Introduction to Flavor Physics
GraSPA 2014

Diego Guadagnoli

LAPTh, Université de Savoie et CNRS, BP110, F-74941 Annecy-le-Vieux Cedex, France
E-mail: diego.guadagnoli@lapth.cnrs.fr

Lecture 1: Introduction to EFTs.

Flavor physics rests on the basic idea of an **effective field theory**, which in turn is one of the most basic guides in modeling physics systems. In this lecture we introduce this idea. We will proceed by examples, introducing the minimal necessary notation only when it's needed.

We will introduce the idea of an EFT by the explicit example of **Rayleigh scattering**, namely elastic scattering of visible light off atoms. Rayleigh scattering explains why the sky is blue, and this will be the prize of this exercise. (Disclaimer: technically, what happens in Rayleigh scattering is that the atom gets polarized by the e.m. wave and emits like a dipole. Therefore, one can make an entirely classical derivation of this effect. We want instead to use the tools and the formalism of particle physics. Therefore, we will need a small detour where we will introduce all the basic concepts. The resulting derivation will be much more fun than the classical one.)

L1.1 Introduction

In order to approach the problem using quantum-physics tools, we need some basic concepts:

- **particles**, and how they are mathematically described (= quantum fields),
- **interactions**, namely products of fields satisfying certain requirements. These products of fields appear in 'Lagrangian' functions, akin to the analogous objects describing the dynamics of classical-mechanical systems,
- **relevant interactions**.

The last concept is at the core of the idea of EFTs. It consists in *identifying the characteristic energy or distance scale of a problem*, and accordingly writing down a sensible set of interactions describing that problem, while discarding irrelevant details (= effects at scales widely different than the characteristic one).

For example, if we are asked to evaluate the amount of heat dissipated by a TGV stopping from full speed, all we need is its kinetic energy $E_{\rm train} = MV^2/2$, with M its mass and V its average speed. We do not need details about its internal structure, such as the number of people going back to their seats during the braking. In fact, one can estimate both effects using

$$M \simeq 250 \text{ tons} = 2.5 \cdot 10^5 \text{ kg} \quad \text{(train mass)}$$

$$V \simeq 250 \text{ km/h} = \frac{2.5 \cdot 10^5}{3.6 \cdot 10^3} \frac{\text{m}}{\text{s}} = 0.7 \cdot 10^2 \frac{\text{m}}{\text{s}} \quad \text{(average train speed)}$$
$$m \sim 70 \text{ kg} \quad \text{(average human weight)}$$
$$v \sim 5 \text{ km/h} = \frac{5 \cdot 10^3}{3.6 \cdot 10^3} \frac{\text{m}}{\text{s}} = 1.4 \frac{\text{m}}{\text{s}} \quad \text{(average walking speed)}$$

With respect to the train kinetic energy, the 'internal structure' effect is a correction of order (we assume $N \sim 20$ for the number of people going back to their seats)

$$\frac{mv^2 \times N}{MV^2} = \frac{0.7 \cdot 10^2 \cdot 1.4^2 \times 20}{2.5 \cdot 10^5 \cdot (0.7 \cdot 10^2)^2} \simeq 2 \times 10^{-6} \, . \tag{1.1}$$

Other checks:

$$N \times E_{\text{human}} \simeq 20 \times \frac{1}{2} \cdot 70 \cdot 1.4^2 \frac{\text{kg} \cdot \text{m}^2}{\text{s}^2} \simeq 1.4 \; kJ$$
$$E_{\text{train}} \simeq \frac{1}{2}(2.5 \cdot 10^5)(0.7 \cdot 10^2)^2 \frac{\text{kg} \cdot \text{m}^2}{\text{s}^2} = 6 \cdot 10^8 \; J = 600 \; MJ \tag{1.2}$$

which is correct, since for example vehicles energies are in the MJ range. So the effect is as large as giving road-sign distances (usually in km) with mm accuracy.

In short, the train can be approximated as a 'material point'. We are very familiar with this approximation in classical mechanics, e.g. it is the same approximation used in describing the motion of the moon around the earth, by namely treating both as point-like objects.

***

Here we have encountered the most elementary example of the idea of **separation of scales** when describing a physics problem. The idea of an effective theory is, correspondingly, the idea of *describing phenomena with finite accuracy*, using their characteristic energy or distance scale to identify the relevant interactions.

Let us now put this idea at work in a quantum-physics system, considering a concrete example, where again we will introduce only the minimal necessary formalism as needed.

L1.2 Rayleigh scattering

PROBLEM: the diffusion of light off the atmosphere, and why the sky looks blue.
Let's analyze the sentence:
• 'diffusion of light': photons with wavelength in the visible: $400 \div 700$ nm $= 4 \div 7 \times 10^{-7}$ m.
• 'off the atmosphere': off atoms. Atom size ~ 1 Å $= 10^{-10}$ m.

The atoms are much smaller than the photon wavelength: their internal structure is not resolved, and they can be treated as point-like objects.

Writing down the interaction

In particle-physics, interactions can be visualized (and in fact also mathematically modeled) using the very intuitive tool of Feynman diagrams.

Feynman diagrams: imagine a process, draw its diagrams, use them to write down the mathematical form of the quantum-mechanical (QM) amplitude for the process.

In our case the diagram would be as depicted in fig. 1.

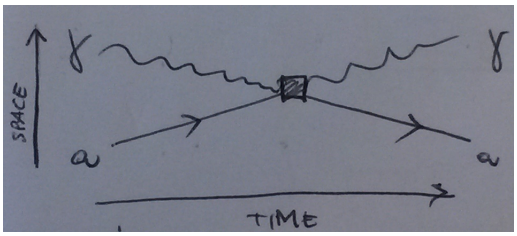


Figure 1: Scattering of photons (γ) off atoms (a).

We need to now write down the QM amplitude corresponding to this diagram. How to mathematically represent particles? In particle-physics, particles are described by **fields**. They are functions of the space-time coordinates, describing the *amplitude* of finding that particle in a given point in space-time. From amplitudes, one can calculates the quantum-mechanical probability $\mathcal{P}$ of a given process as $\mathcal{P} = |\text{amplitude}|^2$. The quantum-mechanical probability is then the analogue of a wave intensity $= |\text{wave amplitude}|^2$ in classical wave mechanics.

Why the field viewpoint

Particle-physics processes are understood (and calculated) using *quantum field theories* (QFTs). QFTs put together physics at very *small scales* (necessitating quantum mechanics) and physics at *high energies* (necessitating relativity). So, QFT marries quantum mechanics with special relativity, and is the mathematical framework for relativistic QM. In relativistic processes there is no conservation of the particle number (because energy can turn into mass and viceversa) and one needs to keep track of the event sequence (causality). Hence the necessity of the 'field' viewpoint.

Very basics on Lagrangians and fields

Interactions, like the one in fig. 1, are built out of products of fields. Interactions are terms of the Lagrangian function. The latter, similarly as in classical mechanics, describes the dynamics of the given physical system.

Lagrangian – In classical mechanics, one introduces the Lagrangian as $S = \int \mathcal{L} dt$, with S the action. The dynamics of the system is obtained by the equations of motion that minimize the action. In QFT the integral is performed over space-time coordinates, d^4x, rather than just over time. Hence $S = \int \mathscr{L} d^4x$. The space-time Lagrangian density is a function of the fields describing the system and their first derivatives: $\mathscr{L} = \mathscr{L}(\phi, d\phi/dx^\mu)$, with μ an index labeling the four space-time coordinates. The symbol ϕ denotes collectively the fields, that are just the 'quantized' analogue of waves in classical mechanics, namely space-time distributions of a certain measurable quantity, like charge, or spin.

Fields – In classical mechanics, we can represent fields in coordinate space as the Fourier transform of the corresponding fields, or amplitudes, in momentum space:

$$\phi(x) \propto \int dp \left(a(p) e^{-ixp} + a^*(p) e^{ixp} \right) \ , \tag{1.3}$$

with $a(p)$ the amplitude for the wave to have momentum p. Quantized fields are 'similar', but for the fact that a are 'quantized', namely they create or destroy one particle with that given momentum.

