \propto e^2 m_e \log \Lambda^2$$

i.e. it is only logarithmically sensitive to the scale Λ. This is because the quantum correction vanishes if $m_e \to 0$, in which case there is an extra chiral symmetry which protects the electron mass.

In the SM, no such symmetry protects the mass of the Higgs boson. This is embarassing. Of course, we can renormalize quantum corrections by redefining the parameters of the Higgs sector, but, at the end of the day, there must be a natural cut-off Λ. What this quantum correction is telling us is that we should expect

$$m_H^2 \sim \Lambda^2$$

which, *a priori*, is much much larger than $1TeV$, perhaps of order of the Planck scale.

There are two solutions to this problem that I want to mention, because there are both of much experimental interest. The first idea is

Supersymmetry

The idea of supersymmetry is that, for each boson in Nature, there is a corresponding fermion, with the same number of degrees of freedom (on-shell). Also, supersymmetry relates their couplings. In a manifestly supersymmetric version of the SM, the quantum correction to the Higgs mass is typically of the form

$$\delta m_H^2 \propto \int d^4k \frac{i}{k^2 - m_s^2} - \int d^4k \frac{i}{k^2 - m_f^2}$$

where m_s is the mass of boson and m_f is the mass of its supersymmetric partner. The crucial point is that a fermion loop comes in with an extra *minus sign.* If supersymmetry is unbroken, $m_s = m_f$ and the quadratic divergence cancels. Hence, supersymmetry plays a role analog of gauge symmetry for gauge bosons or chiral symmetry for fermions. In Nature, certainly supersymmetry must be broken, or there would be a massless photino, companion of the photon, etc. Hence $m_s \neq m_f$. Nevertheless, the quadratic divergence cancels and

$$\delta m_H^2 \propto (m_s^2 - m_f^2) \log \Lambda^2$$

The splitting between m_s and m_f is of order of the scale of supersymmetry breaking. Hence, the mass of the Higgs should not be much different from this scale either. This is the strongest motivation for supersymmetry and also the reason why we may expect to discover supersymmetric partners of the SM particles at the LHC.

The other plausible explanation is to have

Large Extra Dimensions

The philosophy is strikingly different. The "absolute" cut-off Λ is of the order of a few TeV. Hence the Higgs can be rather light, despite quantum corrections. On the other hand the Planck scale is not fundamental, but related to Λ. The way this proposal works is that all the known matter fields and interaction, with the exception of gravity are confined to live in $D = 3 + 1$ dimensions. Gravity on the other can escape to other, compact spatial dimensions, say d of them. Then, at short distances, that is distances small compared to the radius of the extra compact dimensions, Newton's interaction between two massive bodies of masses m_1 and m_2 is not falling in $1/r^2$ but rather like

$$F \sim \frac{m_1 m_2}{M^{2+d} r^{2+d}}$$

If $d = 0$, we recover Newton's law and $M = M_{Planck} \sim 10^{19} GeV$. If $d \neq 0$, *i.e.* if there are extra dimensions, then at long distances compared to the characteristic radius r_c of these extra dimensions, $r \gg r_c$, we recover a law in $1/r^2$,

$$F_{eff} \sim \frac{m_1 m_2}{(M^{2+d} r_c^d)} \times \frac{1}{r^2}$$

with the combination

$$M^{2+d} r_c^d = M_{Planck}$$

to be identified with the Planck scale. Hence, in this picture, the Planck scale is not fundamental but related to M, r_c and the number of extra spatial dimensions. If we set $M \sim TeV$, you can show that

exercise 21 $r_c \sim mm$ *for* $d = 2$.

This is not (totally) in conflict with observation, since, although we have explored parts of the world to very small scales using accelerators, we don't know much about gravity at short distances (yet).

So there could be quite large extra dimensions out there. This picture is not without problems but it has opened new avenues for research.

Review articles

G. Altarelli, *The standard electroweak theory and beyond*, hep-ph/9811456.
Idem.

Appendix A

Natural units

The SI is not very convenient to discuss quantitive aspects of high energy physics processes. Times are very short compared to the second, mass very small compared to the kg, etc. Instead, we use the natural units system (NUS). They are called natural units because they use fundamental constants to eliminate redundancies among SI units. They are thus natural also because they are more physical.

In NUS, all units are expressed in units of energy. A convenient unit for energy is the eV (electron-volt) which is the energy of a particle of charge $q_e = 1.602 \cdot 10^{-19}$ C (coulomb) in a potential of 1 V (volt). So 1eV $= 1.602\,10^{-19}$ J (joule). We can express time in terms of length using the speed of light in vacuum which we set to one,

$$\mathrm{c} = 1$$

Thus $1s \equiv 2.998 \cdot 10^8$ m (3 is a very good approximation). The equivalent is a shortcut for $1\mathrm{s}{\cdot}\mathrm{c} = 3{\cdot}10^8$m. We can then express mass in terms of energy: $1\mathrm{eV} = 1.602{\cdot}10^{-19}\,\mathrm{J} \equiv 1.782{\cdot}10^{-36}$ kg (ie $1\mathrm{eV/c}^2 = \ldots$ kg).

Similarly, we can express a length in terms of energy setting

$$\hbar = 1$$

Since $1\hbar = 1.055 \cdot 10^{-34}$ kg m^2/s $\equiv 3.519 \cdot 10^{-43}$ kg $\cdot$ m $\equiv 1.974 \cdot 10^{-7}\mathrm{eV} \cdot m \approx 200\,\mathrm{MeV} \cdot \mathrm{fm}$.

For cosmological applications, it is useful to set $k_B = 1$, so that temperatures too are expressed in GeV too. In the SI, $k_B = 8.617 \cdot 10^{-5}\mathrm{eV}/K$.

Doing so, we get the following conversion table:

1 J	$6.242 \cdot 10^{11}$ GeV	1 GeV	$1.602 \cdot 10^{-10}$ J
$1\,\mathrm{kg} \cdot \mathrm{c}^2$	$5.608 \cdot 10^{26}$ GeV	1 GeV/c^2	$1.783 \cdot 10^{-27}$ kg
$1\,\mathrm{s}/\hbar$	$1.519 \cdot 10^{24}$ GeV^{-1}	$1\,\hbar$/ GeV	$6.582 \cdot 10^{-25}$ s
$1\,\mathrm{m}/\hbar c$	$5.068 \cdot 10^{15}$ GeV^{-1}	$1\,\hbar$c/ GeV	$1.973 \cdot 10^{-16}$ m
$1\,k_B$K	$8.617 \cdot 10^{-14}$ GeV	1 GeV/k_B	$1.160 \cdot 10^{13}$ K

Notice that the speed of light is no longer measured but (since 1983) is defined to be

$$c = 299\,792\,458 \text{ m/s}$$

This means that the meter is the distance travelled by light in 1/299792458 of a second. The second is 9192631770 transitions of caesium-133 between its ground state and its hyperfine excited state.

A.1 Applications

For conversions, a useful, easy to remember approximation is

$$\hbar\, c \approx 200\,\mathrm{MeV} \cdot 10^{-15}\mathrm{m}.$$

meaning that 1fm is about a fifth of a GeV. This is about the size of a nucleon.

Consider the area (cross section)

$$\sigma = \frac{(\hbar c)^2}{1 GeV^2} \approx 4\,10^{-32}\,\mathrm{m}^2 = 0.4\,\mathrm{mbarn}$$

where 1 barn = $10^{-28}\mathrm{m}^2$. A look at the Particle Data Book shows that cross-sections characteristic of strong interactions are of order few mbarn.

Concerning lifetimes, consider two particles with similar mass. The Δ(1.2 GeV) baryon has a lifetime of about $10^{-24}s$ (width $\Gamma \sim 100\,\mathrm{MeV}$), characteristic of a decay driven by strong interactions. The Λ baryon, of similar mass $\sim$ 1.1 GeV on the other hand has a lifetime of $3\,10^{-10}s$, longer than about 14 orders of magnitude, characteristic of a weak decay!

The Newton constant of gravitation[1] is $G_N \approx 6.674 \times 10^{-11}\mathrm{m}^3\mathrm{kg}^{-1}\mathrm{s}^{-2}$. Using [c] = $\mathrm{m} \cdot \mathrm{s}^{-1}$ and $[\hbar] = J \cdot \mathrm{s} = \mathrm{kg} \cdot \mathrm{m}^2 \cdot \mathrm{s}^{-1}$, we see that

$$1\,\frac{G_N}{\hbar \mathrm{c}} = 2.110 \cdot 10^{15}\,\mathrm{kg}^{-2} = \frac{1}{\mathrm{m}_{\mathrm{pl}}^2}$$

The quantity $\mathrm{m}_{\mathrm{pl}} = 2.176 \cdot 10^{-8}\,\mathrm{kg} = 2.221 \cdot 10^{-19}$ GeV is called the Planck mass. In geodynamics units, this mass is set to

$$G_N = 1$$

meaning that all masses are expressed in units of the Planck mass. Also, it relates mass to distance, through

$$1\,\mathrm{kg}\,(G_N/\mathrm{c}^2) = 7.426 \times 10^{-11}\mathrm{m}$$

which is (twice) the Schwarschild radius of a black hole of mass 1 kg. In this system of units, we may then state that

$$1\,\hbar\,(G_N/c^3) = 2.613 \cdot 10^{-70}\mathrm{m}^2$$

is a Planck size ($\sim$ Planck mass/momentum/energy $\times$ Planck distance) cell in phase space, presumably the minimum possible phase-space area.

