Master's thesis
Degree Programme in Physical Sciences (Theoretical Physics)

Theories of Neutrino Oscillations

Emad Shojaeifar
April 2015

Supervisor: Doc. Anca Tureanu
Reviewers: Doc. Anca Tureanu
Prof. Katri Huitu

UNIVERSITY OF HELSINKI
DEPARTMENT OF PHYSICS

P.O. Box 64 (Gustaf Hällströmin katu 2)
00014 University of Helsinki

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta/Osasto — Fakultet/Sektion — Faculty/section	Laitos — Institution — Department
Faculty of Science	Department of Physics

Tekijä — Författare — Author

Emad Shojaeifar

Työn nimi — Arbetets titel — Title

Theories of Neutrino Oscillations

Oppiaine — Läroämne — Subject

Theoretical Physics

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Master's thesis	April 2015	112

Tiivistelmä — Referat — Abstract

The best renormalizable theory in particle physics is the Standard Model. According to this model, neutrinos must be massless. We know that neutrinos have three different flavors. Owing to the observations, it is seen that there are oscillations between different flavors of neutrinos. This fact forces us to go beyond the Standard Model of Particle Physics, because neutrino oscillations can only happen if the neutrinos are massive. The first solution is to define three mass eigenstates and claim that the flavor eigenstates are superpositions of the mass eigenstates.

However, due to some problems it is concluded that physical states must be replaced with fields. In this thesis we study neutrino oscillations. We begin our work with reviewing fermion fields before the Standard Model. Then the Standard Model and experiments which confirm the existence of different neutrino flavors are shortly discussed. Moreover, we present the observations that show neutrino oscillations phenomenon and after that, the Dirac and Majorana neutrinos are discussed. Finally, oscillations are described theoretically. First, quantum mechanical oscillations are discussed and we find the phase and probability of the flavor transitions. Then due to the problems of this approach, oscillations in quantum field theory are presented. It is seen that, again, the results derived with the quantum mechanical treatment are accepted in the quantum field theory approach.

Avainsanat — Nyckelord — Keywords

Dirac and Majorana neutrinos, neutrino oscillations, dispersion, coherence length

Säilytyspaikka — Förvaringsställe — Where deposited

Kumpula Science Library

Muita tietoja — Övriga uppgifter — Additional information

Contents

Chapter 1

Introduction

Particle physics is a branch of physics in which the behavior of elementary (fundamental) and composite particles is studied and observed. Elementary particles do not have any substructure, whereas composite particles are made from combination of elementary particles. During the history of particle physics, different elementary particles have been discovered in experiments or theoretically (i.e., special theory predicts the existence of an elementary particle) which might be seen in the future in laboratories such as newly discovered **Higgs boson** or have not been observed yet like **graviton**.

In particle physics, some particles have integer spin, i.e., they obey the Bose-Einstein statics and are called *bosons* such as photon, while the semi-integer spin particles obey the Fermi-Dirac statics and are called *fermions* such as electron, proton, quarks etc[1]. The known fundamental fermions are

Leptons: three charged leptons (electron (e), muon (μ), tau (τ)) as well as three *neutral leptons* or *neutrinos* form the flavors of leptons. We set lepton number (L_e, L_μ, L_τ) $+1_{e,\mu,\tau}$ for leptons.

Quarks: there are six types of quarks, known as *flavors* (up (u), charmed (c), top (t), down (d), strange (s) and bottom (b)), up and down quarks are the lightest ones. Top quark is the heaviest one, in general the heavier quarks are unstable and rapidly decay to up and down quarks. Up, charmed and top quarks have electric charge $+2/3$, while down, strange and bottom have electric charge $-1/3$.
Quarks carry color charges: red, green and blue, as well as electric charge, this is the basic assumption of the *color theory*. Therefore when the behavior of the quarks is studied, we should consider *quantum chromodynamics* (QCD) instead of *quantum electrodynamics* (QED) in which existing particles carry, exclusively, electric charge.

[1] The wave functions of identical bosons are symmetric under the exchange of any pair of them or in other words any two identical bosons can occupy the same quantum state. In the case of the fermion, the wave functions of two identical fermions are antisymmetric. This fact is called the **Pauli exclusion principle**.

The antiparticles of the fundamental fermions: for every particle there is an *antiparticle* with the same mass and lifetime, but opposite charge[2]. Hence, antiquarks and antileptons (we set lepton number $-1_{e,\mu,\tau}$ for antileptons) are also among the elementary fermions.

The known fundamental bosons are

Gluons (g): eight massless, spin one bosons which mediate the strong force. They are electrically neutral although carry color charges like quarks and so the corresponding theory will be QCD.

Photon (γ): massless, spin one particle which is its own antiparticle.

$W^{\pm}$ and Z^0 bosons: they are spin one and massive particles, W^+ and W^- are each other's antiparticles, while Z^0 is its own antiparticle.

Higgs boson (H): as it was said, Higgs boson is recently discovered [50], [9] in CERN, it is neutral, spin zero and very massive.

Quarks and gluons form the bound three-quark systems called *baryons*[3] which are fermions and quark-antiquark pairs are called *mesons*, they are bosons. The nucleons (protons and neutrons) are the lightest baryons. Pions ($\pi^{\pm}, \pi^0$) are examples of mesons. Baryons and mesons are also called *hadrons*.

The top quark is too unstable to form observable hadrons. Great efforts have been done to observe isolated free quarks (not in hadrons), however, the result has been fruitless. *Color confinement* or simply *confinement* states that color charged particles such as quarks can exist only when confined in hadrons and cannot be isolated. Therefore they cannot be directly observed [3].

There are four fundamental interactions or forces in particle physics: *strong, electromagnetic, weak* and *gravitational*. Each of these interactions has different ranges and strengths[4] when they are measured at a typical energy scale of 1 GeV [6].

For each of these interactions, there is a mediator. In the language of quantum field theory (QFT), when one electron is repelled by another electron, these two particles have exchanged a photon (the mediator) and this photon is responsible for the repulsion. Photon is the mediator of the electromagnetic interaction. The range of this interaction is infinite, while the strength of it is determined by the *fine structure constant* α where $\alpha \approx 10^{-2}$ [6]. This interaction is responsible for electrically charged particles and is very well described by QED.

When we go inside the atomic nuclei in which there are only nucleons, this force becomes very weak and the behavior of the nucleons is described by the

[2]This can be misleading, since some particles are neutral but their antiparticles are different (we will see later) although there are neutral particles which are their own antiparticles such as photon or neutral pion.

[3]We set baryon number (B) $+1$ for all baryons and -1 for all antibaryons and 0 for the rest.

[4]The field strengths of these interactions over a distance reduce from left to right.

strong force. This is the reason why the electromagnetism governs much of macroscopic physics.

The strong force, which is mediated by gluons, is independent of the electric charge and is responsible for binding protons and neutrons to form the nucleus of an atom (strong interaction is also called strong nuclear force) or keeping quarks together to form hadrons. The range of the strong interaction is about a fermi, i.e., 10^{-15} m [6].

The weak interaction is responsible for radioactive decays such as beta decay. It has a range of 10^{-17} m [6]. Photon and gluons do not experience weak interaction but other particles (such as leptons) do. We know that $W^{\pm}$ and Z bosons are mediators of the weak interaction. The term weak is due to the fact that in comparison to the strong and electromagnetic interactions, weak processes are much weaker.

These three interactions are unified in the Standard Model of particle physics (SM). As for the gravitation, we do not have a quantum theory yet, but only a macroscopic one, i.e., *General Relativity* (GR). However in the frame of QFT, the *graviton* is a hypothetical elementary particle that mediates the gravitational interaction but it has never been observed[5] [1], [3], [4].

In this thesis, we are going to speak about neutrinos. Neutrinos are neutral elementary particles with three different flavors which interact weakly with matter. They have very large penetration length, which is much larger than photon penetration length [1]. Very small cross section of the neutrino interaction with matter is not the only difference between neutrino and other fundamental fermions. The other one is the mass of neutrinos. In fact neutrino mass is one of the crucial questions for neutrino physicists.

The idea of the existence of neutrino was first discussed after the Ellis and Wooster experiment [1]. The experiment was β decay in which we have electron or positron emission from a radioactive nucleus when it is transformed into a slightly lighter nucleus.

For instance if potassium went to calcium

$$^{40}_{19}\mathtt{K} \rightarrow {}^{40}_{20}\mathtt{Ca} + e^{-}, \tag{1.1}$$

without a third neutral particle (see (1.1)), this reaction would definitely violate the energy conservation. As a result, the existence of a neutral particle which avoids this violation is necessary and reaction (1.1) is forbidden. This particle was first postulated by Wolfgang Pauli and called neutron (since it must be neutral) in 1930, and then in 1933 Enrico Fermi called it neutrino[6] [1].

By 1950, the existence of the neutrino had been proved theoretically, although it had not been observed. In fact, as neutrinos interact weakly with matter, they do not leave any footprint. This reason made neutrino detection very complicated.

To detect neutrinos, we needed an extremely intense source. In the mid-1950s

[5] Gravity has an infinite range but a low energy coupling about 10^{-38} [6].

[6] After discovering of the neutron (the particle which forms an atom with electrons and protons) in 1932 by James Chadwick, there were two particles with the same name. Thus, Enrico Fermi changed the name neutron (for the particle released in beta decay) to neutrino.

Clyde Cowan and Frederick Reines, through inverse beta-decay of neutron, confirmed the existence of antineutrino for the first time. This experiment became famous as the Cowan-Reines neutrino experiment. Neutron (n) beta-decay to proton (p^+) is

$$n \to p^+ + e^- + \overline{\nu}, \tag{1.2}$$

where $\overline{\nu}$ is the predicted antineutrino[7]. The inverse beta-decay will be

$$\overline{\nu} + p^+ \to n + e^+. \tag{1.3}$$

Cowan and Reines set up an intense source of protons in front of antineutrinos created in a nuclear reactor by beta decay. The resulted neutron and positron confirmed the existence of the antineutrinos and consequently, neutrino experimentally[8].

Later on, by experiment it was proved that neutrinos have two other flavors besides electron one. They are **muon** (muonic) neutrino (ν_μ) which comes with μ^+ or μ^- and **tau** (tauic) neutrino (ν_τ) that appears with τ^+ or τ^-. These three flavors differ from each other, and states corresponding to the neutrino flavors must be orthogonal,

$$\langle \nu_\alpha \, | \nu_{\alpha'} \rangle = \delta_{\alpha\alpha'}, \tag{1.4}$$

where α and α' refer to the neutrino flavors e, μ, τ [1]. Each charged lepton with its corresponding neutrino (for instance electron and electron neutrino) form the three generations of the leptons.

The Sun produces electron neutrinos. It is seen that some of these electron neutrinos change to other flavors. This phenomenon which is known as *neutrino oscillations* [1], [2], [3], [4] shows that neutrinos must be massive. Massive neutrino goes beyond the SM, since according to the SM neutrinos must be massless.

Another significant issue is the existence of antineutrino because neutrinos are neutral. Davis and Harmer put this question to a test. The Cowan-Reines experiment verified that (1.2) works, hence the crossed reaction, i.e.,

$$\nu + n \to p^+ + e^-, \tag{1.5}$$

must also occur. If neutrinos and antineutrinos were the same, we would have

$$\overline{\nu} + n \to p^+ + e^-. \tag{1.6}$$

Davis and Harmer found that the above reaction would not happen and so neutrinos and antineutrinos are different. If we give lepton number $L = 1$ to

[7] Cowan and Reines did not know that they had detected antineutrino (not neutrino). Actually, due to the conservation of lepton number in (1.2), antineutrinos must appear, the fact which became known later on. Hence, they thought this particle had been neutrino.

[8] The newly discovered particle was called neutrino, however, in fact it was **electron** antineutrino (as it will be seen). There are also other types of neutrinos which are different from electron one but at that time this fact was not discovered yet. Electron neutrino is also called electronic neutrino.

the leptons and $L = -1$ to the antileptons and $L = 0$ for the rest of particles, it can clearly be seen that reaction (1.6) violates the lepton number conservation. Therefore, the *law of conservation of lepton numbers* was suggested by them[9].

Thus, the difference between a neutrino and antineutrino is lepton number which can be determined in a laboratory. This is worth emphasizing that the definition of antiparticles does not just reduce to electric charge.

Strategy of this thesis. The thesis has been written in two parts. The first part consists of chapters 2, 3 and 4 in which the physics of neutrino is discussed.

We start chapter 2 by some short explanation of fermion fields, as well as electromagnetic and weak charges before the SM. In chapter 3, the SM is discussed briefly and we see that neutrinos are massless in the SM. Moreover, we describe how other flavors of neutrinos (i.e., muon and tau) are observed

Then in the next chapter, the experiments which show neutrino oscillations are presented. The Dirac and Majorana neutrinos are discussed and we find the mixing matrix of neutrinos.

In the second part, i.e., chapters 5 and 6, neutrino oscillations phenomenon is explained theoretically. In this thesis, we do not discuss neutrino oscillations experimentally and just use the results or explain very briefly if needed.

In chapter 5, we have neutrino oscillations in quantum mechanics. This chapter contains the first theories by which scientists explained oscillations. We find the oscillation phase and probability amplitude in the standard oscillation as well as wave-packet treatment. However, this is not the end of story. The chapter finishes with the problems of the quantum mechanical approach.

Chapter 6 describes neutrino oscillations from the QFT point of view. It is seen how the problems are solved. The simplest model in QFT with the plane wave particles is presented. The external wave-packet model is discussed in detail and we solve the amplitude integral to find the probability in three different regimes. Furthermore, stationary boundary conditions are explained.

[9] In fact, the SM does not allow any violation of the lepton flavor number, but we will see that this sentence is not correct when we go beyond the SM. Moreover, the answer to the question of the existence of antineutrinos is not just yes. As we go ahead, in the case of the Majorana neutrino, neutrinos and antineutrinos become the same and there will not be any lepton number conservation in that case.

Chapter 2

Before the standard model

2.1 Fermions

2.1.1 The Dirac field

The quantity S, called **action**, which is defined as

$$S = \int L \mathrm{d}t = \int \mathcal{L}\left(\phi, \partial_\mu \phi\right) \mathrm{d}t, \tag{2.1}$$

is very important in field theory. Here, L is the Lagrangian and $\mathcal{L}$ is the Lagrangian density. Lagrangian density is a function of fields ϕ and their derivatives.

According to the principle of least action, a system evolves in a way that the action is extremum. So, the variation of S, i.e., δS must vanish. The solution of $\delta S = 0$, results in the Euler-Lagrange equation of motion as follows

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial \left(\partial_\mu \phi\right)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \tag{2.2}$$

As for the spin 1/2 particles such as neutrinos (or other fermions) we should take the **Dirac Lagrangian** into consideration[1]

$$\mathcal{L}_0 = \overline{\psi}\left(x\right)\left(i\gamma^\mu \partial_\mu - m\right)\psi. \tag{2.3}$$

Here ψ, which is spinor, is the classical Dirac fermion field (i.e., not quantized), γ^μ are the Dirac matrices

$$\gamma^0 = \begin{pmatrix} \mathbb{1}_2 & 0_2 \\ 0_2 & -\mathbb{1}_2 \end{pmatrix} \quad \texttt{and} \quad \gamma^k = \begin{pmatrix} 0_2 & \tau^k \\ -\tau^k & 0_2 \end{pmatrix}, \tag{2.4}$$

[1] In this case we cannot apply the **Klein-Gordon Lagrangian** since it describes exclusively spin-0 particles.

where τ^k are the Pauli matrices and $\overline{\psi} = \psi^\dagger \gamma^0$. Furthermore, we define γ^5 as

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.5}$$

Regarding (2.2) and (2.3), the **Dirac equation** (i.e., equation of motion for spinors) becomes

$$(\not{\partial} - m)\,\psi(x) = 0, \tag{2.6}$$

where $\not{\partial} = \gamma^\mu \partial_\mu$. The Dirac equation can be solved for a plane wave. In the case of particles we obtain

$$\psi(x) = u(p) \exp(-ip \cdot x),$$

and antiparticles

$$\psi(x) = v(p) \exp(+ip \cdot x),$$

where $u(p)$ and $v(p)$ are the Dirac spinors[2],

$$u^{(s)} = \begin{pmatrix} \sqrt{p \cdot \tau}\zeta^s \\ \sqrt{p \cdot \overline{\tau}}\zeta^s \end{pmatrix} \quad \textit{and} \quad v^{(s)} = \begin{pmatrix} \sqrt{p \cdot \tau}\zeta^{-s} \\ -\sqrt{p \cdot \overline{\tau}}\zeta^{-s} \end{pmatrix}, \quad s = 1, 2, \tag{2.7}$$

where taking the Pauli matrices as a three-vector ($\vec{\tau}$), here τ^μ and $\bar{\tau}^\mu$ are supposed to be a four-vector, i.e.,

$$\tau^\mu \equiv (\mathbb{1}, \vec{\tau}), \qquad \bar{\tau}^\mu \equiv (\mathbb{1}, -\vec{\tau})$$

Here ζ^s is a two-component spinor basis. They can be $\zeta^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\zeta^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. In general (with radial angles θ and ϕ) we have

$$\zeta(\uparrow) = \begin{pmatrix} \cos\frac{\theta}{2} \\ \mathrm{e}^{i\phi}\sin\frac{\theta}{2} \end{pmatrix}, \qquad \zeta(\downarrow) = \begin{pmatrix} -\mathrm{e}^{-i\phi}\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \end{pmatrix}. \tag{2.8}$$

In addition we define

$$\zeta^{-s} = -i\tau^2 (\zeta^s)^* = (\zeta(\downarrow), -\zeta(\uparrow)). \tag{2.9}$$

The quantized Dirac field operator (the above $\psi(x)$ is not quantized) becomes

$$\hat{\psi}(x) = \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_\mathbf{p}}} \sum_s \left(\hat{a}^s_\mathbf{p} u^s(p)\, \mathrm{e}^{-ip\cdot x} + \hat{b}^{s\dagger}_\mathbf{p} v^s(p)\, \mathrm{e}^{ip\cdot x} \right), \tag{2.10}$$

where $E_\mathbf{p} = \sqrt{m^2 + |\mathbf{p}|^2}$. Here $\hat{a}^s_\mathbf{p}$ annihilates a fermion (contrary to $\hat{a}^{s\dagger}_\mathbf{p}$ which creates one) while $\hat{b}^{s\dagger}_\mathbf{p}$ creates one antifermion (contrary to $\hat{b}^s_\mathbf{p}$ which annihilates one). By taking hermitian conjugate from the above equation one can easily find $\hat{\overline{\psi}}$.

[2]By taking square root of a matrix, we mean taking the positive root of each eigenvalue.

2.1.2 Helicity and chirality

Now let us write the field ψ as the 2−dimensional representations ψ_L and ψ_R as follows

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}. \tag{2.11}$$

These two-component representations are called left- and right-handed **Weyl spinors**. In the case that ψ_L and ψ_R equal ζ^1 or ζ^2, they become eigenstates of the *helicity operator* h, defined as

$$h \equiv \frac{\vec{\tau} \cdot \mathbf{p}}{2\,|\mathbf{p}|}, \tag{2.12}$$

where $\mathbf{p}$ is along the z axis. Hence there are two eigenvalues, i.e., $h = 1/2$ which corresponds to right-handed particles (spin is in the same direction as the momentum) or $h = -1/2$ which corresponds to left-handed particles (spin is in the opposite direction as the momentum).

While helicity is a property of the 2−component spinors, *chirality* is a property of the 4−component spinors. We define the *projector operators* P_R and P_L as

$$P_R = \frac{1+\gamma^5}{2}, \quad P_L = \frac{1-\gamma^5}{2} \Rightarrow \tag{2.13}$$

$$P_L P_R = 0, \tag{2.14}$$

which project a spinor onto the right- and left-handed spinors, respectively. So,

$$\left.\begin{array}{l} \psi_L = P_L \psi \\ \psi_R = P_R \psi \end{array}\right\} \Rightarrow \left.\begin{array}{l} \overline{\psi}_L = P_R \overline{\psi} \\ \overline{\psi}_R = P_L \overline{\psi} \end{array}\right\}, \tag{2.15}$$

and then (with regard to (2.3)) the free Dirac Lagrangian becomes

$$\begin{aligned} \mathcal{L}_0 = \overline{\psi}\,(x)\,(i\gamma^\mu \partial_\mu - m)\,\psi = \overline{\psi}_L\,(x)\,(i\gamma^\mu \partial_\mu)\,\psi_L + \overline{\psi}_R\,(x)\,(i\gamma^\mu \partial_\mu)\,\psi_R \\ - m\left(\overline{\psi}_L \psi_R - \overline{\psi}_R \psi_L\right). \end{aligned} \tag{2.16}$$

It is seen that the mass term of the Lagrangian must contain both left- and right-handed spinors.

Contrary to the charged leptons which are both left- and right-handed, neutrinos are only left-handed. Thus, we can have left-handed lepton doublets as follows

$$\psi_{eL} = \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}, \psi_{\mu L} = \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}, \psi_{\tau L} = \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix}. \tag{2.17}$$

The right-handed lepton singlets are

$$e_R, \mu_R, \tau_R. \tag{2.18}$$

It should be remarked that:

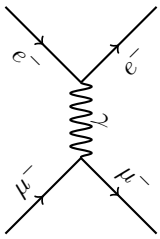


Figure 2.1: The electromagnetic interaction.

- according to the SM, there are (right)left-handed (anti)neutrinos. Actually, at present, we know that (left)right-handed (anti)neutrinos cannot be predominant and if they exist, they interact with matter much weaker than the weak interaction of the (right)left-handed (anti)neutrinos. This is the reason why $\overline{\nu}_L$ and ν_R are called *sterile* or *inert* (anti)neutrinos.
- In doublets (2.17), it is seen that each lepton, e.g., e_L^- comes with the corresponding neutrino, i.e., (in this case) ν_{eL}. This stems from the fact that (as it was briefly discussed above and will be fully discussed later) neutrinos are different and a charged lepton from one generation does not appear with a neutrino from another generations.
- Although the sterile neutrinos do not take part in the weak charges (they do not couple to the weak $W^{\pm}$ and Z^0 bosons), as we shall see later, they can play important roles in the case of neutrino mass generation and neutrino oscillations. In this case, we have gone beyond the SM since we are dealing with massive neutrino [1]. When we go beyond the SM, (left)right-handed (anti)neutrinos may be used.

2.2 Currents

2.2.1 The electromagnetic current

The vertex of the electromagnetic (EM) interaction is $(-ie\gamma^{\nu})$, where e is the coupling constant, i.e., the electric charge of the electron. For instance in the electromagnetic interaction (see Fig (2.1))

$$e^-\mu^- \to e^-\mu^-, \tag{2.19}$$

one can write (from the Feynman rules) for the electron-electron and muon-muon currents

$$j_e^{\nu} = -\overline{u}_e\gamma^{\nu}u_e \quad \texttt{and} \quad j_{\mu}^{\nu} = -\overline{u}_{\mu}\gamma^{\nu}u_{\mu}. \tag{2.20}$$

In general there are three charged leptons ($l = e^-, \mu^-, \tau^-$), let us define $u_l \equiv l$, then

$$j_{\texttt{EM}}^{\mu}(x) = \sum_l -\bar{l}(x)\gamma^{\mu}l(x). \tag{2.21}$$

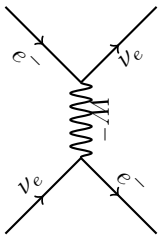


Figure 2.2: Charged weak interaction.

2.2.2 Weak currents

Weak interaction can be present in three different processes:

1. Non-leptonic processes in which no leptons are present (only hadrons are present) in the initial and final states.
2. Semileptonic processes in which both hadrons and leptons take part.
3. Leptonic processes in which only leptons take part, e.g.,

$$\nu_\mu e^- \to \nu_\mu e^-, \quad \texttt{or} \quad \nu_\mu e^- \to \nu_e \mu^-. \tag{2.22}$$

If the charges of initial and final particles in a weak interaction differ by one unit, it is called the *charged current* (CC) and is mediated by $W^\pm$. The vertex of this interaction is

$$\frac{-ig}{2\sqrt{2}}\gamma^\mu\left(1-\gamma_5\right). \tag{2.23}$$

For instance in the weak interaction (see Fig (2.2))

$$e^-\nu_e \to \nu_e e^-, \tag{2.24}$$

the charge difference between e^- and ν_e is one, so it is a charge current. One can write

$$\begin{aligned} j^\mu_{e^-\mu_e} &= \overline{u}_{\nu_e}\left(\frac{-ig}{2\sqrt{2}}\gamma^\mu\left(1-\gamma_5\right)\right)u_e, \\ j^\mu_{\mu_e e^-} &= \overline{u}_{e'^-}\left(\frac{-ig}{2\sqrt{2}}\gamma^\mu\left(1-\gamma_5\right)\right)u_{\nu'_e}. \end{aligned} \tag{2.25}$$

We know that $\overline{u}_{\nu_e}\gamma^\mu u_e$ transforms as a 4−vector while $\overline{u}_{\nu_e}\gamma^\mu\gamma_5 u_e$ transforms as an axial 4−vector. Hence, the spacetime structure of the charged-current is

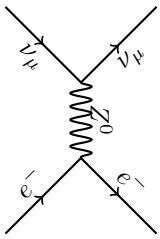


Figure 2.3: Neutral weak interaction.

vector-axial vector or $V - A$. Moreover, in general the charged current can be written as

$$j^\mu = \overline{\psi}\left(\frac{-ig}{2\sqrt{2}}\gamma^\mu\left(1-\gamma_5\right)\right)\psi = \frac{-ig}{\sqrt{2}}\overline{\psi}_L\gamma^\mu\psi_L, \tag{2.26}$$

where in the last equality we used

$$1-\gamma_5 = \frac{1}{2}\left(1-\gamma_5\right)^2; \qquad [\gamma^\mu,\gamma_5] \neq 0. \tag{2.27}$$

Eq. (2.26) shows that the charged current is similar to (2.21) with the difference that in the CC, only left handed particles can take part. In addition to the charged weak current, there is neutral weak current which is mediated by Z^0 boson. The discovery of the neutral weak current happened in two different types of events.

1. Elastic scattering of ν_μ or $\overline{\nu}_\mu$ with electrons [11]. Muon neutrinos are invisible, they come and interact with electrons in liquid, but in the final state there is no μ^-. Since we cannot have ν_μ and e^- at one vertex, it is concluded that this interaction indicates a neutral current (Fig (2.3)).

$$\nu_\mu e^- \to \nu_\mu e^-. \tag{2.28}$$

2. Neutrinos can scatter from a nucleus N,

$$\nu_\mu/\overline{\nu}_\mu + N \to \nu_\mu/\overline{\nu}_\mu + \texttt{hadrons}. \tag{2.29}$$

   It is seen that in the final state there is no μ^-/μ^+. So, it shows that (2.29) happens through neutral currents, i.e., invisible neutrinos come, interact, go invisibly again and we are left with a hadron [2].

2.3 Symmetry and parity violation

Symmetry is a transformation under which equation of motion (2.2) remains invariant. This happens only if the action S is invariant under the field trans-

formations

$$\phi(x) \to \phi'(x) = \phi(x) + \alpha\Delta\phi(x), \tag{2.30}$$

where α is an infinitesimal parameter and $\Delta\phi(x)$ is a deformation of the field configuration.

According to Noether's theorem, symmetries under continuous transformations lead to conservation laws [5]. For instance, the gauge group $U(1)$ is a group of phase transformations. Under the global transformations of $U(1)$ the Dirac Lagrangian is invariant (see section (3.2.1)). As a result of Noether's theorem, the conserved quantity is the electric charge. Thus, no matter we are in what interaction, the sum of electric charges of initial and final particles, taking part in a reaction, must be equal.

Symmetries corresponding to the gauge group such as $U(1)$ or $SU(2)$ (we shall discuss later) are called *fundamental symmetries*. This means that the conserved quantity must be absolutely invariant in all interactions (as a result of Noether's theorem) and this fact cannot be violated at all unless the symmetry is broken by some reason.

On the other hand, there are quantities such as baryon and lepton numbers[3], quark[4] and lepton flavors[5] which must not be necessarily conserved [2].

We have other symmetries such as *parity* $\mathcal{P}$ that sends a particle in a state $(t, \mathbf{x})$ to the state $(t, -\mathbf{x})$, *time reversal* $\mathcal{T}$ which sends $(t, \mathbf{x})$ to $(-t, \mathbf{x})$ and *charge conjugation* $\mathcal{C}$ that exchanges particles with their antiparticles. These symmetries are not conserved by all interactions. Only experiments tell us whether these symmetries are conserved or violated by an interaction.

Under double parity transformations we come back to the first state, so $\hat{\mathcal{P}} = \hat{\mathcal{P}}^{-1}$. The angular momentum ($l$) contribution to parity of a particle in spherically symmetric potential is $(-1)^l$. As well as this contribution, we define *intrinsic parity*. The creation (and similarly annihilation) operators for fermions and antifermions under intrinsic parity transformations become

$$\hat{\mathcal{P}}\hat{a}^\dagger_{\mathbf{p}}\hat{\mathcal{P}} = \hat{a}^\dagger_{-\mathbf{p}}, \;\; \hat{\mathcal{P}}\hat{b}^\dagger_{\mathbf{p}}\hat{\mathcal{P}} = \hat{b}^\dagger_{-\mathbf{p}}. \tag{2.31}$$

As for space and momentum states, under parity transformations we have

$$\begin{aligned} \hat{\mathcal{P}}\,|r\rangle &= |-r\rangle\,, \\ \hat{\mathcal{P}}\,|p\rangle &= |-p\rangle\,, \end{aligned} \tag{2.32}$$

while angular momentum (such as spin) state $|J\rangle$ becomes

$$\hat{\mathcal{P}}\,|J\rangle = |r \times p\rangle = |J\rangle\,. \tag{2.33}$$

[3]Baryon and lepton number conservations, in the Standard Model of particle physics, are called *accidental global symmetries*. All the know interactions conserve baryon and lepton numbers [2]. However, in the future, the violation of them is possible to be observed [5].

[4]Quark flavors are assumed to be conserved in the strong interactions.

[5]Lepton flavor violation, for instance muon to electron conversion is forbidden by the Standard Model [49]. Nonetheless, there are some phenomena in contradiction with the Standard Model. To be able to discuss them, we should go beyond the Standard Model [49].

Hence, parity inverts momentum whereas keeps angular momentum direction. In other words, parity changes (right)left-handedness to (left)right-handedness.

Electromagnetic and strong interactions conserve parity while weak interaction does not. As a result, we observe only ν_L and $\overline{\nu}_R$ take part in reactions. A reaction such as

$$\pi^+ \to \mu^+ \nu_L, \tag{2.34}$$

under parity becomes

$$\pi^+ \not\to \mu^+ \nu_R, \tag{2.35}$$

hence it can be concluded that weak interaction violates parity.

Parity violation. It is worth saying some word about parity violation. Tsung-Dao Lee and Chen-Ning Yang surveyed the experimental information on parity conservation and reached the conclusion that in weak interactions, parity conservation neither can be accepted nor refuted. However, through the beta decay experiment of (spin aligned) cobalt-60 ($^{60}_{27}\mathtt{Co}$) to nickel-60 ($^{60}_{28}\mathtt{Ni}$) (suggested by them and conducted by physicist Chien-Shiung Wu) i.e.,

$$^{60}_{27}\mathtt{Co} \to {}^{60}_{28}\mathtt{Ni} + e^- + \overline{\nu} + 2\gamma, \tag{2.36}$$

the ***parity non-conservation*** was finally ascertained in the weak interaction. This experiment became famous as the Wu experiment[6]. If parity is conserved, it is expected that distribution of the detected electrons at angle (θ) is approximately proportional to the distribution at angle ($\pi - \theta$) or in other words, under parity ($\theta \to \pi - \theta$) we have similar distribution of detected electrons.