Note that xp has the dimensions of an action, which is energy $\times$ time. Since the 'quantum of action' is a universal constant, $\hbar$, one can choose physics units where $\hbar = 1$ (i.e. it is treated as dimensionless). One can do the same with the speed of light in the vacuum: $c = 1$.[1] In these units $[p] = [E] = \text{mass} = \text{length}^{-1}$, and $[x] = [t] = \text{length}$.

[1] Note that, since $c \simeq 3 \cdot 10^8$ m/s, taking $c = 1$ operatively just means that, if I chose the second as unity

Writing down the interaction

The atom corresponds to a field with zero charge and angular momentum: a scalar field, indicated as a in fig. 1. What about the photon? The photon is a 'quantum' of electromagnetic field. But what combination of the E and B fields is good to represent the photon in our problem? We need to start from the two unsourced Maxwell's equations:

$$\begin{aligned} \vec{\nabla} \cdot \vec{B} &= 0 \ , \\ \vec{\nabla} \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t} \ , \end{aligned} \tag{1.4}$$

which, as we know, admit the solution

$$\begin{aligned} \vec{B} &= \vec{\nabla} \times \vec{A} \ , \\ \vec{E} &= -\vec{\nabla}\phi - \frac{\partial}{\partial t}\vec{A} \ , \end{aligned} \tag{1.5}$$

with $\vec{A}$ the vector potential and ϕ the scalar potential. Let's look at these two equations. One has for example $E_x = -\partial_x \phi - \partial_0 A_x$, and this makes it natural to identify $\phi \equiv A^0$, the zeroth component of the four-vector

$$A^\mu \equiv \begin{pmatrix} \phi \\ \vec{A} \end{pmatrix} \ . \tag{1.6}$$

So, in relativistic notation[2], the $\vec{E}$ components become

$$E^i = -\partial^0 A^i + \partial^i A^0 = -F^{0i} \ , \tag{1.7}$$

and we are tempted to define the 'e.m. field tensor'

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu \ , \tag{1.8}$$

with $\partial^\mu \equiv \frac{\partial}{\partial x_\mu}$. What do we get for F^{ij}, with $i, j \neq 0$?

$$F^{ij} = \partial^i A^j - \partial^j A^i = -\partial_i A^j + \partial_j A^i = -\epsilon^{ijk} B^k \ , \tag{1.9}$$

where I introduced the antisymmetric symbol ϵ^{ijk}, with $\epsilon^{123} = 1$. The last equality in eq. (1.9) follows by comparison with the first of eqs. (1.5). We therefore see that $F^{\mu\nu}$ bundles together all the components of the electric and the magnetic fields as

$$F^{\mu\nu} \equiv \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ & 0 & -B_z & B_y \\ & & 0 & -B_x \\ & & & 0 \end{pmatrix} . \tag{1.10}$$

Note that the entries below the diagonal are minus the corresponding ones above the diagonal, because $F^{\nu\mu} = -F^{\mu\nu}$, by its definition in eq. (1.8).[3]

of time, then my unity of length would be 10^8 m/s. In this way measurements of length and measurements of time can be identified with one another, because: length = constant × time.

[2] According to this notation, for $i = 1, 2, 3$ A^i denotes the ith component of $+\vec{A}$, whereas A_i denotes minus the same quantity; ∂_i denotes the ith component of $+\partial/\partial\vec{x}$ and ∂^i denotes minus the same quantity. For $i = 0$, the notation is analogous, but for the fact that there are no minus signs around.

[3] It is easy to check (exercise) that the two unsourced Maxwell's equations (1.4) can be written as

$$\partial_\mu \tilde{F}^{\mu\nu} = 0 \ , \quad \text{with} \quad \tilde{F}^{\mu\nu} \equiv \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma} \ , \tag{1.11}$$

Note that the solution in eq. (1.5) is not unique:

$$\begin{aligned} \vec{A}' &= \vec{A} - \vec{\nabla}\alpha(\vec{x},t) = \vec{A} + ie^{-i\alpha}\vec{\nabla}e^{+i\alpha} , \\ \phi' &= \phi + \frac{\partial\alpha}{\partial t} = \phi - ie^{-i\alpha}\frac{\partial}{\partial t}e^{+i\alpha} , \end{aligned} \tag{1.14}$$

with α an arbitrary function of space-time coordinates, would be equally good solutions of eqs. (1.5). In four-dimensional notation, these 'gauge' transformations become simply

$$A'^{\mu} = A^{\mu} + \partial^{\mu}\alpha . \tag{1.15}$$

The crucial point is that in QFT the presence of an e.m. charge or coupling is formally represented by the non-invariance under these gauge transformations. The other crucial point is that $F^{\mu\nu}$ is *invariant* under this transformation: $F'^{\mu\nu} = F^{\mu\nu}$. Therefore, it is the only e.m.-field combination that can couple to electrically neutral objects.

***

Since the atom is electrically neutral, it can only couple to $F^{\mu\nu}$, not to A^{μ} individually. So, the correct building block to describe the photon field is a combination of $F^{\mu\nu}$. This combination must be invariant under space-time transformations (= relativistic invariance). The latter is achieved by the field combination $F^{\mu\nu}F_{\mu\nu}$, with all indices saturated[4]. Quantities with all indices saturated are invariant under space-time transformations in the same way as the scalar product $\vec{u}\cdot\vec{v} \equiv u_i v_i$ ($i = 1,2,3$ or x,y,z), with $\vec{u}$ and $\vec{v}$ two spatial vectors, is invariant under space rotations, represented by orthogonal matrices.

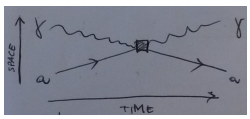
Hence, the atom-photon interaction must be of the form

$$\mathscr{L}_{\rm int} \propto \phi^*\phi\, F^{\mu\nu}F_{\mu\nu} . \tag{1.16}$$

Let us look at fig. 1. In the interaction in eq. (1.16), the field ϕ describes the atom approaching the interaction point (squared box in the figure), the field ϕ^* describes the atom leaving the interaction point, and each of the two powers of $F^{\mu\nu}$ describes the photon field (approaching and respectively leaving the interaction point).

The QM probability $\mathcal{P}$ for the atom-photon scattering is then

$$\mathcal{P} \propto \left| \right. \quad \left. \right|^2 . \tag{1.17}$$



Now, to understand why the sky is blue, we need to work out the dependence of the scattering probability in eq. (1.17) on the photon energy. This is easy to work out. Recall from

---

and that the two sourced Maxwell's equations

$$\begin{aligned} \vec{\nabla}\cdot\vec{E} &= \frac{\rho}{\epsilon_0} , \\ \vec{\nabla}\times\vec{B} &= \frac{\vec{j}}{\epsilon_0} + \frac{\partial\vec{E}}{\partial t} , \end{aligned} \tag{1.12}$$

become

$$\partial_\mu F^{\mu\nu} = j^\nu/\epsilon_0 , \quad \text{with } j^\mu \equiv \begin{pmatrix} \rho \\ \vec{j} \end{pmatrix} . \tag{1.13}$$

[4] $F^{\mu\nu}\tilde{F}_{\mu\nu}$ is an equally good field combination. For simplicity we will however drop this term in our discussion, as this term leads to the very same conclusions.

definition (1.8) that $F^{\mu\nu}$ involves derivatives of the four-vector A^μ. The latter obeys a plane-wave representation entirely similar to eq. (1.3). Therefore:

$$F^{\mu\nu} \propto \partial^\mu e^{ipx} \propto E_\gamma \ . \tag{1.18}$$

As a consequence, the atom-photon diffusion probability will depend on E_γ^4, namely photons with higher energy (close to the blu color) will be diffused much more than those with lower energy (close to the red). This is why the sky looks blue.