Here is the summary of Planck units

[1]What gravity has to do in a course on weak interactions? For one thing, field theories are plagued with infinities. If we have a renormalizable theory, like the SM, we can hide them away but the prejudice is that a more fundamental theory (quantum + gravity) should cut-off infinities. Furthermore, as we will discuss in the last lecture, it is possible there are other (large) dimensions. In that case, the observed value of m_{pl} is not really fundamental and gravity could become strong at much lower energies, as low as a few TeV.

1 $\sqrt{\hbar c/G_N}$	$m_{\rm pl}$	$2.176 \cdot 10^{-8}$ kg	$1.221 \cdot 10^{19}$ GeV
1 $\sqrt{\hbar G_N/c^3}$	$l_{\rm pl}$	$1.616 \cdot 10^{-35}$ m	$8.190 \cdot 10^{-20}$ GeV^{-1}
1 $\sqrt{\hbar G_N/c^5}$	$t_{\rm pl}$	$5.390 \cdot 10^{-44}$ s	—”—

Appendix B

Special Relativity – notations

This appendix gives a quick review of the principles of special relativity. I follow the notations of Jackson's book on electromagnetism, which are also that followed by P.Tinyakov. They however differ from the course of complements of mathematics by M. Henneaux. We write the Minkowski metric tensor in an inertial frame and cartesian spatial coordinates as

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{B.1}$$

with $\mu, \nu = 0, 1, 2, 3$ corresponding to t, x, y, z. These coordinates are used to parametrize events (points) in (flat) space-time also called Minkowski space-time (Minkowski for short). A choice of coordinates corresponds to a choice of frame of reference. Following the tradition sets by Einstein, we will use the word 'observer' to refer to an inertial frame of reference. Like in galilean relativity, inertial observers feel no acceleration and thus move at constant velocity with respect to each others. Unlike in euclidean space however, we have to be careful with the position of the indices, so let us fix our convention. In a given frame, the position of an event will be given by the components of a 4-vector (contravariant vector)

$$x^\mu = (x^0, \mathbf{x}) \equiv (ct, \mathbf{x}) = (ct, x, y, z) \tag{B.2}$$

where c is the speed of light. It is convenient to set $c = 1$ (natural units), so that time and space coordinates have the same dimension. Since the speed of light is the same for all observers this is also physically meaningful.

The metric allows to define an interval s in Minkowski

$$\begin{aligned} s^2 &= \eta_{\mu\nu} x^\mu x^\nu \text{ (components, summation over repeated indices)} \\ &\equiv x \cdot x \quad \text{(a convenient notation for a scalar product)} \\ &\equiv x^2 \quad \text{(another convenient notation)} \end{aligned} \tag{B.3}$$

or explicitly

$$s^2 = (x^0)^2 - \mathbf{x}^2 = t^2 - \mathbf{x}^2 \tag{B.4}$$

Just like the euclidean distance, which is invariant under rotations, intervals are invariant (i.e. an interval is a scalar) under Lorentz transformations, which are isometries of Minkowski that preserve the origin of coordinates. Unlike the euclidean notion of distance, because of the signature of the metric, s^2 is not a definite positive quantity. A vector such that the interval $s^2 > 0$ is said to be time-like, $s^2 < 0$ for space-like vectors, and $s^2 = 0$ for null vectors. The purpose of the metric is of course broader as it allows to define the scalar product of any two 4-vectors,

$$x \cdot y = \eta_{\mu\nu} x^\mu y^\nu = y \cdot x, \tag{B.5}$$

and also affine quantities, like for instance the interval between any two events P and P',

$$\Delta s^2 = (x_P - x_{P'})^2 = (x_{P'} - x_P)^2 \tag{B.6}$$

There is a correspondance between particle trajectories (worldlines) in Minkwoski and the sign of s^2. Free, classical inertial particles follow straight-lines. These are such that $s^2 = 0$ for massless particles (like the photon, the particle of light), which define the so-called light-cone. Massive particles propagate along trajectories such that $s^2 > 0$. Causality prevents (classical) particles from propagating along $s^2 < 0$ trajectories. For time-like particle, one can define the length of the worldline of a material point by integrating along infinitesimal intervals ds,

$$\tau = \int ds = \int \sqrt{\eta_{\mu\nu} dx^\mu dx^\nu} \tag{B.7}$$

The quantity τ is called the proper time. For a linear trajectory (*ie* constant velocity v) passing through the origin of the system of coordinates

$$\tau = \int \sqrt{1 - \mathbf{v}^2} dt = \sqrt{1 - \mathbf{v}^2}\, t \tag{B.8}$$

(origin of proper time when the object crosses the origin), which is the familiar relation between proper time (clock of the moving object) and coordinate time (time of the observer), also called time dilation (*ie* $t > \tau$).

Formally the metric establishes a correspondence between a 4-vector with an upper index and one with a lower index (**covariant vectors**) through

$$x^2 = x_\mu x^\mu \quad \text{with} \quad \boxed{x_\mu = \eta_{\mu\nu} x^\nu = (t, -\mathbf{x})} \tag{B.9}$$

Equivalently

$$x \cdot y = \eta_{\mu\nu} x^\nu y^\mu = x_\mu y^\mu = y_\mu x^\mu \tag{B.10}$$

An instance of a covariant vector is the derivative operator (gradient in Minkowski spacetime) which is naturally written with a lower index

$$\boxed{\partial_\mu = \frac{\partial}{\partial x^\mu} = (\partial_t, \partial_x, \partial_y, \partial_z) = (\partial_t, \nabla)} \tag{B.11}$$

so that

$$\partial_\mu x^\nu = \delta^\nu_\mu \quad \text{and} \quad \partial_\mu x^\mu = 4 \tag{B.12}$$

Then

$$\boxed{\partial^\mu = \eta^{\mu\nu}\frac{\partial}{\partial x^\nu} = (\partial_t, -\partial_x, -\partial_y, -\partial_z) = (\partial_t, -\nabla)} \tag{B.13}$$

Conversely, the inverse metric η^{-1}, defined in components with upper indices, $\eta^{\alpha\mu}\eta_{\mu\beta} = \delta^\alpha_\beta$, where δ^α_β (Kronecker symbol) are the components of the unit matrix $\mathbb{I}_4$ in 4-dimension, can be used to raise indices

$$x^\mu = \eta^{\mu\nu}x_\nu. \tag{B.14}$$

Clearly the matrix η^{-1} is equal to η in components, Eq.(B.1).

Energy and momentum are also components of a 4-vector. For an object of mass m, we may make use of the notion of proper time to define[1]

$$p^\mu \hat{=} m\frac{dx^\mu}{d\tau} = \left(m\frac{dt}{d\tau}, m\frac{d\mathbf{x}}{d\tau}\right) = (E, \mathbf{p}) \tag{B.15}$$

In the particle's rest-frame, $p^\mu = (m, \mathbf{0})$, while a particle moving at velocity $\mathbf{v} = d\mathbf{x}/dt$ with respect to an observer O has

$$(E, \mathbf{p}) = \left(\frac{m}{\sqrt{1-\mathbf{v}^2}}, \frac{m\mathbf{v}}{\sqrt{1-\mathbf{v}^2}}\right) \tag{B.16}$$

using the time-dilation relation $d\tau = \sqrt{1-\mathbf{v}^2}dt$. Finally, we have

$$p^2 \equiv E^2 - \mathbf{p}^2 = m^2 \longrightarrow E = \sqrt{\mathbf{p}^2 + m^2} \tag{B.17}$$

For a massless particle (ie a photon) there seems to be a problem since $ds^2 = 0$ so we need another parameter, say λ (an affine parameter), to compute the derivative along a photon worldline,

$$p^\mu = \frac{dx^\mu}{d\lambda} \tag{B.18}$$

In practice this is quite trivial since we may always write (to the price of giving up manifest Lorentz covariance)

$$\frac{d}{d\lambda} = \frac{dt}{d\lambda}\frac{d}{dt} = E\frac{d}{dt} \tag{B.19}$$

and

$$\mathbf{p} = \frac{d\mathbf{x}}{d\lambda} = E\frac{d\mathbf{x}}{dt} \rightarrow \frac{|\mathbf{p}|}{E} = 1 \tag{B.20}$$

with $d\lambda$ a scalar as it should be (note that under a boost $dt' = \sqrt{(1-v)/(1+v)}dt$ (Doppler effect) and $E' = \sqrt{(1-v)/(1+v)}E$, note also that $d\lambda = dt/E \equiv d\tau/m$ for a massive particle). Regardless, by construction

$$p^2 = 0 \longrightarrow E^2 = \mathbf{p}^2 \tag{B.21}$$

[1] Although we will have little use of this notation, we may also write

$$p^\mu = mu^\mu$$

with $u^2 = 1$, is called the 4-velocity. An interesting point is the following. Consider an observer A with 4-velocity u_A, that observes a particle with 4-momentum P. Then $u_A \cdot P = E_A$ is the energy of the particle as measured by A (see also exercises).

for a particle moving along a light-cone, *ie* a massless particle.

Notice that these definitions are consistent with the correspondence between energy/momentum and the derivative operator from QM,

$$E \leftrightarrow i\partial_t \quad \text{and} \quad \mathbf{p} \leftrightarrow -i\nabla, \tag{B.22}$$

that motivates to define (watch the signs)

$$\boxed{p^\mu = (E, \mathbf{p}) \leftrightarrow i\partial^\mu} \tag{B.23}$$

and

$$\boxed{p_\mu = (E, -\mathbf{p}) \leftrightarrow i\partial_\mu} \tag{B.24}$$

Appendix C

Collisions kinematics

We consider a $2 \to 2$ elastic process $p_1 + p_2 = p'_1 + p'_2$ with $p_i^{(\prime)}$ the 4-momenta, eg $p_1^2 = p_1'^2 = m_1^2$, etc.