Furthermore, the photons released from the return of nickel to its ground state is an EM reaction. We know that EM conserves parity and so it is expected an approximately symmetric distribution of photons in both angles after counting the photon numbers. The same symmetry is expected in the number of detected electrons ***if and only if*** parity is conserved in beta decay.

After carrying out the experiment, the symmetry in the photon numbers (as expected) was seen whereas there was asymmetry in the electron distribution. This means that electrons with a definite spin (since cobalt was spin aligned) are emitted in one direction. Thus, experiment states that parity is violated in weak interactions [29].

Charge conjugation. As already mentioned, under charge conjugation operator particles and antiparticles are exchanged. Under double action of charge conjugation, we come back to the particle again, so $\hat{\mathcal{C}} = \hat{\mathcal{C}}^{-1}$. For the annihilation (and similarly creation) operators under charge conjugation we have

$$\hat{\mathcal{C}} \hat{a}^s_{\mathbf{p}} \hat{\mathcal{C}} = \hat{b}^s_{\mathbf{p}}; \qquad \hat{\mathcal{C}} \hat{b}^s_{\mathbf{p}} \hat{\mathcal{C}} = \hat{a}^s_{\mathbf{p}}. \tag{2.37}$$

[6] In fact, cobalt-60 decays by beta decay to nickel-60, electron and antineutrino. The produced nickel is excited, it comes back to its ground state by emitting two photons.

Whereas strong and electromagnetic interactions conserve charge conjugation, weak interaction violates it. This fact can be easily seen from (2.34), under charge conjugation it becomes

$$\pi^- \not\to \mu^- \overline{\nu}_L, \tag{2.38}$$

which is not correct due to presence of the left-handed antineutrino. However, reaction (2.34) is conserved under $\mathcal{CP}$, i.e.,

$$\pi^- \to \mu^- \overline{\nu}_R. \tag{2.39}$$

Electromagnetic and strong interactions conserve $\mathcal{CP}$. Weak interaction does not conserve it, for example in processes such as kaon decays [1], [12] or kaon mixing [5] in which $\mathcal{CP}$ is violated. All the fundamental interactions conserve $\mathcal{CPT}$ [5].

Chapter 3

The standard model for leptons

3.1 Introduction

Neutrino oscillations phenomenon goes beyond the SM. So, first one should become familiar with the SM.

In perturbation theory when we go to higher orders, divergent integrals appear (loop integrals). As these divergences happen in higher orders, they are called *ultraviolet divergences* and cause a big problem because the reactions happen in reality but solving the loop integrals will result in infinity. Therefore we have to be able to get rid of these divergences, in fact they can be removed systematically. In QED with the help of renormalization methods, one can remove divergences [5].

Until the end of the sixties the only renormalizable theory was QED, the search for a renormalizable weak interaction theory ended in unification of the electromagnetic and weak interactions[1]. A single gauge theory $U(1) \times SU(2)$ which is simply called the *electroweak* interaction.

QCD is the theory which describes the strong interaction, it is a gauge theory $SU(3)$. QCD and electroweak together, i.e., $U(1) \times SU(2) \times SU(3)$ form the Standard Model of fundamental interactions. In this chapter we are going to briefly discuss the SM for leptons, the mass generation predicted by it, the reason why the SM is not able to answer all questions and finally the chapter is going to be finished by some information about different neutrino flavors. This chapter is mainly following Ref. [1].

[1]The fact that this unification leads to a renormalizable theory was proved by t'Hooft and Veltman [39].

3.2 Gauge transformations

3.2.1 Covariant derivative

The group $U(1)$ is a group of phase transformations

$$U(1) = \exp\left(i\Lambda\frac{1}{2}\right), \tag{3.1}$$

where Λ is an arbitrary real number. If Λ is constant, (3.1) is the global transformations and if we have $\Lambda(x)$, it becomes the gauge transformations. This group is clearly Abelian since the group elements commute.

In general, $SU(n)$ groups are represented by $n \times n$ complex, unitary matrices ($UU^{\dagger} = \mathbb{1}_n$) with determinant one. The number of group generators is $n^2 - 1$ and an element of $SU(n)$ is written as

$$U = \exp\left(i\Lambda^a X_a\right) \qquad \texttt{with}\ a = 1, ..., n-1, \tag{3.2}$$

where X_a are the group generators and Λ_a are real parameters[2]. For the special case $SU(2)$, we have three generators, they are $(\tau_j/2)$, where τ_j are the three Pauli matrices. Under the global transformations of $U(1)$ and $SU(2)$ the Dirac Lagrangian is invariant while gauge transformations of $U(1)$ and $SU(2)$ ruin this invariance [1].

To have an invariant Lagrangian in the gauge transformations, an interacting vector field is considered and the *covariant derivative* D_μ is formed. Thus, we obtain for $U(1)$ and $SU(2)$ the covariant derivatives

$$\partial_\mu \to D_\mu = \left(\partial_\mu + ig\frac{1}{2}B_\mu(x)\right), \tag{3.3}$$

$$\partial_\mu \to D_\mu = \left(\partial_\mu + ig'\frac{1}{2}\vec{\tau}\,.\mathbf{A}_\mu(x)\right), \tag{3.4}$$

where $B_\mu(x)$ and $\mathbf{A}_\mu(x)$ are $U(1)$ and $SU(2)$ vector gauge fields, respectively. Moreover, g and g' are dimensionless coupling constants of the gauge fields with the spinors. After these changes we see that the fields B_μ and $\mathbf{A}_\mu$ go under these transformations

$$B_\mu(x) \to B'_\mu(x) = B_\mu(x) - \frac{1}{g}\partial_\mu\Lambda(x), \tag{3.5}$$

$$\mathbf{A}_\mu(x) \to \mathbf{A}'_\mu(x) = \mathbf{A}_\mu(x) - \frac{1}{g}\partial_\mu\vec{\Lambda}(x) - \vec{\Lambda}(x) \times \mathbf{A}_\mu(x). \tag{3.6}$$

Hence, instead of (2.3), the Lagrangian becomes

$$\mathcal{L} = \overline{\psi}(x)\left(i\gamma^\mu D_\mu - m\right)\psi. \tag{3.7}$$

[2]Again here we have global and gauge transformations regarding the dependence of Λ_a on x, exactly such as $U(1)$ case.

Adding any mass term in the form of $m^2 B_\mu B^\mu$ or $m^2 A_\mu A^\mu$ to the above Lagrangian will ruin its invariance since the mass terms are not invariant under (3.5) and (3.6) transformations. As a result, particles corresponding to the fields A and B must be massless.

As for the $U(1)$ symmetry, this result is correct since the particle corresponding to this symmetry (i.e., photon) is massless but in the case of $SU(2)$ there are three massive bosons. This fact causes a serious problem which can be solved with the help of the *spontaneous symmetry breaking* of $SU(2)$ in a mechanism called the *Higgs mechanism* [40], [41].

3.2.2 Again weak charges, but this time with the SM formulation

From the Lagrangian which is invariant under the local $SU(2)$ transformations, we can find the conserved current[3] j_k^μ (according to Noether's theorem) as

$$\begin{aligned}\mathcal{L}_I(x) &= -g'\overline{\psi}(x)\gamma^\mu \frac{1}{2}\tau_k A_\mu^k \psi(x) \Rightarrow j_k^\mu = \overline{\psi}(x)\gamma^\mu \frac{1}{2}\tau_k \psi(x) \\ &= -g' j_k^\mu(x) A_\mu^k(x). \end{aligned} \tag{3.8}$$

It should be noted that j_k^μ with $k = 1, 2, 3$ are conserved and are not the only solutions, but, any combinations of them can also be solutions. Let us define the new quantities $\tau^\pm$ and $W_\mu^\pm$ as follows

$$\tau^\pm = \tau_1 \pm i\tau_2 \quad \texttt{and} \quad W_\mu^\pm = \frac{1}{\sqrt{2}}\left(A_\mu^1 \mp i A_\mu^2\right), \tag{3.9}$$

then, by using the $SU(2)$ doublets (2.17) (e.g., electron doublet) we will have the conserved currents corresponding to τ^+ and τ^-, i.e., j_+^μ and j_-^μ as follows

$$j_+^\mu = 2\overline{\nu}_{eL}\gamma^\mu e_L, \qquad j_-^\mu = 2\overline{e}_{eL}\gamma^\mu \nu_{eL}. \tag{3.10}$$

The charge difference of the two elements (electron and neutrino electron) in j_+^μ is $(+1)$, hence this current changes the charges of particles by one and consequently, the particle corresponding to the W^+ field must have an electric charge equal to $(+1)$. With the same reasoning for the j_-^μ, it is concluded that W^- field corresponds to a particle with electric charge (-1). In general all three neutrinos take part in the charged current, so regarding all three doublets, the charged current may be written as

$$j_+^{\mu CC} = 2\sum_{l=e,\mu,\tau} \overline{\nu}_{lL}\gamma^\mu l_L, \quad \texttt{or} \quad j_-^{\mu CC} = 2\sum_{l=e,\mu,\tau} \overline{l}_L\gamma^\mu \nu_{lL}. \tag{3.11}$$

Moreover, the current j_3^μ becomes

$$j_3^\mu = \frac{1}{2}\left(\overline{\nu}_{eL}\gamma^\mu \nu_{eL} - \overline{e}_L\gamma^\mu e_L\right). \tag{3.12}$$

[3] This conserved current corresponding to the invariance of the Lagrangian under the gauge transformations of $SU(2)$ is called *isovector current*.

This current does not change the charges and so the particle corresponding to A_μ^3 must be neutral, it is called Z^0 boson. In general j_3^μ reads

$$j_3^\mu = \frac{1}{2}\left(\sum_{l=e,\mu,\tau} \overline{\nu}_{lL}\gamma^\mu \nu_{lL} - \sum_{l=e,\mu,\tau} \bar{l}_L \gamma^\mu l_L\right). \tag{3.13}$$

With respect to (3.8), (3.9), (3.11) and (3.13), the Lagrangian $\mathcal{L}_I(x)$ can be written as

$$\mathcal{L}_I(x) = \left(-\frac{g'}{2\sqrt{2}} j_{\mathtt{cc}}^\mu W_\mu + \mathtt{h.c.}\right) - g' j_3^\mu A_\mu^3, \tag{3.14}$$

where

$$j_{\mathtt{cc}}^\mu = 2\sum_{l=e,\mu,\tau} \overline{\nu}_{lL}\gamma^\mu l_L. \tag{3.15}$$

3.2.3 Electroweak unification

As we saw, there is an interaction between the $SU(2)$ doublet and field A_μ. We define *weak isospin* I_W or simply *isospin* I as a quantity that corresponds to $SU(2)$. Each member of a doublet has the same I but different I_3, where I_3 is the third component of the isospin.

The SM unifies the electromagnetic and weak interaction. As already said, electromagnetism conserves parity while weak interaction does not. So, the unification of these two interactions cannot be done on the basis of $SU(2)$ gauge transformations, or in other words we need to enlarge our gauge group.

The group representation of the electromagnetism is $U(1)$ and for the weak interaction is $SU(2)$, so the minimal group of the electroweak (the unified form) is $SU(2)\times U(1)$. This symmetry group includes both electromagnetic and weak interactions and hosts three and one $(3+1)$ gauge fields.

We have defined isospin as the quantity corresponding to $SU(2)$, additionally we need to find some quantity corresponding to $U(1)$. In each doublet there are particles with different charges, we define *weak hypercharge* Y_W or in brief *hypercharge* Y as the quantity related to $U(1)$ which is the group of hypercharge. The hypercharge is connected to the electric charge Q by the Gell-Mann-Nishijima relation

$$Q = I_3 + \frac{1}{2}Y. \tag{3.16}$$

Our Lagrangian must be invariant under this enlarged gauge transformations, so the partial derivative in the free Lagrangian must be replaced by the covariant derivative. Then, in the case of the left-handed doublets we have

$$\partial_\mu \psi_{lL} \to D_\mu \psi_{lL} = \left(\partial_\mu + ig\frac{1}{2}Y_L^{\mathtt{lep}} B_\mu(x) + ig'\frac{1}{2}\vec{\tau}\,.\mathbf{A}_\mu(x)\right)\psi_{lL}. \tag{3.17}$$

Since the right-handed leptons are $SU(2)$ singlets, then their covariant derivative is

$$\partial_\mu l_R \to D_\mu l_R = \left(\partial_\mu + ig\frac{1}{2}Y_R^{\mathtt{lep}} B_\mu(x)\right) l_R. \tag{3.18}$$

For the doublet we have $I = 1/2$, and so the neutrino's third component of isospin becomes $+1/2$ and for the charged leptons it is $-1/2$. Therefore from (3.16), we see that $Y_L^{\texttt{lep}} = -1$. The isospin of the singlets is zero and so their third isospin is zero too. Thus (3.16) says that $Y_R^{\texttt{lep}} = -2$. Now, the total Lagrangian (left- and right-handed) can be written as follows (see (3.14))

$$\mathcal{L}_I(x) = -\frac{g'}{2\sqrt{2}}\left(j_+^\mu W_\mu^+ + j_-^\mu W_\mu^-\right) - g' j_3^\mu A_\mu^3 - g\left(-\frac{1}{2}\left(\sum_{l=e,\mu,\tau}\overline{\nu}_{lL}\gamma^\mu\nu_{lL} + \sum_{l=e,\mu,\tau}\overline{l}_L\gamma^\mu l_L\right) - \sum_{l=e,\mu,\tau}\overline{l}_R\gamma^\mu l_R\right)B_\mu. \tag{3.19}$$

Let us add and subtract this term

$$-\frac{1}{2}gB_\mu\sum_{l=e,\mu,\tau}\overline{l}_L\gamma^\mu l_L$$

from the second line of (3.19). Then,

$$\mathcal{L}_I(x) = -\frac{g'}{2\sqrt{2}}\left(j_+^\mu W_\mu^+ + j_-^\mu W_\mu^-\right) - g' j_3^\mu A_\mu^3 - g\frac{1}{2}j_Y^\mu B_\mu,$$

where $1/2 j_Y^\mu$

$$\frac{1}{2}j_Y^\mu = -\sum_{l=e,\mu,\tau}\overline{l}_L\gamma^\mu l_L - \sum_{l=e,\mu,\tau}\overline{l}_R\gamma^\mu l_R = j_{\texttt{EM}}^\mu - j_3^\mu, \tag{3.20}$$

is the neutral current after unification of the electromagnetic and weak interactions.

3.3 Higgs mechanism

3.3.1 The boson masses

The gauge $U(1)\times SU(2)$ symmetry is valid only when the fields A_μ and B_μ are massless, while we know that $W^\pm$ and Z^0 are massive. These particles become massive after the spontaneous symmetry breaking during which the fields A_μ and B_μ interact with the scalar Higgs boson. Since we are in $U(1)\times SU(2)$, so the Higgs field[4] $\phi(x)$ must have the $U(1)\times SU(2)$ transformations properties and consequently, we suppose that it is an $SU(2)$ doublet. So,

$$\phi(x) = \begin{pmatrix}\phi^\pm(x)\\ \phi^0(x)\end{pmatrix}, \tag{3.21}$$

[4]The origin of the scalar field mass is unclear [10].

where $\phi^{\pm}$ and ϕ^0 are scalar complex fields of particles with electric charges ± 1 and 0, respectively. In fact the symmetry violation happens when the scalar field $\phi(x)$ acquires a *vacuum expectation value*[5] (VEV). From the conservation of the electric charge, the vacuum expectation value of the charged field $\phi^{\pm}$ is zero. Additionally, if we choose the potential part of the Lagrangian as

$$V(\phi) = -\mu^2 \phi^* \phi + \lambda (\phi^* \phi)^2, \tag{3.22}$$

where μ and λ are positive constants, for $\mu^2 > 0$ the field ϕ^0 acquires an expectation value. The minimum of the potential $V(\phi)$ occurs at

$$\langle \phi^0(x) \rangle = \frac{v}{\sqrt{2}}, \tag{3.23}$$

where $v = \left(\mu^2/\lambda\right)^{1/2}$, and so the Higgs doublet at the minimum potential for breaking the symmetry becomes

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{3.24}$$

After the Higgs interaction and breaking the symmetry, the three $W^{\pm}$ and Z^0 receive their mass (their speeds become less than one) whereas photon still stays massless because it does not interact with the Higgs. The Higgs interaction with the boson fields results in the mass terms of $W^{\pm}$ and Z^0 in terms of the coupling constants, i.e.,

$$m_W = g' \frac{v}{2}; \qquad m_Z = \sqrt{g'^2 + g^2}\frac{v}{2}. \tag{3.25}$$

Through various experimental data, the value of these coupling constants can be calculated, and so we are left with [42]

$$m_Z = (91.1876 \pm 0.0021)\texttt{GeV}; \qquad m_W = (80.385 \pm 0.015)\texttt{GeV}. \tag{3.26}$$

3.3.2 The fermion masses

One may ask whether the Higgs mechanism can be responsible for the mass generation of fermions such as quarks, charged and neutral leptons as well as bosons. If the answer is positive, for instance, in the case of a lepton such as electron we can write the mass generation Lagrangian (the Yukawa Lagrangian for electron $\mathcal{L}_e$) with the left- and right-handed lepton fields[6] as follows

$$\mathcal{L}_Y^{\texttt{lep}} = -\sum_{l,l'} \lambda_Y^{\texttt{lep}} \overline{\psi}_{lL} M_{ll'}^{\texttt{lep}} l_{l'R} \phi, \tag{3.27}$$

[5] Vacuum expectation value $\langle \phi(x) \rangle$ occurs when the scalar field potential $V(\phi)$ becomes minimum.

[6] The reason why we need the left- and right-handed lepton fields in the mass Lagrangian is due to the fact that Higgs is a doublet. We set on the left of it the lepton field doublet (in the form of the hermitian conjugate $\overline{\psi}_{eL}$) and on the right of it, the lepton field singlet e_R. Additionally, only in this form hypercharge and isospin are both conserved.

where $M^{\mathtt{lep}}$ is a complex matrix and $\lambda_Y^{\mathtt{lep}}$ is the Yukawa coupling. To find the mass terms we should diagonalize this matrix. After doing so, for instance, for the electron Lagrangian [5] we have

$$\mathcal{L}_e = -\lambda_e \overline{\psi}_{eL} \phi e_R + \mathtt{h.c.} = -\frac{1}{\sqrt{2}} \lambda_e v \bar{e}_L e_R + \mathtt{h.c.} \Rightarrow$$
$$m_e = \frac{1}{\sqrt{2}} \lambda_e v. \tag{3.28}$$

The above relation indicates that the mass of an electron is proportional to the Yukawa coupling. The SM does not put any constraint on the Yukawa coupling (unlike the bosonic case as it was discussed above) and so the mass term (3.28) is a free parameter in the SM. In addition, since neutrinos are not right-handed, then they must be massless in the Standard Model. However, what about the experiment? Are really neutrinos massless?

In fact, in three different types of experiments, we are looking for neutrino masses. First: cosmological observations, second: search for neutrino-less double beta decay, and third: direct determination of the neutrino mass by kinematics (model-independent) such as tritium beta decay which puts limit[7] on the neutrino mass [18].

Therefore, how can we theoretically describe neutrino mass? Let us suppose that there are inert right-handed neutrinos, then the SM can generate neutrino mass through the interaction of the leptons and conjugated Higgs field[8] in the Yukawa Lagrangian

$$\mathcal{L}_Y^{\nu} = \sum_{l,l'} \lambda_Y^{\nu} \overline{\psi}_{lL} M_{ll'}^{\nu} \nu_{l'R} \tilde{\phi}, \tag{3.29}$$

where M^{ν} is a complex matrix and λ_Y^{ν} is the Yukawa coupling for neutrinos. Again after diagonalizing this matrix we can find neutrino mass but it is proportional to the Yukawa coupling. Hence, there is the same problem that we had in the case of electron. Due to this fact, it can be said that in general, it is unnatural that the Higgs mechanism is responsible for the mass generation of the charged and neutral leptons since unlike the boson masses, the Higgs mechanism does not put any constraint on the fermion masses and so they are free parameters [1]. The question about neutrino mass will be discussed in the next chapter.

[7] Different β spectrum of tritium near its endpoint has resulted in a variety of the mass limits for neutrinos. These experiments and their references can be seen in [19].

[8] If neutrinos interact with the Higgs field, hypercharge and isospin will not be conserved, and also the SM cannot generate any mass term for the neutrinos. As a result, we define the conjugated Higgs field as

$$\tilde{\phi} = i\tau_2 \phi^* = \frac{1}{\sqrt{2}} \begin{pmatrix} v \\ 0 \end{pmatrix}.$$

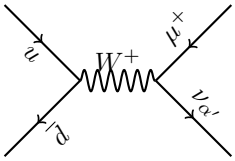


Figure 3.1: The pion decay through weak interaction.

3.4 Discovery of the muon neutrino

Now let us finish this chapter with discussing the reason why neutrinos appear in three different flavors. Before starting, again, it should be emphasized that since neutrinos hardly interact with matter, the scientists have to apply very special methods to detect them (such as in the case of the discovery of the electron neutrino).

3.4.1 The first grounds for the existence of muon neutrino

Neutrinos interact only with very massive $W^{\pm}$ and Z^0 bosons. These massive bosons are unstable. For instance, the lifetime of W is less than 10^{-17} s [26], the existence of $W^{\pm}$ is identified through their decays to lepton (or antilepton) and neutrino (or antineutrino) [26].

In pion ($\pi^+ = u\overline{d}$) decay, which is mediated by W^+, we understand the existence of the virtual[9] W^+ from its decay to antimuon (μ^+) and $\nu_{\alpha'}$[10] (see Fig. (3.1))

$$\pi^+ \to \mu^+ + \nu_{\alpha'}. \tag{3.30}$$

Beta decay (as explained above) for an atom with Z protons is

$$Z \to (Z-1) + e^+ + \nu_{\alpha}. \tag{3.31}$$

Again this process is mediated by W^+ which now decays to positron and ν_{α}. The important question is whether $\nu_{\alpha} = \nu_{\alpha'}$ or not. If the answer to this question is negative, then we must have at least two types of neutrino. Now let us consider the muon decays reactions. The ordinary muon decay is

$$\mu \to e + \nu + \overline{\nu}. \tag{3.32}$$

Hence, on the one hand, from (3.30) and (3.31) we are faced with a W boson decaying into a charged lepton and a neutrino (when there is one charged lepton such as electron or muon we have just one neutrino) while on the other hand, in the ordinary neutrino decay (3.32), when both electron and muon are present

[9]The word virtual means that a particle is off-shell. An on-shell particle obeys energy-momentum conservation rule, $E^2 - p^2 = m^2$, while for an off-shell one, the energy-momentum conservation is violated. Particles in propagators (i.e., mediators) are off-shell.

[10]We shall understand the reason of α' very soon.

there is a pair of neutrinos. From this comparison, the fact that the neutrino appearing with muon is totally different with the one coming with electron, can be correct. Moreover, there is an exotic muon decay as

$$\mu \to e + \gamma. \tag{3.33}$$

The absence of the above process may lead to the fact that neutrinos coupled to muons are different from the ones coupled to electrons[11] [27].

According to this explanation, we can be doubtful about the fact which states $\nu_\alpha = \nu_{\alpha'}$ in (3.30) and (3.31). In other words, the neutrino coupled to the (anti)muon and produced in the pion decay is not the neutrino coupled to the (anti)electron in beta decay. This is not the final result and only is accepted if an experiment confirms it.

3.4.2 The final confirmation of the existence of muon neutrino

The possibility of the existence of at least two types of neutrino was tested in an experiment at the Brookhaven AGS in which the interaction of high-energy neutrinos produced in pion decay reaction (3.30) with matter was observed. Pions are produced from 15-BeV protons striking a target of beryllium, the resulting flux strikes an iron shield wall. Neutrino interactions are observed in an aluminum spark chamber behind this shield [25].

In this experiment single track events were observed. The distinction between single high-energy muon tracks (produced after pion decay to muon and neutrino) of $p_\mu > 300$ MeV and single high-energy electron track (produced at the accelerator Cosmotron) with a mean energy greater than 400 MeV is clear. Produced neutrinos were exposed to the high-energy electrons and the number of electron events were counted and compared with the muon events. If $\nu_e = \nu_\mu$, then from 29 single muon events received[12], $\frac{2}{3} \times 29$ electron shower or events[13] were expected. However there were only 6 electron events.

This result tells that the electrons do not coupled to the produced neutrinos in the pion decay (while electrons or positrons couple to the neutrinos or antineutrinos produced in beta decay), these neutrinos tend to couple with muons. Hence, we must have at least two types of neutrino, i.e., $\nu_e \neq \nu_\mu$.

But what about those 6 shower events? Did these 6 electrons couple to the neutrinos or not. Actually not, these events were due to the background neutrons or electron muon coupling or other sources but not neutrinos produced in the pion decay [25]. Our final conclusion for this part becomes:

neutrinos or antineutrinos produced with electrons or positrons, i.e., $\overline{\nu}_e, \nu_e$ are ***completely*** *different from the neutrinos or antineutrinos produced with muons*

[11] We do not discuss it here, the detailed discussion can be found in [27], [48].

[12] In fact 34 single muon events were observed, however 5 of them were supposed to be due to the cosmic-ray background [25].

[13] Since electrons had greater energy than muons, the cross section of electron events became more than muon events.

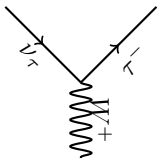


Figure 3.2: The charged current of the tau neutrino interaction.

or antimuons, i.e., $\overline{\nu}_\mu, \nu_\mu$. ***Electron neutrinos do not couple with muons and vice versa.***

3.5 Neutrino tau discovery

The existence of two neutrinos which appear with their corresponding charged leptons was proved. As there are three charged leptons, the question may be asked if there must be another neutrino flavor coming with the third generation of charged lepton τ or not. If it exists, it must be called tau neutrino(ν_τ).

The DONUT experiment (Fermilab E872) was designed to detect the tau neutrino through its charged current by identifying the lepton τ. In this experiment, a beam of neutrinos[14] interacted with emulsion targets[15]. If a tau neutrino interacted with the target, it would be expected to have a tau lepton charged current (see Fig (3.2)).

This is because the number of charged currents of muon and electron neutrinos ($\nu_\mu W^+\mu^-$ and $\nu_e W^+ e^-$) were calculated in this experiment[16]. Hence, if a tau lepton existed it would show itself in a charged current with a tau lepton. Therefore the goal was looking for tau leptons.

At the neutrino energies in this experiment, the created τ lepton decayed within 2 mm of its creation and so the scientists expected a track with a kink for the signature of the tau lepton (if it existed). Among all of the candidates for tau leptons, four events met all the requirements and so the charged current interaction for the tau lepton was also observed. The fact confirmed the existence of the third neutrino flavor, so in summary it can be said that

there are three neutrinos corresponding to the three charged leptons. Each neutrino or antineutrino couples to the corresponding charged or anti-charged lepton and avoids coupling to another charged or anti-charged leptons.

[14] These neutrinos were produced from collision of high energy (800 GeV [28]) proton (deuterium) with antiproton which resulted in charmed meson (D meson) pair $\overline{D}D$, D meson contains charm quark whose decay leads to neutrinos with different flavors.

[15] A nuclear emulsion target is a photographic plate which records the tracks of the charged particles passing through.

[16] As it was said, the high-energy protons gave a beam of neutrinos with different flavors and consequently two charged currents from electron and muon and one (if it existed) for tau were expected. Muons and electrons were detected by different methods and the existence of their charged currents was confirmed. In this experiment, after corrections, there were (94 ± 17) muon neutrino charged currents and (61 ± 14) electron neutrino ones [28].

Chapter 4

Neutrino mixing

4.1 Birth of new physics for the neutrino

In this section we are going to observe that neutrinos oscillate, i.e., if a neutrino is produced in one flavor (e.g., electron), during its travel from the production region to the place where we detect it, this neutrino can change to other flavors (e.g., muon or tau). This phenomenon is called ***neutrino mixing*** or ***oscillations***. Let us speak about observations which show this mixing. As a result of these experiments, a new theory about neutrinos in physics was born. This chapter follows mainly from Ref. [1].

4.1.1 Natural observations

Among these experiments, some of them are natural, i.e., neutrinos that are observed have not been produced in laboratories, but, we receive them from neutrino sources in our Universe, i.e.,

- Solar neutrinos.
- Atmospheric neutrinos.

Let us start from the solar neutrinos.

Solar neutrinos. The first experiment was done by Davis Jr., who detected the solar neutrinos [7], [8] i.e., the Homestake experiment[1]. We know that there are a number of neutrino-producing reactions in the Sun. They can be categorized into two groups, the first group (called proton-proton chain (pp) which produces 98.5% of the solar neutrinos [16]) initiated by $p+p \to d+e^+ + \nu_e$ and continues. This group has three phases, in phase I for instance we have these reactions:

$$p + p \to d + e^+ + \nu_e, \quad p + e^- + p \to d + \nu_e, \ \texttt{etc}, \tag{4.1}$$

[1]The Homestake is a gold mine at Lead, South Dakota, 4850 ft underground.

where d means deuterium ^{2}H. Phase II contains

$$^7\mathtt{Be} + e^- \rightarrow {}^7\mathtt{Li} + \nu_e, \ \ \mathtt{etc}, \tag{4.2}$$

and phase III includes

$$^8\mathtt{B} \rightarrow {}^7\mathtt{Be}^* + e^+ + \nu_e, \ \ \mathtt{etc}. \tag{4.3}$$

In the second group we have carbon-nitrogen cycle (CNO) which produces 1.5% of the solar neutrinos[2] [16]. In total it can be said that in the Sun, four hydrogen nuclei called protons convert into a helium nucleus (^{4}He), two antielectrons and two electron neutrinos, i.e.,

$$4p \rightarrow {}^4\mathtt{He} + 2e^+ + 2\nu_e. \tag{4.4}$$

In the above reactions we see that only electron neutrinos are produced (the Sun is a source of electron neutrinos). Moreover, due to the fact that neutrinos interact weakly with matter, we can expect that these electron neutrinos escape from the Sun easily. On earth our detectors are ready to count the number of these neutrinos.

In the Homestake experiment, the total flux of ***high energy*** neutrinos are tested. These neutrinos are produced either from electron capture on ^{7}Be (4.2) or from beta decay of ^{8}B (4.3) which both result in high energy neutrinos (above 0.814 MeV) [15]. A radiochemical technique based on the inverse beta decay is used for detecting electron neutrinos. In fact, neutrinos are detected from the number of radioactive argon atoms ^{37}Ar produced in a large tank of 390000 liters tetrachloroethylene ($\mathtt{C_2Cl_4}$) containing 520 tons ^{37}Cl [22], i.e.,

$$\nu_e + {}^{37}\mathtt{Cl} \rightarrow {}^{37}\mathtt{Ar} + e^-. \tag{4.5}$$

Now our duty is to remove the produced ^{37}Ar from tetrachloroethylene and observe the decay of ^{37}Ar with half-life of 35 days. It was done through some processes and the result stated that the solar neutrino flux was 3×10^{-36} sec^{-1} per ^{37}Cl atoms [22] or 3 SNU (solar-neutrino units) where a solar neutrino unit is one interaction per 10^{36} target atoms sec^{-1} [15], [16].