Embellishments

In eq. (1.16) we have written a proportionality relation, we have namely omitted to specify the coupling strength. We can actually work it out very simply by just finding its mass dimensions. To this end, let us first rewrite $\mathscr{L}_{\rm int}$ in eq. (1.16) making explicit its coupling strength C

$$\mathscr{L}_{\rm int} = C\phi^*\phi\, F^{\mu\nu}F_{\mu\nu} \ . \tag{1.19}$$

We know that, since the action is dimensionless (with $\hbar = 1$), $[\mathscr{L}] = \text{mass}^4$. Furthermore, $[F^{\mu\nu}] = \text{mass}^2$. What about $[\phi]$? We can work out the a dimension from the Feynman diagram representing the free propagation of the atom, depicted in fig. 2. We know that the atom propagates at non-relativistic speeds, so the diagram in fig. 2 must be of the form $\mathscr{L}_{\rm prop} = \phi^*\,(p^2/2m)\,\phi$, where again ϕ and ϕ^* represent the atom at the beginning and at the end of the propagation path. Since $[\mathscr{L}_{\rm prop}] = \text{mass}^4$ and $[p^2/2m] = \text{mass}$, it follows that $[\phi] = \text{mass}^{3/2}$. Now we have all the ingredients to work out the mass dimension of ϕ. We know that $[C][\phi]^2[F^{\mu\nu}]^2 = \text{mass}^4$, and that $[\phi]^2[F^{\mu\nu}]^2 = \text{mass}^7$. It follows that $[C] = \text{mass}^{-3} = \text{length}^3$. The only length scale in the problem (apart from the photon wavelength, present in $F^{\mu\nu}$) is the atom size a_0, of the order of 10^{-10} m. So the final behavior of the diffusion probability is

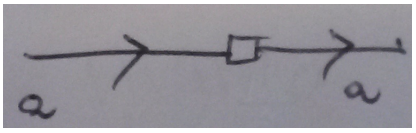


Figure 2: Diagram for the free propagation of an atom a.

$$\mathcal{P} \sim a_0^6 E_\gamma^4 \tag{1.20}$$

which turns out to be right!

***

To recapitulate:

- By general considerations of symmetry and mass dimensions we got the correct photon-energy dependence of the diffusion probability.
- By a simple dimensional argument we also got right the overall normalization: a_0^6. That the scattering probability grows as a power of the atom size is what one expects on the basis of geometrical considerations.
- The whole argument is fully consistent for photon energies much smaller than the inverse atom size a_0^{-1}, which is our case, since we wanted to describe photons in the visible.

- Note that, if the photon wavelength had on the other hand been comparable with the atom size, the photon would have been able to resolve the internal constituents of the atom itself. In this case, new scales (those of the internal atom constituents) would have entered the game, and the question would have arised, which of these scales determines the size of the coupling C. The answer is that, barring dynamical reasons, all of them contribute to C.

Lecture 2: Basics about flavor physics

The example in sec. L1.2 is a metaphor of the way of reasoning in theoretical physics at large, and in flavor physics in particular. In fact, flavor physics makes inherent use of the idea of EFT. The reason is the following: flavor physics is concerned with processes of the kind depicted in fig. 3. In this process, the initial state includes a quark of mass m_i, which interacts in the black box with some other particle (not depicted) to yield a final-state quark of mass m_j *different* than m_i. Flavor violation is exactly the fact that $m_j \neq m_i$. (In general there will be other quarks participating in the process as 'spectators', that namely do not change flavor. These quarks are likewise depicted in the figure.) The mass scales involved

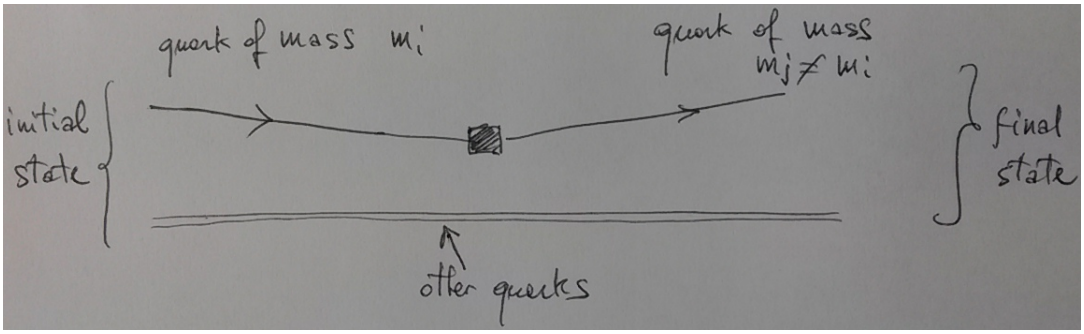


Figure 3: Schematic representation of a flavor-violating process.

can be paralleled to those entering the discussion of Rayleigh scattering. The crucial point is, in particular, that the energy scale m_W involved in the interaction represented by the black box is much larger (i.e. the associated distance m_W^{-1} much shorter) than the energy scale of the external states, of order $m_q = \max(m_i, m_j)$, in the same way as the atom energy scale a_0^{-1} is much larger than the photon energy E_γ. To perform calculations in flavor physics it is therefore necessary to use lines of reasoning similar to that used in the previous section, i.e. it is necessary to build an EFT.

***

Before going in more detail into flavor-physics interactions, we should address the question **why is flavor physics interesting** at all. The reason is that the initial and final quarks give rise to external states (the simplest ones known as mesons, bound states of a quark and an anti-quark, the binding due to strong interactions) that one can produce copiously at colliders, at relatively cheap costs. Their decay rates can thereby be accurately measured. Many of them can also be accurately calculated by the techniques of EFTs. *One can in this way probe 'indirectly' the high-energy interactions represented by the black box.* Indirectly means that one does not need to directly reach the energy scales (indicated by m_W) involved in the black-box interaction.

The rest of this lecture is devoted to understanding more closely how the process in fig. 3 can arise at all. To make the discussion self-contained, we will first make a short presentation of the elementary constituents of matter, and then of their flavor-violating interactions.

L2.1 Matter constituents and interactions

As mentioned, the process in fig. 3 involves quarks. Quarks and what else? And what are quarks at all? Quarks are among the 'basic' constituents of matter. Basic means here elementary, namely structure-less, or point-like, for what we know. The only matter constituents known so far are quarks and leptons.

Quarks: particles of electric charge $+2/3$ (up-type quarks: up, charm, top) or $-1/3$ (down-type quarks: down, strange, bottom). They also have intrinsic angular momentum, or spin, $1/2$. They are schematized in fig. 4. The different instances of quarks u, d, s, c, b, t are called 'flavors', and they differ only by their mass, increasing in the list u, d, s, c, b, t. The heavier flavors tend to decay to the lighter ones, so only u and d quarks, the lightest, can give rise to stable states.

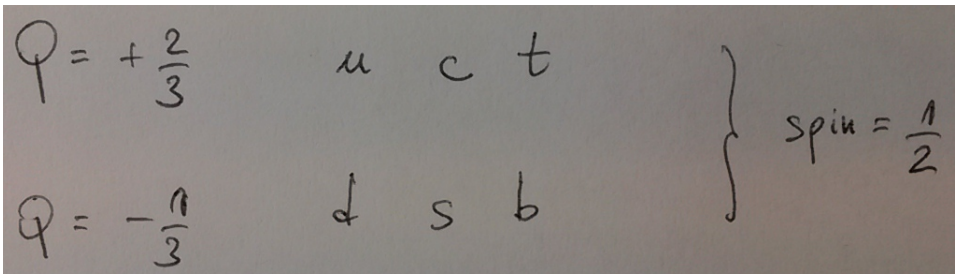


Figure 4: Quarks

For completeness, besides quarks there are leptons, which are entirely analogous to quarks apart from:

- different charge assignments: up-type leptons (neutrinos) have charge 0, down-type leptons (electron e, muon μ, tau τ) have charge -1;
- lepton masses are different than quark masses, and flavor by flavor they are smaller (e.g. $m_e < m_d$ etc.);
- quarks interact strongly, weakly and electromagnetically. Leptons do not interact strongly.

While also leptons have a vast flavor phenomenology, in these lectures we will focus on quarks.

Vector bosons. Strong, weak and electromagnetic interactions are represented by interactions where quarks meet with 'vector bosons', respectively gluons g (strong), massive vectors W, Z (weak) and photons γ (e.m.).

Higgs boson. Within our current understanding, any elementary particle that has mass, receives it via its interaction with the Higgs boson. (Neutrinos are a possible exception, but this is irrelevant here.) The Higgs boson is described by a scalar field, the only field that is invariant under space-time transformations.