In the CM frame, we have

$$p_1 \cdot p'_1 = E_{1,c}E'_{1,c} - \vec{p}_{1,c} \cdot p'_{1,c} = E_{1,c}^2 - p_c^2 \cos\chi = p_c^2(1 - \cos\chi) + m_1^2$$

from energy conservation.

Our goal is to express E'_1 and E'_2 in the lab frame (which we take to be the initial rest frame of particle 2, with $E_2 = m_2$), as a function of E_1. It is convenient to use the scattering angle in the CM frame, χ.

Consider the following identities:

$$p_1 \cdot p_2 = \frac{1}{2}(\,(p_1 + p_2)^2 - m_1^2 - m_2^2) = \frac{1}{2}(\,(p'_1 + p'_2)^2 - m_1^2 - m_2^2) = p'_1 \cdot p'_2 = E_1 m_2$$

where for the last identity, we took the initial rest frame of particle 2. Similarly,

$$p'_1 \cdot p_2 = -\frac{1}{2}(\,(p'_1 - p_2)^2 + m_1^2 + m_2^2) = \frac{1}{2}(\,(p_2 - p'_2)^2 - m_1^2 - m_2^2) = p_1 \cdot p'_2 = E'_1 m_2$$

Finally

$$E'_1 m_2 = p_1 \cdot p'_2 = p_1 \cdot (-p'_1 + p_1 + p_2) = -p_1 \cdot p'_1 + m_1^2 + E_1 m_2$$

gives

$$E_1 - E'_1 = \frac{p_c^2}{m_2}(1 - \cos\chi)$$

The next step is to express p_c^2 in terms of E_1. Take

$$p_1 \cdot p_2 = E_{1,c}E_{2,c} - \vec{p}_1 \cdot \vec{p}_2 = \sqrt{p_c^2 + m_1^2}\,\sqrt{p_c^2 + m_2^2} + p_c^2 = E_1 m_2$$

which gives

$$p_c^2 = \frac{m_2^2(E_1^2 - m_1^2)}{m_1^2 + m_2^2 + 2m_2 E_1}$$

so, finally,

$$E_1' = E_1 - \frac{m_2(E_1^2 - m_1^2)}{m_1^2 + m_2^2 + 2m_2E_1}(1 - \cos\chi)$$

and

$$E_2' = m_2 + \frac{m_2(E_1^2 - m_1^2)}{m_1^2 + m_2^2 + 2m_2E_1}(1 - \cos\chi)$$

where the second terms represent the energy lost (gained) by particle 1 (resp. 2).

As an application, consider the scattering of a neutrino ($m_1 = 0$) onto an electron. In the limit $E_1 \gg m_e$, we get from the expression for E_1' that

$$1 - \frac{E_1'}{E_1} \approx \frac{1 - \cos\chi}{2}$$

so that maximal energy loss of the neutrino corresponds to back scattering, $\chi = \pi$.

Appendix D

Distributions and Fourier transform

Heaviside distribution

The Heaviside or Theta function (or better distribution) is defined by

$$\theta(\tau) = \begin{cases} 1 & \tau \geq 0 \\ 0 & \tau < 0 \end{cases} \tag{D.1}$$

In calculations, we often meet the following integral representation of $\theta(\tau)$:

$$\boxed{\theta(\tau) = \lim_{\epsilon \to 0^+} \frac{-1}{2\pi i} \int_{-\infty}^{+\infty} d\omega \, \frac{e^{-i\omega\tau}}{\omega + i\epsilon}} \tag{D.2}$$

Indeed, the integrand has a pôle in the lower plane $\omega = -i\epsilon$, so that $\tau > 0$, we may evaluate the integral by deforming the contour around this pôle, and get 1, while for $\tau < 0$, we close the contour in the upper ω plane, and get 0.

Since this is a step function, the derivative of θ gives the Dirac delta:

$$\frac{d\theta(\tau)}{d\tau} = \delta(\tau) \tag{D.3}$$

Hence

$$\delta(\tau) = \lim_{\epsilon \to 0^+} \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega \, \frac{\omega}{\omega + i\epsilon} \; e^{-i\omega\tau} \tag{D.4}$$

or

$$\boxed{\delta(\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega \, e^{-i\omega\tau}} \tag{D.5}$$

Also,

$$\lim_{\epsilon \to 0^+} \frac{i}{\omega + i\,\epsilon} = \int_{-\infty}^{+\infty} d\tau \, \theta(\tau) e^{i\omega\tau} = \int_0^{\infty} d\tau \, e^{i\omega\tau} \tag{D.6}$$

is the Fourier transform of $\theta(\tau)$. The latter expression is equal to i/ω for $\omega \neq 0$. Indeed, for $\omega > 0$, we may deform the contour from 0 to $+\infty$ into 0 to $e^{i\phi}\infty$ (with $\phi > 0$) in the upper τ complex plane, to get

$$\int_0^{\infty} d\tau \, e^{i\omega\tau} = \int_0^{\infty} d\rho e^{i\phi} e^{i\omega\rho e^{i\phi}} = \frac{i}{\omega}. \tag{D.7}$$

For $\omega < 0$, we take $\phi < 0$, with the same answer. For $\omega = 0$, we see from (D.5) we see that the integral is $\pi\delta(\omega)$. Hence

$$\lim_{\epsilon\to 0^+} \frac{i}{\omega + i\epsilon} = P\frac{i}{\omega} + \pi\delta(\omega) \tag{D.8}$$

where P stands for the principal part of i/ω (*i.e.* $= i/\omega$ for all ω but $\omega = 0$).

A very useful generalization of this formula is

$$\boxed{\lim_{\epsilon\to 0^+} \frac{1}{\omega - \omega_0 + i\epsilon} = P\frac{1}{\omega - \omega_0} - i\pi\delta(\omega - \omega_0)} \tag{D.9}$$

Appendix E

Perturbation theory and Fermi's Golden rule

E.1 Time-dependent perturbation theory

In this appendix, we review the standard framework of time-dependent perturbation theory in quantum mechanics. It makes the connection with QM more transparent. However, the modern approach rest on the S-matrix and the direct use of covariant amplitudes, based on Feynman diagrams. The latter approach is more powerful.

Concretally, we will interested in evaluating matrix elements like

$$\langle f|H_{int}|i\rangle \tag{E.1}$$

where H_{int} is an interaction Hamiltonian, that we will treat as a small perturbation. This perturbation may be constant, or have some time dependence. We start in the Schroedinger picture and study the evolution of a state

$$|\psi(t)\rangle \tag{E.2}$$

under the effect of interactions.

The full Hamiltonian is decomposed a free part Hamiltonion, H_0, a problem we assume we may solve, with eigenstates

$$H_0|n\rangle = E_n|n\rangle \tag{E.3}$$

and an interaction part

$$H = H_0 + H_{\text{int}}. \tag{E.4}$$

The sum above is a shorthand for either a sum over discret states, or an integral over a continuum of states.

The state $|\psi(t)\rangle$ can now be expressed in the basis of states $|n\rangle$. We write this as

$$|\psi(t)\rangle = \sum_n c_n(t)e^{-iE_nt}|n\rangle. \tag{E.5}$$

Then, the time evolution of $\psi(t)\rangle$ is given by

$$i\partial_t|\psi(t)\rangle = (H_0 + H_{\rm int})|\psi(t)\rangle = \sum_n (E_n c_n(t) + i\dot{c}_n(t))e^{-iE_n t}|n\rangle. \tag{E.6}$$

The first term on the RHS is just free evolution, so we may rewrite this equation as an equation for the evolution of the coefficients $c_n(t)$,

$$i\dot{c}_m(t) = \sum_n e^{i(E_m - E_n)t} c_n(t)\langle m|H_{\rm int}|n\rangle \tag{E.7}$$

where we have used $\langle m|n\rangle = \delta_{n,m}$.

This is still a complicated equation, so we assume that the effect of $H_{\rm int}$ is small, and may treated as perturbation $\mathcal{O}(\lambda)$ where λ is a small parameter. Accordingly, we write

$$c_n(t) = c_n^{(0)}(t) + c_n^{(1)}(t) + c_n^{(2)}(t) + \ldots \tag{E.8}$$

with $c_n^{(k)} = \mathcal{O}(\lambda^k)$. To leading order (zeroth order) we have

$$\dot{c}_m^{(0)} = 0 \quad \to \quad c_n^{(0)} = \text{constant} \tag{E.9}$$

Interactions to first order:

To first order in λ we have in full generality

$$i\dot{c}_m^{(1)}(t) = \sum_n c_n^{(0)} e^{i(E_m - E_n)t}\langle m|H_{\rm int}|n\rangle \tag{E.10}$$

Now, we specifically choose $|\psi(t)\rangle$ to be initially (way before the interaction is active) in a state $|i\rangle$, with energy E_i, so that $c_n(0) = \delta_{i,n}$. We then write $\mathcal{A}_{fi}$ the amplitude $c_f^{(1)}(t)$ to be in the state $|f\rangle$ at time t, starting from the state $|i\rangle$. Explicitely,

$$\boxed{\mathcal{A}_{fi}^{(1)}(t) = -i\int^t dt' e^{i(E_f - E_i)t'}\langle f|H_{\rm int}|i\rangle} \tag{E.11}$$

From this expression, we may consider different possible physical problems. For instance a typical problem in NR QM is the evolution of a state in presence of a periodic perturbation, $H_{\rm int} = V(t) \propto \sin(\omega t)$. As you may know, in such a case, transitions are only allowed to states with $E_f \approx E_i \pm \omega$ (equally holds if the perturbation is active for a large enough time). Here we will be more interested in transitions between different particle states, typically a collision, or a desintegration. Typically the interaction Hamiltonian is time-independent (in the Schroedinger picture), although, for practical and physical reasons, it is usually assumed that the interactions are negligible in the past and in the future. This may seem reasonable for collisions, as it amounts to assume that the particles do not interact when they are at large distances, *ie* they are asymptotically free. However for decay it is not clear what we mean. Also a particle may interact with itself and it is not clear what it means by isolating a particle from its own interactions... Notice that this is why were are a bit loose, as we have not specified the lower bound in the integral.