On the other hand, theoretically, the number of neutrinos produced in the Sun can be calculated. The number of detected solar neutrinos in the experiment was around one third of the neutrinos predicted by theories, the fact which became famous as the ***solar neutrino problem*** or mystery of the missing neutrinos [8], [15].

There were three possibilities, 1- error in the theoretical calculations, 2- the experiment was done wrongly, 3- one should have searched for new physics for the neutrinos. After checking again and doing a series of tests, it became clear that there was no error in neither theoretical nor experimental issues [8], [15]. Therefore only the third possibility stayed.

Additionally, in another experiment called Kamiokande II (Kamioka Nucleon

[2]To see a list of all of the nuclear reactions in the Sun, see Ref. [15].

Decay Experiment in its second version) solar neutrino problem was tested. Again, for high energy neutrinos but this time exclusively neutrinos from beta decay (4.3).

Kamiokande II is an imaging water Cherenkov detector located 1000 m underground containing 2142 metric tons of water in which the energetic solar neutrinos coming from the Sun collide with the water's electrons in an elastic scattering $\nu_e e \to \nu_e e$.

Cherenkov radiation means that if a charged particle (e.g., an electron) moves in a dielectric medium (e.g., water) at a speed greater than the speed of light in that medium[3] we will have a blue glow.

This is because the charged particle interacts with the medium and excites the molecules. When the excited molecules return to the ground state, they release energy with emitting a photon. The intensity of the photon is not considerable. However, since our charged particle moves faster than the speed of light in that medium, behind the charged particle a light cone forms where the intensity of light becomes very high (as an aircraft moving at the speed greater than the sound speed) and we receive a blue radiation.

In Kamiokande II, energetic neutrinos (coming from the Sun) give their energy to the electrons and the result of this exchange of energy will be Cherenkov radiation. With the help of the photomultiplier tubes (PMT's), the intensity of this radiation is measured[4] and the result, in comparison to the theoretical calculation, confirms the solar neutrino problem [16].

In fact, in this experiment, solar neutrino flux was measured based on 1040 days of data collection, 450 days from January 1987 until May 1988 for $E_e > 9.3$ MeV and 590 days from June 1988 until April 1990 for $E_e > 7.5$ MeV. The ratios of the experimental data to the theoretical calculation from SSM for both energy threshold cases were

$$\begin{array}{ll} \texttt{Data/SSM} = 0.46 \pm 0.05\texttt{(stat)} \pm 0.06\texttt{(syst)} & E_e > 9.3, \\ \texttt{Data/SSM} = 0.70 \pm 0.08\texttt{(stat)} \pm 0.09\texttt{(syst)} & E_e > 7.5. \end{array} \tag{4.6}$$

Since the ratios are not around one, the solar neutrino problem was confirmed [16]. Here (stat) and (syst) refer to the statical and systematic uncertainties.

All of the above experiments have confirmed the Solar neutrino problem for high-energy neutrinos. This phenomenon has also been observed for the low-energy neutrinos (produced in the (pp) reactions) in the gallium experiments GALLEX[5] [17] and SAGE[6] [23]. In both experiments the rates of the detected neutrinos are well below the SSM predictions [17], [23].

[3] According to the relativity no particle can move faster than light, but as light passes through a medium its speed reduces. Here, we mean that a particle moves at a speed greater than light in that medium but still less than light in vacuum.

[4] The angle of the incident neutrinos coming from the Sun with our detectors is important and must be taken into account.

[5] GALLEX II received neutrinos from 19 August 1992 until 22 June 1994, before that we had GALLEX I which started its detection from 14 May 1991 until 29 April 1992. In [17] the results of the GALLEX II experiment and analysis of the received data in GALLEX I can be found.

[6] In both of these experiments (GALLEX and SAGE), very low-energy neutrinos with energy threshold 233 KeV are monitored through counting of the germanium (Ge) in the

Atmospheric neutrinos. Cosmic rays are highly energetic particles such as protons or atomic nuclei in our Universe. These high-energy particles collide with the earth's atmosphere, the result of this impact being showers of hadrons (baryons and mesons). The hadrons decay and produce atmospheric neutrinos. For instance a process such as

$$\pi^+ \to \mu^+ + \nu_\mu, \tag{4.7}$$

followed by

$$\mu^+ \to e^+ + \overline{\nu}_\mu + \nu_e, \tag{4.8}$$

produce neutrinos. In reaction (4.7), there is a pion (meson) π^+ which results in muonic (4.7) and electronic (4.8) neutrinos. The ratio of the number of muon-like ($\mu-$like) to electron-like ($e-$like) events (μ/e) has been calculated both from the received data and Monte Carlo (MC) simulation (a method for calculating theoretically), i.e.,

$$R \equiv (\mu/e)_{\rm DATA} \,/\, (\mu/e)_{\rm MC}\,.$$

It is expected that R must be one, however, from experimental measurement one obtains significantly small values for R. Therefore again we encounter the mystery of the missing neutrinos [14].

4.1.2 Solution

The solution of this problem is described by neutrino oscillations which was first proposed by Pontecorvo [47]. In fact, neutrinos are created at one especial flavor for instance electron neutrinos in the Sun, but as they propagate (from the Sun to the earth), their flavor states can change to muon or tau neutrinos and so our detectors on earth detect less electron neutrinos [47].

In other words, if we sensitize our detectors to muon and tau neutrinos as well as electron, it is expected to receive the total flux of neutrinos coming from the Sun (or everywhere else) because neutrino oscillations theory believes that we do not miss any neutrino, but, there are just transformations between neutrino flavors.

So far, we have discussed that neutrinos have been detected either through an inverse beta decay (such as the Homestake, GALLEX and SAGE) or elastic scattering in Kamiokande II. In the case of the inverse beta decay, we are detecting neutrinos through charge current (CC), i.e., ($\nu_e W^+ e^-$) and consequently, we miss other neutrino flavors since they do not take part in the CC with electron. Hence, to detect other flavors we have to look for the neutral current reactions (Fig. (4.1)).

A very important point. As it was said in the Kamiokande II experiment, we use electron neutrino scattering to detect neutrinos and so Fig. (4.1) is possible. Therefore, elastic scattering (i.e., Kamiokande II experiment) is sensitive

inverse beta decay reaction of ^{71}Ga to electron, electron neutrino and Ge. The target of the GALLEX experiment is $GaCl_3$ while the target of the SAGE is gallium metal.

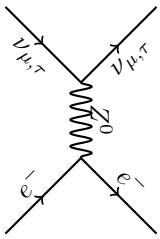


Figure 4.1: The muon and tau neutrino scatterings with electron.

to all flavors of neutrinos.

However, the differential cross section of $(\nu_e e)$ is approximately six times larger than $(\nu_{\mu,\tau}, e)$ [20]. As a result, sensitivity of the elastic scattering to electron neutrino is bigger than muon and tau neutrinos. This leads to this fact that elastic scattering in Kamikande II experiment still shows the solar neutrino problem or in other words, the elastic scattering will not give us the total flux of the coming neutrinos.

In the Sudbury Neutrino Observatory [7] (SNO) the ^{8}B solar neutrinos flux is measured through the following CC, NC and ES reactions [21]

$$\begin{array}{ll} \nu_e + d \to p + p + e^- & \text{(CC)} \\ \nu_x + d \to p + n + \nu_x & \text{(NC)} \\ \nu_x + e^- \to \nu_x + e^- & \text{(ES)} \end{array} \tag{4.9}$$

where x refers to the three neutrino flavors. The CC reaction is exclusively sensitive to the electron neutrino but ES is sensitive to all flavors (although the sensitivity to muon and tau neutrinos are less than electron). Therefore if

$$\phi^{\mathtt{CC}}(\nu_e) < \phi^{\mathtt{ES}}(\nu_x),$$

then it is proved that all neutrinos received at the detectors must not be electron although they were produced in the Sun in electronic neutrino states. From the measured ^{8}B neutrino fluxes in SNO for CC and ES we have [21]

$$\phi^{\mathtt{CC}}_{\mathtt{SNO}}(\nu_e) = 1.75 \pm 0.07(\mathtt{stat})^{+0.12}_{-0.11}\mathtt{syst} \pm 0.05(\mathtt{theor}) \times 10^6 \mathtt{cm}^{-2}\mathtt{s}^{-1}, \tag{4.10}$$

$$\phi^{\mathtt{ES}}_{\mathtt{SNO}}(\nu_x) = 2.39 \pm 0.34(\mathtt{stat})^{+0.16}_{-0.14}\mathtt{syst} \times 10^6 \mathtt{cm}^{-2}\mathtt{s}^{-1}, \tag{4.11}$$

where (theor) refers to the theoretical uncertainty of the CC cross section. Therefore (4.10) and (4.11) confirm that neutrinos oscillate. Furthermore, regarding NC, we obtain [24]

$$\phi^{\mathtt{NC}}_{\mathtt{SNO}}(\nu_x) = 5.09^{+0.44}_{-0.43}(\mathtt{stat})^{+0.46}_{-0.43}\mathtt{syst} \times 10^6 \mathtt{cm}^{-2}\mathtt{s}^{-1}, \tag{4.12}$$

[7] SNO is an imaging water Cherenkov detector located at a depth of 6010 m of water equivalent in the INCO, Ltd. Creighton mine near Sudbury, Ontario. It contains 1000 metric tons of ultrapure D_2O with PMT's to measure produced light intensity, the data have been reported between November 2, 1999 and January 15, 2001 [21].

which again proves the existence of another neutrino flavors on earth although they are just produced in electron flavors. Neutrino oscillations have also been observed in the atmospheric neutrinos in the Super-Kamiokande experiment [14].

4.1.3 What do neutrino oscillations say about the neutrino mass?

Without a doubt if neutrinos oscillate (as they do), they must be massive. Because neutrinos are produced in one special flavor in the Sun or everywhere else but as they propagate in vacuum, they change to another flavors. In other words, neutrinos ***feel time***. According to the relativity, if a particle such as photon moves at the speed of light, it must not feel time whereas neutrinos do. Therefore, they must propagate at the speed of less than light and are massive, i.e.,

$$\texttt{Neutrino Oscillations} \rightarrow \texttt{Massive Neutrino}.$$

4.2 Neutrino behavior

4.2.1 Dirac neutrino

As mentioned in the previous chapter, neutrinos do not receive their mass in interaction with the Higgs fields. If they are massive, the mass term of the free Lagrangian for the neutrinos is (see (2.3))

$$\hat{\mathcal{L}} = -\sum_{\alpha',\alpha} \hat{\bar{\nu}}_{\alpha' L}(x)\, M^{D}_{\alpha',\alpha} \hat{\nu}_{\alpha R}(x) + \texttt{h.c.}, \tag{4.13}$$

where α and α' refer to the neutrino flavors, M^D is a complex 3×3 matrix[8] which makes the Lagrangian a scalar and h.c. refers to the hermitian conjugate. It was shown in (2.16) that the Lagrangian which describes a fermion mass must include both left- and right-handed fermions.

The Lagrangian (4.13) is a an operator, therefore we can write the matrix representation of this operator in a basis of eigenvectors. For instance, for flavor eigenstates as a basis we have

$$\hat{\mathcal{L}} = \begin{pmatrix} \langle \nu_{eL}| \hat{\mathcal{L}} |\nu_{eR}\rangle & \langle \nu_{eL}| \hat{\mathcal{L}} |\nu_{\mu R}\rangle & \langle \nu_{eL}| \hat{\mathcal{L}} |\nu_{\tau R}\rangle \\ \langle \nu_{\mu L}| \hat{\mathcal{L}} |\nu_{eR}\rangle & \langle \nu_{\mu L}| \hat{\mathcal{L}} |\nu_{\mu R}\rangle & \langle \nu_{\mu L}| \hat{\mathcal{L}} |\nu_{\tau R}\rangle \\ \langle \nu_{\tau L}| \hat{\mathcal{L}} |\nu_{eR}\rangle & \langle \nu_{\tau L}| \hat{\mathcal{L}} |\nu_{\mu R}\rangle & \langle \nu_{\tau L}| \hat{\mathcal{L}} |\nu_{\tau L}\rangle \end{pmatrix}. \tag{4.14}$$

This matrix is not diagonal because the non-diagonal terms such as $\langle \nu_{eL}| \hat{\mathcal{L}} |\nu_{\tau R}\rangle$ do not vanish, this is due to the non-zero off-diagonal mass terms in the matrix M.

In fact, interactions between the left- and right-handed neutrino field operators (see (2.10)), i.e., $\hat{\bar{\nu}}_{\alpha' L}(x)$ and $\hat{\nu}_{\alpha R}(x)$ with the left- and right-handed

[8] The superscript D refers to the Dirac neutrino.

neutrino states, i.e., $\langle \nu_{\alpha' L}|$ and $|\nu_{\alpha R}\rangle$ result in two Dirac deltas which remove integrals, and then we are left with off-diagonal terms $M_{\alpha'\alpha}$. As a result, eigenvectors $|\nu_{\alpha L}\rangle$ and $|\nu_{\alpha R}\rangle$ are not the Lagrangian eigenvectors.

Now, let us find the basis in which the Lagrangian matrix becomes diagonal, i.e., we should diagonalize (4.14). To do so, it is enough to find the basis in which the matrix M becomes diagonal. Let us denote this basis by $|\nu_i\rangle$ for $i = 1, 2, 3$. Then,

$$\langle \nu_{iL}|\, \hat{\mathcal{L}}\, |\nu_{jR}\rangle = m_i \delta_{ij}, \tag{4.15}$$

where the matrix m is diagonal,

$$m = \begin{pmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{pmatrix}. \tag{4.16}$$

From the completeness relation one can write

$$|\nu_{\alpha L}\rangle = \sum_i |\nu_{iL}\rangle\, U_{i\alpha}; \quad \text{and} \quad |\nu_{\alpha R}\rangle = \sum_i |\nu_{iR}\rangle\, V_{i\alpha}, \tag{4.17}$$

where $U_{i\alpha} = \langle \nu_{iL}\, |\nu_{\alpha L}\rangle$ and $V_{i\alpha} = \langle \nu_{iR}\, |\nu_{\alpha R}\rangle$ are *unitary mixing matrices*[9] [1]. These matrix elements are the projections of the basis $|\nu_i\rangle$ into $|\nu_\alpha\rangle$. Now, regarding (4.17) we obtain

$$\langle \nu_{\alpha L}|\, \hat{\mathcal{L}}\, |\nu_{\beta R}\rangle = M_{\alpha\beta} = \sum_{i,j} U^*_{\alpha i} \underbrace{\langle \nu_{iL}|\, \hat{\mathcal{L}}\, |\nu_{jR}\rangle}_{m_i \delta_{ij}} V_{j\beta}, \tag{4.18}$$

or in the matrix form

$$M^D = U^\dagger m V. \tag{4.19}$$

It is clear that, now, $|\nu_i\rangle$ are the eigenvectors of Lagrangian (4.13) (contrary to the flavor eigenstates which are not). By putting (4.19) into (4.13), it is obtained that

$$\hat{\mathcal{L}} = -\sum_{\alpha',\alpha} \hat{\bar{\nu}}_{\alpha' L}(x) \left[\sum_{i',i} m_i U^*_{\alpha' i'} \delta_{i' i} V_{i\alpha} \right] \hat{\nu}_{\alpha R}(x) + \texttt{h.c.} \tag{4.20}$$

Now since $\hat{\bar{\nu}}_{\alpha' L}(x)$ and $\hat{\nu}_{\alpha R}(x)$ are dependent on α' and α, we can bring them into the sigma over i', i. Therefore $\hat{\bar{\nu}}_{i' L}(x)$ and $\hat{\nu}_{iR}(x)$ read

$$\hat{\bar{\nu}}_{i' L} = \sum_{\alpha'} \hat{\bar{\nu}}_{\alpha' L} U^*_{\alpha' i'} \Rightarrow \hat{\nu}_{i' L} = \sum_{\alpha'} U_{i'\alpha'} \hat{\nu}_{\alpha' L} \tag{4.21}$$

[9] Here, we have supposed that left- and right-handed neutrinos are transformed with different unitary operators U and V, respectively.

and

$$\hat{\nu}_{iR} = \sum_{\alpha} V_{i\alpha} \hat{\nu}_{\alpha R}. \tag{4.22}$$

Relations (4.21) and (4.22) can also be written as

$$\hat{\nu}_{\alpha L}(x) = \sum_{i=1}^{3} U^{*}_{\alpha i} \hat{\nu}_{iL}(x), \tag{4.23}$$

and

$$\hat{\nu}_{\alpha R}(x) = \sum_{i=1}^{3} V^{*}_{\alpha i} \hat{\nu}_{iR}(x). \tag{4.24}$$

The unitarity of the mixing matrices U and V [1] confirms orthogonality of the mass eigenstates, because for example from (4.17) we have

$$|\nu_{iL}\rangle = \sum_{\alpha} U^{*}_{i\alpha} |\nu_{\alpha L}\rangle \Rightarrow \langle \nu_{jL} | \nu_{iL} \rangle = \sum_{\alpha,\alpha'} U^{*}_{i\alpha} U_{\alpha' j} \underbrace{\langle \nu_{\alpha' L} | \nu_{\alpha L} \rangle}_{\delta_{\alpha\alpha'}} \Rightarrow$$

$$\langle \nu_{jL} | \nu_{iL} \rangle = \delta_{ij}. \tag{4.25}$$

It is obvious that we can derive a similar relation for the right-handed neutrinos.

From (4.21) and (4.22) one can write Lagrangian (4.20) as follows

$$\begin{aligned} \hat{\mathcal{L}} &= -\sum_{i',i} m_i \left[\hat{\overline{\nu}}_{i'L}(x) \delta_{i'i} \hat{\nu}_{iR}(x) + \hat{\overline{\nu}}_{iR}(x) \delta_{ii'} \hat{\nu}_{i'L}(x) \right] \\ &= -\sum_{i=1}^{3} m_i \hat{\overline{\nu}}_i(x) \hat{\nu}_i(x). \end{aligned} \tag{4.26}$$

Lagrangian (4.13) contains the Dirac mass term, i.e., we have been discussing the Dirac neutrinos, this is the reason why the superscript D was used. In the case of the Dirac neutrino, (4.26) is invariant under the global transformations, i.e.,

$$\nu_i \to \nu'_i = \exp(i\Lambda) \nu_i, \tag{4.27}$$

for a constant Λ. The result of this symmetry must be a conserved quantum number, i.e., lepton number. As already said, if we set neutrino lepton number 1_α for ($\alpha = e, \mu$ and τ) and -1_α for each antineutrino, there must be lepton number conservation. We will see that in another case (Majorana neutrino) lepton number is not conserved.

Why do neutrinos oscillate? In fact oscillations stem from the fact that (4.14) (i.e., the matrix M) is not diagonal in the basis of flavor eigenstates and becomes diagonal in the basis of $|\nu_i\rangle$. Since (4.13) is the free (non-interacting) Lagrangian, i.e., mass of neutrino is not produced in an interaction[10], we call

[10] Contrary to the bosons that receive their masses in interactions with the Higgs boson field.

the eigenstates $|\nu_i\rangle$ the ***propagation eigenstates*** or ***mass eigenstates***[11]. The mass and flavor eigenstates are totally different, and this difference causes oscillations phenomenon.

As it was said, Lagrangian (4.13) is not the interaction Lagrangian. Flavor neutrinos are born in a weak interaction. The interaction Lagrangian $\mathcal{L}_{int}$ is

$$\hat{\mathcal{L}}_{int} = \overline{\nu}_\alpha W^{\pm} l_\alpha, \tag{4.28}$$

where $W^{\pm}$ are the weak boson fields, and l_α is the charged lepton field. It is completely clear that this Lagrangian is diagonal in the flavor basis only if the neutrino and charged lepton taking part in this interaction have the same flavor. Exactly what we have in the nature. Hence, if we define a quantum number called flavor, it is conserved by $\hat{\mathcal{L}}_{int}$ while violated by (4.13).

From now on, we call (4.13) the propagation Lagrangian and show it by $\hat{\mathcal{L}}_{\text{propag}}$. These two operators $\left(\hat{\mathcal{L}}_{in}, \hat{\mathcal{L}}_{\text{propag}}\right)$ do not commute with each other and consequently, they do not share the same eigenstates.

Let us elaborate more on this issue. Flavor neutrinos are produced in an interaction, this born flavor neutrino is a vector in another Hilbert space whose components are the mass eigenstates. This means that this vector is the superposition of the mass eigenstates, i.e.,

$$|\nu_{eL}\rangle = U_{e1}\,|\nu_{1L}\rangle + U_{e2}\,|\nu_{2L}\rangle + U_{e3}\,|\nu_{3L}\rangle\,, \tag{4.29}$$

$$|\nu_{\mu L}\rangle = U_{\mu 1}\,|\nu_{1L}\rangle + U_{\mu 2}\,|\nu_{2L}\rangle + U_{\mu 3}\,|\nu_{3L}\rangle\,, \tag{4.30}$$

$$|\nu_{\tau L}\rangle = U_{\tau 1}\,|\nu_{1L}\rangle + U_{\tau 2}\,|\nu_{2L}\rangle + U_{\tau 3}\,|\nu_{3L}\rangle\,. \tag{4.31}$$

Similarly, we can write for the right-handed flavors but with the help of the matrix V. In chapters 5 and 6, it will be seen how oscillations happen when neutrinos propagate.

4.2.2 Majorana neutrino

To discuss Majorana neutrino, we should first find charge conjugated Dirac field ψ_c, i.e., how the Dirac field (2.10) transforms under the charge conjugation operator. With the help of (2.7), (2.8) and (2.9) we see that the charge conjugated Dirac field becomes[12]

$$\left.\begin{array}{l} u^s(p) = -i\gamma^2 \left(v^s(p)\right)^* \\ v^s(p) = -i\gamma^2 \left(u^s(p)\right)^* \end{array}\right\} \Rightarrow \psi_c = \mathcal{C}\psi\mathcal{C}^{-1} = -i\left(\overline{\psi}\gamma^0\gamma^2\right)^T. \tag{4.32}$$

Now let us define the matrix C as

$$C = i\gamma^0\gamma^2 \Rightarrow C^T = -C \Rightarrow \psi_c = C\overline{\psi}^T = C\gamma^0\psi^*, \tag{4.33}$$

and also from matrix C definition it is seen that

$$C\gamma_\mu^T C^{-1} = -\gamma^\mu, \tag{4.34}$$

[11] Owing to this fact we may show them by $|m_i\rangle$ as well as $|\nu_i\rangle$.

[12] For the detailed calculation, see Ref. [5].

while

$$\left[C, \gamma^5\right] = 0. \tag{4.35}$$

Let us denote neutrino field ψ (for each flavor α) by ν_α. Then

$$(\nu_\alpha)^c = C\overline{\nu}_\alpha^T, \tag{4.36}$$

and consequently, we have

$$(\nu_{\alpha L})^c = C\overline{\nu}_{\alpha L}^T, \tag{4.37}$$

and

$$(\nu_{\alpha R})^c = C\overline{\nu}_{\alpha R}^T. \tag{4.38}$$

If we write $\nu_{\alpha L} = \begin{pmatrix} \nu_{\alpha L} \\ 0 \end{pmatrix}$ and $\nu_{\alpha R} = \begin{pmatrix} 0 \\ \nu_{\alpha R} \end{pmatrix}$, then

$$\gamma^5 \nu_{\alpha L} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \nu_{\alpha L} \\ 0 \end{pmatrix} = \begin{pmatrix} -\nu_{\alpha L} \\ 0 \end{pmatrix} \Rightarrow \gamma^5 \nu_{\alpha L} = -\nu_{\alpha L} \tag{4.39}$$

$$\Rightarrow \overline{\nu}_{\alpha L}\gamma^5 = +\overline{\nu}_{\alpha L} \Rightarrow \gamma^5 \overline{\nu}_{\alpha L}^T = +\overline{\nu}_{\alpha L}^T$$

$$\Rightarrow \gamma^5 (\nu_{\alpha L})^c = (\nu_{\alpha L})^c, \tag{4.40}$$

where (4.35) and (4.37) have been used. With the exactly similar reasoning, it is obtained that

$$\gamma^5 \nu_{\alpha R} = \nu_{\alpha R} \Rightarrow \tag{4.41}$$

$$\gamma^5 (\nu_{\alpha R})^c = -(\nu_{\alpha R})^c. \tag{4.42}$$

By comparing (4.40) with (4.41) it is understood that $(\nu_{\alpha L})^c$ must be right-handed component of the charged conjugated neutrino field, whereas from (4.39) and (4.42) we see that $(\nu_{\alpha R})^c$ must be left-handed component.

This result allows us to form another Lagrangian for the neutrino mass term, here we enter $\overline{\nu}_{\alpha L}$ (i.e., the left-handed field, exactly such as Dirac case) and $(\nu_{\alpha L})_c$ (i.e., the right-handed field). The new Lagrangian is called the **Majorana Lagrangian**[13] $\left(\mathcal{L}^M\right)$

$$\mathcal{L}^M = -\frac{1}{2} \sum_{\alpha', \alpha} \overline{\nu}_{\alpha' L} M^M_{\alpha' \alpha} (\nu_{\alpha L})_c + \texttt{h.c.}, \tag{4.43}$$

where M^M is the complex non-diagonal **Majorana mass matrix**. If a neutrino is described by the Majorana Lagrangian, we call it the **Majorana neutrino**[14].

Same as the Dirac mass term, here we should diagonalize M^M to find each

[13] In other words, this Lagrangian is a mixture of the neutrino field and neutrino conjugated field which may result in the neutrino mass.

[14] We do not know if neutrinos are Dirac or Majorana yet [1].

flavor mass. First, it can be easily shown that M^M is a symmetric matrix. Since the Lagrangian is scalar, we then have

$$\mathcal{L}^M = \left(\mathcal{L}^M\right)^T \Rightarrow \overline{\nu}_{\alpha' L} M^M_{\alpha'\alpha} C \overline{\nu}^T_{\alpha L} = -\overline{\nu}_{\alpha L} \left(M^M_{\alpha'\alpha}\right)^T C^T \overline{\nu}^T_{\alpha' L} = \\ -\overline{\nu}_{\alpha L} M^M_{\alpha\alpha'} C^T \overline{\nu}^T_{\alpha' L} = \overline{\nu}_{\alpha L} M^M_{\alpha\alpha'} C \overline{\nu}^T_{\alpha' L} \Rightarrow M^M = \left(M^M\right)^T, \tag{4.44}$$

and so, the Majorana mass matrix is symmetric. Here, in the first line one minus sign was entered, due to the fact that when we transpose the Lagrangian, $\overline{\nu}^T_{\alpha L}$ is brought to the left of $\overline{\nu}_{\alpha' L}$. Neutrinos are fermions and when the orders are changed (owing to the anticommutation of fermions) a minus sign appears. However, in the second line this minus sign vanishes because $C^T = -C$. A symmetric matrix can be diagonalized with the help of a unitary matrix[15] such as U, i.e.,

$$M^M = U m U^T, \tag{4.45}$$

where the matrix m is the diagonalized matrix

$$m = \begin{pmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{pmatrix}. \tag{4.46}$$

We know that $\nu_{\alpha L}$ is a 4×1 matrix. Let us define a 3×1 matrix ν_L (whose each component is 4×1) as

$$\nu_L = \begin{pmatrix} \nu_{eL} \\ \nu_{\mu L} \\ \nu_{\tau L} \end{pmatrix}, \tag{4.47}$$

so, Lagrangian (4.43) can be written as

$$\mathcal{L}^M = -\frac{1}{2} \overline{\nu}_L M^M \left(\nu_L\right)_c + \texttt{h.c.} \tag{4.48}$$

By plugging (4.45) into the above Lagrangian, it becomes

$$\mathcal{L}^M = -\frac{1}{2} \overline{\nu}_L U m U^T \left(\nu_L\right)_c + \texttt{h.c.} \tag{4.49}$$

Now we write

$$\overline{\nu}_L U = \nu^\dagger_L \gamma^0 U, \tag{4.50}$$

the matrix U is 3×3, it can be exchanged with γ^0, hence

$$\overline{\nu}_L U = \nu^\dagger_L \gamma^0 U = \nu^\dagger_L \left(U^\dagger\right)^\dagger \gamma^0 = \left(U^\dagger \nu_L\right)^\dagger \gamma^0 = \overline{U^\dagger \nu_L}. \tag{4.51}$$

In addition, we can exchange the 4×4 matrix C with the 3×3 matrix U^T, therefore

$$U^T \left(\nu_L\right)_c = U^T C \overline{\nu}^T_L = C U^T \overline{\nu}^T_L = C \left(\overline{\nu}_L U\right)^T = C \left(\nu^\dagger_L \gamma^0 \left(U^\dagger\right)^\dagger\right)^T \\ = C \left(\left(U^\dagger \nu_L\right)^\dagger \gamma^0\right)^T = C \overline{U^\dagger \nu_L}^T = \left(U^\dagger \nu_L\right)_c. \tag{4.52}$$

[15] The detailed calculation has been done in Ref. [1].

From (4.51) and (4.52), Lagrangian (4.49) reads

$$\begin{aligned}\mathcal{L}^M &= -\frac{1}{2}\overline{U^\dagger \nu_L}\; m\; \left(U^\dagger \nu_L\right)_c + \texttt{h.c.} \\ &= -\frac{1}{2}\overline{U^\dagger \nu_L}\; m\; \left(U^\dagger \nu_L\right)_c - \frac{1}{2}\overline{(U^\dagger \nu_L)_c}\; m\, U^\dagger \nu_L. \end{aligned} \tag{4.53}$$

Now we would like to define ν_L^M and ν^M as

$$\left.\begin{aligned} \nu_L^M &= U^\dagger \nu_L \\ \nu^M &= \nu_L^M + \left(\nu_L^M\right)_c \end{aligned}\right\} \Rightarrow \tag{4.54}$$

$$\begin{aligned} &-\frac{1}{2}\overline{\nu}^M m \nu^M = \\ &-\frac{1}{2}\overline{U^\dagger \nu_L}\; m\; \left(U^\dagger \nu_L\right)_c - \frac{1}{2}\overline{(U^\dagger \nu_L)_c}\; m\, U^\dagger \nu_L + \overline{\nu_L^M}\nu_L^M + \left(\overline{\nu_L^M}\right)_c \nu_L^M. \end{aligned} \tag{4.55}$$

From (4.33) we have

$$\left(\nu_L^M\right)_c = C\gamma^0 \left(\nu_L^M\right)^* \Rightarrow \overline{\nu_L^M}\nu_L^M + \left(\overline{\nu_L^M}\right)_c \nu_L^M = \overline{\nu_L^M}\nu_L^M - \left(\overline{\nu_L^M}\nu_L^M\right)^* = 0. \tag{4.56}$$

We know that $\overline{\nu_L^M}\nu_L^M$ is a real scalar, thus its complex conjugate equals itself. Therefore, the above result becomes zero and consequently, the last two terms in (4.55) vanish. Finally, Lagrangian (4.53) takes this form

$$\mathcal{L}^M = -\frac{1}{2}\overline{\nu}^M m \nu^M. \tag{4.57}$$

Note. Before continuing, it is worth noting that according to our definition for ν^M, we see that it consists of two terms, the first term is left-handed and the second one is right-handed (as discussed above). Now, we define

$$\nu_R^M = \left(\nu_L^M\right)_c, \tag{4.58}$$

and

$$\nu^M = \nu_L^M + \nu_R^M. \tag{4.59}$$

In the case of the Majorana neutrino, left- and right-handed components are connected through (4.59). This is a characteristic of the Majorana neutrino, because in the Dirac Lagrangian we define right- and left-handed neutrinos independently[16].