***

In short, quarks have the following interactions:

$$\begin{array}{ll} \text{e.m.} & \bar{Q}\gamma Q \ , \\ \text{strong} & \bar{Q}gQ \ , \\ \text{weak} & \bar{Q}WQ' \ , \bar{Q}ZQ \ , \\ \text{Higgs} & \bar{Q}HQ' \ , \end{array} \tag{2.1}$$

where the field $\bar{Q}$ creates a quark or annihilates an anti-quark, and the field Q does the opposite. Note that each interaction with one boson (γ, g, W, Z, H) involves two fermions. This is in order to globally conserve charge and angular momentum. *Flavor physics arises from the interplay between the last two classes of interactions:* weak and with the Higgs. Let us see this in more detail.

L2.2 Flavor-violating interactions

Weak and Higgs interactions – Let us write down the quark-W interactions:

$$\mathscr{L}_{qqW} \;\propto\; \bar{U} W^{(-)} D \;+\; \bar{D} W^{(+)} U \;, \tag{2.2}$$

where D denotes either of d, s, b quarks and U either of u, c, t quarks. So for example the first interaction corresponds to the diagram in fig. 5, where an initial d yields a u and a W. The parentheses indicate the electric charges. Similarly, there are quark-quark-Z interactions:

$$\mathscr{L}_{qqZ} \;\propto\; \bar{U} Z^{(0)} U \;+\; \bar{D} Z^{(0)} D \;, \tag{2.3}$$

as well as quark-quark-Higgs interactions. Avoiding unnecessary details,[5] the latter interactions have the basic form

$$\mathscr{L}_{qqH} \;\propto\; \bar{U}_i (y_u)_{ij} H U_j \;+\; \bar{D}_i (y_d)_{ij} H D_j \;, \tag{2.4}$$

where y_u, y_d are two 3×3 complex matrices, and i, j label the flavor: $D_1 = d, U_2 = c$, etc. This is the crucial difference with respect to the W, Z interactions in eqs. (2.2)-(2.3): qqW and qqZ interactions weigh equally quarks of different flavors, because of a 'gauge' symmetry similar to that in electromagnetism. On the other hand, no analogous symmetry is known for qqH, so the latter can have the most general couplings y_u and y_d, that weigh differently different combinations of quarks.

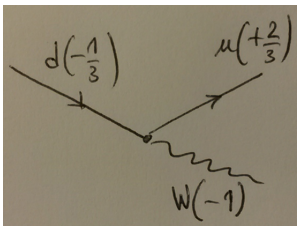


Figure 5: Example of quark-quark-W interaction.

Quark masses – As mentioned, quark receive their masses from their interactions with the Higgs, so mass terms must arise from eq. (2.4). Similarly as in the diagram of fig. 2, quark masses are defined at the Lagrangian level by terms of the kind $\bar{u} m_u u$, $\bar{d} m_d d$, etc.[6] Recalling that the Higgs is a scalar, we know how to obtain these mass terms from the interactions in eq. (2.4):

- Allow the Higgs field to take a non-null energy density even in the vacuum. We can do this only with the Higgs, because it is a scalar (the only fundamental scalar):

  $$H \to v + h \;, \tag{2.5}$$

  with v the vacuum density and h the fluctuations around this density. So, eq. (2.4) contains

  $$\mathscr{L}_{qqm} \;\propto\; \bar{U}_i (y_u)_{ij} v U_j \;+\; \bar{D}_i (y_d)_{ij} v D_j = \bar{U}_i (M_u)_{ij} U_j \;+\; \bar{D}_i (M_d)_{ij} D_j \;. \tag{2.6}$$

- Diagonalize the matrices M_u and M_d. Quark masses will then be:

  $$\hat{M}_{u,d} = \{\text{eigenvalues of } M_{u,d}\} \;, \tag{2.7}$$

  where

  $$\hat{M}_u = \text{diag}(m_u, m_c, m_t) \;, \qquad \hat{M}_d = \text{diag}(m_d, m_s, m_b) \;. \tag{2.8}$$

Diagonalization of $M_{u,d}$ – For any complex matrix M, there exist two unitary matrices X_L and X_R such that

$$X_L^\dagger M X_R = \hat{M} \tag{2.9}$$

with $\hat{M}$ diagonal and real.

[5] Like spin projectors and charge-conjugation operators for the Higgs.

[6] We associated to fig. 2 the non-relativistic kinetic energy term $\phi^* \left(p^2/2m\right) \phi$. Here terms like $\bar{u} m_u u$ are associated to the rest energy of the given particle, namely for the u quark $m_u c^2$, with $c = 1$.

Proof by explicit construction.

- Rewrite $M = HU$, with H Hermitian and U unitary (polar decomposition theorem, like writing a complex number z as $|z|e^{i\arg(z)}$).
- Let then $H = Z^\dagger \hat{H} Z$, with $\hat{H}$ diagonal and Z unitary (this decomposition exists by definition, since H is Hermitian).
- Then eq. (2.9) becomes $X_L^\dagger \underbrace{(Z^\dagger \hat{H} Z U)}_{M} X_R = \hat{M}$. Choosing the two unitary matrices X_L and X_R as $X_L = Z^\dagger$ and $X_R = U^\dagger Z^\dagger$ one sees that $\hat{M} = \hat{H}$.

Using eq. (2.9), we can rewrite the up-quark mass term in eq. (2.6) as (I suppress flavor indices for simplicity here)

$$\bar{U} M_u U = \bar{U} X_L \hat{M}_u X_R^\dagger U \ . \tag{2.10}$$

Therefore, we just need to redefine the up-quark fields as

$$U = X_R U' \ , \qquad \bar{U} = \bar{U}' X_L^\dagger \tag{2.11}$$

and perform analogous transformations for the D fields, and we'll get, for eq. (2.10)

$$\bar{U} M_u U = \bar{U}' \hat{M}_u U' \tag{2.12}$$

where, we recall, $\hat{M}_u$ is the diagonal matrix containing the u, c, t masses as entries (see the first of eqs. (2.8)). The primed quark fields are eigenstates of mass.

***

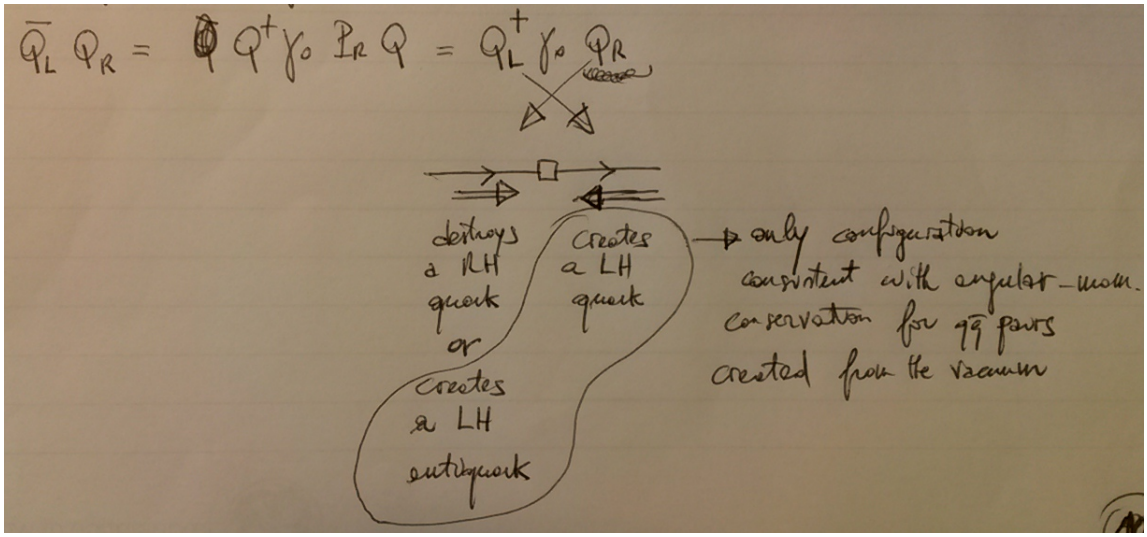


Figure 6: Spin configurations for quark-quark pairs in the interactions (2.6).

The careful reader may wonder whether we are allowed to redefine the fields U and $\bar{U}$ by two independent matrices, as in eq. (2.11). The answer is yes. The reason is the following. Recall that quarks have spin 1/2. This quantum number is such that U and D quarks get a further degree of freedom depending on whether their spin is oriented parallel or antiparallel to their direction of motion. These two possibilities are labelled as R (for right-handed) and L (for left-handed) respectively. The quark fields are then U_R, U_L, D_R, D_L. Each of them is an independent field.