There is a systematic way of dealing with all this, but at the end it all rests in the fact that we can solve exactly free propagation, and may only (in most non-trivial cases) treat interactions as perturbations. So we consider an Hamiltonian interaction that may switched on and off adiabatically at large times (past and future). By adiabatic we mean that the evolution of the interaction hamiltonian is so slow that it may be taken to be essentially constant (!) for most of the range of integration, still taking it to be zero at infinity. Physically this ensures conservation of energy between the initial and final states, $E_f = E_i$. Indeed, taking the limit of Eq.(E.11) from $-\infty$ and $+\infty$ for $H_{\rm int}$ time independent, we get

$$\boxed{\mathcal{A}_{fi}^{(1)} = -i(2\pi)\delta(E_f - E_i)\langle f|H_{\rm int}|i\rangle} \tag{E.12}$$

Before discussing further this result, it is interesting to rewrite Eq.(E.11) in a different way:

$$\begin{aligned}\mathcal{A}_{fi}^{(1)} &= -i\int^t dt'\langle m|e^{iE_m t'}H_{\rm int}e^{-iE_i t'}|i\rangle = -i\int^t dt'\langle m|e^{iH_0 t'}H_{\rm int}e^{-iH_0 t'}|i\rangle \\ &= -i\int^t dt'\langle m|H_I(t')|i\rangle\end{aligned} \tag{E.13}$$

where it is natural to introduce the so-called Dirac or interaction picture operator

$$H_I(t) = e^{iH_0}H_{\rm int}(0)e^{-iH_0 t}. \tag{E.14}$$

The subscript I stands for interaction in the Dirac picture. The interaction picture is somewhat intermediate between Schrodinger and Heisenberg, in that the time dependence of the operators, like the interaction Hamiltonian, is associated to the free Hamiltonian. This suits well with the fact that in particle physics, we start with free states (particles far away from each others), and consider interactions (between the particles) as perturbations.

Now, for a quantum field,

$$H_{\rm int}(0) = \int d^3x\mathcal{H}_{\rm int}(\mathbf{x}, 0) \tag{E.15}$$

so that Eq.(E.12) takes the neat form

$$\boxed{\mathcal{A}_{fi}^{(1)} = -i\int d^4x\langle f|\mathcal{H}_I(\mathbf{x}, t)|i\rangle} \tag{E.16}$$

Here

$$\mathcal{H}_{\rm I}(\mathbf{x}, t) \tag{E.17}$$

means that it is expressed in terms of field operators in the Dirac/interaction picture, for instance for a scalar field,

$$\phi_I(\mathbf{x}, t) = e^{iH_0(t-t_0)}\phi(\mathbf{x}, t_0)e^{-iH_0(t-t_0)} \tag{E.18}$$

At this level (meaning here in these lectures) there is no practical distinction between ϕ_I and free fields. What is useful however is that the expression (E.16) is fully Lorentz covariant, as it is expressed as an integral over space-time.

Transition to a continuum: Fermi Golden rule

If we have discrete states, we may right-away obtain the probability $p_{fi}(t)$ that a system undergoes a transition form a state i to a state f. Let us do this in two ways. First consider Eq.(E.11), for a time-independent interaction (here we re-instate an initial time t_0 for clarity)

$$\begin{aligned} p_{fi}(t) &= |\mathcal{A}_{fi}|^2 = \frac{|e^{i(E_f-E_i)t} - e^{-i(E_f-E_i)t_0}|^2}{(E_f-E_i)^2}|\langle f|H_{\text{int}}|i\rangle|^2 \\ &= 4\frac{\sin^2\left[(E_f-E_i)(t-t_0)/2\right]}{(E_f-E_i)^2}|\langle f|H_{\text{int}}|i\rangle|^2 \end{aligned} \tag{E.19}$$

From the probability, we define the mean rate (probability of transition per unit of time)

$$\mathcal{W}_{fi} = \frac{p_{fi}}{T} \tag{E.20}$$

with $T = t - t_0$. Now consider the function $g(\omega_{fi}, t)$ of $E_f - E_i = \omega_{fi}$

$$g(\omega_{fi}, t) = \frac{\sin^2 \omega_{fi} t/2}{\omega_{fi}^2 t} \tag{E.21}$$

This function is strongly peaked at $\omega_{fi} = 0$, with a maximum $g(0) = t/4$, and has first zero's at $\omega_{fi} = \pm(2\pi/t)$. Hence, as t increases, the peak increases, and the width decreases, while the area under the curve, which we estimate to be $\approx \pi/2$, stays roughly constant. This is a neat illustration of the so-called uncertainty relation between energy and time which we read here as: if we perturbed a quantum system with energy E_i for a time t, the outcome might be a state with energy within the range $E_f - E_i \sim \pi/t$.[1] Also the function $g(\omega)$ smells like a delta function. Incidentally,

$$\lim_{t\to\infty} g(\omega_{fi}, t) = \frac{\pi}{2}\delta(\omega_{fi}). \tag{E.22}$$

Hence, in the limit of a large interaction time, in practice $t \gtrsim 1/\omega_{fi}$, the transition rate becomes

$$\boxed{\mathcal{W}_{fi} = 2\pi\delta(E_f - E_i)|\langle f|H_{\text{int}}|i\rangle|^2} \tag{E.23}$$

The interaction Hamiltonian is time-independent for the processes we will consider, hence this is the formula that we need.

What if had started right-away from (E.12)? The transition probability would involve the square of a delta function, a pretty singular object. However, from the previous derivation we know that we have actually just have

$$p_{fi} = \lim_{T\to\infty} T(2\pi)\delta(E_f - E_i)|\langle f|H_{\text{int}}|i\rangle|^2 \tag{E.24}$$

with $T = t - t_0$ the total interaction time, so we had better consider the transition rate, $\lim_{T\to\infty}(p_{fi}/T) = \mathcal{W}_{fi}$.

[1] This is no energy violation, of course, as energy is anyway not conserved for a time-dependent Hamiltonian (in other word, the energy required for the transition is provided by an external source).

Now, in general there will be a number of states $\{|f\rangle\}$ consistent with energy conservation, so that

$$\mathcal{W}_{fi}^{\text{tot}} = \sum_{\{f\}} \mathcal{W}_{fi} \tag{E.25}$$

In the case of continuum, the number of states in a range dE_f around E_f is given by

$$dn(E_f) = \rho(E_f)dE_f \tag{E.26}$$

where $\rho(E)$ is the density of states, generally a smooth function of E. Starting from finite time we have

$$\mathcal{W}_{fi} = \lim_{t\to t} \frac{1}{t} \int dE_f \rho(E_f) 4 \frac{\sin^2(E_f - E_i)t/2}{(E_f - E_i)^2} |\langle f|H_{\text{int}}|i\rangle|^2 \tag{E.27}$$

or finally

$$\boxed{\mathcal{W}_{fi} = 2\pi\delta(E_f - E_i)|\langle f|H_{\text{int}}|i\rangle|^2 \rho(E_f)} \tag{E.28}$$

This formula is known as the Fermi's Golden Rule. It gives us the transition rate to a continuum of states, and is the formula we will use in the lectures.

Interactions to second order

First order interactions are relevant for simple processes, typically the calculation of a decay rate, or some specific scattering amplitudes (point-like interactions), but for some processes we need to go to second order. Following the development started above, we get, to second order

$$i\dot{c}_f^{(2)}(t) = \sum_n c_n^{(1)}(t) e^{i(E_f - E_n)t} \langle f|H_{\text{int}}|n\rangle \tag{E.29}$$

Using Eq(E.11) for $c_n^{(1)}(t)$, we rewrite this expression as

$$\dot{c}_f^{(2)}(t) = (-i)^2 \sum_n e^{i(E_f - E_n)t} \left(\int_{t_0}^{t} dt' e^{i(E_n - E_i)t'} \right) \langle f|H_{\text{int}}|n\rangle\langle n|H_{\text{int}}|i\rangle \tag{E.30}$$

You may check, by direct derivation, that the solution of this equation is given by

$$\boxed{c_f^{(2)}(t) = (-i)^2 \sum_n \left(\int_{t_0}^{t} dt'' e^{i(E_f - E_n)t''} \int_{t_0}^{t''} dt' e^{i(E_n - E_i)t'} \right) \langle f|H_{\text{int}}|n\rangle\langle n|H_{\text{int}}|i\rangle} \tag{E.31}$$

Let us focus on the integrals. We have

$$\begin{aligned} f(t - t_0) &= \int_{t_0}^{t} dt'' e^{i(E_f - E_n)t''} \int_{t_0}^{t''} dt' e^{i(E_n - E_i)t'} \\ &= \int_{t_0}^{t} dt'' e^{i(E_f - E_n)t''} \left(\frac{e^{i(E_n - E_i)t''} - e^{i(E_n - E_i)t_0}}{i(E_f - E_i)} \right) \\ &= \int_{t_0}^{t} dt'' \frac{e^{i(E_f - E_i)t''}}{i(E_n - E_i)} \left(1 - e^{-i(E_n - E_i)t'' + i(E_f - E_i)t_0} \right) \end{aligned} \tag{E.32}$$