Let us come back to our discussion. In the case of the Dirac neutrinos, we define ν_i (for $i = 1, 2, 3$) which are different from ν_α. These two were related

[16]There will be another difference between the Majorana and Dirac cases which will be discussed very soon.

to each other in (4.21) and (4.22) via unitary matrices U and V. Similarly, in the case of the Majorana neutrinos we define ν_i for the field of neutrino with mass m_i. From (4.54), it is seen that how ν_L goes under a unitary matrix transformations. Thus same as (4.21) and (4.22) the result of these transformations should bring us the fields of the massive neutrinos, i.e.,

$$\nu^M = U^\dagger \nu_L + \left(U^\dagger \nu_L\right)_c = \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \Rightarrow \mathcal{L}^M = -\frac{1}{2}\sum_{i=1}^{3} m_i \overline{\nu}_i \nu_i. \tag{4.60}$$

What is $\left(\nu^M\right)_c$? To find it we use (4.58) and (4.59), then

$$\nu^M = \nu_L^M + \left(\nu_L^M\right)_c \Rightarrow \left(\nu^M\right)_c = \left(\nu_L^M\right)_c + \nu_L^M = \nu^M \Rightarrow \left(\nu_i\right)_c = \nu_i. \tag{4.61}$$

This important result is called the **Majorana condition**. Let us see what is the result of this condition in more detail. From the quantized Dirac field (2.10) for the neutrino we have

$$\nu_i(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_\mathbf{p}}} \sum_s \left(a_\mathbf{p}^s u^s(p) \exp(-ip\cdot x) + b_\mathbf{p}^{s\dagger} v^s(p) \exp(+ip\cdot x) \right), \tag{4.62}$$

then regarding (4.33), $\left(\nu_i\right)_c$ becomes

$$\left(\nu_i\right)_c = \int \frac{d^3\mathbf{p}}{(2\pi)^3} \frac{1}{\sqrt{2E_\mathbf{p}}} \sum_s \left(b_\mathbf{p}^s u^s(p) \exp(-ip\cdot x) + a_\mathbf{p}^{s\dagger} v^s(p) \exp(+ip\cdot x) \right). \tag{4.63}$$

From the Majorana condition (4.61), it is concluded that

$$a_\mathbf{p}^s = b_\mathbf{p}^s, \tag{4.64}$$

which means that the Majorana neutrino and antineutrino are not two distinct particles, but one particle

$$\texttt{Majorana neutrino} \equiv \texttt{Majorana antineutrino}. \tag{4.65}$$

There is still one unanswered question! Does the Majorana Lagrangian satisfy lepton number conservation? To answer it, we should know whether or not (4.53) is invariant under the global transformations

$$\nu_L \to \nu_L' = \exp(i\Lambda)\,\nu_L, \tag{4.66}$$

for constant Λ. Under global transformations we have

$$\left.\begin{aligned} \left(U^\dagger \nu_L\right)_c &\to \left(U^\dagger \nu_L'\right)_c = \exp(-i\Lambda)\left(U^\dagger \nu_L\right)_c \\ \overline{U^\dagger \nu_L} &\to \overline{U^\dagger \nu_L'} = \exp(-i\Lambda)\,\overline{U^\dagger \nu_L} \end{aligned}\right\} \Rightarrow \mathcal{L}'^M \neq \mathcal{L}^M, \tag{4.67}$$

and so, the answer of the above question (conservation of lepton number in the Majorana case) is NO.

This result is consistent with what we already said in the introduction chapter. There, it was said that Davis and Harmer concluded that neutrinos differ with antineutrinos and so they claimed that we could not have lepton number violation. But now, we claim that ***if and only if*** *neutrinos are Dirac, Davis and Harmer's claim is acceptable and if we go beyond the Dirac neutrinos, we expect lepton number violation, in this case there is no difference between neutrinos and antineutrinos at all.*

4.3 Neutrino mixing matrix

4.3.1 Number of angles and phases in the Dirac and Majorana neutrino mixing

Consider the unitary mixing matrix U. This matrix belongs to $U(n)$ group. As for an orthogonal matrix $O(n)$ we have

$$O^T(n)O(n) = \mathbb{1}. \tag{4.68}$$

For instance in the case $n = 2$, from the above condition we see

$$O = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \mathbb{1} \Rightarrow \begin{cases} a^2 + b^2 = c^2 + d^2 = 1 \\ \\ ac + bd = 0. \end{cases} \tag{4.69}$$

While we are left with three equations, there are four quantities and so there is just one independent quantity. From (4.69) it is seen that this matrix can be represented by one Euler-type angle (the only independent quantity) θ. In the general case, we derive the number of independent quantities (i.e., the number of angles) as follows

$$n_\theta = \frac{n\,(n-1)}{2}. \tag{4.70}$$

There is the same story for the unitary matrices. Here our condition becomes $UU^\dagger = \mathbb{1}$. With the help of this condition, for a 2×2 unitary matrix we derive again $a^2 + b^2 = 1$ but here a and b are complex. Let us write $a = re^{i\alpha}$ and $b = r'e^{i\beta}$ where r and r' are real. As a result, we reach $r^2 + r'^2 = 1$, so again this matrix can be characterized with one Euler-type angle, and in general the number of angles are (4.70).

As for the unitary matrices, apart from the so-called Euler-type angles we have phases. Phases are parameters in a unitary matrix which are not angles. In section (3.2.1), it was mentioned that in a unitary matrix there are n^2 parameters. Some of these parameters are angles. The rest are the number of phases, i.e., n_ϕ,

$$n_\phi = n^2 - \frac{n\,(n-1)}{2} = \frac{n\,(n+1)}{2}. \tag{4.71}$$

For instance in a 2×2 unitary matrix, there is one angle and three phases. However, one can be rid of the phases if the matrix goes under special transfor-

mations[17]. Every unitary matrix U' can be related to another unitary matrix U, as

$$U_{jl'} = \sum_{i,l} S_{jl}(\beta) U'_{li} \left(S_{l'i}(\alpha)\right)^{\dagger}, \tag{4.72}$$

where $S(\beta)$ and $S(\alpha)$ are both diagonal phase matrices,

$$S_{jl}(\beta) = \mathrm{e}^{i\beta_j}\delta_{jl}; \quad \text{and} \quad S_{l'i}(\alpha) = \mathrm{e}^{i\alpha_{l'}}\delta_{l'i}. \tag{4.73}$$

With the help of (4.72) and (4.73), it can be shown that [1] in the case of the Dirac neutrinos, we can reduce $(2n-1)$ of the total phases. Thus, the number of physical phases, which cannot be removed at all, become

$$n_{\phi}^{\texttt{physical}} = \frac{n(n+1)}{2} - (2n-1) = \frac{1}{2}(n-1)(n-2). \tag{4.74}$$

In the case of the Majorana neutrino, we should take the Majorana condition (4.61) into consideration. Then, it is found that [1] the Majorana mixing matrix U^M is

$$U^M = US^M(\overline{\alpha}), \tag{4.75}$$

where U is the Dirac mixing matrix and

$$S^M(\overline{\alpha}) = \mathrm{e}^{-i\alpha_n} S(\alpha). \tag{4.76}$$

Additionally, the number of Majorana phases are [1]

$$n_{\phi}^{M} = \frac{n(n-1)}{2}. \tag{4.77}$$

4.3.2 Constraints on the neutrino mixing matrices under CP invariance

Now, let us find constraints of the neutrino mixing matrices (Dirac and Majorana) in the case of the CP invariance. First the Dirac neutrino. The CP operator is V_{CP}, the charged current Lagrangian $\mathcal{L}_I^{CC}(x)$

$$\mathcal{L}_I^{CC}(x) = -\frac{g}{\sqrt{2}}\sum_{\alpha,i}\bar{l}_L(x)\gamma^{\mu}U_{li}\nu_{iL}(x)W_{\mu}^{\dagger} - \frac{g}{\sqrt{2}}\sum_{l,i}\overline{\nu}_{iL}(x)\gamma^{\mu}U_{il}^{*}l_L(x)W^{\mu},$$

becomes $\mathcal{L}_I^{CC}(x')$ under CP invariance (under parity we have: $\mathbf{x}' = -\mathbf{x}$). So,

$$\mathcal{L}_I^{CC}(x') = V_{CP}\mathcal{L}_I^{CC}(x)V_{CP}^{-1}. \tag{4.78}$$

Our strategy is to find each of the lepton field under CP transformations. Under parity and charge conjugation transformations we have

$$\left.\begin{array}{l} \mathcal{P}l(x)\mathcal{P}^{-1} = \eta_p\gamma^0 l(x') \\ \\ \mathcal{C}l(x)\mathcal{C}^{-1} = \eta_c C\bar{l}^T(x) \end{array}\right\} \Rightarrow V_{CP}l(x)V_{CP}^{-1} = \eta_{pc}\gamma^0 C\bar{l}^T(x'), \tag{4.79}$$

[17]When a phase can be removed, it means that the phase is not physical.

where η_p, η_c and η_{pc} are arbitrary phases. Similarly, one can find the transformations of the neutrino field as follows

$$V_{CP}\nu_{iL}(x)V_{CP}^{-1} = \eta'_{pc}\gamma^0 C\bar{\nu}_{iL}^T(x'), \tag{4.80}$$

with arbitrary phase η'_{pc}. Under the CP transformations the vector field of W boson is transformed as

$$V_{CP}W_\mu(x)V_{CP}^{-1} = -\eta''_{pc}\delta_\mu W_\mu^\dagger(x'), \tag{4.81}$$

with arbitrary phase η''_{pc}. Here, δ_μ is the sign factor $\delta = (1, -1, -1, -1)$. We assume that due to the arbitrariness of phases

$$\eta_{pc} = \eta'_{pc} = \eta''_{pc} = 1. \tag{4.82}$$

Hence from (4.78), (4.79), (4.80) and (4.81), we get this result:

$$V_{CP}\mathcal{L}_I^{CC}(x)V_{CP}^{-1} = -\frac{g}{\sqrt{2}}\sum_{l,i}\bar{l}_L(x')\gamma^\mu U_{li}^*\nu_{iL}(x')W_\mu^\dagger \tag{4.83}$$

$$-\frac{g}{\sqrt{2}}\sum_{l,i}\overline{\nu}_{iL}(x')\gamma^\mu U_{il}l_L(x')W^\mu, \tag{4.84}$$

and so by comparing the above result with $\mathcal{L}_I^{CC}$ it is understood that under the CP invariance, we have

$$U_{il} = U_{li}^*. \tag{4.85}$$

The above equation shows that under CP invariance (as for the Dirac neutrino), there must not be any phase in the mixing matrix. This situation is different for the Majorana neutrino. The CP phase factor in this case is not arbitrary, but, it can be $\pm i$. In addition, it is derived that the Majorana mixing matrix satisfies the condition

$$\rho_i U_{li}^M = U_{li}^{M*}; \qquad \rho_i = \pm 1, \tag{4.86}$$

if CP is invariant[18].

4.3.3 Standard neutrino mixing parametrization

Let us come back to the mixing matrix U. From (4.70) and (4.74) we see that in the case of $n = 2$, there is one angle and zero phase. Thus, in the two-generation Dirac neutrinos, the mixing matrix U becomes

$$U = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{4.87}$$

[18]The result discussed in this section will be used later on. We have not explained the process for derivation of these results with detail. The detailed calculations can be found in Ref. [1].

This matrix is real since there is no phase in it. So, automatically CP is invariant. If our neutrinos are Majorana, from (4.77) one phase is expected. Then, regarding (4.75) and (4.76), the Majorana mixing matrix becomes

$$U = \begin{pmatrix} \cos(\theta)\, e^{i\overline{\alpha}_1} & \sin(\theta) \\ -\sin(\theta)\, e^{i\overline{\alpha}_1} & \cos(\theta) \end{pmatrix}. \tag{4.88}$$

Under CP invariance, from condition (4.86), it is obtained that

$$\rho_1 = e^{-2i\overline{\alpha}_1}; \quad \text{for} \quad \rho_1 = \pm 1, \tag{4.89}$$

and $\rho_2 = 1$.

Now, let us move to the case of three generations, i.e., $n = 3$. Due to (4.70) and (4.74), we are left with three angles and one phase. As a result of the three angles, there must be three rotations around each axis: $x = 1, y = 2$ and $z = 3$. We put each neutrino vector $|\nu_i\rangle$ for $i = 1, 2, 3$ along each axis $1, 2, 3$ respectively, so the orthogonality condition (4.25) is satisfied.

To find the mixed vector, the first rotation is around $|\nu_3\rangle$ at the angle θ_{12}. Then we obtain

$$|\nu_1\rangle' = \overbrace{\cos\theta_{12}}^{c_{12}} |\nu_1\rangle + \overbrace{\sin\theta_{12}}^{s_{12}} |\nu_2\rangle \tag{4.90}$$

$$|\nu_2\rangle' = -s_{12} |\nu_1\rangle + c_{12} |\nu_2\rangle \tag{4.91}$$

$$|\nu_3\rangle' = |\nu_3\rangle, \tag{4.92}$$

and so

$$\begin{pmatrix} |\nu_1\rangle' \\ |\nu_2\rangle' \\ |\nu_3\rangle' \end{pmatrix} = \underbrace{\begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{U_1} \begin{pmatrix} |\nu_1\rangle \\ |\nu_2\rangle \\ |\nu_3\rangle \end{pmatrix}. \tag{4.93}$$

The second rotation is around $|\nu_2\rangle'$, the angle will be θ_{13}. As it was said, there is one phase. Let us define the phase δ and take the rotation at the angle θ_{13} as well as the phase transformations $e^{\pm i\delta}$ into consideration. Therefore,

$$|\nu_1\rangle'' = c_{13} |\nu_1\rangle' + s_{13} e^{-i\delta} |\nu_3\rangle' \tag{4.94}$$

$$|\nu_2\rangle'' = |\nu_2\rangle' \tag{4.95}$$

$$|\nu_3\rangle'' = -s_{13} e^{i\delta} |\nu_1\rangle' + c_{13} |\nu_3\rangle', \tag{4.96}$$

and so

$$\begin{pmatrix} |\nu_1\rangle'' \\ |\nu_2\rangle'' \\ |\nu_3\rangle'' \end{pmatrix} = \underbrace{\begin{pmatrix} c_{13} & 0 & s_{13} e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13} e^{i\delta} & 0 & c_{13} \end{pmatrix}}_{U_2} \begin{pmatrix} |\nu_1\rangle' \\ |\nu_2\rangle' \\ |\nu_3\rangle' \end{pmatrix}. \tag{4.97}$$

It is worth recalling that the non-zero phase δ, represents the CP violation effects. Finally, the last rotation will be around $|\nu_1\rangle''$ at the angle θ_{23}, the rotation results in

$$|\nu_1\rangle''' = |\nu_1\rangle'' \tag{4.98}$$
$$|\nu_2\rangle''' = c_{23}\,|\nu_2\rangle'' + s_{23}\,|\nu_3\rangle'' \tag{4.99}$$
$$|\nu_3\rangle''' = -s_{23}\,|\nu_2\rangle'' + c_{23}\,|\nu_3\rangle'', \tag{4.100}$$

and consequently,

$$\begin{pmatrix} |\nu_1\rangle''' \\ |\nu_2\rangle''' \\ |\nu_3\rangle''' \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix}}_{U_3} \begin{pmatrix} |\nu_1\rangle'' \\ |\nu_2\rangle'' \\ |\nu_3\rangle'' \end{pmatrix}. \tag{4.101}$$

Now, one can write the mixing matrix U as

$$U = U_1U_2U_3 = \begin{pmatrix} c_{13}c_{12} & c_{13}s_{12} & c_{13}\mathrm{e}^{-i\delta} \\ -c_{23}s_{12} - s_{23}c_{12}s_{13}\mathrm{e}^{i\delta} & c_{23}c_{12} - s_{23}s_{12}s_{13}\mathrm{e}^{i\delta} & c_{13}s_{23} \\ s_{23}s_{12} - c_{23}c_{12}s_{13}\mathrm{e}^{i\delta} & -s_{23}c_{12} - c_{23}s_{12}s_{13}\mathrm{e}^{i\delta} & c_{13}c_{23} \end{pmatrix}. \tag{4.102}$$

The above relation is the *standard parametrization* of the 3×3 mixing matrix. It is clear that

$$\begin{pmatrix} |\nu_1\rangle''' \\ |\nu_2\rangle''' \\ |\nu_3\rangle''' \end{pmatrix} = U \begin{pmatrix} |\nu_1\rangle \\ |\nu_2\rangle \\ |\nu_3\rangle \end{pmatrix}. \tag{4.103}$$

Here, our duty becomes finding the values of these three angles $(\theta_{12}, \theta_{23}, \theta_{13})$ and one phase δ. On theoretical grounds, it is not possible and we should seek them by experiments.

In the case of the Majorana neutrino, from (4.75) and (4.102) we see easily that

$$U^M = \begin{pmatrix} c_{13}c_{12}\mathrm{e}^{i\overline{\alpha}_1} & c_{13}s_{12}\mathrm{e}^{i\overline{\alpha}_2} & c_{13}\mathrm{e}^{-i\delta} \\ \left(-c_{23}s_{12} - s_{23}c_{12}s_{13}\mathrm{e}^{i\delta}\right)\mathrm{e}^{i\overline{\alpha}_1} & \left(c_{23}c_{12} - s_{23}s_{12}s_{13}\mathrm{e}^{i\delta}\right)\mathrm{e}^{i\overline{\alpha}_2} & c_{13}s_{23} \\ \left(s_{23}s_{12} - c_{23}c_{12}s_{13}\mathrm{e}^{i\delta}\right)\mathrm{e}^{i\overline{\alpha}_1} & \left(-s_{23}c_{12} - c_{23}s_{12}s_{13}\mathrm{e}^{i\delta}\right)\mathrm{e}^{i\overline{\alpha}_2} & c_{13}c_{23} \end{pmatrix}, \tag{4.104}$$

where $\overline{\alpha}_1$ and $\overline{\alpha}_2$ are the two extra phases in the Majorana case. Under CP invariance, i.e., condition (4.86), it can easily be seen that the three phases become

$$\rho_1 = \mathrm{e}^{-2i\overline{\alpha}_1}; \quad \rho_2 = \mathrm{e}^{-2i\overline{\alpha}_2}, \quad \text{and} \quad \rho_3 = \mathrm{e}^{-2i\delta}, \tag{4.105}$$

where $\rho_i = \pm 1$.

Chapter 5

Quantum mechanical oscillations

Until now, it has been stated that neutrinos are created at $t = 0$ (e.g., in the Sun) in an especial flavor state $|\nu_\alpha\rangle$, but as they travel (e.g., on earth), some of these neutrinos are received in other flavors, say, β, i.e.,

$$|\nu_\alpha\rangle \to |\nu_\beta\rangle \quad \text{where} \quad \alpha \neq \beta. \tag{5.1}$$

In this section we are going to discuss, theoretically, neutrino oscillations in quantum mechanics. This means that the transition probability of flavors will be derived with the help of the quantum mechanics.

5.1 Plane wave treatment

Let us start with the simplest model in QM, i.e., plane wave massive neutrinos. In this model [30] we assume that the massive neutrinos are created with definite mass m_i and momentum $\mathbf{p}_i$. In fact, the mass eigenstates $|\nu_i\rangle$ are in the mass Hilbert space $\mathcal{H}_{\text{mass}}$ and momentum states $|\mathbf{p}_i\rangle$ are in the momentum Hilbert space $\mathcal{H}_{\text{mom}}$. Therefore, the full Hilbert space is

$$\mathcal{H} = \mathcal{H}_{\text{mom}} \times \mathcal{H}_{\text{mass}}. \tag{5.2}$$

So a state describing a massive neutrino, $|\psi_i\rangle$, is supposed to be

$$|\psi_i\rangle = |\nu_i, \mathbf{p}_i\rangle \equiv |\nu_i\rangle\, |\mathbf{p}_i\rangle\,, \tag{5.3}$$

and

$$H\,|\psi_i\rangle = E_i\,|\psi_i\rangle\,, \tag{5.4}$$

where H is the free neutrino Hamiltonian and E_i is the energy of the i^{th} massive neutrino. What is the time evolution of the state $|\psi_i\rangle$? To answer this question,

we use time evolution of states in the Schrödinger picture. Therefore

$$|\psi_i(t)\rangle = \mathrm{e}^{-iE_i t}\,|\nu_i(0)\rangle\,|p_i(0)\rangle \Rightarrow \tag{5.5}$$

$$\langle \mathbf{x}\,|\psi_i(t)\,\rangle = |\psi_i(x)\rangle = \exp\left(-iE_i t + i\mathbf{p}_i\cdot\mathbf{x}\right)|\nu_i\rangle\,, \tag{5.6}$$

which shows that our eigenstates are plane waves[1]. Regarding what has been said, the state vector describing ν_α, i.e., a produced (in a weak interaction) neutrino at $(t_p = 0, \mathbf{x}_p = 0)$ in a given flavor α can be easily found. This state vector, denoted by $|\psi^p_\alpha\rangle$, must be a superposition of $|\psi_i\rangle$, so

$$|\psi^p_\alpha\rangle = \sum_{i=1}^{3} U_{\alpha i}\,|\psi_i\rangle\,. \tag{5.7}$$

From (5.6) and (5.7), it is seen that at creation $(t_p = 0, \mathbf{x}_p = 0)$, the state $|\psi^p_\alpha\rangle$ becomes $|\nu_\alpha\rangle$.

Oscillations happen due to the interferences of $|\psi_i\rangle$. Do these interferences happen at the same time? When the interferences occur, the three eigenstates $|\psi_i\rangle$ reach the same point. If each of these three eigenstates has a different phase velocity, it is concluded that the interferences must not occur at the same time. Owing to this fact, we write

$$\langle \mathbf{x}\,|\psi^p_\alpha(t)\,\rangle = |\psi^p_\alpha(x)\rangle = \sum_{i=1}^{3} U_{\alpha i}\exp\left(-iE_i t_i + i\mathbf{p}_i\cdot\mathbf{x}\right)|\nu_i\rangle\,. \tag{5.8}$$

From (4.17) we have

$$|\nu_i\rangle = \sum_{\alpha'} U^*_{i\alpha'}\,|\nu_{\alpha'}\rangle\,, \tag{5.9}$$

so (5.8) reads

$$|\psi^p_\alpha(x)\rangle = \sum_{\alpha'}\sum_{i=1}^{3} U_{\alpha i}\exp\left(-iE_i t_i + i\mathbf{p}_i\cdot\mathbf{x}\right)U^*_{i\alpha'}\,|\nu_{\alpha'}\rangle\,. \tag{5.10}$$

From the above equation, it is understood that $|\psi^p_\alpha(x)\rangle$ can be written as a superposition of the flavor states. With respect to (5.10), the transition amplitude $\mathcal{A}(\nu_\alpha \to \nu_\beta)$ can be found as follows

$$\langle \nu_\beta\,|\psi^p_\alpha(x)\,\rangle = \mathcal{A}(\nu_\alpha \to \nu_\beta) = U_{\alpha i}\mathrm{e}^{-i\phi_i}U^*_{i\beta}, \tag{5.11}$$

where

$$\phi_i = E_i t_i - \mathbf{p}_i.\mathbf{x} = \sqrt{\mathbf{p}_i^2 + m_i^2}\,t_i - \mathbf{p}_i.\mathbf{x}. \tag{5.12}$$

[1] According to the Heisenberg uncertainty principle, due to the zero momentum uncertainty the spatial uncertainty becomes infinite. As a result, plane waves are delocalized in space. This is one of the disadvantages of the plane wave treatment.

From amplitude (5.11), one can easily find the transition probability $\mathcal{P}(\nu_\alpha \to \nu_\beta)$ as follows

$$\mathcal{P}(\nu_\alpha \to \nu_\beta) = |\mathcal{A}(\nu_\alpha \to \nu_\beta)|^2 = \left| \sum_{i=1}^{3} U_{\alpha i} \mathrm{e}^{-i\phi_i} U_{i\beta}^* \right|^2 \tag{5.13}$$

$$= \sum_{i,j=1}^{3} \mathrm{e}^{-i(\phi_i - \phi_j)} U_{\beta j} U_{j\alpha}^* U_{\alpha i} U_{i\beta}^*. \tag{5.14}$$

As a result of (5.14), it is obtained that

$$\sum_\beta \mathcal{P}(\nu_\alpha \to \nu_\beta) = \sum_{i,j=1}^{3} U_{j\alpha}^* \mathrm{e}^{i\phi_j} \underbrace{\sum_\beta U_{i\beta}^* U_{\beta j}}_{\delta_{ji}} \mathrm{e}^{-i\phi_i} U_{\alpha i}$$

$$= \sum_{i=1}^{3} U_{i\alpha}^* U_{\alpha i} = 1, \tag{5.15}$$

where we have used unitarity of the matrix U. This result is natural because the created neutrino at $t = 0$, either stays at the same or oscillates to other flavors. Similarly for $\mathcal{A}(\overline{\nu}_\alpha \to \overline{\nu}_\beta)$, it is derived that[2]

$$\mathcal{A}(\overline{\nu}_\alpha \to \overline{\nu}_\beta) = \sum_{i=1}^{3} U_{i\alpha}^* \mathrm{e}^{-i\phi_i} U_{\beta i} \Rightarrow$$

$$\mathcal{P}(\overline{\nu}_\alpha \to \overline{\nu}_\beta) = \sum_{i,j=1}^{3} \mathrm{e}^{-i(\phi_i - \phi_j)} U_{j\beta}^* U_{\alpha j} U_{i\alpha}^* U_{\beta i} \Rightarrow \tag{5.16}$$

$$\sum_\beta \mathcal{P}(\overline{\nu}_\alpha \to \overline{\nu}_\beta) = 1. \tag{5.17}$$

By comparing (5.14) with (5.16), it is understood that if $U_{\alpha i}^* = U_{i\alpha}$ (i.e., CP invariance condition in the Dirac neutrino case (see (4.85))), then,

$$\mathcal{P}(\nu_\alpha \to \nu_\beta) = \mathcal{P}(\overline{\nu}_\alpha \to \overline{\nu}_\beta). \tag{5.18}$$

Note. What we have found here is the Dirac transition amplitude, if U is the Dirac mixing matrix. So, let us denote the amplitude by $\mathcal{A}^D(\nu_\alpha \to \nu_\beta)$ where superscript D refers to Dirac. As it was said in the previous section, Majorana matrix elements are

$$U_{\alpha i}^M = U_{\alpha i} \mathrm{e}^{i\bar{\alpha}_i}.$$

[2]We should follow exactly the same process as $\mathcal{A}(\nu_\alpha \to \nu_\beta)$, with the difference that

$$\overline{\nu}_{\alpha L} = \sum_{i=1}^{3} \overline{\nu}_{iL} U_{i\alpha}^*,$$

has been used.

Then, Majorana transition amplitude $\mathcal{A}^M(\nu_\alpha \to \nu_\beta)$ becomes

$$\mathcal{A}^M(\nu_\alpha \to \nu_\beta) = \sum_{i=1}^{3} U^M_{\alpha i} \mathrm{e}^{-i\phi_i} U^{M*}_{\beta i} = \sum_{i=1}^{3} U_{\alpha i} \mathrm{e}^{-i\phi_i} U^*_{\beta i} = \mathcal{A}^D(\nu_\alpha \to \nu_\beta). \tag{5.19}$$

Therefore since the Majorana and Dirac amplitudes (and consequently, their probabilities) are equal, from now on, we just compute the probabilities for the Dirac neutrinos.

Now let us come back to the transition probability (5.14). Our goal is to determine the phase difference $\phi_i - \phi_j$. To do so, we are faced with some problems as follows.

In neutrino oscillations experiments, propagation time (between when our neutrino is produced, e.g., in the Sun and when we receive it on earth) is not measured. However, the distance of the neutrino propagation is obvious (e.g., the distance between the Sun and earth). Therefore, it is needed to convert time into distance to make our calculation much easier.

Additionally, massive neutrinos are not observed. In other words, the only particles which can be detected are flavor neutrinos which lie in the superposition of the massive neutrino eigenstates. Hence, we need to relate the massive neutrinos' momenta to measured momenta of the flavor.

5.2 Oscillation phase

5.2.1 Standard oscillation phase

Our derived phase (5.12) is

$$\phi_i = E_i t_i - \mathbf{p}_i.\mathbf{x} = \sqrt{\mathbf{p}_i^2 + m_i^2}\, t_i - \mathbf{p}_i.\mathbf{x}. \tag{5.20}$$

To relate momenta of the mass eigenstates to the flavor's momentum, let us Taylor expand the phase around averages $\mathbf{p}$, m. Here, $\mathbf{p}$, m are the momentum and mass of the flavor. Since we are in the plane wave treatment, the average momentum must not be very different from $\mathbf{p}_i$. The Taylor expansion of the energy brings us

$$E_i = \sqrt{\mathbf{p}_i^2 + m_i^2} = E + \mathbf{v}.\delta\mathbf{p}_i + \frac{m}{E}\delta m_i, \tag{5.21}$$

where $E = \sqrt{\mathbf{p}^2 + m^2}$, $\mathbf{v} = \mathbf{p}/E$, $\delta\mathbf{p}_i = \mathbf{p} - \mathbf{p}_i$ and $\delta m_i = m_i - m$. The above expansion has been done to first orders in the momentum and mass. As for the momentum, it is acceptable since $\delta\mathbf{p}_i$ must be very small in the plane wave treatment. But, what about the mass? In the ultra-relativistic limit we have $|\mathbf{v}| \approx 1$, then from (5.21), it is seen that

$$\delta E_i - \delta p_i \approx \frac{m}{E}\delta m_i, \tag{5.22}$$

where $\delta E_i = E_i - E$. In this limit (ultra-relativistic), we can say that $\delta E_i \approx \delta p_i$ and $m/E \ll 1$. Accordingly, one can conclude that (although not necessarily) δm_i is not very small. Hence, we are not (necessarily) allowed to ignore higher orders of δm_i in (5.21). If instead of m_i, we expand ϕ_i around $\mathbf{p}_i$ and ***mass squared***, m_i^2, the above problem will be solved. This mean that as long as $\delta m_i^2 = m_i^2 - m^2$ is small with respect to E, the expansion reads

$$\phi_i \approx Et - \mathbf{p}.\mathbf{x} + \frac{\delta m_i^2}{2E} t + (\mathbf{v}t - \mathbf{x}) .\delta \mathbf{p}_i + E\delta t_i, \tag{5.23}$$

where we have also expanded the phase around the average time t, with $\delta t = t_i - t$. As a result,

$$\phi_i - \phi_j \approx \frac{\delta m_{ij}^2}{2E} + (\mathbf{v}t - \mathbf{x}) \left(\delta \mathbf{p}_i - \delta \mathbf{p}_j\right) + E \left(\delta t_i - \delta t_j\right), \tag{5.24}$$

with

$$\delta m_{ij}^2 = m_i^2 - m_j^2. \tag{5.25}$$

Very important note. Since the eigenstates are approximated by plane waves, so there are infinite uncertainties on position and time or in other words, the plane waves are delocalized in space. Hence, the classical relation $\mathbf{x} = \mathbf{v}t$, cannot be used for a delocalized particle (the velocity cannot be defined because to find it, we should first localize a particle[3]).