It turns out that the quarks annihilated by U and those created by $\bar{U}$, respectively, in the interaction in eq. (2.6), need to have opposite projections of their spin along their direction of motion. A more in-depth motivation for this statement is depicted in fig. 6.

Flavor-changing interactions, finally –

- The interactions (2.6) are actually (we suppress the explicit indices i, j for simplicity)

$$\bar{U}_L(M_u)U_R \;+\; \bar{D}_L(M_d)D_R \tag{2.13}$$

because they are interactions with a scalar, and the argument given in fig. 6 holds. So, combinations are of the kind LR or RL;

- Conversely, interactions with vector bosons, like the W, Z, γ, g will involve $\bar{Q}Q$ pairs in combinations LL or RR (or both).
- To diagonalize M_u, M_d in eq. (2.13), redefine:

$$U_L = X_L^{(u)} U'_L \quad \text{and} \quad U_R = X_R^{(u)} U'_R \tag{2.14}$$

and analogous for the $D_{L,R}$ fields, this time with $X_{L,R}^{(d)}$ matrices. Then mass terms will be

$$\bar{U}'_L \underbrace{\left[\left(X_L^{(u)}\right)^\dagger M_u X_R^{(u)}\right]}_{\hat{M}_u} U'_R \;+\; \text{analogous for } D \text{ quarks} \;. \tag{2.15}$$

- What happens to the other interactions in eq. (2.1)? Strong, e.m. and weak interactions with the Z, that are all charge-neutral, involve quarks in the combinations $\bar{U}_L U_L$ or $\bar{U}_R U_R$ or the analogous with the D quarks. The redefinitions in eq. (2.14) will then yield products $(X_L^{(u)})^\dagger X_L^{(u)} = \mathbb{1}$ or $(X_R^{(u)})^\dagger X_R^{(u)} = \mathbb{1}$, hence they will have no effect.
- On the other hand, in quark interactions with the W, one will have

$$\bar{U}_L W^{(-)} D_L = \bar{U}'_L \underbrace{\left[\left(X_L^{(u)}\right)^\dagger X_L^{(d)}\right]}_{V \neq \mathbb{1}!} W^{(-)} D'_L \;. \tag{2.16}$$

The matrix $V \equiv \left(X_L^{(u)}\right)^\dagger X_L^{(d)}$ is a 3×3 unitary matrix, and is in general not diagonal.

This matrix is known as the Cabibbo-Kobayashi-Maskawa matrix. If it were diagonal, the interaction (2.16) would allow only 'vertical' transitions (see fig. (4)): $d \leftrightarrow u$, $s \leftrightarrow c$ and $b \leftrightarrow t$. Instead, because of its off-diagonal entries, also any 'oblique' transition in this figure is allowed, e.g. $d \leftrightarrow t$ or $s \leftrightarrow u$.

- As a result, $\bar{U} W^{(-)} D$ interactions will be of the form depicted in fig. 7, namely interactions with any down-type quark as initial state and any up-type quark as final state (or viceversa), with coupling strengths ruled by the V-matrix entries.
- *All of flavor-physics phenomena involving quarks arise from the basic interaction in fig. 7.*

In the next section we will play with this interaction to build actual flavor processes.

L2.3 Examples of flavor-violating processes

Example 1 – A first example of flavor-physics phenomena that can be built out of the basic interaction in fig. 7 is depicted in fig. 8. In this reaction, an initial-state b-quark undergoes an interaction where two 'virtual' particles, a W and a t quark, are emitted and reabsorbed, and the final-state quark is not a b, but an s. Namely the initial- and final-state down-type

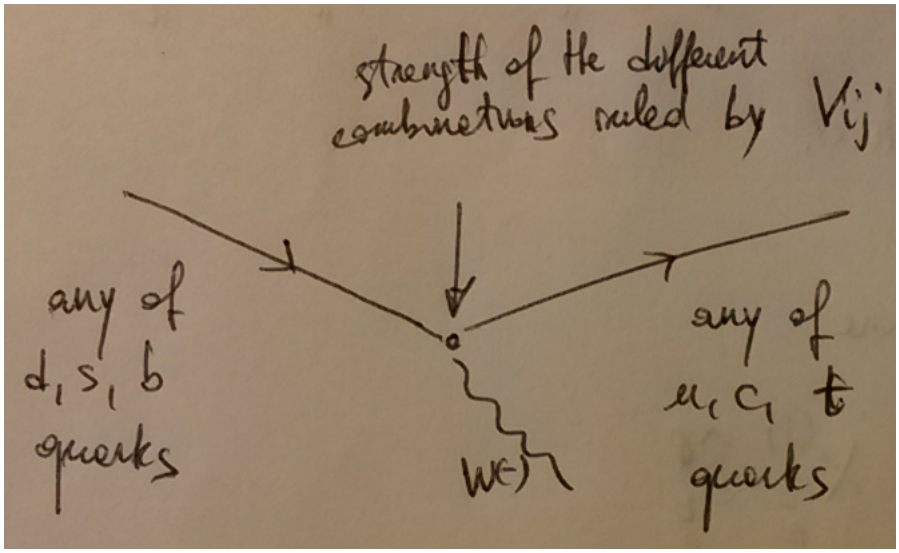


Figure 7: The interaction responsible for all flavor processes involving quarks.

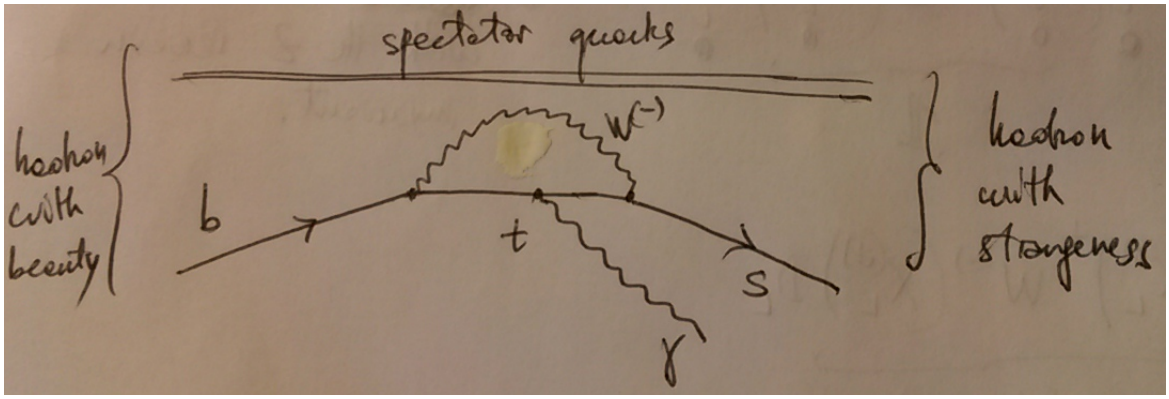


Figure 8: Diagram for the process $b \to s\gamma$.

quarks have different masses, namely flavors. The virtual top (or also the W) can emit a photon, also detected in the final state. The initial b and the final s will form bound states with other 'spectator' quarks, giving rise to hadrons with 'beauty' (containing the quark b) or with 'strangeness' (containing the quark s). An example is $B \to K^*\gamma$, very well measured.

The peculiarity of this kind of processes is that they are *electrically neutral* (in the above example the b and the s have the same charge $-1/3$) but they involve different 'flavors' (i.e. different masses) in the initial and final states. These processes are called **flavor-changing neutral currents** (FCNC). They can occur only at the quantum level (= involve the emission and reabsorption of 'virtual' states, as in fig. 8). In the Standard Model (SM), they are *forbidden at the classical level*. As such, these processes are generally very rare, and excellent to probe possible effects beyond the SM ones.

Example 2: particle-antiparticle oscillations – Another, even more peculiar flavor phenomenon is depicted in fig. 9. According to this diagram, one can observe a meson containing a $d, \bar{s}$ pair, known as K^0 meson, oscillate into its own anti-particle, containing namely a $\bar{d}, s$ pair, and known as $\bar{K}^0$. Again, it involves the emission and reabsorption of virtual t and W particles. Such closed diagrams are known as 'loops'. This process will be the topic of lecture 3.