If we take the limit $t_0 \to -\infty$, and since $E_n - E_i \neq 0$, the second term between brackets is rapidly oscillating, so that, on average, we expect it to gives zero. Alternatively we invoke the adiabatic condition on the interaction Hamiltonian. To do so, we introduce a factor $e^{\lambda t}$, with $\lambda > 0$, so that $\lim_{t_0 \to -\infty} e^{\lambda t_0} = 0$. Eventually we let $\lambda \to 0$. Then

$$\lim_{t_0 \to -\infty} f(t - t_0) = \int_{-\infty}^{t} dt'' \frac{e^{i(E_f - E_i)t'' + 2\lambda t''}}{i(E_n - E_i - i\lambda)} \tag{E.33}$$

Now we may take the limit $\lambda \to 0$ and let $t \to +\infty$, to get

$$\lim_{t \to +\infty} f(t) = i(2\pi)\delta(E_f - E_i)\frac{1}{E_i - E_n} \tag{E.34}$$

where, implicitely $E_n \neq E_i$. Finally, to second order we have the following simple expression

$$\boxed{\mathcal{A}^{(2)}_{fi} = -i(2\pi)\delta(E_f - E_i)\sum_{n \neq i}\langle f|H_{\rm int}|n\rangle \frac{1}{E_i - E_n}\langle n|H_{\rm int}|i\rangle} \tag{E.35}$$

for the amplitude of transition $\mathcal{A}^{(2)}_{fi}$ from i to f. Notice that although energy is conserved between the initial and final states, no restriction is set on the energy of the intermediate states n.

Finally we notice that, keeping in mind the domains of integration over time, we may write Eq.(E.31) in a quite suggestive way

$$\begin{aligned} c^{(2)}_f(t) &= (-i)^2 \sum_n \int d^4x \,\langle f|\mathcal{H}_I|n\rangle \int d^4y \,\langle n|\mathcal{H}_I|i\rangle \\ &= (-i)^2 \langle f| \left(\int d^4x \,\mathcal{H}_I \int d^4y \,\mathcal{H}_I \right) |i\rangle. \end{aligned} \tag{E.36}$$

If we compare to Eq.(E.16), we see that a simple pattern is emerging. The full expansion of the transition amplitude, to all order in the interactions, is called the Dyson series. Here we have the first two terms, which is all we need at the level of these lectures.

E.2 Generalities on elementary processes (Yukawa interactions)

We will consider a very simple Lagrangian, made of the Lagrangian of a free massive scalar field, a free massive Dirac field, and a Yukawa coupling between the two formers,

$$\mathcal{L} = \mathcal{L}_\phi + \mathcal{L}_\psi + \mathcal{L}_{\rm Yukawa} \tag{E.37}$$

where

$$\begin{aligned} \mathcal{L}_\phi &= \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{1}{2}M^2\phi^2, & \text{(E.38)} \\ \mathcal{L}_\psi &= i\bar{\psi}\partial\!\!\!/\psi - m\bar{\psi}\psi, & \text{(E.39)} \\ \mathcal{L}_Y &= -g\phi\bar{\psi}\psi & \text{(E.40)} \end{aligned}$$

The first two Lagrangians are simple, in the sense that we know the spectrum (states and eigenvalues) of their respective Hamiltonians. In the absence of interactions, $g \to 0$, the scalar and fermionic sector of the theory are completely independent. The rational for most of calculations involving relativistic particles is that we may start with free solutions, and study the effect of interactions as a perturbation. This is precisely how we solve all (but the simplest) problems in non-relativistic quantum mechanics. Since the coupling y is dimensionless, we may anticipate that it is this parameter that will control the effectiveness of the perturbative expansion. As we will briefly mention at the end of the lecture, this is correct, but only to some extent, due to the emergence of infinities in the calculation.

For the time being, we ignore this and we separate the Hamiltonian of our theory into a free Hamiltonian H_0, sum of the free Hamiltonians for the scalar and the fermionic field, and an interaction Hamiltonian H_{int}

$$H = H_0 + H_{\text{int}}. \tag{E.41}$$

Since the Yukawa coupling involves no derivative of the fields, we have simply $\mathcal{H}_{int} = -\mathcal{L}_Y$, so

$$H_{\text{int}} = \int d^3x\, \mathcal{H}_{\text{int}} = \int d^3x\, g\phi\bar{\psi}\psi \tag{E.42}$$

Now consider an elementary process, in which initial state $|i\rangle$, that involves a number of scalar and fermionic particles, goes through the Yukawa interaction, into a final state $|f\rangle$. We would like the compute the amplitude to go from i to f. Which processes are possible? From the formalism of perturbation theory, reviewed in the appendix, to leading order, we have for the amplitude Eq.(E.12),

$$\mathcal{A}^{(1)}_{fi} = -i(2\pi)\delta(E_f - E_i)\langle f|H_{\text{int}}|i\rangle \tag{E.43}$$

where H_{int} is the interaction Hamiltonian in the Schrödinger picture (ie it does not depend on time). The Dirac delta tells us that energy is conserved between the initial and final states.

Now, each of our fields ϕ, $\bar{\psi}$ and ψ may at best create/destroy one particle (or antiparticle, we called them particles generically), so, to leading order, the simplest processes involve three particles at a time. For instance we may have initially one fermion, and in the final state one fermion and one scalar. Since the fermion number is conserved, the number of fermion minus the number of anti-fermions is unchanged between the initial and final states. We could have the reverse, a scalar and a fermion in the initial state, and a fermion in the final state. Finally we may have initially a scalar and in the final state a fermion and its anti-fermion, and the reverse. Certainly we cannot have a process with particles initial states and no particles in the final states (or the reverse): this would violate the conservation of energy. Incidentally, relativistic kinematics (conservation of both energy and momentum) tells us eventually what is possible and what is not. Remark that momentum conservation is implicit, just from the fact that the interaction Hamiltonian is space-independent.

Let us consider the emission of a scalar particle by a fermion. Up to spin aspects, this is like the emission of a photon by an electron, say, provided we take $M \to 0$. Let us have a quick look at the kinematics of this processus. We have initially an electron with energy E and momentum $\mathbf{p}$. Let's work in the rest-frame of this particle. Then its energy is equal to m and energy conservation asks for

$$m \stackrel{?}{=} E' + \omega \tag{E.44}$$

where $E' = \sqrt{\mathbf{p}'^2 + m^2}$ is the energy of the electron in the final state, and $\omega = \sqrt{\mathbf{k}^2 + M^2}$ is the energy of the scalar particle. Also $\mathbf{p}' = -\mathbf{k}$ in the CM frame. Clearly the RHS of relation Eq.(E.44) is always larger than the LHS, for all momenta. Hence it is not possible the fermion can not emit a real (in the litteral sense) scalar particle. If we go to the massless limit, $M \to 0$, the situation is unchanged, unless $\mathbf{k} = 0$, but that means no particle. Since the kinematics is the same as for photon emission, we conclude that the emission of a real photon is also excluded. This is of course because the mass of the electron is the same before and after the emission.

Now, in the case of a bound state electron the situation is different, as one has to take into account the binding energy of the electron. In this case the emission of a photon is generally possible, provided we take into account the fact that a photon is a particle of spin, and angular momentum must be conserved in the process. Alternatively, we may consider a Yukawa coupling between a scalar and two *different* fermions. Such a possibility is realized in the SM. In particular we may envision the decay of a top quark into a bottom quark, with emission of a scalar particle. In that set up the scalar particle is a bit special (see next year) but what is important is this process is allowed kinematically and actually exists in Nature.

Let us go on a bit with the problem of emission and consider the following quantity

$$q^2 = \omega^2 - \mathbf{k}^2 = (E - E')^2 - (\mathbf{p} - \mathbf{p}')^2. \tag{E.45}$$

which for a real scalar particle emission, we should have $q^2 = m^2$. This combination is Lorentz invariant, so let us evaluate it in a simple frame. Again we take the CM frame $\mathbf{p} = 0$. Then

$$q^2 = 2m^2 + \mathbf{p}^2 - 2m\sqrt{m^2 + \mathbf{p}^2} - \mathbf{p}^2 < 0 \tag{E.46}$$

which means that the four momentum q^μ is space-like. Certainly $q^2 \neq M^2$.

Now we consider a process of annihilation of a fermion and antifermion into the scalar particle, then

$$q^2 = \omega^2 - \mathbf{k}^2 = (E + E')^2 - (\mathbf{p} + \mathbf{p}')^2 \tag{E.47}$$

It is now convenient to work in the CM of the initial particles, so that the last term is vanishing. We conclude that for this process

$$q^2 > 0 \tag{E.48}$$

so that the four momentum of the putative scalar is time-like. Notice that this parameter is so common that it has its own letter, $s = q^2$. This is one of the Mandelstam variables. The square-root of s is the center-of-mass energy.

More importantly note that unless $E + E' = M$, $q^2 \neq M^2$. Why do we bother if these processes are kinematically forbidden for particles? Well, it turns out that these processes, with $q^2 \neq M^2$ are actually very relevant, even thought the particle involved (here the scalar) may only exists as intermediate states, not real particles. These are call *virtual* particles. It is also said that such a virtual particle (here the scalar) is off-mass-shell (off-shell for brief), meaning $q^2 \neq M^2$. By way of comparison, a real, true bona fide particle is said to be on-mass-shell (on-shell for brief), meaning $q^2 = M^2$. The purpose of this chapter is to see explicitly why and how such states emerge in calculation of simple processes. Basically we will see that interactions between particles occur through the exchange of virtual particles (actually just one to leading-order).