In spite of this fact, in the standard prescription it is supposed that[4] $\mathbf{x} = \mathbf{v}t$. This condition is called the *classical propagation*. Therefore in comparison with $\delta m_i^2 t/2E$, $(\mathbf{x} - \mathbf{v}t)\, \delta p_i$ becomes negligible and since δp_i is supposed to be of order δm_i^2, then the classical propagation condition imposes that

$$|\mathbf{x} - \mathbf{v}t| \ll t. \tag{5.26}$$

Due to the classical propagation condition, the second term in (5.24) vanishes. Here, it is worth saying that in some papers such as [31] or [35], we apply the *equal momentum prescription* in the standard approach. This means that $\delta \mathbf{p}_i = 0$. This assumption also leads to disappearing of the second term of (5.24). However, as it will be seen in an experiment in section (5.4), setting momentum uncertainty equal to zero results in no oscillations[5].

The second condition in the standard approach is *equal time prescription*. This condition says that interferences between the mass eigenstates only happen in equal time[6], in other words, $\delta t_i = 0$. Equality in time brings with itself

[3] We shall see that due to the fact, we should move from the plane wave treatment to the wave packet approach in which the massive neutrinos become localized in space.

[4] This is one of the drawbacks of the plane wave approach.

[5] Zero momentum uncertainty or applying classical propagation brings us the same result although both of them cannot be justified by the plane wave approach. The reason which pushes us into using the wave packet treatment rather than the plane wave.

[6] If this prescription is not satisfied, we are in the non-standard oscillation. We are not going to discuss it in this thesis, see Ref. [13].

equality in the **phase velocities** of the massive neutrinos. It should be noted that due to the delocalization of the massive neutrinos, *equal velocity prescription* cannot be concluded. Furthermore, equality in velocities is not likely [37], since if $\mathbf{v}_i = \mathbf{v}_j$, then $\gamma_i = \gamma_j$, where $\gamma = \left(1 - |\mathbf{v}|^2\right)^{-1/2}$. Hence,

$$E = \gamma m \Rightarrow \frac{E_i}{E_j} = \frac{m_i}{m_j}. \tag{5.27}$$

The ratio E_i/E_j (for $i \neq j$) is approximately one (see (5.21)), so

$$\frac{E_i}{E_j} \approx 1 \Rightarrow m_i \approx m_j \Rightarrow \delta m_{ij}^2 \approx 0, \tag{5.28}$$

where in this case oscillations would be destroyed. Therefore, equal velocity prescription is not possible. Regarding the standard assumptions, (5.24) reads

$$\phi_i - \phi_j \approx \frac{\delta m_{ij}^2}{2E} t \approx \frac{\delta m_{ij}^2}{2\,|\mathbf{p}|} L, \tag{5.29}$$

where $L \equiv |\mathbf{L}|$, and in the last approximation, again, we have used the classical propagation condition. The above result is called the ***standard oscillation phase***. By defining the oscillation length L_{ij}^{osc} as follows

$$L_{ij}^{\text{osc}} = \frac{4\pi\,|\mathbf{p}|}{\delta m_{ij}^2}, \tag{5.30}$$

relation (5.29) becomes

$$\phi_i - \phi_j \approx 2\pi \frac{L}{L_{ij}^{\text{osc}}}. \tag{5.31}$$

The phase has been expanded around the average momentum $\mathbf{p}$. Since for each mass eigenstate there is a momentum uncertainty, then the point around which we expand the phase must have a momentum spread δp. The distance at which a neutrino of momentum p goes under one oscillations more than a neutrino of momentum $p + \delta p$ is called the *coherence length*[7] L_{ij}^{coh} [38]. So,

$$\frac{L_{ij}^{\text{coh}} \delta m_{ij}^2}{2\,(p + \delta p)} \approx \frac{L_{ij}^{\text{coh}} \delta m_{ij}^2}{2\delta p} - 1 \Rightarrow L_{ij}^{\text{coh}} \approx L_{ij}^{\text{osc}} \frac{p}{\delta p}. \tag{5.32}$$

After this length, oscillations vanish because there are not any interferences between mass eigenstates anymore.

[7] In Ref. [38], this length has been derived in terms of energy, while here we have obtained it in terms of momentum.

5.2.2 Another prescriptions

The phase (5.20) can also be written as (suppose that $\mathbf{x}$ and $\mathbf{p}$ are in the same direction)

$$\phi_i = E_i t_i - x\sqrt{E_i^2 - m_i^2}. \tag{5.33}$$

By expanding the above phase around averages E and m^2, we obtain for the phase difference, again, (5.29) if *equal energy prescription*, i.e., $\delta E_i = 0$, is taken into consideration. Regarding equal energy prescription, we do not even need to use the classical propagation condition (although the equal time is still needed).

Equal energy was suggested by Lipkin [43]. In the next chapter (Eq. (6.34)) we shall see that, in principle, this prescription is acceptable. However, experimentally, energy uncertainty is far from being zero [13].

Conservation of energy-momentum at production, is another prescription which was discussed by Winter [44]. In this case, neither energies nor momenta of the mass eigenstates are equal.

We use the energy-momentum conservation to find the exact values of the energy and momentum of the neutrino, e.g., in the pion (π) decay to muon (μ) and neutrino ($\pi \to \mu\nu$). However, since by this exact computation, we have found which mass eigenstate has taken part in oscillations, there will be no quantum interferences anymore and so oscillations vanish (in section (5.4), this fact is going to be discussed more).

5.2.3 Oscillation probability

Now let us come back to the transition probability (5.14) and apply the standard oscillation phase (5.31), then

$$\begin{aligned}\mathcal{P}\left(\nu_\alpha \to \nu_\beta\right) &= \sum_{i,j=1}^{3} \mathrm{e}^{-i(\phi_i - \phi_j)} U_{\beta j} U_{j\alpha}^* U_{\alpha i} U_{i\beta}^* \\ &= \sum_{i,j=1}^{3} \exp\left(-2i\pi \frac{L}{L_{ij}^{\mathrm{osc}}}\right) U_{\beta j} U_{j\alpha}^* U_{\alpha i} U_{i\beta}^*. \end{aligned} \tag{5.34}$$

If oscillations happen due to the interferences of relativistic and non-relativistic massive neutrinos, it is expected to have incoherence at some distance (see (5.32)). In this case, there is the incoherent probability $\mathcal{P}^{\mathrm{inc}}\left(\nu_\alpha \to \nu_\beta\right)$ as follows

$$\mathrm{e}^{-i(\phi_i - \phi_j)} \to \delta_{ij} \Rightarrow \mathcal{P}\left(\nu_\alpha \to \nu_\beta\right) = \sum_{i=1}^{3} U_{\beta j} U_{j\alpha}^* U_{\alpha i} U_{i\beta}^*. \tag{5.35}$$

But in general, relation (5.34) can be explicitly written in this form:

$$\mathcal{P}\left(\nu_\alpha \to \nu_\beta\right) = \sum_{i(i=j)} U_{\beta i} U_{i\alpha}^* U_{\alpha i} U_{i\beta}^* + \sum_{i \neq j} \exp\left(-2i\pi \frac{L}{L_{ij}^{\mathrm{osc}}}\right) U_{\beta j} U_{j\alpha}^* U_{\alpha i} U_{i\beta}^*. \tag{5.36}$$

The second term of this equation is

$$\sum_{i\neq j}\exp\left(-2i\pi\frac{L}{L_{ij}^{\rm osc}}\right)U_{\beta j}U_{j\alpha}^{*}U_{\alpha i}U_{i\beta}^{*}=\exp\left(-2i\pi\frac{L}{L_{12}^{\rm osc}}\right)\underbrace{U_{\beta 2}U_{2\alpha}^{*}U_{\alpha 1}U_{1\beta}^{*}}_{a}$$
$$+\exp\left(-2i\pi\frac{L}{L_{21}^{\rm osc}}\right)\underbrace{U_{\beta 1}U_{1\alpha}^{*}U_{\alpha 2}U_{2\beta}^{*}}_{a^{*}}+\exp\left(-2i\pi\frac{L}{L_{13}^{\rm osc}}\right)\underbrace{U_{\beta 3}U_{3\alpha}^{*}U_{\alpha 1}U_{1\beta}^{*}}_{b}$$
$$+\exp\left(-2i\pi\frac{L}{L_{31}^{\rm osc}}\right)\underbrace{U_{\beta 1}U_{1\alpha}^{*}U_{\alpha 3}U_{3\beta}^{*}}_{b^{*}}+\exp\left(-2i\pi\frac{L}{L_{23}^{\rm osc}}\right)\underbrace{U_{\beta 3}U_{3\alpha}^{*}U_{\alpha 2}U_{2\beta}^{*}}_{c}$$
$$+\exp\left(-2i\pi\frac{L}{L_{32}^{\rm osc}}\right)\underbrace{U_{\beta 2}U_{2\alpha}^{*}U_{\alpha 3}U_{3\beta}^{*}}_{c^{*}},\qquad(5.37)$$

where a, b and c are complex numbers. Therefor, (5.37) becomes

$$\sum_{i\neq j}\exp\left(-2i\pi\frac{L}{L_{ij}^{\rm osc}}\right)U_{\beta j}U_{j\alpha}^{*}U_{\alpha i}U_{i\beta}^{*}=$$
$$2\left[{\rm Re}\left(\sum_{i>j}\exp\left(-2i\pi\frac{L}{L_{ij}^{\rm osc}}\right)U_{\beta j}U_{j\alpha}^{*}U_{\alpha i}U_{i\beta}^{*}\right)\right],\qquad(5.38)$$

where (Re) means the real part of a complex number. With a similar calculation, we derive for the first term of (5.36)

$$\sum_{i}U_{\beta j}U_{j\alpha}^{*}U_{\alpha i}U_{i\beta}^{*}=\delta_{\beta\alpha}-2\left[{\rm Re}\left(\sum_{i>j}U_{\beta j}U_{j\alpha}^{*}U_{\alpha i}U_{i\beta}^{*}\right)\right],\qquad(5.39)$$

and so (5.36) becomes

$$\mathcal{P}\left(\nu_{\alpha}\to\nu_{\beta}\right)=\delta_{\alpha\beta}-2\left\{{\rm Re}\left[\sum_{i>j}U_{\beta j}U_{j\alpha}^{*}U_{\alpha i}U_{i\beta}^{*}\left(1-\exp\left(-2i\pi\frac{L}{L_{ij}^{\rm osc}}\right)\right)\right]\right\}$$
$$=\delta_{\beta\alpha}+R_{\beta\alpha}+\frac{1}{2}A_{\beta\alpha}^{\rm CP},\qquad(5.40)$$

where

$$R_{\beta\alpha}=-2\sum_{i>j}{\rm Re}\left(U_{\beta j}U_{j\alpha}^{*}U_{\alpha i}U_{i\beta}^{*}\right)\left(1-\cos\left(2\pi\frac{L}{L_{ij}^{\rm osc}}\right)\right),\qquad(5.41)$$

and

$$A_{\beta\alpha}^{\rm CP}=4\sum_{i>j}{\rm Im}\left(U_{\beta j}U_{j\alpha}^{*}U_{\alpha i}U_{i\beta}^{*}\right)\left(\sin\left(2\pi\frac{L}{L_{ij}^{\rm osc}}\right)\right).\qquad(5.42)$$

By doing the similar calculations, but this time using (5.16) we obtain

$$\mathcal{P}\left(\overline{\nu}_\alpha \to \overline{\nu}_\beta\right) = \delta_{\beta\alpha} + R_{\beta\alpha} - \frac{1}{2}A^{\mathrm{CP}}_{\beta\alpha}. \tag{5.43}$$

Relations (5.40) and (5.43) result in

$$R_{\beta\alpha} = \frac{1}{2}\left(\mathcal{P}\left(\nu_\alpha \to \nu_\beta\right) + \mathcal{P}\left(\overline{\nu}_\alpha \to \overline{\nu}_\beta\right)\right) - \delta_{\beta\alpha}, \tag{5.44}$$

and

$$A^{\mathrm{CP}}_{\beta\alpha} = \mathcal{P}\left(\nu_\alpha \to \nu_\beta\right) - \mathcal{P}\left(\overline{\nu}_\alpha \to \overline{\nu}_\beta\right). \tag{5.45}$$

Under CP transformations, $\mathcal{P}\left(\nu_\alpha \to \nu_\beta\right)$ converts into $\mathcal{P}\left(\overline{\nu}_\alpha \to \overline{\nu}_\beta\right)$. Consequently, it is clear that $R_{\beta\alpha}$ is CP-even, i.e., $R_{\beta\alpha}$ does not change under CP transformations while $A^{\mathrm{CP}}_{\beta\alpha}$ gains a negative sign under CP transformations, i.e., it is CP-odd. If there is CP invariance, both probabilities are equal as shown in (5.18) and therefore $A^{\mathrm{CP}}_{\beta\alpha}$ vanishes. Owing to this fact, $A^{\mathrm{CP}}_{\beta\alpha}$ is called the CP asymmetry.

All of the interactions conserve CPT (see introduction). Under time reversal, we have

$$t_i \to -t_i \quad \text{and} \quad \mathbf{p}_i \to -\mathbf{p}_i. \tag{5.46}$$

Hence

$$\hat{T}\mathcal{A}\left(\nu_\alpha \to \nu_\beta\right)\hat{T}^{-1} = \sum_{i=1}^{3} U_{\alpha i}\mathrm{e}^{i\phi_i}U^*_{i\beta} = \mathcal{A}\left(\nu_\beta \to \nu_\alpha\right) \Rightarrow \tag{5.47}$$

$$\hat{T}\mathcal{P}\left(\nu_\alpha \to \nu_\beta\right)\hat{T}^{-1} = \mathcal{P}\left(\nu_\beta \to \nu_\alpha\right), \tag{5.48}$$

where $\hat{T}$ is the time reversal operator. Then, as a result of the CPT invariance, for relation (5.45) we have

$$A^{\mathrm{CP}}_{\beta\alpha} = -A^{\mathrm{CP}}_{\alpha\beta} \Rightarrow A^{\mathrm{CP}}_{\alpha\alpha} = 0. \tag{5.49}$$

Furthermore, by subtracting (5.17) from (5.15), it is obtained that

$$\sum_\beta \mathcal{P}\left(\nu_\alpha \to \nu_\beta\right) - \mathcal{P}\left(\overline{\nu}_\alpha \to \overline{\nu}_\beta\right) = \sum_\beta A^{\mathrm{CP}}_{\beta\alpha} = 0 \Rightarrow$$
$$A^{\mathrm{CP}}_{e\mu} = A^{\mathrm{CP}}_{\tau e} = -A^{\mathrm{CP}}_{\tau\mu}. \tag{5.50}$$

From the above relation, it is understood that by finding one CP asymmetry, the other ones will be derived, i.e., there is only one independent CP asymmetry.

5.3 Oscillation probability in two- and three-flavor cases

5.3.1 Two-flavor oscillations

First, let us consider neutrino oscillations in two-flavor case ($n = 2$), e.g., $\nu_e \leftrightarrow \nu_\mu$. The mixing matrix for $n = 2$ was found in (4.87), i.e.,

$$U^D = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{5.51}$$

For $\alpha \neq \beta$ we have

$$U_{\beta 1} U^*_{1\alpha} U_{\alpha 2} U^*_{2\beta} = -\cos^2\theta \sin^2\theta, \tag{5.52}$$

while

$$\mathrm{Im}\left(U_{\beta j} U^*_{j\alpha} U_{\alpha i} U^*_{i\beta}\right) = 0. \tag{5.53}$$

Using (5.40), (5.41), (5.42), (5.52) and (5.53) we obtain

$$\mathcal{P}\left(\nu_\alpha \to \nu_\beta\right) = \frac{1}{2}\sin^2\left(2\theta\right)\left(1 - \cos\left(2\pi\frac{L}{L^{\mathrm{osc}}_{ij}}\right)\right), \qquad \alpha \neq \beta. \tag{5.54}$$

Moreover, very easily we see that

$$\mathcal{P}\left(\nu_\alpha \to \nu_\alpha\right) = 1 - \underbrace{\mathcal{P}\left(\nu_\alpha \to \nu_\beta\right)}_{\alpha\neq\beta} = 1 - \frac{1}{2}\sin^2\left(2\theta\right)\left(1 - \cos\left(2\pi\frac{L}{L^{\mathrm{osc}}_{ij}}\right)\right). \tag{5.55}$$

5.3.2 Three-flavor oscillations

In the case of the three-flavor oscillations, if the CP violation phase δ is equal to zero, i.e., $A^{\mathrm{CP}}_{\beta\alpha} = 0$, from (5.40) we have[8]

$$\begin{aligned}\mathcal{P}\left(\nu_\alpha \to \nu_\beta\right) &= \delta_{\beta\alpha} - 4U_{\beta 1}U_{1\alpha}U_{\alpha 2}U_{2\beta}\sin^2\left(2\pi\frac{L}{L^{\mathrm{osc}}_{21}}\right) \\ &- 4U_{\beta 1}U_{1\alpha}U_{\alpha 3}U_{3\beta}\sin^2\left(2\pi\frac{L}{L^{\mathrm{osc}}_{31}}\right) - 4U_{\beta 2}U_{2\alpha}U_{\alpha 3}U_{3\beta}\sin^2\left(2\pi\frac{L}{L^{\mathrm{osc}}_{32}}\right).\end{aligned} \tag{5.56}$$

For more discussion on the three-flavor oscillations with relativistic and non-relativistic mass eigenstates[9] see Ref. [36]. Furthermore, for the case in which $\delta \neq 0$, see Ref. [1].

[8] This is valid if there are interferences between all three mass eigenstates.

[9] In this case, the interferences between some of the mass eigenstates are incoherent and the story becomes different.

5.4 Problems of the plane wave approach

According to (5.31) we have

$$\phi_i - \phi_j \approx 2\pi \frac{L}{L_{ij}^{\text{osc}}}. \tag{5.57}$$

From the above relation we see that if interferences between two eigenfunctions i and j happen in $L = mL_{ij}^{\text{osc}}$ (where $m \in \mathbb{Z}$), there are constructive interferences, while in $L = \frac{m}{2} L_{ij}^{\text{osc}}$ we have destructive interferences. Therefore, the oscillation length is the distance between two constructive interferences.

Now, let us compare neutrino oscillations with Young's double-slit experiment in which the distance between two constructive interferences (bright fringes) is one wavelength[10] λ. The light diffraction happens in the double-slit experiment ***if and only if*** $\lambda \gtrsim \sigma_x$, where σ_x is the spatial uncertainty (or the length of the slits).

Same as Young's experiment in which we have interferences of coherent beams of light, in neutrino oscillations phenomenon (where there are interferences of the coherent mass eigenstates), it is expected that oscillations happen ***if and only if***

$$L_{ij}^{\text{osc}} \gtrsim \sigma_x. \tag{5.58}$$

Let us see if the above condition is violated in the plane wave treatment or not. In an experiment [31], we aim at finding neutrino oscillations $\nu_\mu \leftrightarrow \nu_e$. The muon neutrino is produced through pion (in flight and not at rest) decay, i.e., $\pi^+ \to \mu^+ + \nu_\mu$. First, when the distance between source to detector L is varied, the $L-$dependence of the probability (see (5.54)) is observed.

Second, we add apparatus to find the produced neutrino mass very precisely. By precisely, it is meant

$$\Delta m_\nu^2 < \left| m_1^2 - m_2^2 \right|, \tag{5.59}$$

where Δm_ν^2 is the produced neutrino mass uncertainty. In other words, we know that in an event such as $\nu_\mu \leftrightarrow \nu_e$, which mass eigenstates are contributing.

The produced neutrino mass with the above uncertainty is obtained because our apparatus measures the $3-$momenta and energies of the pion and muon with uncorrelated uncertainties Δp and ΔE, respectively. Then, from the energy momentum conservation, the neutrino energy (with uncertainty ΔE_ν), momentum (with uncertainty Δp_ν) and consequently, the neutrino mass with uncertainty Δm_ν^2 is derived.

As the measurements are done, it is observed that event rates at the detector do not oscillate with change of L anymore. The reason can be explained as follows:

1. when with our accurate measurements, we determine which massive neutrino state is involved, there is no longer a flavor which is a superposition

[10] The waves that come from both openings in the double-slit experiment are coherent, so their wavelengths are equal.

of the mass states, or in other words there are no interferences between mass states anymore. Flavor neutrino is set in one of the mass state and due to the fact oscillations wash out.

2. The accurate measurements of the pion and muon momenta will put infinite uncertainty on their positions. When there is an infinite uncertainty on the position of pion, consequently there will be an infinite uncertainty on the neutrino production region.

In the language of mathematics, it can be said that

$$\begin{aligned} m_\nu^2 &= E_\nu^2 - p_\nu^2 \Rightarrow \\ \Delta m_\nu^2 &= \left[\left(2E_\nu\right)^2 \left(\Delta E_\nu\right)^2 + \left(2p_\nu\right)^2 \left(\Delta p_\nu\right)^2 \right]^{1/2} < \left| m_1^2 - m_2^2 \right| \Rightarrow \\ \Delta p_\nu &< \underbrace{\frac{\left| m_1^2 - m_2^2 \right|}{2p_\nu}}_{\approx \left(L_{12}^{\text{osc}}\right)^{-1}} \Rightarrow \Delta x > L_{12}^{\text{osc}}, \end{aligned} \tag{5.60}$$

which, of course, contradicts (5.58). So, mass eigenstates with definite momentum results in no oscillations. This is the reason why we go towards the wave packet, i.e, momentum uncertainty for the mass eigenstates increases to satisfy $\Delta p_\nu \gtrsim \left(L_{ij}^{\text{osc}}\right)^{-1}$. We should remember that the classical propagation condition discussed in section (5.2.1) is another problem of the plane wave approach. These are not the only problems of the plane wave treatment. We shall see more, later on.

5.5 Intermediate wave packet treatment

5.5.1 Transition probability and decoherence

Contrary to the plane wave approach, in the intermediate wave packet treatment, it is assumed that the massive neutrinos are localized everywhere in space. This approach has been discussed in Ref. [31] generally. Here, for simplicity, we follow Ref. [35] and suppose that the mass eigenstates in momentum space, i.e., $\psi_a\left(p\right)$, are Gaussian[11]. So

$$\psi_a\left(p\right) = \left(\sqrt{2\pi}\sigma_p\right)^{1/2} \exp\left[-\frac{\left(p - \langle p_a \rangle\right)^2}{4\sigma_p^2} \right], \tag{5.61}$$

where σ_p is the momentum spread[12] and the Gaussian wave packets are sharply peaked around $\langle p_a \rangle$. Regarding (5.61), the mass eigenstates in configuration

[11] It is supposed that the momentum spread is along the propagation direction x. The orthogonal momentum spreads are ignored and so we do not use the vector symbol.

[12] Here, it is supposed that the widths at production and detection are equal, i.e., σ_p. In [33], the probability is derived for different widths.

space, $\tilde{\psi}_a(t,x)$, become $\left(E=\sqrt{p^2+m^2}\right)$

$$\begin{aligned}\tilde{\psi}_a(t,x) &= \int \frac{\mathrm{d}p}{2\pi}\psi_a(p)\,\mathrm{e}^{-iEt+ipx} \\ &= \left(\sqrt{2\pi}\sigma_x\right)^{-1/2}\exp\left[i\left(\langle p_a\rangle x-\langle E_a\rangle t\right)-\frac{(x-v_a t)^2}{4\sigma_x^2}\right], \end{aligned} \tag{5.62}$$

where

$$\langle E_a\rangle=\sqrt{\langle p_a\rangle^2+m_a^2};\quad v_a=\frac{\langle p_a\rangle}{\langle E_a\rangle};\quad \sigma_x\sigma_p=\frac{1}{2}. \tag{5.63}$$

Now, the transition amplitude for a neutrino created at $t=0,x=0$ in flavor ν_α, and detected on earth, $t=T$ and $x=X$, as ν_β becomes

$$\begin{aligned}&\mathcal{A}_{\alpha\to\beta}(X,T)=\sum_{a=1}^{3}U_{\beta a}\psi_a(X,T)\,U_{\alpha a}^*\Rightarrow \\ &\mathcal{P}_{\alpha\to\beta}(X,T)=\left|\mathcal{A}_{\alpha\to\beta}(X,T)\right|^2=\frac{1}{\sqrt{2\pi\sigma_x}}\sum_{a,b=1}^{3}U_{\beta a}U_{\alpha a}^*U_{\beta b}^*U_{\alpha b} \\ &\times\exp\left[i\left(\langle p_a\rangle-\langle p_b\rangle\right)X-i\left(\langle E_a\rangle-\langle E_b\rangle\right)T\right] \\ &\times\exp\left[-\frac{(X-v_aT)^2}{4\sigma_x^2}-\frac{(X-v_bT)^2}{4\sigma_x^2}\right]. \end{aligned} \tag{5.64}$$

In practical experiments, we do not measure time T. Therefore, we take integral over T from (5.64) to derive the time average of the probability at X, which is known, e.g., the distance between the Sun and earth[13]. By completing the squares inside the exponentials of (5.64) and applying

$$\int_{-\infty}^{+\infty}\mathrm{e}^{-ax^2}\mathrm{d}x=\sqrt{\frac{\pi}{a}}, \tag{5.65}$$

it is obtained that

$$\begin{aligned}\mathcal{P}_{\alpha\to\beta}(X,T) &= \left[\sum_{a'=1}^{3}\frac{\left|U_{\alpha a'}\right|^2}{|v_{a'}|}\right]^{-1}\sum_{a,b=1}^{3}\left[\frac{2}{v_a^2+v_b^2}\right]^{1/2}U_{\beta a}U_{\alpha a}^*U_{\beta b}^*U_{\alpha b} \\ &\times\exp\left\{i\left[\left(\langle p_a\rangle-\langle p_b\rangle\right)-\left(\langle E_a\rangle-\langle E_b\rangle\right)\left[\frac{v_a+v_b}{v_a^2+v_b^2}\right]\right]X\right\} \\ &\times\exp\left[-\frac{X^2}{4\sigma_x^2}\frac{(v_a-v_b)^2}{v_a^2+v_b^2}-\frac{\left(\langle E_a\rangle-\langle E_b\rangle\right)^2}{4\sigma_p^2\left(v_a^2+v_b^2\right)}\right]. \end{aligned} \tag{5.66}$$

[13]Since we are integrating over time, it has been assumed that each massive neutrino feels the same time.

In the above relation, all of the terms have been derived from solving the time integral except the term

$$\left[\sum_{a'=1}^{3}\frac{|U_{\alpha a'}|^2}{|v_{a'}|}\right]^{-1},$$

which has been put by hand as a normalization factor. The phase can be seen in the second line of (5.66) in the imaginary exponential. The oscillation length is defined as follows

$$L_{ij}^{\rm osc}=\frac{2\pi}{|\langle E_a\rangle-\langle E_b\rangle|}\left[\frac{v_a+v_b}{v_a^2+v_b^2}-\frac{\langle p_a\rangle-\langle p_b\rangle}{\langle E_a\rangle-\langle E_b\rangle}\right]^{-1}. \tag{5.67}$$

Therefore, the imaginary exponential becomes

$$\exp\left(2\pi i\frac{X}{L_{ij}^{\rm osc}}\right). \tag{5.68}$$

Let us write the term

$$\exp\left[-\frac{X^2}{4\sigma_x^2}\frac{(v_a-v_b)^2}{v_a^2+v_b^2}\right]$$

in (5.66) as

$$\exp\left[-\left(\frac{X}{L_{ab}^{\rm coh}}\right)^2\right], \tag{5.69}$$

where

$$L_{ab}^{\rm coh}=2\sigma_x\left[\frac{v_a^2+v_b^2}{(v_a-v_b)^2}\right]^{1/2}. \tag{5.70}$$

The factor (5.69) is a damping factor, i.e., coherent interferences vanish if $X\gtrsim L_{ab}^{\rm coh}$. This is the reason why $L_{ab}^{\rm coh}$ is called the coherence length. It is clear that to have oscillations, the coherence length must be much bigger than the spatial width σ_x. This condition is satisfied if (see (5.70)) $|v_a-v_b|\ll 1$. In other words, two massive neutrinos must be nearly degenerate[14]. The term

$$\exp\left[\frac{(\langle E_a\rangle-\langle E_b\rangle)^2}{4\sigma_p^2\,(v_a^2+v_b^2)}\right],$$

imposes that to have oscillations, we must have

$$\langle E_a\rangle-\langle E_b\rangle\lesssim\sigma_p\sqrt{v_a^2+v_b^2}. \tag{5.71}$$

[14]This result, i.e., nearly degenerate massive neutrinos, will be seen as we go to neutrino oscillations in quantum field theory too.

Figure 5.1: Decoherence due to the mass eigenstates with different group velocities. The figure has been taken from Ref. [13].

Let us briefly comment on the origin of the decoherence. We know that oscillations happen due to the interferences of the mass eigenstates. Suppose that there are two mass eigenstates with different group velocities $\mathbf{v}_i$ and $\mathbf{v}_j$. If there is no overlap between these two mass eigenstates, i.e., when these two have crossed each other completely, then there will not be any interferences and oscillations wash out (Fig. (5.1)).

5.5.2 Wave packet treatment is still problematic

In paper [33], it has been shown that the coherence length must be smaller than a limit (see equation (14) in [33]). However, an accurate measurement of energy at detector increases the coherence length [33]. This fact contradicts equation (14) in [33]. This is one of the problems of the wave packet treatment which leads to the approach of neutrino oscillations in the framework of quantum field theory (next chapter).

Another problem of this treatment is that we cannot define a Fock space of flavor states which are superpositions of the mass eigenstates (unless the massive neutrinos are extremely relativistic or nearly degenerate) [34]. However, this problem can be solved by defining *weak-process states*. In fact neutrinos are not observed directly, but particles which interact with neutrinos are detected. In the weak-process states strategy, we write the probability of mixing as a transition amplitude between observable particles. In the limit of relativistic or nearly degenerate massive neutrinos, the newly defined states (i.e., weak-process) reduce to the weak states [34].

In general, we can say that neglecting observable particles or processes of neutrino creation and detection is a problem of the wave packet treatment[15]. In addition, there are more questions which cannot be answered in the wave packet approach, e.g., unstable neutrino oscillations etc. For a list of these problems, see [13]. It is also worth noting that there are some quantum mechanical models which solve some of these problems such as weak-process states or models stated in [13].

[15]Of course, whatever is said as problems of the wave packet treatment, should be added to the plane wave approach problems.