Why these processes are calculable – An important point about processes like in fig.

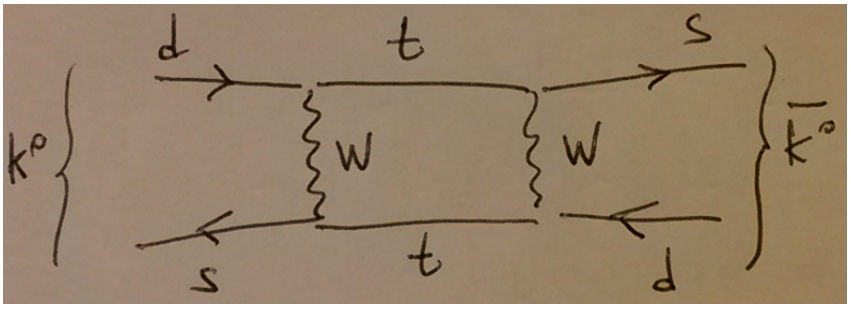


Figure 9: Oscillation of a K^0 meson into its antiparticle, the $\bar{K}^0$.

8 and 9 is the fact that the 'virtual' particles in the loops are much, much heavier than the external particles. For example, in $K^0 - \bar{K}^0$ mixing, one has:

$$\begin{aligned} m_d &= 8 \text{ MeV} , \quad m_s = 100 \text{ MeV} , \\ m_t &= 175 \cdot 10^3 \text{ MeV} , \quad m_W = 80 \cdot 10^3 \text{ MeV} , \end{aligned} \tag{2.17}$$

where the MeV is a unit of energy, or mass (recall $E = mc^2$, with $c = 1$) convenient in particle physics: 1 MeV $\simeq 2 \times 10^{-30}$ kg!

Therefore, at the energy scale (or 'speed') of the external particles (the d and s, or actually the K^0 and $\bar{K}^0$, which are the observable states), the internal particles t and W are motionless to an excellent approximation (try and compare with the examples of the train or of Rayleigh scattering).

As a consequence, the 'box' loop in fig. 9 can be approximated with a point. The plausibility of this situation is depicted in fig. 10.

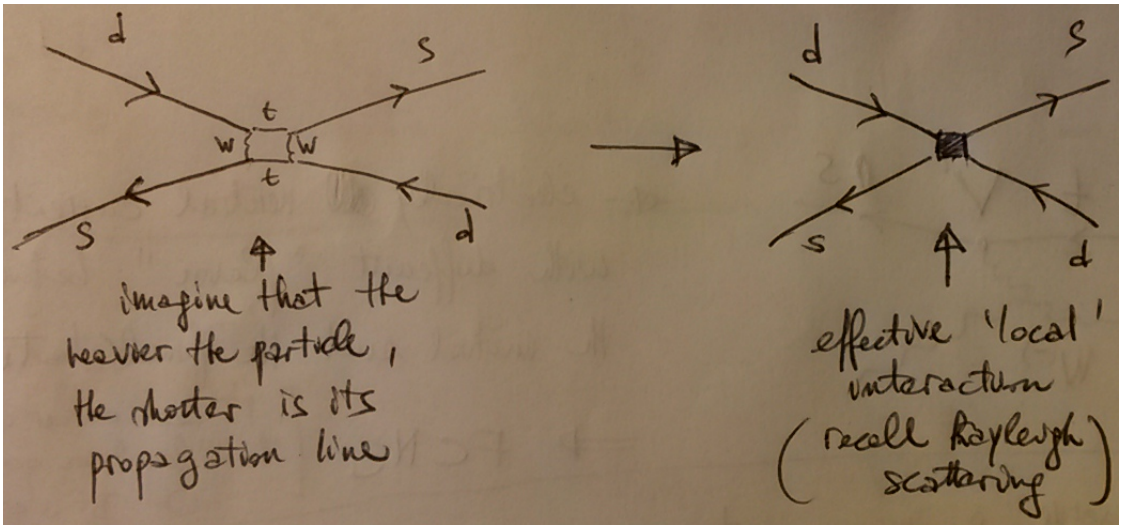


Figure 10: Modeling $K^0 - \bar{K}^0$ oscillations as a point-like interaction.

***

Take-home message – Flavor processes are calculable *if very massive dynamics in the loops can be bundled in 'local' or point-like interactions.* The overall strength of these interactions can in these cases be either calculated, or measured from experimental processes.

On the other hand, non-local loops are very, very difficult objects to calculate in quantum field theory.

Fermi theory – The above reasoning is exactly the one that was originally followed by Enrico Fermi, who pioneered the understanding of weak interactions.

Example: Fermi's theory of β decay (1933, as he was professor at the Rome University). To understand the radioactive reaction $n \to pe^-\bar{\nu}$, he wrote down an 'effective' interaction with 4 fermions at the vertex, exactly as the rightmost diagram in fig. 10. Thereafter, he determined the overall coupling strength from data. We modernly understand the Fermi interaction as depicted in fig. 11.

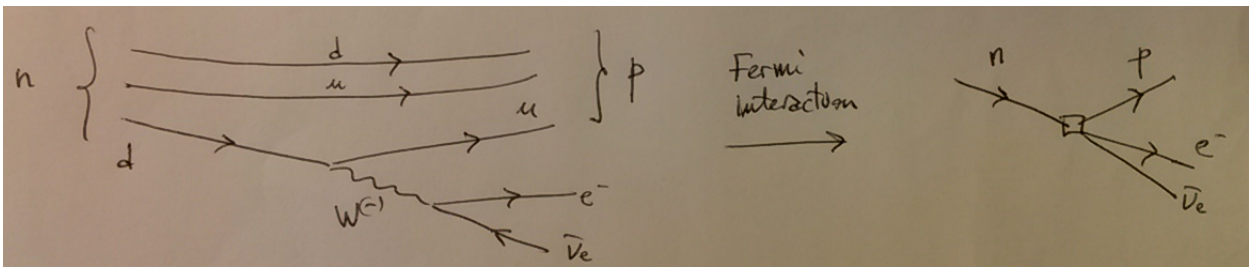


Figure 11: The β-decay of a neutron within the Fermi theory.

Lecture 3: Violation of the CP symmetry

The previously introduced matrix V (= Cabibbo-Kobayashi-Maskawa, or CKM matrix) is actually responsible for:

- all the processes with flavor violation[7];
- all the (so far observed) processes that violate a discrete symmetry called CP.

The CP operation on a physical object stands for charge conjugation (C) plus a parity transformation (P).

- **C interchanges particles and antiparticles.** So C conservation means that the rate for a process equals the rate for the same process with particles replaced by their respective antiparticles. C is violated in weak interactions,[8] but conserved in e.m. and strong ones.
- **P transforms $\vec{x} \to -\vec{x}$, if $\vec{x}$ is a vector.** On the other hand, pseudovectors, like angular momenta (hence also spins), are left unchanged. Therefore, P exchanges left-handed (LH) with right-handed (RH) particles and viceversa.
- **CP performs the two operations together.** So, in particular it exchanges a right-handed e^- into a left-handed e^+, and a left-handed e^- into a right-handed e^+.

Why CP violation is important. Note that, in the absence of CP violation, LH protons would always balance against RH antiprotons, and viceversa. Therefore, CP violation is required at some point in the history of the universe to generate the observed matter-antimatter imbalance.

As a matter of fact, CP violation is one of the three conditions enunciated by Sakharov in 1967 for matter-antimatter asymmetry to be possible. He was inspired by the experimental discovery of CP violation in the $K^0 - \bar{K}^0$ system, that we now have the instruments to explore.

L3.1 The $K^0 - \bar{K}^0$ system

• The $K^0 - \bar{K}^0$ system is a system of two mesons. As mentioned at the beginning of Lecture 2, mesons are bound states of a quark and an antiquark, bound together by strong interactions. One has:

$$K^0 \sim \begin{pmatrix} d \\ \bar{s} \end{pmatrix} , \qquad \bar{K}^0 \sim \begin{pmatrix} \bar{d} \\ s \end{pmatrix} . \tag{3.1}$$

We have previously drawn the diagram in fig. 9. So the K^0 can become its antiparticle and viceversa. Note that this can only happen via weak interactions (only at the 'loop' level). No other SM interaction is able to generate this kind of process.