Consider finally the decay of the scalar into fermions. Again we work in the rest frame of the scalar, $\omega = M$. Then energy conservation demands

$$M = E_p + E_{p'} \tag{E.49}$$

and momentum conservation $0 = \mathbf{p} + \mathbf{p}'$ implies that $E_p = E_{p'} = \sqrt{m^2 + \mathbf{p}^2}$. Then

$$M^2 = 4E_p^2 = 4(m^2 + \mathbf{p}^2) \tag{E.50}$$

so that decay is possible provided that $M > 2m$. This is the topic of the next section.

E.2.1 The amplitude or Matrix element

In this section we will obtain the amplitude for the decay of the scalar into a pair fermion-antifermion. We first rewrite the amplitude (E.12) as

$$\begin{aligned}\mathcal{A}_{fi} &= -i(2\pi)\delta(E_f - E_i)\langle f|H_{int}|i\rangle \\ &= -ig(2\pi)\delta(E_f - E_i)\int d^3x\,\langle \mathbf{p}, s; \mathbf{p}', s'|\phi(\mathbf{x})\bar{\psi}(\mathbf{x})\psi(\mathbf{x})|\mathbf{k}\rangle\end{aligned} \tag{E.51}$$

where we have expressed the fact that we have in the initial state a scalar particle with momentum $\mathbf{k}$ and in the final state a fermion with momentum $\mathbf{p}$ and helicity s, and an antifermion with momentum $\mathbf{p}'$ and helicty s'. Hence our first task is to evaluate the matrix element

$$\langle \mathbf{p}, s; \mathbf{p}', s'|\phi(\mathbf{x})\bar{\psi}(\mathbf{x})\psi(\mathbf{x})|\mathbf{k}\rangle. \tag{E.52}$$

Let us begin with the easy part, the action of the ϕ field. For the scalar, we adopt the following convention

$$\langle 0|\phi(\mathbf{x}, 0)|\mathbf{k}\rangle = e^{i\mathbf{k}\cdot\mathbf{x}} \tag{E.53}$$

This is indeed what we need as in the final state of our process we are in the vacuum state of the scalar field. Remember that we work in a Fock space.

Similarly, the initial state we have the vacuum of the fermion and anti-fermion, hence we need to evaluate

$$\langle \mathbf{p}, s; \mathbf{p}', s'|\bar{\psi}(x)\psi(x)|0\rangle \equiv \langle \mathbf{p}, s|\bar{\psi}(x)|0\rangle\langle \mathbf{p}', s'|\psi(x)|0\rangle \tag{E.54}$$

Indeed, we have interpreted $\psi(\mathbf{x})$ as the operator that destroys an incoming fermion/creates an incoming anti-fermion, and conversely for $\bar{\psi}$. We write down the matrix elements as

$$\langle \mathbf{p}, s; \mathbf{p}', s'|\bar{\psi}(x)\psi(x)|0\rangle = \bar{u}^s(p)v^{s'}(p')e^{-i(\mathbf{p}+\mathbf{p}')\cdot\mathbf{x}} \tag{E.55}$$

That's OK, since we have the wave-function of an outgoing fermion ($\bar{u}$) and outgoing antifermion (v), and their helicity is captured by the plane wave spinors u and v.

Putting (E.53) and (E.55) in (E.51), we have

$$\mathcal{A}_{fi} = (-ig)(2\pi)\delta(E_f - E_i)\left(\bar{u}_s(p)v_{s'}(p')\right)\int d^3x\, e^{i(\mathbf{k}-\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} \tag{E.56}$$

or

$$\boxed{\mathcal{A}_{fi} = (-ig)(2\pi)^4\delta^4(p + p' - k)\left(\bar{u}_s(p)v_{s'}(p')\right)} \tag{E.57}$$

We have a manifestly Lorentz covariant expression, so we should be on the right track. (To be sure, notice that there is a hidden summation here, $\bar{u}^s v^{s'} \equiv \sum_{k=1...4} \bar{u}^s_k v^{s'}_k$ since the u and v are actually 4-components column vectors.)

The next thing we need to do is to square the amplitude. Now they are two things we need to care of. The first thing is that we have to interpret of the square of the delta function. Here we just need to remember that the delta comes from integration of the four-volume, so we may just replace one factor of $(2\pi)^4\delta^4$ by $T \cdot V$. At the end of the day we will divide the probability by this factor, so as to get the probability of transition per unit time (the volume will factor out). The other thing is that we want the full decay probability, that is we have to sum over all the final helicity state of the fermion and of the antifermion. Hence we want to evaluate

$$\begin{aligned} |\mathcal{A}_{fi}|^2 &= g^2 VT\,(2\pi)^4\delta^4(p + p' - k) \sum_{s,s'} \left|\bar{u}^s(p)v^{s'}(p')\right|^2 \\ &= g^2 VT\,(2\pi)^4\delta^4(p + p' - k) \sum_{s,s'} \left(\bar{u}^s(p)v^{s'}(p')\bar{v}^{s'}(p')u^s(p)\right) \end{aligned} \tag{E.58}$$

Here we recognize something we have already met in the QFT course,

$$\sum_s u^s(p)\bar{u}^s(p) = \not{p} + m \tag{E.59}$$

and

$$\sum_{s'} v^{s'}(p')\bar{v}^{s'}(p') = \not{p}' - m \tag{E.60}$$

Beware of the notation. We have 4×4 matrices, so what we are actually trying to evaluate is

$$\sum_{s',s} \left\{ \sum_{i,j=1...4} \left(\bar{u}^s_i v^{s'}_i \bar{v}^{s'}_j u^s_j\right) \right\} = \sum_{i,j=1...4} \left((\not{p} + m)_{ji}(\not{p}' - m)_{ij}\right) = \mathrm{Tr}\left[(\not{p} + m)(\not{p}' - m)\right] \tag{E.61}$$

Hence

$$|\mathcal{A}_{fi}|^2 = g^2 VT\,(2\pi)^4\delta^4(p + p' - k)\,\mathrm{Tr}\left[(\not{p} + m)(\not{p}' - m)\right]. \tag{E.62}$$

Using finally the properties of the Dirac gamma matrices, we may evaluate the trace,

$$\mathrm{Tr}\left[(\not{p} + m)(\not{p}' - m)\right] = \mathrm{Tr}\left[\gamma^\mu\gamma^\nu\right] p_\mu p'_\nu - m^2\mathrm{Tr}\left[1\right] = 4(p \cdot p' - m^2)$$

so we finally get

$$\boxed{|\mathcal{A}_{fi}|^2 = g^2 VT\,(2\pi)^4\delta^4(p + p' - k)\,4\,(p \cdot p' - m^2)} \tag{E.63}$$

This seems like a lot of work, and considering that we are computing the simplest process one can think of, imagine the work that would be required to compute more complex processes. Much of these dreadful tasks has been to some extend simplified by Feynman who discovered simple patterns in the construction of amplitudes. This is summarized, for each theories, by a set of rules called Feynman rules that give rightaway Lorentz covariant expressions (as Schwinger put it: “Feynman brought QFT to the masses"). This is however an advanced topic that to be addressed properly requires a more sophisticated approach to perturbations in quantum field theories. This topic is covered in the courses of the Master program.

Let us go on and express this in the rest frame of the scalar and use the conservation of energy and momentum implemented by the delta functions. Then $k = (M, \mathbf{0})$ while $p = (E_p, \mathbf{p})$ and $p' = (E_p, -\mathbf{p})$, and so $\mathbf{p}^2 = M^2/4 - m^2$. Hence

$$|\mathcal{A}_{fi}|^2 = g^2 VT\,(2\pi)^4 \delta^4(p + p' - k)\, 2M^2 \left(1 - \frac{4m^2}{M^2}\right) \tag{E.64}$$

The decay rate

We are almost done. Let us first get rid of the volume of space-time, VT. This factor occured because really we are considering a process that could take place anywhere in space-time, since we are working with plane waves, while we are interested in the rate for decay of our scalar particle sitting somewhere. Strictly speaking we should built up a localized particle, with a wave-packet, etc. This is cumbersome, so much that nobody does that. Still we should take into account the normalization of our initial state, which is not normalized to “one particle in the box”. Indeed remember that

$$\langle k|k'\rangle = 2E_k (2\pi)^3 \delta^3(\mathbf{k} - \mathbf{k}') \tag{E.65}$$

in the infinite volume limit. Remember also this implies that density of states $\mathbf{k}$ is given by

$$\frac{d^3k}{(2\pi)^3} \frac{1}{2E_k} \tag{E.66}$$

so that

$$\int \frac{d^3k}{(2\pi)^3} \frac{1}{2E_k} 2E_k (2\pi)^3 \delta^3(\mathbf{k} - \mathbf{k}') = 1 \tag{E.67}$$

The reason for this apparently peculiar normalization is that it is Lorentz invariant.

Now for a finite box, Eq.(E.65) is equivalent to

$$\langle \mathbf{k}|\mathbf{k}'\rangle = 2E_k V \delta_{\mathbf{k},\mathbf{k}'} \tag{E.68}$$

Hence, when we consider the decay of our scalar particle, we should divide by a factor of $2E_k V$ to have a state normalized to “one particle in the box”.