5.5.3 Equal energy prescription

As a final word in this section, let us say that the wave packet approach tells us that only interferences between wave packets with the same energy bring us non-zero probability. This fact can be seen if we write the wave packet (5.61) as a function of energy in configuration space, i.e.,

$$\begin{aligned}\tilde{\psi}_a(t,x) &= \int \mathrm{d}E_a \psi_a(E_a) \exp(-iE_a t + ip_a x) \Rightarrow \int \mathrm{d}t \tilde{\psi}_a^*(t,x)\, \tilde{\psi}_b(t,x) \\ &= \int \mathrm{d}t \mathrm{d}E_a \mathrm{d}E_b \psi_a^*(E_a)\, \psi_b(E_b)\, \mathrm{e}^{-i(E_b-E_a)t}\, \mathrm{e}^{i(p_b-p_a)x} \qquad (5.72) \\ &= 2\pi \int \mathrm{d}E_a \psi_a^*(E_a)\, \psi_b(E_b)\, \mathrm{e}^{i(p_b-p_a)x}. \qquad (5.73)\end{aligned}$$

In the second line, we have used the Dirac delta. This delta can be used if the time interval ΔT in (5.72) is infinite, or in other words, energy uncertainty is zero. This is the case called the stationary boundary condition which is going to be explained later on. If ΔT is finite, the result of the time integral in (5.72) becomes

$$\frac{1}{i(E_a - E_b)} \exp(\Delta T (E_a - E_b)).$$

By expanding energies around an average momentum and mass squared (same as before), it is seen that as

$$\Delta T \gg T_{ij}^{\mathrm{osc}} \quad \text{where} \quad T_{ij}^{\mathrm{osc}} = \frac{L_{ij}^{\mathrm{osc}}}{v}, \qquad (5.74)$$

the infinite ΔT becomes reasonable and (5.73) is accepted.

Chapter 6

Oscillations in quantum field theory

6.1 Introduction

As it was discussed in the previous chapter, neutrino oscillations in quantum mechanics is confronted with some problems. For example, it is not possible (except in the extremely relativistic case) to define flavor operators in quantum mechanics approach. Here, in this chapter we are going to see how this problem (and also the others) are going to be addressed in quantum field theory approach. In this chapter we have followed Ref. [13].

In general, all of the models which describe neutrino oscillations phenomenon can be grouped into four categories as follows:

1. the external wave-packet models,
2. the stationary boundary-conditions models,
3. the source-propagator models,
4. the Blasone-Vitiello models.

Let us first start from the first category. In the external wave-packet models, it is supposed that neutrinos are not created in isolation, this means that they are not free particles. We know that neutrinos are created in a weak interaction such as beta decay, where this interaction is mediated by W boson as a propagator. The produced neutrinos are free particles (i.e., they are on-shell or in other words they obey the energy-momentum conservation rule, $p^2 = m^2$) whereas the W boson which acts as a propagator is not free (it is a virtual particle, i.e., off-shell or in other words, it does not obey the energy-momentum conservation rule).

In the external wave packet models, it is assumed that the oscillating particle (i.e., neutrino) must be a propagator and not a free particle. By this we mean

that neutrinos are created (as propagators) as a result of interactions of some fields, then propagate to the earth where our detectors are set up. There, they interact with the fields of particles in the detector. Hence, production, propagation and detection are, all, considered as one process and intermediate state which is the oscillating neutrino cannot be observed directly (see Fig. (6.1)). So, the oscillation process is considered globally.

Actually, neutrino is created (observed) due to the interactions of some particles at the source (detector). These particles at the source and detector are described by wave packets and are called external (in comparison to the oscillating particle which is internal). This is the reason why this model is called the external wave packet model. This model was first proposed by Giunti et al. [32] for neutrinos.

In the simplest case, the propagating neutrino is stable. This means that the oscillating particle does not decay during its travel from the source to the detector. However, it is also possible to analyze an unstable neutrino as the propagator[1].

The second category is the stationary boundary conditions. These models are, actually, a special case of the external wave packet. In section (6.2.2) we will discuss the simplest consistent model of this category proposed by Kobzarev et al. [45]. In the source-propagator models, neutrinos are described by a propagator coupled to the source but not to the detector. This approach results in the nonstandard oscillation length or recoil oscillations. The last option is the Blasone-Vitiello models in which completely different way (in comparison with the first three ones) has been used for analyzing neutrino oscillations. In this thesis we are going to discuss the first approach (in the stable case) in detail.

6.2 The external wave packet

In quantum field theory, particles are created and annihilated by field operators. For instance, if a field $\hat{\phi}(x)$ acts on the vacuum state $|0\rangle$, the result will be one-particle state $|\mathbf{k}\rangle$ which is a state of one particle with momentum $\mathbf{k}$. In general, a wave packet representing a given state $|\psi\rangle$ with mass m can be written as the superposition of the single-particle states of momentum $\mathbf{k}$, as follows

$$|\psi\rangle = \int [\mathrm{d}\mathbf{k}]\, |\mathbf{k}\rangle \langle \mathbf{k}\, |\psi\rangle = \int [\mathrm{d}\mathbf{k}]\, \psi(\mathbf{k})\, |\mathbf{k}\rangle\,, \tag{6.1}$$

where state $|\mathbf{k}\rangle$ is taken at time $t = 0$, $\psi(\mathbf{k})$ is the wave function in momentum space, again at time $t = 0$ and

$$[\mathrm{d}\mathbf{k}] = \frac{\mathrm{d}\mathbf{k}}{(2\pi)^3} \frac{1}{\sqrt{2E(\mathbf{k})}}; \qquad E(\mathbf{k}) = \sqrt{\mathbf{k}^2 + m^2}. \tag{6.2}$$

[1] We are not going to discuss this case in this thesis.

The reason of $\sqrt{2E(\mathbf{k})}$ in the above integral is that the normalization of the free states $|\mathbf{k}\rangle$ and $|\mathbf{p}\rangle$ is set as

$$\langle \mathbf{k} | \mathbf{p} \rangle = 2E(\mathbf{k}) (2\pi)^3 \delta^3 (\mathbf{p} - \mathbf{k}), \tag{6.3}$$

and hence, since

$$\langle \psi | \psi \rangle = 1 \Rightarrow \int \frac{\mathrm{d}\mathbf{k}}{(2\pi)^3} |\psi(\mathbf{k})|^2 = 1. \tag{6.4}$$

Now, let us write the wave function $\psi(\mathbf{k})$ in configuration space as

$$\tilde{\psi}(\mathbf{x}, t) = \int \frac{\mathrm{d}\mathbf{k}}{(2\pi)^3} \psi(\mathbf{k}) \, \mathrm{e}^{-iE(\mathbf{k})t + i\mathbf{k}\cdot\mathbf{x}}. \tag{6.5}$$

If it is assumed that our wave packet is sharply peaked at $\mathbf{k} = \mathbf{K}$, then we can expand $\psi(\mathbf{k})$ around $\mathbf{K}$ and ignore the small terms. Hence, (6.5) becomes

$$\tilde{\psi}(\mathbf{x}, t) = \int \frac{\mathrm{d}\mathbf{k}}{(2\pi)^3} \left[\psi(\mathbf{K}) + \frac{\partial \psi}{\partial \mathbf{k}} (\mathbf{k} - \mathbf{K}) + \left(\mathcal{O} (\mathbf{k} - \mathbf{K})^2 \right) \right] \mathrm{e}^{-iE(\mathbf{k})t + i\mathbf{k}\cdot\mathbf{x}}. \tag{6.6}$$

From the above equation it is seen that $\tilde{\psi}(\mathbf{x}, 0)$ has a maximum at the point $\mathbf{x} = 0$ if

$$\psi(\mathbf{K} + \mathbf{k}) = \psi(\mathbf{K} - \mathbf{k}) \tag{6.7}$$

is satisfied. This is because in this case, the position derivative of $\tilde{\psi}(\mathbf{x}, 0)$ at $\mathbf{x} = 0$ becomes an integral of an odd function and so vanishes. Therefore regarding condition (6.7), the wave function in momentum space is denoted by $\psi(\mathbf{k}, \mathbf{K})$. If the wave function in configuration space has a maximum in $\mathbf{x}_0 \neq 0$ and $t_0 \neq 0$, then the wave function in momentum space is denoted by $\psi(\mathbf{k}, \mathbf{K}, \mathbf{x}_0, t_0)$. In this case, the wave function in configuration space becomes

$$\tilde{\psi}(\mathbf{x}, t, \mathbf{K}, \mathbf{x}_0, t_0) = \int \frac{\mathrm{d}\mathbf{k}}{(2\pi)^3} \psi(\mathbf{k}, \mathbf{K}) \, \mathrm{e}^{-iE(\mathbf{k})(t - t_0) + i\mathbf{k}\cdot(\mathbf{x} - \mathbf{x}_0)}. \tag{6.8}$$

Now according to (6.1), let us assume a wave packet state $|P_I\rangle$ (which is sharply peaked at momentum $\mathbf{Q}$ and centered at $(t_p, \mathbf{x}_p)$) by,

$$|P_I\rangle = \int [\mathrm{d}\mathbf{q}] \psi_{P_I}(\mathbf{q}, \mathbf{Q}, \mathbf{x}_p, t_p) \, |P_I(\mathbf{q})\rangle, \tag{6.9}$$

which represents an incoming particle. This particle interacts with an outgoing particle represented by another wave packet state $|P_F\rangle$ (which is sharply peaked at momentum $\mathbf{K}$ and again centered at $(t_p, \mathbf{x}_p)$), i.e.,

$$|P_F\rangle = \int [\mathrm{d}\mathbf{k}] \psi_{P_F}(\mathbf{k}, \mathbf{K}, \mathbf{x}_p, t_p) \, |P_F(\mathbf{k})\rangle. \tag{6.10}$$

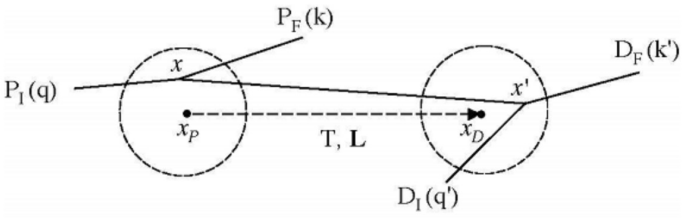


Figure 6.1: Feynman diagram for neutrino creation, propagation and annihilation. The creation and annihilation points are not only x_P and x_D, but, we should sum over all points where this process can occur. This is due to the superposition principle of quantum mechanics which states that if a process can happen in different ways (i.e., all the points in the dashed circles, centered in x_P and x_D), all the amplitudes should be added together. The figure has been taken from Ref. [13].

The result of this interaction is creation of a propagator that, here, is the intermediate neutrino, i.e., $|\nu_\alpha\rangle$. Same as the previous chapter, we use the Greek index for flavor neutrinos. The produced neutrino propagates from its production region and reaches our detector on earth in state $|\nu_\beta\rangle$. Due to the interaction between this neutrino and a particle at detector, for instance represented by the wave packet $|D_I\rangle$ (which is sharply peaked at momentum $\mathbf{Q}'$ and centered at detection point $(t_D, \mathbf{x}_D)$),

$$|D_I\rangle = \int [\mathrm{d}\mathbf{q}'] \psi_{D_I}\left(\mathbf{q}', \mathbf{Q}', \mathbf{x}_D, t_D\right) |D_I\left(\mathbf{q}'\right)\rangle\,, \tag{6.11}$$

we obtain another particle represented by wave packet $|D_F\rangle$ (which is sharply peaked at momentum $\mathbf{K}'$ and centered at $(t_D, \mathbf{x}_D)$)

$$|D_F\rangle = \int [\mathrm{d}\mathbf{k}'] \psi_{D_F}\left(\mathbf{k}', \mathbf{K}', \mathbf{x}_D, t_D\right) \left|D_F\left(\mathbf{k}'\right)\right\rangle. \tag{6.12}$$

All of what has been discussed, can be easily seen in a Feynman diagram (see Fig. (6.1)). The amplitude, $\mathcal{A}$, for this process is [5]

$$\mathcal{A} = \langle P_F D_F | \hat{T} \left\{ \exp\left(-ig \int \mathrm{d}^4 x \mathcal{L}_{int}\left(x\right)\right)\right\} - \mathbb{1}\, |P_I D_I\rangle\,, \tag{6.13}$$

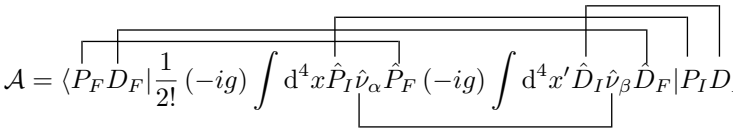
where $\mathcal{L}_{int}$ is the interaction Lagrangian, $\hat{T}$ is the time ordering operator and g is the coupling constant. This Lagrangian annihilates $|P_I\rangle$ whereas creates $|P_F\rangle$ and the neutrino $|\nu_\alpha\rangle$ at spacetime point x. Hence, we show the interaction Lagrangian operator as the multiplication of the three fields

$$\mathcal{L}_{int}\left(x\right) = -\hat{P}_I \hat{\nu}_\alpha \hat{P}_F. \tag{6.14}$$

Expanding the amplitude (6.13) to order g^2 means that we must have another interaction but this time at spacetime point x'. In other words, the created neutrino at point x propagates to the point x', then it interacts with $|D_I\rangle$, both are annihilated and state $|D_F\rangle$ is created[2]. So, the interaction Lagrangian at x' is an operator as the multiplication of the three fields

$$\mathcal{L}_{int}\left(x'\right) = -\hat{D}_I \hat{\nu}_\beta \hat{D}_F. \tag{6.15}$$

Now, let us show the amplitude and interactions in this form:



$$\mathcal{A} = \langle P_F D_F | \frac{1}{2!}\left(-ig\right) \int \mathrm{d}^4 x \hat{P}_I \hat{\nu}_\alpha \hat{P}_F \left(-ig\right) \int \mathrm{d}^4 x' \hat{D}_I \hat{\nu}_\beta \hat{D}_F | P_I D_I \rangle.$$

We call the neutrino field contraction in the above amplitude, propagator. It is denoted by $G\left(x - x'\right)$ as follows

$$G\left(x - x'\right) = \overbrace{\hat{\nu}_\alpha\left(x\right) \ \ \hat{\nu}_\beta\left(x'\right)}. \tag{6.16}$$

It is worth saying a word about the propagator. As we said, the intermediate flavor state $|\nu_\alpha\left(x\right)\rangle$ is created, at x, by a field operator such as $\hat{\nu}_\alpha\left(x\right)$ which acts on the vacuum state $|0\rangle$, i.e.,

$$\hat{\nu}_\alpha\left(x\right)|0\rangle = |\nu_\alpha\left(x\right)\rangle. \tag{6.17}$$

Same as the state $|\nu_\alpha\left(x\right)\rangle$, we can have state $|\nu_\beta\left(x'\right)\rangle$ at the spacetime point x'. The transition amplitude for a particle, which was born in the state $|\nu_\alpha\left(x\right)\rangle$, to be in the state $|\nu_\beta\left(x'\right)\rangle$ is called the propagator $G\left(x' - x\right)$. Simply, it can be said that $G\left(x' - x\right)$ is the amplitude for a particle to propagate from x' to x, i.e.,

$$G_{\alpha\beta}\left(x' - x\right) = \langle \nu_\alpha\left(x\right)|\nu_\beta\left(x'\right)\rangle = \langle 0|\,\hat{\nu}^*_\alpha\left(x\right)\hat{\nu}_\beta\left(x'\right)|0\rangle \ \ \text{if} \ \ x^0 > x'^0, \tag{6.18}$$

and the amplitude for a particle to propagate from x to x' is

$$G_{\beta\alpha}\left(x' - x\right) = \langle 0|\,\hat{\nu}_\beta\left(x'\right)\hat{\nu}^*_\alpha\left(x\right)|0\rangle \ \ \text{if} \ \ x^0 < x'^0. \tag{6.19}$$

Hence, we can write the propagator as

$$G_{\alpha\beta}\left(x' - x\right) = \langle 0|\,T\left[\hat{\nu}_\beta\left(x'\right)\hat{\nu}^*_\alpha\left(x\right)\right]|0\rangle, \tag{6.20}$$

where $T\left[\hat{\nu}_\beta\left(x'\right)\hat{\nu}^*_\alpha\left(x\right)\right]$ is the so-called time-ordering operator defined as

$$\begin{aligned} &T\left[\hat{\nu}_\beta\left(x'\right)\hat{\nu}^*_\alpha\left(x\right)\right] = \\ &\theta\left(x^0 - x'^0\right)\hat{\nu}^*_\alpha\left(x\right)\hat{\nu}_\beta\left(x'\right) + \theta\left(x'^0 - x^0\right)\hat{\nu}_\beta\left(x'\right)\hat{\nu}^*_\alpha\left(x\right). \end{aligned} \tag{6.21}$$

[2]This is the reason why we use $\hat{T}$. In fact, time ordering operator puts t and t' in order.

Propagator (6.16) is called the *mixed propagator*. The mixed propagator is non-diagonal in the basis of flavor eigenstates, because the flavor transformation is violated, this is the reason of neutrino mixing.

Here, the problem becomes similar to what was done in the quantum mechanical approach. It means that by a unitary transformation on the fields $\hat{\nu}_\alpha$, we obtain new fields, say, $\hat{\nu}_i$. Furthermore, another Lagrangian called $\mathcal{L}_{prop}$ which is diagonal in the mass basis ν_i and violates flavor transformation is defined. Propagator (6.16) is built from $\mathcal{L}_{prop}$ and is diagonal in the mass basis. The process of diagonalization is same as what we did in section (4.2.1). However, the following points should be considered

1. In the quantum mechanical oscillations, we claimed that first: there are flavor states which are created with the interaction Lagrangian, and second: there are mass eigenstates which are created with the free Dirac Lagrangian.
   It was seen that with regard to these assumptions, the flavor eigenstates would be ill-defined. In the QFT approach this problem is overcome since we assume that both the flavor and mass eigenstates are not free or in other words they cannot be defined in isolation. So, mixing in QFT applies to fields and not physical states.
   To be more exact, in QFT the flavor eigenstate is created with the interaction Lagrangian $\mathcal{L}_{int}$ and the massive neutrinos with $\mathcal{L}_{prop}$ which includes the mass matrix generated by the Yukawa interactions.

2. Contrary to the quantum mechanical approach in which we supposed that right- and left-handed neutrinos must take part into the free Lagrangian (see (4.13)), here in QFT, we just have neutrino fields. So, we use only one unitary matrix in comparison to (4.13) where two different unitary matrices U and V were used for left- and right-handed neutrinos, respectively.

Regarding these points[3], we can easily diagonalize propagator (6.16) with the help of the unitary matrix transformation U, i.e.,

$$\hat{\nu}_\alpha = \sum_{i=1}^{3} U^*_{\alpha i} \hat{\nu}_i \Rightarrow \tag{6.22}$$

$$G_{\beta\alpha}\left(x' - x\right) = \sum_i U^*_{\beta i} G_{D,ii}\left(x' - x\right) U_{i\alpha}, \tag{6.23}$$

where $G_{D,jj}\left(x' - x\right)$ are elements of the diagonalized propagator. The above relation can also be written in the matrix form as

$$G\left(x' - x\right) = U^\dagger G_D\left(x' - x\right) U. \tag{6.24}$$

[3]We are not going to repeat the process of diagonalization. It is same as what has been done in section (4.2.1). Moreover, see Ref. [1] for more detail.

With the help of the Fourier transformations, one can write the diagonal propagator $G_D\left(x'-x\right)$ as

$$G\left(x'-x\right)=\int\frac{\mathrm{d}^4p}{\left(2\pi\right)^4}\mathrm{e}^{-ip\cdot\left(x'-x\right)}\tilde{G}\left(p^2\right), \tag{6.25}$$

where $\tilde{G}\left(p^2\right)$ is [5]

$$\tilde{G}\left(p^2\right)=\frac{i}{p^2-m_i^2+i\epsilon}. \tag{6.26}$$

Now, let us stop speaking about the propagator and come back to the interaction amplitude. We should apply the field operators on their corresponding wave packet states. For instance for $|P_I\rangle$ we obtain

$$\hat{P}_I\left|P_I\right\rangle=\int\left[\mathrm{d}\mathbf{q}\right]\psi_{P_I}M_P\left(q\right)\mathrm{e}^{-iq\cdot x}, \tag{6.27}$$

where $M_P\left(q\right)$ is a term which includes the polarization (for vector particles) or spinors (for spin 1/2 particles). The subscript P refers to the production. We can do the same calculations for the other states, so the amplitude becomes

$$\begin{aligned}\mathcal{A}=\int\left[\mathrm{d}\mathbf{q}\right]\psi_{P_I}\int\left[\mathrm{d}\mathbf{k}\right]\psi_{P_F}^*\int\left[\mathrm{d}\mathbf{q}'\right]\psi_{D_I}\int\left[\mathrm{d}\mathbf{k}'\right]\psi_{D_F}^*\times\\ \int\mathrm{d}^4xM_P\left(q,k\right)\mathrm{e}^{-i(q-k)\cdot x}\int\mathrm{d}^4x'M_D\left(q',k'\right)\mathrm{e}^{-i\left(q'-k'\right)\cdot x'}G\left(x-x'\right),\end{aligned} \tag{6.28}$$

where $M_P\left(q,k\right)$ and $M_D\left(q',k'\right)$ are the interaction amplitudes at production and detection. Performing the changes of variables,

$$x\rightarrow x+x_P;\qquad x'\rightarrow x'+x_D, \tag{6.29}$$

we are going to plug the diagonal propagator (6.25) into (6.28). Then, the interaction amplitude (for the massive neutrinos as propagator) $\mathcal{A}_j$ becomes

$$\mathcal{A}_j=\int\frac{\mathrm{d}^4p}{\left(2\pi\right)^4}\psi\left(p^0,\mathbf{p}\right)\frac{i}{p^2-m_j^2+i\epsilon}\mathrm{e}^{-ip\cdot\left(x_D-x_P\right)}, \tag{6.30}$$

where

$$\begin{aligned}\psi\left(p^0,\mathbf{p}\right)=\int\mathrm{d}^4x\mathrm{e}^{ip\cdot x}\int\mathrm{d}^4x'\mathrm{e}^{-ip\cdot x'}\int\left[\mathrm{d}\mathbf{q}\right]\psi_{P_I}\left(\mathbf{q},\mathbf{Q}\right)\mathrm{e}^{-iq\cdot x}\times\\ \int\left[\mathrm{d}\mathbf{q}\right]\psi_{P_F}^*\left(\mathbf{k},\mathbf{K}\right)\mathrm{e}^{ik\cdot x}\int\left[\mathrm{d}\mathbf{q}'\right]\psi_{D_I}\left(\mathbf{q}',\mathbf{Q}'\right)\mathrm{e}^{-iq'\cdot x'}\times\\ \int\left[\mathrm{d}\mathbf{k}\right]\psi_{D_F}^*\left(\mathbf{k}',\mathbf{K}'\right)\mathrm{e}^{ik'\cdot x'}M_P\left(q,k\right)M_D\left(q',k'\right).\end{aligned} \tag{6.31}$$

By using relation (6.23), the transition amplitude for a neutrino created as ν_α and received, on earth, as ν_β, i.e., $\mathcal{A}(\alpha \to \beta, T, \mathbf{L})$ becomes

$$\mathcal{A}(\alpha \to \beta, T, \mathbf{L}) = \sum_j U^*_{\beta j} \left\{ \int \frac{\mathrm{d}^4 p}{(2\pi)^4} \psi\left(p^0, \mathbf{p}\right) \frac{i}{p^2 - m_j^2 + i\epsilon} \mathrm{e}^{-ip\cdot(x_D - x_P)} \right\} U_{j\alpha}. \tag{6.32}$$

To find the transition amplitude, (6.30) should be solved. In principle, we can integrate over p^0 and $\mathbf{p}$ in any order. However, to get rid of the poles, it is better to first integrate over p^0.

The integration over p^0 is done by using the **Jacob-Sachs theorem** [46]. In this theorem, it is assumed that the energy spectra of all incident particles are limited to a finite range. As a result of this assumption, the overlap function $\psi\left(p^0, \mathbf{p}\right)$ becomes non-zero only if $p^2 = E^2 - \mathbf{p}^2$ (i.e., $p^0 = E$) within certain bounds for positive energy. Therefor, we may show the overlap function by $\psi\left(E\left(\mathbf{p}\right), \mathbf{p}\right)$. Then, according to the Jacob-Sachs theorem we have the following asymptotic behavior[4] ($T \to \infty$)

$$\int \mathrm{d}E \psi\left(E\left(\mathbf{p}\right), \mathbf{p}\right) G\left(E, \mathbf{p}^2\right) \mathrm{e}^{-iET} \to \frac{\pi}{E\left(\mathbf{p}\right)} \psi\left(E\left(\mathbf{p}\right), \mathbf{p}\right) \mathrm{e}^{-iE(\mathbf{p})T}. \tag{6.33}$$

Thus, (6.30) becomes

$$\mathcal{A}_j = \frac{\pi}{(2\pi)^4} \int \frac{\mathrm{d}^3 \mathbf{p}}{E_j\left(\mathbf{p}\right)} \psi\left(E_j\left(\mathbf{p}\right), \mathbf{p}\right) \mathrm{e}^{-i\phi_j(\mathbf{p})}, \tag{6.34}$$

where $\phi_j\left(\mathbf{p}\right) = E_j\left(\mathbf{p}\right) T - \mathbf{p}.\mathbf{L}$.

Now, everything reduces to finding the overlap function $\psi\left(E, \mathbf{p}\right)$. In other words, the forms of the wave packets in (6.31) must be determined. In the following sections, we are going to become familiar with different wave packet forms and see if they result in oscillations or not.

6.2.1 Plane wave in- and outgoing states

In the simplest case, let us assume that all the external states (in- and outgoing states) were approximated by plane waves. Then, we would have

$$\begin{aligned} &\psi_{P_I}\left(\mathbf{q}, \mathbf{Q}\right) \sim \delta^{(3)}\left(\mathbf{q} - \mathbf{Q}\right) && \psi_{P_F}\left(\mathbf{k}, \mathbf{K}\right) \sim \delta^{(3)}\left(\mathbf{k} - \mathbf{K}\right) \\ &\psi_{D_I}\left(\mathbf{q}', \mathbf{Q}'\right) \sim \delta^{(3)}\left(\mathbf{q}' - \mathbf{Q}'\right) && \psi_{D_F}\left(\mathbf{k}', \mathbf{K}'\right) \sim \delta^{(3)}\left(\mathbf{k}' - \mathbf{K}'\right). \end{aligned} \tag{6.35}$$

By plugging (6.35) into (6.31), we obtain

$$\psi\left(p^0, \mathbf{p}\right) \sim \delta^{(4)}\left(P - Q + K\right) \delta^{(4)}\left(K' - Q' - P\right) M_P\left(Q, K\right) M_P\left(Q', K'\right). \tag{6.36}$$

[4] For proof of the Jacob-Sachs theorem, see Ref. [46].

These delta functions say that:

$$Q = P + K; \quad \text{and} \quad P + Q' = K', \tag{6.37}$$

where these equalities can also be seen from Fig. (6.1). Regarding (6.37), (6.36) may be written as

$$\begin{aligned}&\psi\left(p^0, \mathbf{p}\right) \sim \\ &\delta^{(4)}\left(P - Q + K\right) \delta^{(4)}\left(K + K' - Q - Q'\right) M_P\left(Q, K\right) M_P\left(Q', K'\right).\end{aligned} \tag{6.38}$$

By plugging (6.38) into (6.34), it is obtained that

$$\mathcal{A}_i = \delta^{(4)}\left(K + K' - Q - Q'\right) \frac{i}{\left(Q - K\right)^2 - m_j^2 + i\epsilon} \mathrm{e}^{-i(Q-K)\cdot(x_D - x_P)}. \tag{6.39}$$

It is clear that the above amplitude will not result in oscillations, because there is no mass in the phase of oscillations. Hence, external states with plane wave approximation will ruin oscillations. It means that we should look for models in which the overlap function is not proportional to the Dirac delta.

6.2.2 The simplest consistent model

In a simple model proposed by Kobzarev et al. [45], a neutrino (as a propagator) is produced at $\mathbf{x}_P$ due to the interaction of a charged lepton with an infinitely heavy nucleus. It is assumed that the charged lepton wave function $\psi_{P_I}\left(\mathbf{q}, \mathbf{Q}\right)$ is plane wave, i.e., same as the previous section. However, since the infinitely heavy nucleus is situated at $\mathbf{x}_P$, then momentum uncertainty becomes infinite and so its wave function is constant

$$\psi_{P_F}\left(\mathbf{k}, \mathbf{K}\right) \sim \text{constant}. \tag{6.40}$$

We have the same situation at the detection point, i.e., the incoming neutrino interacts with another infinitely heavy nucleus, and as a result of this scattering a charged lepton (approximated with plane wave) $\psi_{D_F}\left(\mathbf{k}', \mathbf{K}'\right)$ is emitted. Hence,

$$\psi_{D_I}\left(\mathbf{q}', \mathbf{Q}'\right) \sim \text{constant}. \tag{6.41}$$

By substituting (6.40) and (6.41) (for the heavy nuclei) and the delta functions in (6.35) (for the plane waves $\psi_{P_I}\left(\mathbf{q}, \mathbf{Q}\right)$ and $\psi_{D_F}\left(\mathbf{k}', \mathbf{K}'\right)$) into (6.31), the overlap function reads

$$\begin{aligned}\psi\left(p^0, \mathbf{p}\right) \sim \int \left[\mathrm{d}\mathbf{q}'\right] \int \left[\mathrm{d}\mathbf{k}'\right] \delta^{(4)}\left(p - Q + k\right) \delta^{(4)}\left(-p + K' - q'\right) \\ M\left(Q, k\right) \times M\left(q', K'\right) \sim \delta\left(p^0 - E_{\text{in}}\right) \delta\left(p^0 - E_{\text{out}}\right),\end{aligned} \tag{6.42}$$

where E_{in} and E_{out} are defined as follows

$$E_{\text{in}} = Q^0 - k^0; \quad E_{\text{out}} = k'^0 - Q'^0. \tag{6.43}$$

Therefore, we must have the condition (as expected)

$$E_{\text{in}} = E_{\text{out}}. \tag{6.44}$$

Regarding (6.42) and (6.34), $\mathcal{A}_j$ becomes

$$\begin{aligned}\mathcal{A}_j &\sim \delta\left(E_{\text{in}} - E_{\text{out}}\right) \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{i}{E_{\text{in}}^2 - \mathbf{p}^2 - m_j^2 + i\epsilon} \mathrm{e}^{i\mathbf{p}\cdot\mathbf{L}} \\ &\sim \delta\left(E_{\text{in}} - E_{\text{out}}\right) \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3} \frac{i}{\mathbf{p}_j^2 - \mathbf{p}^2 + i\epsilon} \mathrm{e}^{i\mathbf{p}\cdot\mathbf{L}},\end{aligned} \tag{6.45}$$

where $\mathbf{p}_j = \sqrt{E_{\text{in}}^2 - m_j^2}$. We should note that in the above integral (i.e., in the propagator), $\mathbf{p}_j \neq \mathbf{p}$. This integral can be solved with the help of the contour integration, the result becomes

$$\mathcal{A}_j \sim \frac{1}{|L|} \delta\left(E_{\text{in}} - E_{\text{out}}\right) \mathrm{e}^{i\mathbf{p}_j\cdot\mathbf{L}}. \tag{6.46}$$

As a result of (6.46), the oscillation probability becomes

$$\mathcal{A}_i \mathcal{A}_j^* \sim \mathrm{e}^{i\left(\mathbf{p}_i - \mathbf{p}_j\right)\cdot\mathbf{L}}. \tag{6.47}$$

This model leads to oscillations because if we expand $\mathbf{p}_j$ around an average m^2 to first order, we obtain

$$\mathbf{p}_j = \sqrt{E_{\text{in}}^2 - m_j^2} \cong \mathbf{p}_m - \frac{\delta m_j^2}{2\mathbf{p}_m} \Rightarrow \mathcal{A}_i \mathcal{A}_j^* \sim \frac{\delta m_{ij}^2}{2p_m} L, \tag{6.48}$$

where $\delta m_{ij}^2 = m_i^2 - m_j^2$ and $p_m = \sqrt{E_{\text{in}}^2 - m^2}$. We also derived the above relation in quantum mechanical approach, it was called the standard oscillations. The difference is that in the QFT approach, the "classical propagation condition" has not been used and so, the result becomes consistent. This is the advantage of this simple QFT model over the standard model in QM. However, this model cannot be considered as the final model because of the following issues:

- it cannot describe unstable oscillations, i.e., if a neutrino decays into another particle before reaching the earth. This is because it is not possible to suppose that neutrinos can decay into an infinitely heavy state during their travel (before reaching our detectors on earth).
- In this model, two of the external states were approximated by plane waves. Plane waves are delocalized in space.
- The amplitude is independent of the direction of $\mathbf{L}$.
- Neutrinos are produced and detected with infinitely heavy nuclei whose spatial uncertainties are zero. This is not physical because it is not possible to measure the position of a particle more precisely than its Compton wavelength.