• Because the K^0 and the $\bar{K}^0$ oscillate into each other, they are not 'well defined' particles. In fact, a very interesting quantum-mechanics (QM) problem was the one of defining the physical particles associated with the $K^0 - \bar{K}^0$ system.

• A good candidate for a particle is an object with definite mass and certain 'quantum numbers', which express its properties under certain conservation laws. E.g. an electron has a mass of 0.5 MeV, has charge -1 (in units of e), spin 1/2 (in units of $\hbar$), etc.

[7] By this we mean either 'oblique' transitions in fig. 4 or FCNCs.
[8] E.g. left-handed e^- and left-handed e^+ don't behave in the same way in the SM.

• In QM, each of a particle's properties is mathematically described by an eigenvalue of a suitable operator, of which the particle 'state' is an eigenstate. For example, for the electron mass:

$$\mathscr{H}\,|e(\vec{p}=0)\rangle \;=\; m_e\,|e(\vec{p}=0)\rangle \tag{3.2}$$

where $\mathscr{H}$ is the Hamiltonian operator: as we know, it is the function expressing the energy of a system. The system is an electron at rest, $|e(\vec{p}=0)\rangle$. So its eigenvalue is the rest energy of the electron, its mass.

• We may then say that a good 'particle' ($\Rightarrow$ definite mass) is an eigenstate of the relevant $\mathscr{H}$.

• How to define good physical states for the $K^0 - \bar{K}^0$ system then? Gell-Mann and Pais (those who solved this problem in 1955) started from observing that C exchanges K^0 and $\bar{K}^0$:

$$\begin{aligned} \mathrm{C}|K^0\rangle &= |\bar{K}^0\rangle \ , \\ \mathrm{C}|\bar{K}^0\rangle &= |K^0\rangle \ . \end{aligned} \tag{3.3}$$

Then, the states defined as

$$|K_\pm\rangle \equiv \frac{|K^0\rangle \pm |\bar{K}^0\rangle}{\sqrt{2}} \tag{3.4}$$

would be states of definite eigenvalue ± 1 under C, respectively, and *maybe* a good basis of physical states.

• However, the relevant Hamiltonian is the weak Hamiltonian (the diagram in fig. 9 arises from weak interactions), and charge conjugation is *not a good conserved property* of weak interactions.

• A better choice for the operator exchanging $K^0 \leftrightarrow \bar{K}^0$ is CP, which is (almost) conserved by weak interactions. Note: since K mesons have zero spin, CP is basically like C. Namely one can always write (by a suitable redefinition of the K^0, $\bar{K}^0$ fields)

$$\begin{aligned} \mathrm{CP}|K^0\rangle &= |\bar{K}^0\rangle \ , \\ \mathrm{CP}|\bar{K}^0\rangle &= |K^0\rangle \ , \end{aligned} \tag{3.5}$$

and define

$$|K_\pm\rangle \equiv \frac{|K^0\rangle \pm |\bar{K}^0\rangle}{\sqrt{2}} \tag{3.6}$$

such that

$$\mathrm{CP}|K_\pm\rangle \equiv \pm|K_\pm\rangle \ . \tag{3.7}$$

The $|K_\pm\rangle$ is thus said to be CP-even and CP-odd, respectively.

• Now, in weak interactions CP is not exact, it is slightly violated. Hence the $|K_\pm\rangle$ are approximate, but not perfect physical eitenstates. We can represent this imperfection by defining

$$\begin{aligned} |K_S\rangle \;&\sim\; |K_+\rangle \;+\; \bar{\epsilon}\,|K_-\rangle \\ |K_L\rangle \;&\sim\; |K_-\rangle \;+\; \bar{\epsilon}\,|K_+\rangle \end{aligned} \tag{3.8}$$

where the $|K_S\rangle$ and $|K_L\rangle$ are the true physical eigenstates. As we send $\bar{\epsilon} \to 0$ the $|K_+\rangle$ and $|K_-\rangle$ coincide with the $|K_S\rangle$ and $|K_L\rangle$, respectively. This $\bar{\epsilon}$ quantifies the amount of CP violation. (The origin of the suffix S and L will soon be clear.)

• We can distinguish the $|K_+\rangle$ and $|K_-\rangle$ components inside the physical states by looking at *decays to states that are CP eigenstates*. E.g.

$$\begin{aligned} |2\pi\rangle &= \text{CP even} \ , \\ |3\pi\rangle &= \text{CP odd} \ . \end{aligned} \tag{3.9}$$

Recalling the CP properties of $|K_+\rangle$ and $|K_-\rangle$ we expect[9]

$$|K_+\rangle \ \to \ |2\pi\rangle \quad \text{but} \quad |K_+\rangle \ \not\to \ |3\pi\rangle \ , \tag{3.10}$$

and conversely

$$|K_-\rangle \ \to \ |3\pi\rangle \quad \text{but} \quad |K_-\rangle \ \not\to \ |2\pi\rangle \ . \tag{3.11}$$

• By using eqs. (3.10) and (3.11) into the definition of the physical states $|K_{S,L}\rangle$, we then expect that

$$\textit{the } |K_S\rangle \textit{ will decay most of the time to } |2\pi\rangle, \textit{ and sometimes to } |3\pi\rangle, \tag{3.12}$$

whereas conversely

$$\textit{the } |K_L\rangle \textit{ will decay most of the time to } |3\pi\rangle, \textit{ and sometimes to } |2\pi\rangle. \tag{3.13}$$

In these sentences, 'sometimes' quantifies the amount of CP violation.

• Therefore, the ratio

$$\frac{N_{\text{events}}(K_L \to 2\pi)}{N_{\text{events}}(K_L \to 3\pi)} \tag{3.14}$$

will quantify the amount of CP violation. This is indeed the way CP violation was originally discovered for the first time.

L3.2 The discovery of CP violation (1964): the Cronin-Fitch experiment

CP violation was experimentally discovered in 1964 (and awarded the Nobel prize in 1980) in an experiment that measured the rate in eq. (3.14).

• First problem: how to distinguish the K_S from the K_L experimentally? The answer is in eqs. (3.12) and (3.13). The K_S decays most of the time into 2π, whereas the K_L 'has to' decay most of the time into 3π. Therefore, the K_L decay is 'less easy'. In QM decays to fewer bodies are kinematically easier, and this is intuitively understandable. So the K_S is short-lived whereas the K_L is comparably long-lived, and now we understand the S and L subscripts.

• It is now clear how this observation may be used experimentally. This is depicted in fig. 12.

L3.3 More on CP violation within the SM

In connection with fig. 7 we have stated that all flavor-physics phenomena involving quarks arise from the basic interaction in eq. (2.16). We would like to now expand on this statement.

• By tracing back the line of reasoning leading to eq. (2.16), we see that, in order to have a non-trivial CKM matrix, $V \neq 1$, we need off-diagonal entries in the y_u and y_d matrices. In other words

$$y_{u,d} \text{ off-diagonal} \ \Leftarrow \ \text{flavor violation} \ . \tag{3.15}$$

[9] This statement is not strictly true, but the fully correct statement is irrelevant here (and too advanced).