We also divide by T, since we are looking for the decay rate, ie the probability of decay per unit time, $dP_{\text{decay}}/dt \equiv P_{\text{decay}}/T = \Gamma$. We also have to consider all the possible configurations of the final particles. Clearly there is a continuum, as the final fermions may go in any direction, modulo momentum and energy conservation. This is symbolically characterized by the density of final states, $\rho(E_f) = dn/dE_f$, so that the decay rate Γ takes finally the form[2]

$$\begin{aligned}\Gamma &= \frac{1}{2E_k V}\frac{1}{T}\int |A_{fi}|^2\rho(E_f)dE_f \\ &\equiv g^2 M\left(1-\frac{4m^2}{M^2}\right)\int \frac{d^3\mathbf{p}}{(2\pi)^3 2E_p}\frac{d^3\mathbf{p}'}{(2\pi)^3 2E_{p'}}(2\pi)^4\delta^4(p+p'-k)\end{aligned} \tag{E.69}$$

where in the second line we are in the center-of-mass frame, and we have used the density of states of the fermion and its antiparticle, given by

$$\rho(E_f)dE_f \equiv \frac{d^3\mathbf{p}}{(2\pi)^3 2E_p}\frac{d^3\mathbf{p}'}{(2\pi)^3 2E_{p'}} \tag{E.70}$$

With the delta function on the spatial momentum, we may eliminate the integral of say over $\mathbf{p}'$. We may also integrate over the direction of $\mathbf{p}$, since the process is invariant under rotations, which gives a factor of 4π. Then

$$\Gamma = \frac{(4\pi)(2\pi)}{4(2\pi)^3}g^2 M\left(1-\frac{4m^2}{M^2}\right)\int_0^\infty dp\ \frac{p^2}{E_p^2}\delta(2E_p - M) \tag{E.71}$$

Using

$$\delta(2E_p - M) = \frac{1}{2dE_p/dp}\delta\Big(p - \sqrt{M^2/4 - m^2}\Big) = \frac{E_p}{2p}\delta\Big(p - \sqrt{M^2/4-m^2}\Big) \tag{E.72}$$

we finally get

$$\boxed{\Gamma = \frac{g^2}{8\pi}M\left(1-\frac{4m^2}{M^2}\right)^{3/2}} \tag{E.73}$$

That's it. We have computed our first truly quantum relativistic process. The number of particles is changing, and matter (here the fermions) are created out of the energy of the initial particle. Notice that the end result is quite simple. Dimensionally, we would have expected that $\Gamma \propto M$. Also this happens to leading order in the coupling g, hence the rate is $\Gamma \sim g^2/4\pi M$. We put in the factor of 4π (I know, it is 8π in our calculation but $1 = 2$ in my approximation) because it reminds us of the fine structure constant $\alpha = e^2/4\pi \approx 1/137$ of electromagnetism. We will take as a rule of thumb that perturbation theory works provided $g^2/4\pi \lesssim 1$. Finally there is a so-called phase-factor[3], $(1-4m^2/M^2)^{3/2}$. Clearly we need $M > 2m$ for the decay to take place, but we also see how the decay is suppressed if M is close to the treshold. For $M \gg m$ however, this suppression is small so that, at the end of the day, if you want to estimate a decay rate of a massive particle of mass M into lighter particles, through a coupling g, then

$$\Gamma \approx \frac{g^2}{4\pi}M. \tag{E.74}$$

Easy. In the next section we consider a process that is second order in the perturbation, $O\Big((g^2/4\pi)^2\Big)$.

[2] The decay rate give the mean life-time of a particle, $\tau_{1/2} = 1/\Gamma$, so that if there is an initially N_0 scalar, after a time t, there will be $N(t) = N_0\exp(-\Gamma t)$.

[3] The strange power of $3/2$ is because the final state must be a p-wave, proportional to $|\mathbf{p}|$. This is a bit advanced, but basically it has to do with the fact that our initial state is a spin-0 scalar, of even parity, while a pair of fermion-antifermion is a parity odd state in the s-wave. The p-wave provides an extra $(-1)^l = -1$ factor to restore parity. Advanced, as I said, but basic at the same time. Beautyis in the details.

E.3 Scattering of fermions

Another important class of processes in particle physics is scattering. In NR QM we typically consider the scattering of, say, a non-relativistic electron in the Coulomb potential of some heavy, charged target particle. A fundamental difference in the framework of relativistic particles is the fact that the nature and the number of the particles may changed in the process. However here we will consider only the simplest process, that is the scattering of two distinct fermions. Moreover we only consider this in the case of a Yukawa interactions. Nevertheless, even in this very simple set-up, we will see that the requirement of Lorentz covariance will lead to a re-interpretation of the intermediate states n in terms of virtual particles.

For simplicity, we consider the Yukawa interaction of one fermion with charge g_a with another of charge g_b. The particle are distinguishable, so we do not have to worry about the symmetry aspect of the initial and final states. The first thing we will need is the amplitude for this process. We will derive this quantity in two ways. First using second-order perturbation theory, the second using the notion of Green functions.

The amplitude for scattering

The process involves two fermions in the initial state and two in the final state, hence there is no contribution to leading order in PT, so we try the next contribution. To second order we have the general expression from Eq.(E.35),

$$\mathcal{A}_{fi}^{(2)} = -i(2\pi)\delta(E_f - E_i) \sum_{n\neq i} \langle f|H_{\rm int}|n\rangle \frac{1}{E_i - E_n} \langle n|H_{\rm int}|i\rangle \tag{E.75}$$

For the process $a + b \to a' + b'$, we have

$$\begin{aligned} \mathcal{A}_{fi}^{(2)} &= -i(2\pi)\delta(E_f - E_i) \sum_{E_n\neq E_a+E_b} \langle p'_a, s'_a; p'_b, s'_b| \left\{ \int d^3\mathbf{y}\mathcal{H}_{\rm int}(\mathbf{y}, 0) \right. \\ &\quad \left. \times|n\rangle \frac{1}{E_a + E_b - E_n} \langle n| \int d^3\mathbf{x}\mathcal{H}_{\rm int}(\mathbf{x}, 0) \right\} |p_a, s_a; p_b, s_b\rangle \end{aligned} \tag{E.76}$$

We notice that the delta function enforces energy conservation between the initial and final states, but not on the intermediate states n (otherwise we would get $E_n = E_a + E_b$ and an infinite result, which makes no sense. On the other hand the integrals over space ($\mathbf{x}$ and $\mathbf{y}$) will impose conservation of momentum at the matrix elements. There is thus an asymmetry between the treatement of energy conservation and momentum conservation, so we should worry about the Lorentz covariance of our calculation. Of course this is because our formalism is not manifestly Lorentz covariant. However we will see that this tension is resolved in a very interesting way at the end of the day.

To start with, we notice that two amplitudes contribute to the same process. The first contribution corresponds to the emission of a scalar by the initial a fermion, which later is destroyed by the b. The other contribution is the opposite: the scalar is emitted by the b fermion, and absorbed by the a. Notice that in the two cases the intermediate states are distinct.

Let us begin with the first amplitude. The first matrix element corresponds to the emission of a scalar with momentum k by the fermion a,

$$\int d^3\mathbf{x}\langle p_a', s_a'; p_b, s_b; k|\mathcal{H}_{\rm int}(\mathbf{x}, 0)|p_a, s_a; p_b, s_b\rangle \tag{E.77}$$

Note that b fermion is unaffected, so effectively we are dealing with the matrix element

$$\int d^3\mathbf{x}\langle p_a', s_a'; k|\mathcal{H}_{\rm int}(\mathbf{x}, 0)|p_a, s_a\rangle. \tag{E.78}$$

Accordingly the other matrix elements is given by

$$\int d^3\mathbf{y}\langle p_b', s_b'|\mathcal{H}_{\rm int}(\mathbf{y}, 0)|p_b, s_b; k\rangle. \tag{E.79}$$

The energy of the intermediate state is given by

$$E_n = E_a' + E_b + E_k \tag{E.80}$$

and we have to sum over all the possible k states. Hence we need to evaluate

$$\begin{aligned}\mathcal{A}_{fi}^{(2)} &= -i(2\pi)\delta(E_f - E_i)\sum_k \int d^3\mathbf{y}\langle p_b', s_b'|\mathcal{H}_{\rm int}(\mathbf{y}, 0)|p_b, s_b; k\rangle \\ &\qquad \frac{1}{(E_a - E_a' - E_k)}\int d^3\mathbf{x}\langle p_a', s_a'; k|\mathcal{H}_{\rm int}(\mathbf{x}, 0)|p_a, s_a\rangle\end{aligned} \tag{E.81}$$

where we have used

$$\frac{1}{E_a + E_b - E_n} = \frac{1}{(E_a + E_b) - (E_a' + E_b + E_k)} = \frac{1}{E_a - E_a' - E_k}. \tag{E.82}$$

Let us first evaluate the matrix element (E.78),

$$\int d^3\mathbf{x}\langle p_a', s_a'; k|\mathcal{H}_{\rm int}(\mathbf{x}, 0)|p_a, s_a\rangle = g_a\int d^3\mathbf{x}\langle p_a', s_a'; k|\phi\bar{\psi}_a\psi_a(\mathbf{x}, 0)|p_a, s_a\rangle \tag{E.83}$$

We have a scalar in the final state, which, from Eq.(E.53) gives a factor of $e^{-i\mathbf{k}\cdot\mathbf{x}}$. The operator ψ_a destroys the initial a fermion gives and $\bar{\psi}_a$ destroys the a in the final state, hence, from the analog of Eq.(E.55) we have

$$g_a\int d^3\mathbf{x}\langle p_a', s_a'; k|\phi\bar{\psi}_a\psi_a(\mathbf{x}, 0)|p_a, s_a\rangle = g_a\bar{u}^{s_a'}(p_a')u^{s_a}(p_a)\int d^3\mathbf{x}\, e^{i(\mathbf{p_a}-\mathbf{k}-\mathbf{p_a'})\cdot\mathbf{x}} \tag{E.84}$$

or, after integration over $\mathbf{x}$,

$$g_a\int d^3\mathbf{x}\langle p_a', s_a'; k|\phi\bar{\psi}_a\psi_a(\mathbf{x}, 0)|p_a, s_a\rangle = g_a(2\pi)^3\delta^3(\mathbf{p_a} - \mathbf{p_a'} - \mathbf{k})\bar{u}^{s_a'}(p_a')u^{s_a}(p_a) \tag{E.85}$$