As a result of these drawbacks, we should choose another overlap function and again solve (6.34). Until now, we have been working with the external wave packets whose energy and 3−momentum uncertainties were zero or infinite. Now, let us work with an overlap function which depends explicitly on the energy and 3−momentum uncertainty at the source and at the detector.

Giunti, Kim and Lee proposed for the first time [32] Gaussian external wave packets[5]. We will see how applying the Gaussian wave packets with width σ_x in configuration space leads to a localization term (no oscillations if $L^{\text{osc}} \lesssim \sigma_x$). Moreover, in this approach a coherence length (after which oscillations vanish) is derived. Both of these conditions agree with the intermediate wave packet treatment in quantum mechanics.

6.3 Gaussian overlap function

A Gaussian wave packet in momentum space, centered around $\mathbf{X}_P$ and T_P is:

$$\psi_\chi\left(\mathbf{p};\mathbf{X}_P,T_P,\langle\mathbf{p}\rangle\right) = \left[\sqrt{2\pi}\sigma_{p\chi}\right]^{-3/2}\exp\left[-\frac{\left(\mathbf{p}-\langle\mathbf{p}\rangle\right)^2}{4\sigma_{p\chi}^2} - i\mathbf{p}.\mathbf{X}_P + iE\left(\mathbf{p}\right)T_P\right], \tag{6.49}$$

where the subscript χ says that the system is in state $|\chi\rangle$. The Gaussian wave packet is peaked around the average momentum $\langle\mathbf{p}\rangle$,

$$\langle\mathbf{p}\rangle = \langle\chi|\,\hat{\mathbf{p}}\,|\chi\rangle = \int\frac{\mathrm{d}^3\mathbf{p}}{\left(2\pi\right)^3}\psi_\chi\left(\mathbf{p};\mathbf{X}_P,T_P,\langle\mathbf{p}\rangle\right)\mathbf{p}\,\psi_\chi^*\left(\mathbf{p};\mathbf{X}_P,T_P,\langle\mathbf{p}\rangle\right), \tag{6.50}$$

and $\sigma_{p\chi}$ is width of the wave packet in momentum space, i.e.,

$$\sigma_{p\chi}^2 = \langle\chi|\left(\hat{\mathbf{p}}-\langle\mathbf{p}\rangle\right)^2|\chi\rangle = \int\frac{\mathrm{d}^3\mathbf{p}}{\left(2\pi\right)^3}\psi_\chi\left(\mathbf{p};\mathbf{X}_P,T_P,\langle\mathbf{p}\rangle\right)\left\{\langle\chi|\,\hat{\mathbf{p}}^2\,|\chi\rangle - \left[\langle\chi|\,\hat{\mathbf{p}}\,|\chi\rangle\right]^2\psi_\chi^*\left(\mathbf{p};\mathbf{X}_P,T_P,\langle\mathbf{p}\rangle\right)\right\}. \tag{6.51}$$

Here, for simplicity, we assume that $\sigma_{p\chi}$ has the same amount along the three directions. Now, we define the width $\sigma_{x\chi}$ in configuration space. Therefore, from the Heisenberg uncertainty we have

$$\sigma_{p\chi}\sigma_{x\chi} = \frac{1}{2}. \tag{6.52}$$

As it will be seen, the spatial uncertainty does depend on time and so, the shape of the wave packet differs as time goes on or in other words the wave packet spreads. This fact is called *dispersion*. In the external wave packets, dispersion

[5] For simplicity [32] it is assumed that wave packets are Gaussian. It is clear that any other approximation can be used, however, Gaussian distribution has been chosen since it results in analytical integrations.

is ignored.

Now let us write the wave packet in configuration space, then with respect to (6.8) we have

$$\tilde{\psi}\left(\mathbf{x},t;\mathbf{X}_P,T_P,\langle\mathbf{p}\rangle\right)= \left[\sqrt{2\pi}\sigma_{p\chi}\right]^{-3/2}\int\frac{\mathrm{d}\mathbf{p}}{(2\pi)^3}\exp\left[-\sigma_{x\chi}^2\mathbf{q}^2+i\mathbf{p}\cdot\left(\mathbf{x}-\mathbf{X}_P\right)-iE\left(\mathbf{p}\right)\left(t-T_P\right)\right], \tag{6.53}$$

where $\mathbf{q}=\mathbf{p}-\langle\mathbf{p}\rangle$.

Since the integral is Gaussian, i.e., the momentum is peaked around the average $\langle\mathbf{p}\rangle$, then we expand the energy to the second order around $\langle\mathbf{p}\rangle$ and ignore the higher orders. Hence,

$$E\left(\mathbf{p}\right)=\sqrt{\mathbf{p}^2+m^2}\cong\langle E\rangle+\mathbf{v}.\mathbf{q}+\frac{1}{2\langle E\rangle}\left[\mathbf{q}^2-\left(\mathbf{v}.\mathbf{q}\right)^2\right], \tag{6.54}$$

where

$$\langle E\rangle\equiv E\left(\langle\mathbf{p}\rangle\right)=\sqrt{\langle\mathbf{p}\rangle^2+m^2}\quad\text{ and }\quad\mathbf{v}=\frac{\langle\mathbf{p}\rangle}{\langle E\rangle}. \tag{6.55}$$

By plugging $E\left(\mathbf{p}\right)$ from (6.54) into (6.53), we obtain

$$\tilde{\psi}\left(\mathbf{x},t;\mathbf{X}_P,T_P,\langle\mathbf{p}\rangle\right)=\left[\sqrt{2\pi}\sigma_{p\chi}\right]^{-3/2}\exp\left\{i\langle\mathbf{p}\rangle\cdot\left(\mathbf{x}-\mathbf{X}_P\right)-i\langle E\rangle\left(t-T_P\right)\right\}\times \int\frac{\mathrm{d}\mathbf{p}}{(2\pi)^3}\exp\left\{-\sigma_{x\chi}^2\Sigma^{ij}q_{ij}+iq_i\left[\left(x-X_P\right)^i-iv^i\left(t-T_P\right)\right]\right\}, \tag{6.56}$$

where

$$\Sigma^{ij}=\delta^{ij}+\left(\delta^{ij}-v^iv^j\right)\frac{2i\left(t-T_P\right)\sigma_p^2}{\langle E\rangle}. \tag{6.57}$$

With regard to the matrix Σ, and using

$$\int_{-\infty}^{+\infty}\mathrm{e}^{-ax^2}\mathrm{d}x=\sqrt{\frac{\pi}{a}}, \tag{6.58}$$

the solution of integral (6.56) becomes:

$$\tilde{\psi}\left(\mathbf{x},t;\mathbf{X}_P,T_P,\langle\mathbf{p}\rangle\right)= \frac{\left(2\pi\sigma_x^2\right)^{-3/4}}{\sqrt{\det\Sigma}}\left[\sqrt{2\pi}\sigma_{\mathbf{p}}\right]^{-3/2}\exp\left\{i\langle\mathbf{p}\rangle\cdot\left(\mathbf{x}-\mathbf{X}_P\right)-i\langle E\rangle\left(t-T_P\right)\right\}\times \exp\left\{-\frac{\left(\left(x-X_P\right)-v\left(t-T_P\right)\right)_i\left(\Sigma^{-1}\right)^{ij}\left(\left(x-X_P\right)-v\left(t-T_P\right)\right)_j}{4\sigma_x^2}\right\}. \tag{6.59}$$

With the help of (6.59), the average position $\langle \mathbf{x} \rangle$ can be easily found as follows

$$\begin{aligned}\langle \mathbf{x} \rangle = \langle \chi | \, \hat{\mathbf{x}} \, | \chi \rangle &= \int \mathrm{d}^3\mathbf{x} \tilde{\psi}_\chi \left(\mathbf{x}, t; \mathbf{X}_P, T_P, \langle \mathbf{p} \rangle \right) \mathbf{x} \, \tilde{\psi}^*_\chi \left(\mathbf{x}, t; \mathbf{X}_P, T_P, \langle \mathbf{p} \rangle \right) \\ &= \mathbf{v} \left(t - T_P \right). \end{aligned} \tag{6.60}$$

Additionally, the spatial uncertainty in configuration space[6] can be computed as follows

$$\begin{aligned}\left\langle \left(\Delta x^i \right)^2 \right\rangle_\chi = \langle \chi | \left(\Delta x^i \right)^2 | \chi \rangle = \langle \chi | \left(\hat{x}^i - \langle x^i \rangle \right)^2 | \chi \rangle = \\ \sigma^2_{x\chi} \left[\left(\mathrm{Re} \Sigma^{-1} \right)^{-1} \right]^{ii}. \end{aligned} \tag{6.61}$$

If we assume that the velocity is along the z axis (i.e., x^3), then

$$\left\langle \left(\Delta x^1 \right)^2 \right\rangle_\chi = \left\langle \left(\Delta x^2 \right)^2 \right\rangle_\chi = 1 + \frac{4\sigma^4_{p\chi} \left(t - T_P \right)^2}{\langle E \rangle^2}, \tag{6.62}$$

and

$$\left\langle \left(\Delta x^3 \right)^2 \right\rangle_\chi = 1 + \frac{4m^4_\chi \sigma^4_{p\chi} \left(t - T_P \right)^2}{\langle E \rangle^6}. \tag{6.63}$$

From the above relations we see that the spatial uncertainties are time dependent and so, the wave packet begins to spread. The fact which was spoken about as we were discussing the momentum uncertainty, and it was said that this time-dependence of the spatial uncertainties is known as dispersion. Equations (6.62) show the *transversal spreadings* i.e., spreadings in the directions transverse to z axis (as it was assumed that the velocity is along the z axis) begin at

$$t - T_P \approxeq \frac{\langle E \rangle}{2\sigma^2_{p\chi}}, \tag{6.64}$$

while *longitudinal spreading* (spreading in the direction along $v\hat{z}$) starts at the later time (see (6.63))

$$t - T_P \approxeq \frac{\langle E \rangle^3}{2m^2_\chi \sigma^2_{p\chi}}. \tag{6.65}$$

In spite of (6.64) and (6.65), we neglect dispersion in the external wave packets. Thus (see relations (6.62) and (6.63)), we approximate

$$\frac{2 \left(t - T_P \right) \sigma^2_p}{\langle E \rangle} \approxeq 0 \Rightarrow \Sigma^{ij} \approxeq \delta^{ij}. \tag{6.66}$$

In the next section, we are going to start discussing the non-spreading wave packets.

[6] We have not done the calculations here. It is same as what we did in (6.51), with the difference that in the case of the spatial uncertainty we should use the position state completeness, and $\tilde{\psi}_\chi \left(\mathbf{x}, t; \mathbf{X}_P, T_P, \langle \mathbf{p} \rangle \right)$.

It is worth saying that the reason of $\mathrm{Re}\Sigma^{-1}$ in relation (6.61) is that the imaginary parts of $\tilde{\psi}_\chi$ and $\tilde{\psi}^*_\chi$ cancel each other.

6.4 Non-spreading Gaussian wave packets

6.4.1 The transition amplitude integral

To find the transition amplitude (6.34) for the non-spreading Gaussian wave packets, we should first find the overlap function (6.31). It is assumed that the factors $M_P(q,k)$ and $M_D(q',k')$ in (6.31) vary slowly over the widths of the wave packets and so we approximate them as

$$M_P(q,k) \approxeq M_P(Q,K); \qquad M_P(q',k') \approxeq M_D(Q',K'), \tag{6.67}$$

then, they can be taken out of integral (6.31). Moreover, the wave packet (6.49) and energy expansion (6.54) should be put into the terms such as

$$\int [\mathrm{d}\mathbf{q}]\, \psi_{P_I}(\mathbf{q},\mathbf{Q})\, \mathrm{e}^{-iq\cdot x} = \int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3} \frac{1}{\sqrt{2E_{P_I}(\mathbf{q})}} \psi_{P_I}(\mathbf{q},\mathbf{Q})\, \mathrm{e}^{-iq\cdot x}, \tag{6.68}$$

in (6.31). Then, the result will be what we derived in (6.59). It should be remarked that, here, first: we have $\Sigma^{ij} = \delta^{ij}$ (see (6.66)) and second: there is the extra term $(2E_{P_I}(\mathbf{q}))^{-1/2}$. If the energy expansion (6.54) is plugged into the expression, $(2E_{P_I}(\mathbf{q}))^{-1/2}$, $(2E_{P_I}(\mathbf{Q}))^{-1/2}$ is obtained. Actually, the other terms of (6.54) have been ignored because the energy is in the denominator and the terms containing $(\mathbf{p} - \langle\mathbf{p}\rangle)$ are negligible. Regarding these two points we have

$$\int [\mathrm{d}\mathbf{q}]\, \psi_{P_I}(\mathbf{q},\mathbf{Q})\, \mathrm{e}^{-iq\cdot x} = N_{P_I} \exp\left(-iE_{P_I}(\mathbf{Q})\, t + i\mathbf{Q}\cdot\mathbf{x} - \frac{(\mathbf{x}-\mathbf{v}_{P_I})^2}{4\sigma_{xP_I}^2}\right), \tag{6.69}$$

where $N_{P_I} = \left(2\pi\sigma_{xP_I}^2\right)^{-3/4} (2E_{P_I}(\mathbf{Q}))^{-1/2}$. By doing similar calculations for the term $\int [\mathrm{d}\mathbf{k}]\, \psi^*_{P_F}(\mathbf{k},\mathbf{K})\, \mathrm{e}^{ik\cdot x}$ in (6.31), an overlap function at the source, $\psi_P\left(p^0,\mathbf{p}\right)$, is defined as

$$\begin{aligned}&\psi_P\left(p^0,\mathbf{p}\right) = \\ &N_P \int \mathrm{d}^4x \exp\left(i\left(p^0 - E_P\right)t - i\left(\mathbf{p}-\mathbf{p}_P\right)\cdot\mathbf{x} - \frac{\mathbf{x}^2 - 2\mathbf{v}_P\cdot\mathbf{x}t + \Sigma_P t^2}{4\sigma_{xP}^2}\right).\end{aligned} \tag{6.70}$$

Here, $N_P = N_{P_F}\left(2\pi\sigma_{xP_I}^2\right)^{-3/4}(2E_{P_I}(\mathbf{K}))^{-1/2}$, and

$$E_P = E_{P_I} - E_{P_F}; \;\; \mathbf{p}_P = \mathbf{Q} - \mathbf{K}, \tag{6.71}$$

$$\frac{1}{\sigma_{xP}^2} = \frac{1}{\sigma_{xP_I}^2} + \frac{1}{\sigma_{xP_F}^2}, \tag{6.72}$$

$$\mathbf{v}_P = \sigma_{xP}^2\left(\frac{\mathbf{v}_{P_I}}{\sigma_{xP_I}^2} + \frac{\mathbf{v}_{P_F}}{\sigma_{xP_F}^2}\right); \qquad \Sigma_P = \sigma_{xP}^2\left(\frac{\mathbf{v}_{P_I}^2}{\sigma_{xP_I}^2} + \frac{\mathbf{v}_{P_F}^2}{\sigma_{xP_F}^2}\right). \tag{6.73}$$

Note. We call σ_{xP}, the *spatial uncertainty at production*. If we assume that $\sigma_{xP_I} > \sigma_{xP_F}$, from (6.72), it is seen that $\sigma_{xP} < \sigma_{xP_F}$. Here, we approximate[7]

$$\sigma_{xP} \approx \sigma_{xP_F}. \tag{6.74}$$

Therefore, σ_{xP} is principally determined by the external particle with the smallest width. Additionally, *momentum uncertainty at production*, σ_{pP}, is defined and from the Heisenberg uncertainty principle we have $\sigma_{xP}\sigma_{pP} = 1/2$. By using (6.72), the *production velocity* $\mathbf{v}_P$ reads

$$\mathbf{v}_P = \frac{\mathbf{v}_{P_I}\sigma_{xP_F}^2 + \mathbf{v}_{P_F}\sigma_{xP_I}^2}{\sigma_{xP_I}^2 + \sigma_{xP_F}^2}. \tag{6.75}$$

Regarding (6.74) (i.e., if the spatial width is dominated by the external particle with the smallest width), we see that $\mathbf{v}_P$ is nearly equal to velocity of the external particle with the smallest width, i.e., here $\mathbf{v}_{P_F}$. With similar reasoning, it is obtained for Σ_P that

$$\Sigma_P = \frac{\mathbf{v}_{P_I}^2\sigma_{xP_F}^2 + \mathbf{v}_{P_F}^2\sigma_{xP_I}^2}{\sigma_{xP_I}^2 + \sigma_{xP_F}^2}, \tag{6.76}$$

and using (6.74) we see that $\Sigma_P \approx \mathbf{v}_{P_F}^2$ (i.e., squared of velocity of the external particle with the smallest spatial width). Hence,

$$0 \leq \Sigma_P \leq 1. \tag{6.77}$$

Let us come back to our discussion and define the function $\psi_D^*\left(p^0, \mathbf{p}\right)$. This function is derived by changing index P (production) to D (detection) in the complex conjugated form of (6.70). Therefore, the overlap function (6.31) reads

$$\psi\left(p^0, \mathbf{p}\right) = N\psi_P\left(p^0, \mathbf{p}\right)\psi_D^*\left(p^0, \mathbf{p}\right), \tag{6.78}$$

where $N = N_P N_D M_P\left(Q, K\right) M_D\left(Q', K'\right)$.

The last step becomes integrating over t and $\mathbf{x}$ from (6.70) to derive $\psi_P\left(p^0, \mathbf{p}\right)$ (and of course, $\psi_D^*\left(p^0, \mathbf{p}\right)$ in a similar process). To do so, first, the variable $\mathbf{u}$ is defined as follows,

$$\begin{aligned}\mathbf{u} =& \mathbf{x} - \mathbf{v}_p t \Rightarrow \psi_P\left(p^0, \mathbf{p}\right) = \\ & N_P \int \mathrm{d}t \exp\left(\frac{t^2\left(\mathbf{v}_p^2 - \Sigma_p\right)}{4\sigma_{xp}^2} + i\left(p^0 - E_P - \left(\mathbf{p} - \mathbf{p}_p\right)\cdot\mathbf{v}_p\right)t\right) \times \\ & \int \mathrm{d}^3\mathbf{x} \exp\left(-\frac{\mathbf{u}^2}{4\sigma_{xp}^2} - i\left(\mathbf{p} - \mathbf{p}_p\right)\cdot\mathbf{u}\right).\end{aligned} \tag{6.79}$$

The above integral can be easily calculated by completing squares and using (6.58), so

$$\psi_P\left(p^0, \mathbf{p}\right) = \pi^2\sigma_{pP}^{-3}\sigma_{eP}^{-1}\exp\left(-f_P\left(p^0 - \mathbf{p}\right)\right), \tag{6.80}$$

[7]This approximation works very well especially when $\sigma_{xP_I} \gg \sigma_{xP_F}$.

where

$$\sigma_{eP}^2 = \sigma_{pP}^2 \left(\Sigma_P - \mathbf{v}_P^2\right), \tag{6.81}$$

and

$$f_P\left(p^0 - \mathbf{p}\right) = \frac{\left(\mathbf{p} - \mathbf{p}_p\right)^2}{4\sigma_{pP}^2} + \frac{\left(p^0 - E_P - \left(\mathbf{p} - \mathbf{p}_p\right).\mathbf{v}_p\right)^2}{4\sigma_{eP}^2}. \tag{6.82}$$

Similarly for $\psi_D^*\left(p^0, \mathbf{p}\right)$, we have

$$\begin{aligned}\psi_D^*\left(p^0, \mathbf{p}\right) =& \pi^2 \sigma_{pD}^{-3} \sigma_{eD}^{-1} \exp\left(-f_D\left(p^0, \mathbf{p}\right)\right) \Rightarrow \psi\left(p^0, \mathbf{p}\right) = \\ & N\pi^4 \sigma_{pP}^{-3} \sigma_{eP}^{-1} \sigma_{pD}^{-3} \sigma_{eD}^{-1} \exp\left(-f_j\left(p^0, \mathbf{p}\right)\right),\end{aligned} \tag{6.83}$$

where $f_j\left(\mathbf{p}\right) = f_P\left(p^0, \mathbf{p}\right) + f_D\left(p^0, \mathbf{p}\right)$. As a result of (6.77) and (6.81), we see that

$$\sigma_{eP,D} \leq \sigma_{pP,D}. \tag{6.84}$$

With respect to (6.75), (6.76) and (6.81), σ_{eP}^2 becomes (we use $\sigma_{xP}\sigma_{pP} = 1/2$)

$$\sigma_{eP}^2 = \frac{\sigma_{xP}^2}{4\sigma_{xP_I}^2 \sigma_{xP_F}^2} \left(\mathbf{v}_{P_I} - \mathbf{v}_{P_F}\right)^2. \tag{6.85}$$

If it is supposed that there are more than two external wave packets in the production process, we should sum over all of them. Then (6.85) becomes

$$\sigma_{eP}^2 = \sum_{\alpha<\beta} \frac{\sigma_{xP}^2}{4\sigma_{x\alpha}^2 \sigma_{x\beta}^2} \left(\mathbf{v}_\alpha - \mathbf{v}_\beta\right)^2, \tag{6.86}$$

where α and β refer to all of the external particles. As we know, σ_{xP} is dominated by the particle with the smallest spatial width. Thus, the above sum is dominated by the terms including the two smallest widths unless their velocities are approximately equal. Let us call the smallest width, σ_{x1} (with the corresponding velocity $\mathbf{v}_1$) and the second smallest one, σ_{x2} (with the corresponding velocity $\mathbf{v}_2$). Therefore, (6.86) becomes

$$\sigma_{eP} \sim \frac{|\mathbf{v}_1 - \mathbf{v}_2|}{\sigma_{x2}} \sim \frac{1}{T_P^{\text{overlap}}}, \tag{6.87}$$

where T_P^{overlap} is time of the production process. The above relation shows that σ_{eP} is proportional to the inverse of time. This means that σ_{eP} can be considered as the energy uncertainty at the source. Similarly, σ_{eD} is defined as the energy uncertainty at the detector.

Now we would like to solve the oscillation amplitude (6.34) with the Gaussian wave packets. By plugging (6.83) into (6.34), we obtain

$$\mathcal{A}_j = N\pi\sigma_{pP}^{-3}\sigma_{eP}^{-1}\sigma_{pD}^{-3}\sigma_{eD}^{-1} \int \frac{\mathrm{d}^3\mathbf{p}}{E_j\left(\mathbf{p}\right)} \exp\left(-f_j\left(E, \mathbf{p}\right)\right) \mathrm{e}^{-i\phi_j(\mathbf{p})}. \tag{6.88}$$

In section (6.5), the methods of integration from the above integral are going to be discussed. Before ending the current section, it should be noted that in the wave packet approach, there is no energy-momentum conservation at each vertex anymore. By this we mean that

$$\mathbf{p}_{P,D} \neq \mathbf{p}; \qquad E_{P,D} \neq p^0. \tag{6.89}$$

This non-conservation stems from the fact that we expanded the energy of the wave packet to the second order around the peak value $\langle \mathbf{p} \rangle$. Hence, we are not left anymore with

$$\begin{aligned} \int \mathrm{d}^4 x \mathrm{e}^{i(p+k-q)\cdot x} &= \delta\left(p+k-q\right) \\ \int \mathrm{d}^4 x' \mathrm{e}^{i\left(k'-q'-p\right)\cdot x'} &= \delta\left(k'-q'-p\right), \end{aligned} \tag{6.90}$$

which guarantee the energy-momentum conservation at each vertex. However, as the propagation is macroscopic, we can assume an average momentum $\mathbf{p}_0$, and energy E_0 as follows

$$\mathbf{p}_P = \mathbf{p}_D \equiv \mathbf{p}_0; \qquad E_P = E_D \equiv E_0, \tag{6.91}$$

and do expansions (around $\mathbf{p}_0$ to the second order and around m_0 to the first order). Therefore,

$$\begin{aligned} E_j\left(\mathbf{p}\right) =& \sqrt{\left|\mathbf{p}\right|^2 + m_j^2} = \\ & E_0 + \mathbf{v}_0 . \left(\mathbf{p}-\mathbf{p}_0\right) + \frac{1}{2E_0}\left(p-p_0\right)^a R^{ab}\left(p-p_0\right)^b + \frac{\delta m_j^2}{2E_0}, \end{aligned} \tag{6.92}$$

where

$$\mathbf{v}_0 = \frac{\mathbf{p}_0}{E_0}; \qquad R^{ab} = \delta^{ab} - v_0^a v_0^b; \qquad \delta m_j^2 = m_j^2 - m^2. \tag{6.93}$$

Expansion (6.92) will be used when the solutions of integral (6.88) is discussed.

6.4.2 Stationary boundary conditions

We are going to end section (6.4) with a discussion about the Stationary boundary limit. This condition is derived when

$$\mathbf{v}_{P,D} \to 0. \tag{6.94}$$

Then, we see that $\Sigma_{P,D}$ goes to zero too. In this case due to (6.81), we have

$$\sigma_{eP,D} \to 0. \tag{6.95}$$

It should be noted that $\sigma_{p_{P,D}}$ are not zero. Since $\mathbf{v}_{P,D}$ are the velocities of the production and detection regions and $\sigma_{eP,D}^{-1}$ are the temporal duration of the production and detection, then the ratios $(\mathbf{v}_{P,D}/\sigma_{eP,D})$ are bounded by the macroscopic sizes of the production and detection regions, $S_{P,D}$, i.e.,

$$\frac{\left|\mathbf{v}_{P,D}\right|}{\sigma_{eP,D}} \lesssim S_{P,D}. \tag{6.96}$$

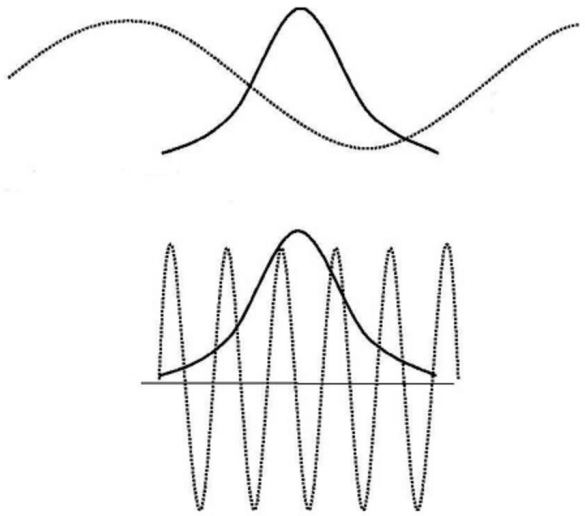


Figure 6.2: The first drawing is related to Laplace's method, i.e., the phase (dotted line) varies slowly in comparison to the overlap function (line). The second drawing corresponds to the stationary phase in which the phase varies rapidly. The drawings have been taken from Ref. [13].

6.5 The methods of integration

Integral (6.88) cannot be solved analytically. However, in this integral there are two types of large parameters (inside the exponentials) with the help of which we can approximate the solutions. These two kinds of parameters are as follows:

1. the large quantities[8] $\sigma_{pP,D}^{-2}$ and $\sigma_{eP,D}^{-2}$ make $f_j\left(p^0,\mathbf{p}\right)$ very large (see (6.82)).

2. Moreover, when time increases, the large parameters T and $\mathbf{L}$ appear in the phase.

Laplace's method. Now, the important question is this: how do these large parameters help approximate the solutions of (6.88)? First, let us start with the case in which the phase $\phi_j\left(\mathbf{p}\right)$ varies much more slowly than $f_j\left(p^0,\mathbf{p}\right)$ appearing in the overlap function (the first drawing in Fig. (6.2)). Under this circumstance, we see that the function inside integral (6.88) has a maximum around where $f_j\left(p^0,\mathbf{p}\right)$ is minimum, i.e., $\mathbf{p}_j$.

As we go further from this point in any direction, the amplitude of the function inside the integral will be vanishingly small and consequently, negligible. Accordingly, we see that for a significantly non-zero amplitude it is sufficient to

[8]This is due to the so-called assumption that our wave packets are sharply peaked, so $\sigma_{pP,D}$ and $\sigma_{eP,D}$ are very small.

expand the very large function $f_j\left(p^0, \mathbf{p}\right)$, just, to the second order around the point $\mathbf{p}_j$ (because expansions to higher orders bring about terms nearly equal to zero or at least much smaller than significantly non-zero terms).

It was assumed that the phase varies slowly over the bump of the overlap function. Therefore, we expand the phase around $\mathbf{p}_j$ too, and keep only the consistent terms and ignore the rest (it will be seen what consistent terms mean as we go further). This method is called **Laplace's method**.

It is clear that we are allowed to apply Laplace's approach when the particle is sufficiently near the source because only in this case T and $\mathbf{L}$ are still not large parameters and the approximation of a slowly-varying phase is acceptable. Furthermore, since time is small, wave packets have not started to spread[9]. It means that Laplace's method is good when dispersion of the wave packets is negligible.

The method of stationary phase. As time increases, the parameters T and $\mathbf{L}$ become large and so we are faced with a fast oscillating (in comparison with the overlap function) phase (see the second drawing in Fig. (6.2)). In this case, we are in a situation exactly contrary to the previous one, i.e, phase varies much faster than the overlap function.

As a result, fast oscillating phase (since phase is imaginary) leads to (nearly) zero amplitude except in a neighborhood around the stationary point $\mathbf{p}_{\mathrm{cl},j}$, which brings us a significantly non-zero amplitude. The point $\mathbf{p}_{\mathrm{cl},j}$ is the point where the first 3−momentum derivatives of the phase vanish. The index, cl, refers to classical and will be explained later on. Same as before, we should expand the phase to the second order and ignore the rest terms. This method which is called **stationary phase** will be applied when dispersion is taken into consideration.

6.6 No-dispersion regime

6.6.1 The transition amplitude

In this section we are going to solve (6.88) with Laplace's method. As explained above, in the first step we need to expand $f_j\left(E_j\left(\mathbf{p}\right), \mathbf{p}\right)$ and the phase around $\mathbf{p}_j$ (where $f_j\left(E_j\left(\mathbf{p}\right), \mathbf{p}\right)$ is minimum) to order $\left(\mathbf{p}-\mathbf{p}_j\right)^2$. Here, there is an important fact.