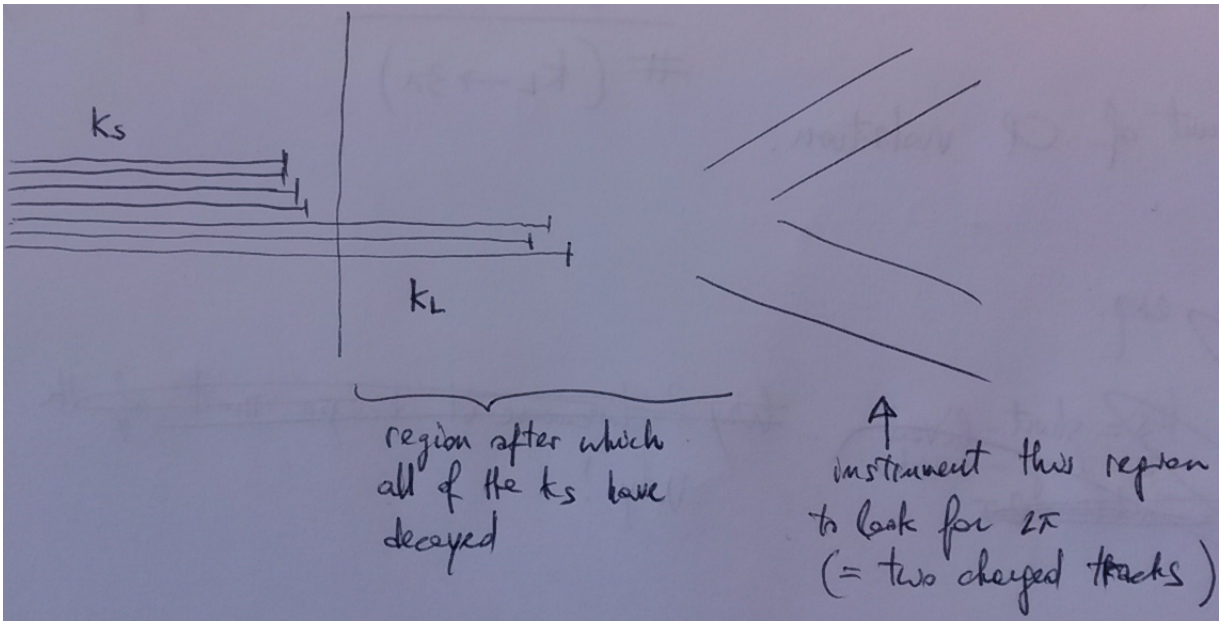


Figure 12: How to measure CP violation at home.

Also all the observed CP violation is, within the SM, caused by the y_u and y_d matrices. In particular[10]

$$y_{u,d} \text{ complex } \Leftarrow \text{ CP violation } . \tag{3.16}$$

• **Unphysical CP violation** – The previous sentence needs actually some qualifications. In fact, in Lecture 2 we have seen that one can redefine quark fields by unitary transformations, as in eq. (2.14). So one may wonder whether phases in y_u and y_d may always be reabsorbed by suitable redefinitions of the quark fields.

Let us make an example. Suppose only the quarks u, d, c, s existed, namely that there were only two generations (see fig. 4). Then, consider the interaction in eq. (2.16) for just two generations, namely with $V = V_{2\times 2}$, and with $U = (u, c)$ and $D = (d, s)$. Let us write the CKM matrix explicitly as

$$V_{2\times 2} = \begin{pmatrix} V_{ud} & V_{us} \\ V_{cd} & V_{cs} \end{pmatrix} . \tag{3.17}$$

Then the $\bar{U}D$ combinations present in the interaction in eq. (2.16) can be schematically written as

$$\begin{pmatrix} \bar{u} & \bar{c} \end{pmatrix} \cdot \begin{pmatrix} V_{ud} & V_{us} \\ V_{cd} & V_{cs} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \end{pmatrix} . \tag{3.18}$$

Now, even if V_{ud} and V_{us} are complex, their phases can be moved into the definition of the d and s fields, respectively. Therefore, we can take V_{ud} and V_{us} as real. Furthermore, $V_{2\times 2}$ is unitary, namely

$$V^\dagger \cdot V = \begin{pmatrix} V_{ud} & V_{cd}^* \\ V_{us} & V_{cs}^* \end{pmatrix} \cdot \begin{pmatrix} V_{ud} & V_{us} \\ V_{cd} & V_{cs} \end{pmatrix} = \mathbb{1}_{2\times 2} . \tag{3.19}$$

[10] Showing the last statement requires somewhat too advanced tools for this course, so we will leave this statement without proof (but ask if you feel like).

(Note that we have dropped complex conjugation signs in V_{ud} and V_{us} as they are real.) The 12 relation reads $V_{ud}V_{us} + V_{cd}^* V_{cs} = 0$, implying that V_{cd} and V_{cs} must have equal phases, because the product $V_{cd}^* V_{cs}$ must be real. Both these V entries multiply $\bar{c}$ in eq. (3.18), so their common phase can be moved into the definition of the c quark.

So we see that, for two generations, V can always be defined as real, hence there is no physical CP violation.

• **Conditions for physical CP violation** – Let us generalize the previous reasoning to N generations, namely to the case of $2N$ quarks. Within the SM (see fig. 4), $N = 3$.

Recalling that it is unitary, the CKM matrix V has N^2 parameters in the $N \times N$ case. Out of them, $N(N-1)/2$ are Euler angles.[11] The rest, $N^2 - N(N-1)/2 = N(N+1)/2$, are phases. However, not all of them are physical. Generalizing the 2×2 example to N generations, one sees that $2N-1$ phases can be moved into the definition of all but one quark fields. So, for $N = 3$ one will have 6 phases, but one can make 5 quark-field redefinitions, implying one single physical phase. (This argument earned the Nobel prize to Kobayashi and Maskawa in 2008.)

In conclusion

$$\begin{aligned} &\text{CP violation is only possible for } N \geq 3. \\ &\text{Within the SM (3 generations) CP violation is due to one single phase.} \end{aligned} \tag{3.20}$$

***

The statement in (3.20) implies that the amount of CP violation predicted by the SM can (in principle) be univocally determined by measuring one single CP-violating quantity. Any other CP-violating quantity can then be predicted. Hence, measuring several CP-violating observables allows to test the SM mechanism of CP violation mentioned above.

As a matter of fact, CP violation has been determined in several other processes after the process in fig. 9. For example, in $B^0 - \bar{B}^0$ oscillations, that are completely analogous to fig. 9, but for the fact that the s quark is replaced by a b quark.

The test of CP violation mentioned above can be visualized in the plots of fig. 13. Each of the bands on these plots represents an observable. If the amounts of CP violation predicted by each observable are compatible with each other, then all the curves should intersect at one point. And it looks they do! This is a very non-trivial test of validity of the SM mechanism of CP violation, to the current level of accuracy of about 15%.

[11] Recall: for 2×2 matrices, there is one rotation, i.e. one such angle; for 3×3 matrices, the angles are 3; for $N \times N$, they are as many as the number of elements above the diagonal, namely $N(N-1)/2$.

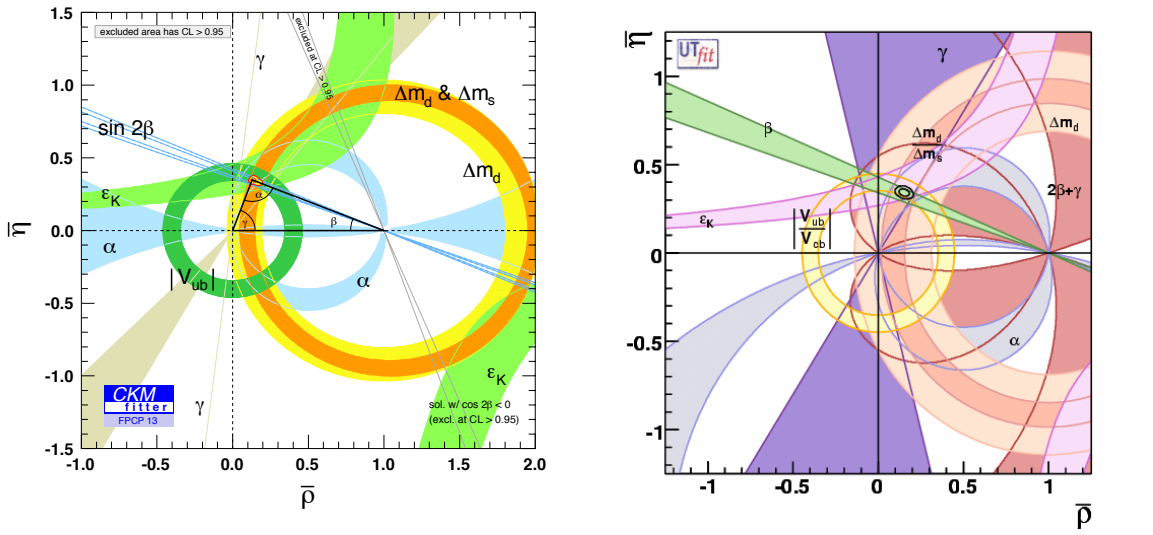


Figure 13: Status of the test of the SM mechanism of CP violation (and of flavor violation in general). Figures taken from `ckmfitter.in2p3.fr` and `utfit.org`, respectively.