Similarly, the second matrix element gives

$$g_b \int d^3\mathbf{y} \langle p_b', s_b' | \phi \bar{\psi}_b \psi_b(\mathbf{y}, 0) | p_b, s_b; k \rangle = g_b (2\pi)^3 \delta^3(\mathbf{p_b'} - \mathbf{p_b} - \mathbf{k}) \bar{u}^{s_b'}(p_b') u^{s_b}(p_b) \tag{E.86}$$

As promised, we see that the integrals over space have imposed conservation of momentum with the intermediate states. The summation over the k states being given by Eq.(**??**),

$$\sum_k \to \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{1}{2E_k} \tag{E.87}$$

we may do the integration over $\mathbf{k}$ with the help of one of the dirac delta function over momenta, to get finally

$$\begin{aligned} \mathcal{A}_{fi}^{(2)}\Big|_{\mathrm{I}} &= -ig_a g_b (2\pi)^4 \delta^4(p_a' + p_b' - p_a - p_b) \frac{1}{2E_k(E_a - E_a' - E_k)} \\ &\quad \times \left(\bar{u}^{s_a'}(p_a') u^{s_a}(p_a)\right) \left(\bar{u}^{s_b'}(p_b') u^{s_b}(p_b)\right) \end{aligned} \tag{E.88}$$

This is quite nice, and almost covariant, but for the denominator in the energy of the intermediate states.

Fortunately we do not have to repeat all these steps to get the second amplitude, as the only difference is that the role of the a and b particles is exchanged $a \leftrightarrow b$, so that

$$\begin{aligned} \mathcal{A}_{fi}^{(2)}\Big|_{\mathrm{II}} &= -ig_a g_b (2\pi)^4 \delta^4(p_a' + p_b' - p_a - p_b) \frac{1}{2E_k(E_b - E_b' - E_k)} \\ &\quad \times \left(\bar{u}^{s_a'}(p_a') u^{s_a}(p_a)\right) \left(\bar{u}^{s_b'}(p_b') u^{s_b}(p_b)\right) \end{aligned} \tag{E.89}$$

Adding the two amplitude we get the factor

$$\frac{1}{2E_k} \left(\frac{1}{E_a - E_a' - E_k)} + \frac{1}{E_b - E_b' - E_k} \right) \tag{E.90}$$

We are going to rewrite this by using the global energy conservation,

$$E_a - E_a' = -E_b + E_b' \tag{E.91}$$

and momentum conservation at the matrix elements,

$$\mathbf{k}^2 = (\mathbf{p_a} - \mathbf{p_a'})^2 = (\mathbf{p_b} - \mathbf{p_b'})^2 \tag{E.92}$$

As $E_k = \sqrt{\mathbf{k}^2 + M^2}$ we get

$$\begin{aligned} \frac{1}{2E_k} \left(\frac{1}{E_a - E_a' - E_k)} + \frac{1}{E_b - E_b' - E_k} \right) &= \frac{1}{(E_a - E_a')^2 - (\mathbf{p_a} - \mathbf{p_a'})^2 - M^2} \\ &= \frac{1}{q^2 - M^2} \end{aligned} \tag{E.93}$$

where we have introduced $q = p_a - p'_a = -(p_b - p'_b)$. Finally, the total amplitude for the scattering is given by

$$\boxed{\mathcal{A}^{(2)}_{fi} = (2\pi)^4 \delta^4(p_f - p_i) \mathcal{M}_{fi}} \tag{E.94}$$

where the so-called Matrix Element is given by

$$\boxed{\mathcal{M}_{fi} = (-ig_a) \left(\bar{u}^{s'_a}(p'_a) u^{s_a}(p_a) \right) \frac{i}{q^2 - M^2} (-ig_b) \left(\bar{u}^{s'_b}(p'_b) u^{s_b}(p_b) \right)} \tag{E.95}$$

This is our final result. The important thing is that the amplitude is now manifestly covariant. Notice also that both energy and momentum are conserved, including in the factor

$$\frac{i}{q^2 - M^2}. \tag{E.96}$$

since $q^2 = (p_a - p'_a)^2 = (p_b - p'_b)^2$. We recognized in this factor the so-called Feynman propagator for the scalar particle. It express the fact that the particles a and b exchange energy and momentum through the intermediate scalar particle. The price to pay is that, unlike for a real scalar, $q^2 \neq M^2$. We say that the scalar exchanged is a virtual particle.

An important contribution from Feynman was to recognize that there was a way to directly express amplitudes in a manifest Lorentz covariant way. Here we took a path that we took us from standard results in QM to show that explicitly. We have expressed the expression for the matrix element $\mathcal{M}_{fi}$ as we would have done it if we had followed a set of rules, called Feynman rules and that you may find in many (advanced or not that advanced) books.

Appendix F

Dirac matrices

Gamma matrices

$$\sum_{s=1,2} u_s(k)\bar{u}_s(k) = \not{k} + m$$

$$\sum_{s=1,2} v_s(k)\bar{v}_s(k) = \not{k} - m$$

$$\left\{\gamma_\mu, \gamma_\nu\right\} = 2g_{\mu\nu}$$

$$\{\gamma_5, \gamma_\nu\} = 0$$

with

$$\gamma_5^2 = 1$$

$$\mathrm{Tr}\left[\gamma_\mu\gamma_\nu\right] = 4g_{\mu\nu}$$

$$\mathrm{Tr}\left[\gamma_\alpha\gamma_\beta\gamma_\mu\gamma_\nu\right] = 4(g_{\mu\nu}g_{\alpha\beta} - g_{\alpha\mu}g_{\beta\nu} + g_{\alpha\nu}g_{\beta\mu})$$

$$\mathrm{Tr}\left[\gamma_5\gamma_\mu\gamma_\nu\right] = 0$$

$$\mathrm{Tr}\left[\gamma_5\gamma_\mu\gamma_\nu\gamma_\alpha\gamma_\beta\right] = 4i\epsilon_{\mu\nu\alpha\beta}$$

Traces over an odd number of gamma matrices vanish identically. Traces with γ_5 vanish if some of the momenta that are to be contracted on the Lorentz indices are identical or collinear.

Decay rates and scattering cross-sections

We only discuss two-body decay processes here. In the rest frame of the initial particle of mass M the total decay rate is

$$\Gamma = \frac{1}{2M}\int \frac{d^3k_1}{(2\pi)^3 2\omega_1}\frac{d^3k_2}{(2\pi)^2 2\omega_2}|\mathcal{M}|^2$$

where $\mathbf{k_1} + \mathbf{k_2} = 0$ and $M = \omega_1 + \omega_2$, and $\mathcal{M}$ is the decay amplitude.

Fields normalization and Feynman rules

All the Feynman rules of the SM can be found in the Appendix of Cheng and Li and (sorry) scattered in the bulk of the text. In this appendix, we only compile some informations not provided elsewhere in the text.

An incoming fermion of momentum k^μ is represented by the spinor $u(k)$. Its antiparticle by $\bar{v}(k)$. An outgoing fermion is $\bar{u}(k)$, while its antiparticle is $v(k)$. A fermion propagator is

$$\frac{i}{\not{p} - m}$$

An incoming gauge boson and momentum k^μ is represented by $\epsilon^\mu(k)$. An outgoing gauge boson is $\epsilon^*(k)$. The propagator of a massless gauge boson (in Feynman gauge) is

$$\frac{-ig_{\mu\nu}}{k^2}$$

For a massive gauge boson of mass M we will often use

$$-i\left(g_{\mu\nu} - \frac{k_\mu k_\nu}{M^2}\right)\frac{1}{k^2 - M^2}$$

Couplings can, in general, be read from the interaction part of the Lagrangian. For instance

$$\mathcal{L}_I = -g\bar{\psi}\psi\phi$$

represent the Yukawa coupling of a fermion to a scalar boson. The Feynman for this coupling is

$$-ig$$

Incidentally, the propagator of a scalar boson of mass m is

$$\frac{i}{k^2 - m^2}$$

If the interaction Lagrangian is

$$\mathcal{L}_I = e\bar{\psi}\gamma^\mu\psi A_\mu$$

the corresponding vertex is

$$ie\gamma^\mu$$

In drawing Feynman diagrams, one must take into account that particles may be indiscernible. The total amplitude must be compatible with either Bose-Einstein or Fermi-Dirac statistic. In the former case, which correspond the permutation of two bosons, the amplitudes must be added. In the latter case, which would correspond to the permutation of two fermions, there is a minus sign.

Loops are not discussed systematically here. Suffices to say that for each loop, we must integrate over the internal momentum with

$$\int \frac{d^4k}{(2\pi)^4}$$

Also, each fermion loop comes with a minus sign, because of Fermi-Dirac statistic. For instance, the one-loop contribution to the polarization tensor of the photon from a fermion of mass m and electric charge e is given by

$$i\Pi_{\mu\nu} = (-)(ie)^2 \int \frac{d^4k}{(2\pi)^4} \mathrm{Tr}\left[\frac{i}{(\not{k} - m)}\gamma_\mu \frac{i}{((\not{k} + \not{q}) - m)}\gamma_\nu\right]$$