We call the fraction $\delta m_j^2/2E_0$, ϵ, where $\delta m_j^2 = m_j^2 - m_0^2$. The point $\mathbf{p}_j$ will be derived to $\mathcal{O}\left(\epsilon\right)$. The transition amplitude is going to be calculated to $\mathcal{O}\left(\epsilon^2\right)$ in the real part (i..e, $f_j\left(E_j\left(\mathbf{p}\right), \mathbf{p}\right)$) and to order $\mathcal{O}\left(\epsilon\right)$ in the phase (i.e., $\phi_j\left(\mathbf{p}\right)$). Furthermore, the first derivative of the phase at $\mathbf{p}_j$ will be computed to

[9] In section (6.3), it was said that the spatial uncertainties are time-dependent and so there must be dispersion, although in our calculations we neglected it. However as we go forward, it will be seen that the wave packets associated to the mass eigenstates start to spread as time increases.

order $\mathcal{O}(\epsilon)$ and the second derivatives of the phase and $f_j(E_j(\mathbf{p}),\mathbf{p})$ to order[10] $\mathcal{O}(\epsilon^0)$.

Let us first find $\mathbf{p}_j$. By plugging $E_j(\mathbf{p})$ from (6.92) into $f_j(E_j(\mathbf{p}),\mathbf{p})$ (with regard to (6.91)), the first derivative of $f_j(E_j(\mathbf{p}),\mathbf{p})$ is

$$\nabla f_j(E_j(\mathbf{p}),\mathbf{p}) = 2\frac{(\mathbf{p}-\mathbf{p}_0)}{\sigma_p^2} + 2\{[(\mathbf{p}-\mathbf{p}_0).\mathbf{u}_P]\,\mathbf{u}_P + [(\mathbf{p}-\mathbf{p}_0).\mathbf{u}_D]\,\mathbf{u}_D\} + \left(\frac{\mathbf{u}_P}{\sigma_{eP}} + \frac{\mathbf{u}_D}{\sigma_{eD}}\right)\frac{\delta m_j^2}{2E_0}, \tag{6.97}$$

where

$$\frac{1}{\sigma_p^2} = \frac{1}{\sigma_{pP}^2} + \frac{1}{\sigma_{pD}^2}, \tag{6.98}$$

and

$$\mathbf{u}_{P,D} = \frac{\mathbf{v}_0 - \mathbf{v}_{P,D}}{2\sigma_{eP,D}}. \tag{6.99}$$

From Eq. (6.98), we see that σ_p must be approximately equal to the smallest width among the production and detection widths. In addition, σ_x is defined in configuration space and we have $\sigma_x\sigma_p = 1/2$. By setting (6.97) zero, $\mathbf{p}_j$ reads

$$\mathbf{p}_j = \mathbf{p}_0 + (\alpha\mathbf{u}_p + \beta\mathbf{u}_D)\frac{\delta m_j^2}{2E_0} + \mathcal{O}(\epsilon^2), \tag{6.100}$$

where $\delta m_j^2 = m_j^2 - m_0^2$. Moreover, α and β are two dimensionless coefficients which can be specified by solving $\nabla f_j(E_j(\mathbf{p}),\mathbf{p}) = 0$. However, the exact forms of them are not needed. Now after finding $\mathbf{p}_j$, we can easily find $f_j(\mathbf{p}_j)$ as follows

$$f_j(\mathbf{p}_j) = \left(\frac{\delta m_j^2}{4\tilde{\sigma}_m E_0}\right)^2 + \mathcal{O}(\epsilon^3), \tag{6.101}$$

where $\tilde{\sigma}_m$ is a parameter with the dimension of width whose exact form is not needed. It is seen that $f_j(\mathbf{p}_j)$ is in $\mathcal{O}(\epsilon^2)$ (as explained above). Furthermore, by using (6.100) we can derive E_j to $\mathcal{O}(\epsilon)$ as

$$E_j = \sqrt{\mathbf{p}_j^2 + m_j^2} = E_0 + \tilde{\rho}\frac{\delta m_j^2}{2E_0}, \tag{6.102}$$

with

$$\tilde{\rho} = 1 + \alpha\mathbf{v}_0.\mathbf{u}_P + \beta\mathbf{v}_0.\mathbf{u}_D. \tag{6.103}$$

To compute probability, (6.100) and (6.102) will be used.

[10] We shall understand later why these orders have been chosen.

Important point. It should be remarked that we have ignored the terms proportional to $\mathcal{O}\left(\epsilon^2\right)$ and higher in the phase because they are very small and negligible. However in the modulus of the amplitude, i.e., in the real part of the amplitude, (6.101) is the leading term and yields the observability conditions (as we will see). In other words, since there is a term such as $\exp\left(-f_j\left(\mathbf{p}_j\right)\right)$, the less $f_j\left(\mathbf{p}_j\right)$ is, the more $\exp\left(-f_j\left(\mathbf{p}_j\right)\right)$ becomes and vice versa. This is the important difference in approximation between the imaginary phase and real overlap function.

Now, we should find the second derivatives of $f_j\left(E_j\left(\mathbf{p}\right),\mathbf{p}\right)$ at $\mathbf{p}_j$. The Hessian matrix Σ^{ab} is defined as follows

$$\Sigma^{ab}=\frac{1}{2}\frac{\partial^2 f_j\left(E_j\left(\mathbf{p}\right),\mathbf{p}\right)}{\partial p^a\partial p^b}\left(\mathbf{p}_j\right)=\frac{\delta^{ab}}{4\sigma_p^2}+u_P^a u_P^b+u_D^a u_D^b. \tag{6.104}$$

From the above relation, it is seen that each element of the Hessian matrix can be written as

$$\Sigma^{ab}=\frac{1}{4\left(\sigma_p^{ab}\right)^2}=\left(\sigma_x^{ab}\right)^2, \tag{6.105}$$

where in the last equality we have used the Heisenberg uncertainty principle. In fact, σ_p^{ab} is the momentum uncertainty along the mixed axis ab, e.g., $\left(\sigma_p^{11}\right)$, is the momentum width along the x direction and so on. We have the same story as for the spatial widths. As a result, the elements of the matrix Σ^{ab} specify the range of $\mathbf{x}$ (and consequently, $\mathbf{p}$) for which the overlap function is non-negligible. In other words, one can write

$$\left|p^{ab}-p_j^{ab}\right|\lesssim\sigma_p^{ab}. \tag{6.106}$$

It is seen that the elements of the Hessian matrix in the Cartesian basis is not diagonal. For simplicity, we would like to diagonalize the Hessian matrix. However, diagonalizing of a matrix with elements (6.104) is very complicated. Thus, first let us define a new basis as $\left(\hat{\mathbf{x}}',\hat{\mathbf{y}}',\hat{\mathbf{z}}'\right)$ where $\hat{\mathbf{x}}'$ is along $\mathbf{u}_P\times\mathbf{u}_D$ while $\hat{\mathbf{y}}'$ and $\hat{\mathbf{z}}'$ belong to the plane defined by $\mathbf{u}_P$ and $\mathbf{u}_D$. In this new basis, the Hessian matrix reads

$$\Sigma=\begin{pmatrix}\left(4\sigma_p^2\right)^{-1} & 0 & 0\\ 0 & \left(4\sigma_p^2\right)^{-1}+\left(u_P^2\right)^2+\left(u_D^2\right)^2 & \left(u_P^2\right)\left(u_P^3\right)+\left(u_D^2\right)\left(u_D^3\right)\\ 0 & \left(u_P^2\right)\left(u_P^3\right)+\left(u_D^2\right)\left(u_D^3\right) & \left(4\sigma_p^2\right)^{-1}+\left(u_P^3\right)^2+\left(u_D^3\right)^2\end{pmatrix}. \tag{6.107}$$

Now by using the symmetry of Σ, we can diagonalize it. The eigenvalue corresponding to the eigenvector along $\mathbf{u}_P\times\mathbf{u}_D$, i.e., σ_x^2 is

$$\sigma_x^2=\frac{1}{4\sigma_p^2}. \tag{6.108}$$

Moreover, $\sigma_{x\pm}^2$, i.e., the eigenvalues corresponding to $\hat{\mathbf{y}}'$ and $\hat{\mathbf{z}}'$ are

$$\sigma_{x\pm}^2 = \frac{1}{4\sigma_p^2} + \frac{1}{2}\left(|\mathbf{u}_P|^2 + |\mathbf{u}_D|^2\right) \pm \frac{1}{2}\sqrt{\left(|\mathbf{u}_P|^2 + |\mathbf{u}_D|^2\right)^2 - 4\,|\mathbf{u}_P \times \mathbf{u}_D|^2}. \tag{6.109}$$

Now, the conditions (6.106) can be simply written as

$$\left|p^{x'} - p_j^{x'}\right| \lesssim \sigma_p; \qquad \left|p^{y'} - p_j^{y'}\right| \lesssim \sigma_{p^-}; \qquad \left|p^{z'} - p_j^{z'}\right| \lesssim \sigma_{p^-}. \tag{6.110}$$

From (6.108) and (6.109), we have

$$\sigma_{x+}^2 \geqslant \sigma_{x-}^2 \geqslant \sigma_x^2 \Rightarrow \sigma_p^2 \geqslant \sigma_{p-}^2 \geqslant \sigma_{p+}^2. \tag{6.111}$$

Stationary boundary limit again. In the limit $|\mathbf{u}_P| \gg |\mathbf{u}_D|$ (resp. $|\mathbf{u}_P| \ll |\mathbf{u}_D|$) the term $|\mathbf{u}_P \times \mathbf{u}_D|^2$ can be ignored with respect to $\left(|\mathbf{u}_P|^2 + |\mathbf{u}_D|^2\right)^2$, and so we obtain $\sigma_x = \sigma_{x-}$. If $|\mathbf{u}_P| \gg 1/4\sigma_p^2$ (resp. $|\mathbf{u}_D| \gg 1/4\sigma_p^2$), then we see that the eigenvector corresponding to σ_{x+} becomes aligned with $\mathbf{u}_P$ (resp. $\mathbf{u}_D$). We have the same situation in the limit of parallel $\mathbf{u}_P$ and $\mathbf{u}_D$. These conditions are satisfied in the stationary boundary condition discussed in section (6.4.2). By using (6.94), (6.95), (6.96) and (6.99) we can write

$$\mathbf{u}_P = \frac{\mathbf{v}_0}{2\sigma_{eP}} - S_P, \tag{6.112}$$

and

$$\mathbf{u}_D = \frac{\mathbf{v}_0}{2\sigma_{eD}} - S_D. \tag{6.113}$$

Here, there are macroscopic regions S_P and S_D. If we assume that S_P is nearly zero while S_D is much bigger than zero, we can easily satisfy $|\mathbf{u}_P| \gg |\mathbf{u}_D|$. Thus, in this case the velocity $\mathbf{v}_0$ becomes aligned with $\mathbf{u}_P$ (from (6.112) we see that $\mathbf{u}_P$ goes to infinity) and consequently, it is obtained that

$$v_0^{x,y} \sim v_{P,D}^{x,y} \to 0; \qquad \sigma_{x-}^2 \to \sigma_x^2; \qquad \sigma_{x+}^2 \to \frac{1}{4\sigma_p^2} + \mathbf{u}_P^2 + \mathbf{u}_D^2 \to \infty. \tag{6.114}$$

Let us come back to our discussion. All the necessary terms for expansion of $f_j\left(E_j\left(\mathbf{p}\right), \mathbf{p}\right)$ have been found. It is now the phase's turn. As already said, this phase must vary slowly (in Laplace's method). As a result of this assumption we expand the phase to the second order around $\mathbf{p}_j$, then

$$\phi_j\left(\mathbf{p}\right) = \left(\sqrt{|\mathbf{p}|^2 + m_j^2}\right) T - \mathbf{p}.\mathbf{L} \cong \phi_j\left(\mathbf{p}_j\right) + \left(\mathbf{v}_j T - \mathbf{L}\right)\left(\mathbf{p} - \mathbf{p}_j\right) + \frac{T}{2E_0}\left(p^a - p_j^a\right) R^{ab}\left(p^b - p_j^b\right), \tag{6.115}$$

where $\mathbf{v}_j = \mathbf{p}_j/E_j$. It should be noted that the first derivative (i.e., the second term of (6.115)) should be calculated to $\mathcal{O}(\epsilon)$, so we expand $\mathbf{v}_j$ to order $\mathcal{O}(\epsilon)$. On the other hand, the second derivative (the third term) must be computed to $\mathcal{O}(\epsilon^0)$, so $\mathbf{p}_j = \mathbf{p}_0$ in the third term.

For a slowly varying phase we assume that

$$(\mathbf{v}_j T - \mathbf{L})(\mathbf{p} - \mathbf{p}_j) \lesssim 1, \tag{6.116}$$

and

$$\frac{T}{2E_0}(p^a - p_j^a) R^{ab} (p^b - p_j^b) \lesssim 1. \tag{6.117}$$

By putting conditions (6.110) into (6.116) and (6.117) we obtain, respectively

$$\left|v_j^{x'} T - L^{x'}\right| \sigma_p \lesssim 1; \quad \left|v_j^{y'} T - L^{y'}\right| \sigma_{p^-} \lesssim 1; \quad \left|v_j^{z'} T - L^{z'}\right| \sigma_{p^+} \lesssim 1, \tag{6.118}$$

and

$$\frac{T}{2E_0} \sigma_p^a R^{ab} \sigma_p^b \lesssim 1. \tag{6.119}$$

From the definition of R^{ab} (see (6.93)), it is seen that

$$\begin{aligned} \sigma_p^a R^{ab} \sigma_p^b = \sum_a \left(\sigma_p^a\right)^2 - \sum_{a \neq b} \left(\sigma_p^a v^a v^b \sigma_p^b\right) \leqslant \sum_a \left(\sigma_p^a\right)^2 \Rightarrow \\ \frac{T}{E_0} \sum_a \left(\sigma_p^a\right)^2 \lesssim 1 \Rightarrow \\ \frac{T}{E_0}\sigma_p^2 \lesssim 1; \quad \frac{T}{E_0}\sigma_{p-}^2 \lesssim 1; \quad \frac{T}{E_0}\sigma_{p+}^2 \lesssim 1. \end{aligned} \tag{6.120}$$

If

$$\frac{T}{E_0}\sigma_p^2 \lesssim 1 \tag{6.121}$$

is satisfied, then the other terms in (6.120) become automatically satisfied (see (6.111)). By applying

$$\mathbf{L} \approxeq \mathbf{v}_0 T, \ \text{ if } \quad \sigma_{x+} \ll |\mathbf{L}|, \tag{6.122}$$

from (6.121), it is obtained that

$$|\mathbf{L}| \lesssim \frac{|\mathbf{p}_0|}{\sigma_p^2}. \tag{6.123}$$

As we go forward, it will be seen why condition (6.122) is valid. In addition, it will become clear that Laplace's method criterion is just (6.121) or equivalently (6.123). This means that if $T \lesssim E_0/\sigma_p^2$ is met, then conditions (6.118) must be

satisfied, otherwise the amplitude becomes negligible.

In (6.122), we assumed that $\sigma_{x+} \ll |\mathbf{L}|$ and Laplace's condition (6.123) was derived. What about the stationary boundary limit in which $\sigma_{x+} \gtrsim |\mathbf{L}|$? In this case due to zero energy uncertainty (6.95), time becomes indeterminate and dispersion loses its meaning. However, (6.123) can be derived without applying (6.121).

From (6.82), it is seen that when (6.94) and (6.95) are satisfied, then $E = E_0$ and so $|\mathbf{p}| = \sqrt{E_0^2 - m_j^2}$. It means that the absolute value of $\mathbf{p}$ becomes constant and we are left with an angular integration in (6.88). As a result, from $(\mathbf{p} - \mathbf{p}_0)^2$ we just keep the term $\mathbf{p}.\mathbf{p}_0$ (the other terms are constant and are taken out of the integral). So in the stationary boundary limit, we should integrate

$$\exp\left(\frac{\mathbf{p}\cdot\mathbf{p}_0}{2\sigma_p^2} + i\mathbf{p}\cdot\mathbf{L}\right). \tag{6.124}$$

If condition (6.123) is valid, then variation of the phase $(i\mathbf{p}.\mathbf{L})$ is more slowly with respect to the angular variation of the overlap function in (6.124). Therefore, Laplace's method is still useful.

Now, everything which is needed (i.e., the expansions) to solve (6.88) with Laplace's method has been obtained. Then, by plugging the exponentials into (6.88), we reach a Gaussian integral. With respect to (6.58), the amplitude reads

$$\mathcal{A}_j = N\sigma_p\sigma_{p-}\sigma_{p+}\exp\left(-iE_jT + i\mathbf{p}_j\cdot L - f_j\left(\mathbf{p}_j\right) - F_j\left(T\right)\right), \tag{6.125}$$

where N absorbs all of the constants and $f_j\left(\mathbf{p}_j\right)$ was found in (6.101). Furthermore, $F_j\left(T\right)$ is

$$F_j\left(T\right) = \frac{1}{4}\left(\mathbf{v}_jT - \mathbf{L}\right)\left(\Sigma + i\frac{T}{2E_0}R\right)^{-1}\left(\mathbf{v}_jT - \mathbf{L}\right), \tag{6.126}$$

where Σ and R are matrices (see (6.104) and (6.93)). In the above function there is the term $i\left(T/E_0\right)R$ which is time-dependent. The function $\exp\left(-F_j\left(T\right)\right)$ is the spacetime envelope of the wave packet with mass eigenstate m_j. Therefore, when time increases and the term $i\left(T/E_0\right)R$ becomes comparable (and not negligible anymore) with the time-independent term Σ, then the wave packets start to spread and so we are faced with dispersion[11].

Accordingly, as time approaches the threshold (6.121) (i.e., the term $i\left(T/E_0\right)R$ becomes comparable with Σ), we are going to have dispersion of the wave packets or in other words, Laplace's method is not applicable anymore. Hence, if we assume that this time-dependent term is negligible with respect to Σ, then Laplace's method is useful. This is the reason that when condition (6.121) or (6.123) is valid, we are in the *no-dispersion regime*.

[11]Thus, despite the fact that the dispersion resulted from the time-dependent spatial uncertainties (see (6.64) and (6.65)) were neglected, again there is spreading of the wave packets due to the function $F_j\left(T\right)$.

Why do we expand the first derivative of the phase to $\mathcal{O}(\epsilon)$, while the second derivatives to $\mathcal{O}(\epsilon^0)$? This issue will be understood when we start discussing transversal and longitudinal dispersions in the following sections. But, here, we explain very briefly.

In fact, as a result of these orders we obtain a coherence length beyond which oscillations vanish. However, in the no-dispersion regime there is no coherence length, in other words, oscillations do not vanish. Thus, in the no-dispersion regime one may assume that $\mathbf{v}_j = \mathbf{v}_0$ (i.e., the first derivative is also in $\mathcal{O}(\epsilon^0)$) and write (6.126) as

$$F(T) = \frac{(v_0^x T - L^x)^2}{4\sigma_x^2} + \frac{(v_0^y T - L^y)^2}{4\sigma_{x-}^2} + \frac{(v_0^z T - L^z)^2}{4\sigma_{x+}^2}. \tag{6.127}$$

To have an amplitude significantly different from zero, we must have $F(T) \lesssim 1$. Hence, it is seen that automatically conditions (6.118) are satisfied[12]. In addition, if we divide both sides of the third condition in (6.118) by $|\mathbf{L}|$, we have

$$|v_0^z T - L^z| \ll 1, \tag{6.128}$$

if $\sigma_{x+} \ll |\mathbf{L}|$. Since σ_{x+} is the largest among the widths (see (6.111)), then we have the same situations for the first and second conditions in (6.118). As a result, (6.122) is proved. The fact which tells us for a non-zero amplitude, $\mathbf{p}_0$ and $\mathbf{L}$ must be nearly parallel. As we go ahead, it will be seen that even in the stationary limit where $\sigma_{x+} \ll |\mathbf{L}|$ is violated, $\mathbf{p}_0$ and $\mathbf{L}$ must stay nearly parallel for a non-zero probability.

6.6.2 The transition probability

From (6.125), the transition probability can be obtained by squaring the amplitude and averaging over the macroscopic time T. This is because time is not observed in the experiment (same as what was discussed in the previous chapter). Before doing so, we first expand the function $F(T)$ around T_0, the point where the first $T-$derivative of $F(T)$, i.e., $F'(T)$ vanishes. Hence from (6.127), we have

$$F'(T_0) = 0 \Rightarrow T_0 = \frac{\tilde{\mathbf{v}}_0.\tilde{\mathbf{L}}}{\tilde{\mathbf{v}}_0^2}, \tag{6.129}$$

where

$$\tilde{\mathbf{v}}_0 = \sigma_x \left(\frac{v_0^x}{\sigma_x}\hat{\mathbf{x}} + \frac{v_0^y}{\sigma_{x-}}\hat{\mathbf{y}} + \frac{v_0^z}{\sigma_{x+}}\hat{\mathbf{z}} \right) = \sigma_x \sqrt{\Sigma^{-1}}\mathbf{v}_0, \tag{6.130}$$

[12]It is worth repeating that if condition (6.121) is valid, then conditions (6.118) must be satisfied, otherwise we will have nearly zero amplitude.

and

$$\tilde{\mathbf{L}} = \sigma_x \left(\frac{L^x}{\sigma_x} \hat{\mathbf{x}} + \frac{L^y}{\sigma_{x-}} \hat{\mathbf{y}} + \frac{L^z}{\sigma_{x+}} \hat{\mathbf{z}} \right) = \sigma_x \sqrt{\Sigma^{-1}} \mathbf{L}, \tag{6.131}$$

with $\Sigma^{-1} = \mathrm{diag}\left(\sigma_x^{-2}, \sigma_{x-}^{-2}, \sigma_{x+}^{-2}\right)$. With respect to (6.130) and (6.131), one can write (6.127) as

$$F\left(T\right) = \frac{\left| \tilde{\mathbf{v}}_0 T - \tilde{\mathbf{L}} \right|^2}{4\sigma_x^2}. \tag{6.132}$$

By using the following identity,

$$\begin{aligned} &\left| \tilde{\mathbf{v}}_0 \times \left(\tilde{\mathbf{v}}_0 \times \tilde{\mathbf{L}} \right) \right| = \left| \tilde{\mathbf{v}}_0 \left(\tilde{\mathbf{v}}_0 . \tilde{\mathbf{L}} \right) - \tilde{\mathbf{L}} \tilde{\mathbf{v}}_0^2 \right| = \left| \tilde{\mathbf{v}}_0 \right| \left| \tilde{\mathbf{v}}_0 \times \tilde{\mathbf{L}} \right| \Rightarrow \\ &4\sigma_x^2 \tilde{\mathbf{v}}_0^2 \underbrace{\frac{1}{4\sigma_x^2} \left| \frac{\tilde{\mathbf{v}}_0 \left(\tilde{\mathbf{v}}_0 . \tilde{\mathbf{L}} \right)}{\tilde{\mathbf{v}}_0^2} - \tilde{\mathbf{L}} \right|^2}_{F(T_0)} = \left| \tilde{\mathbf{v}}_0 \times \tilde{\mathbf{L}} \right|^2 \Rightarrow \\ &F\left(T_0\right) = \frac{\left(\tilde{\mathbf{v}_0} \times \tilde{\mathbf{L}} \right)^2}{4\sigma_x^2 \tilde{\mathbf{v}_0^2}}. \end{aligned} \tag{6.133}$$

Moreover, the second $T-$derivative of $F\left(T\right)$, $F''\left(T_0\right)$, is

$$F''\left(T_0\right) = \frac{\tilde{\mathbf{v}}_0^2}{2\sigma_x^2}. \tag{6.134}$$

Now, if we put the expanded $F\left(T\right)$ into (6.125) and complete the squares, we obtain

$$\begin{aligned} &\int \mathrm{d}T' \mathcal{A}_i \mathcal{A}_j^* \sim \\ &\exp\left[-i\phi_{ij}\left(T_0\right) - \frac{\left(\delta m_i^2\right)^2 + \left(\delta m_j^2\right)^2}{16\tilde{\sigma}_m^2 E_0^2} - \frac{\left(E_i - E_j\right)^2}{4F''\left(T_0\right)} - 2F\left(T_0\right) \right] \times \\ &\int \mathrm{d}T' \exp\left[-F''\left(T_0\right) T'^2 \right], \end{aligned} \tag{6.135}$$

where

$$\phi_{ij}\left(T_0\right) = \left(E_i - E_j\right) T_0 - \left(\mathbf{p}_i - \mathbf{p}_j\right) \cdot \mathbf{L}. \tag{6.136}$$

The time interval ΔT of the above integral is finite and does not go to infinity. However, one can use (6.58) if (compare (6.135) with a Gaussian integral)

$$\Delta T \gtrsim \frac{2}{\sqrt{F''\left(T_0\right)}}, \tag{6.137}$$

because out of the above range the amplitude is nearly zero and thus negligible. Now, from (6.134) and (6.137) we can write

$$|\mathbf{v}_0|\,\Delta T \gtrsim \tilde{\sigma}_{x\,\mathrm{eff}}, \tag{6.138}$$

where

$$\tilde{\sigma}_{x\,\mathrm{eff}} = \frac{|\mathbf{v}_0|}{|\tilde{\mathbf{v}}_0|}\sigma_x. \tag{6.139}$$

The width of the amplitude (6.125) is the width of the overlap function in Laplace's method. The width of the squared amplitude is still the width of the overlap function. So, $\tilde{\sigma}_{x\,\mathrm{eff}}$ (known as effective width) can be interpreted as the width of the overlap function. Regarding this approximation, one can write (6.135) as

$$\int \mathrm{d}T\mathcal{A}_i\mathcal{A}_j^* = N_{\tilde{g}} \exp\left[-i\phi_{ij}\left(T_0\right) - \frac{\left(\delta m_i^2\right)^2 + \left(\delta m_j^2\right)^2}{16\tilde{\sigma}_m^2 E_0^2} - \frac{\left(E_i - E_j\right)^2}{4F''\left(T_0\right)} - 2F\left(T_0\right)\right], \tag{6.140}$$

where $N_{\tilde{g}}$ absorbs all of the non-exponential constants (independent of $\mathbf{L}$) inside itself.

As it was already said for a non-zero probability, $F\left(T_0\right)$ must be very small. Therefore, (6.133) imposes that $\tilde{\mathbf{v}}_0$ and $\tilde{\mathbf{L}}$ are nearly parallel. What about non-tilted $\mathbf{L}$ and $\mathbf{v}_0$? In the case $\sigma_{x+} \ll |\mathbf{L}|$ we saw that $\mathbf{L}$ and $\mathbf{v}_0$ are nearly parallel. Due to this parallel, we can write

$$\mathbf{L} = \frac{\mathbf{v}_0}{|\mathbf{v}_0|}\,|\mathbf{L}| + \mathcal{O}\left(\sigma_{x+}\right). \tag{6.141}$$

Regarding (6.129), (6.130), (6.131) and (6.141), it is derived that

$$T_0 = \frac{|\mathbf{L}|}{|\mathbf{v}_0|} + \mathcal{O}\left(\sigma_{x+}\right). \tag{6.142}$$

Now, by using (6.100), (6.102), (6.103) and (6.142), the phase (6.136) reads

$$\phi_{ij}\left(T_0\right) = \frac{\delta m_{ij}^2}{2\,|\mathbf{p}_0|}\left(|\mathbf{L}| + \mathcal{O}\left(\sigma_{x+}\right)\right), \tag{6.143}$$

which is the standard phase (5.29) or (5.31).

The other case is $\sigma_{x+} \gg \sigma_x$. This case can be (but not only) stationary limit boundary. If so, $\tilde{\mathbf{v}}_0 \to 0$ and $\tilde{\mathbf{L}}^z \to 0$ (see (6.114), (6.130) and (6.131)). Accordingly, the point T_0 (see (6.129)) becomes indeterminate and we cannot have any information about $\mathbf{v}_0$ and $\mathbf{L}$.

This problem can be overcome if we write $F\left(T_0\right)$ with regard to $\mathbf{v}_0$ and $\mathbf{L}$

and not tilted quantities. So, first, we find T_0 by plugging (6.130) and (6.131) into (6.129). Then from (6.132), $F(T_0)$ reads

$$F(T_0) = \left(\sigma_{x+}^2 \left(v_0^y\right)^2 + \left(v_0^z\right)^2\right)\left(L^x\right)^2 - 2\sigma_{x+}^2 v_0^x v_0^y L^x L^y - 2v_0^x v_0^z L^x L^z + \left(\sigma_{x+}^2 \left(v_0^x\right)^2 + \left(v_0^z\right)^2\right)\left(L^y\right)^2 - 2v_0^y v_0^z L^y L^z + \left(\left(v_0^x\right)^2 + \left(v_0^y\right)^2\right)\left(L^z\right)^2. \quad (6.144)$$

It is important to note that the above equation is not valid only in the stationary limit, but, it is accepted in all limits.

Actually, $F(T_0) \lesssim 1$. At the most, we assume that $F(T_0) = 1$. Therefore, (6.144) becomes an ellipse in space (L^x, L^y, L^z). Due to the non-quadratic terms such as $L^x L^y$, the diameters of this ellipse might be not aligned with L^x, L^y and L^z. We should diagonalize the ellipse to find the corresponding eigenvectors and eigenvalues. This is done if the matrix, whose elements are the coefficients of $L^i L^j$ in (6.144), is diagonalized.

By doing so, it is found that there is an eigenvalue $s_3 = 0$ associated to an eigenvector along $\mathbf{v}_0$. The other two eigenvalues are positive s_1 and s_2 (where $s_1 \geqslant s_2$) corresponding to two eigenvectors in the plane $L^z = 0$. Hence, there is a cylinder with an axis along $\mathbf{v}_0$ and elliptical section. If the stationary boundary conditions are taken into consideration (see (6.114)), s_1 and s_2 become

$$s_1 \to \frac{1}{4\sigma_x^2},$$
$$s_2 \to \frac{1}{4\sigma_x^2} \frac{\sigma_x^2 \left(v_0^z\right)^2}{\sigma_{x+}^2 \left(v_0^x\right)^2 + \sigma_{x+}^2 \left(v_0^y\right)^2 + \sigma_x^2 \left(v_0^z\right)^2}.$$

The ellipse equation in the new basis $\left(L^{x'}, L^{y'}, L^{z'}\right)$ becomes

$$\frac{\left(L^{x'}\right)^2}{(1/s_1)} + \frac{\left(L^{y'}\right)^2}{(1/s_2)} + \frac{\left(L^{z'}\right)^2}{(1/s_3)} = 1. \quad (6.145)$$

The ellipse's diameters along the axes $L^{x'}, L^{y'}$ and $L^{z'}$ are $\sqrt{1/s_1}$, $\sqrt{1/s_2}$ and $\sqrt{1/s_3}$, respectively. The diameter along $L^{z'}$ is infinite. Let us assume we are in the stationary limit (6.114): $\sigma_{x+} \to \mathbf{u}_P \to \infty$ and $v_0^{x,y} \sim v_P^{x,y}$. Therefore, $\sqrt{1/s_2}$ can be written as follows

$$\sqrt{1/s_2} = 2\left|\mathbf{u}_P\right| \sqrt{\frac{\left(v_0^x\right)^2 + \left(v_0^y\right)^2}{\left(v_0^z\right)^2}} \lesssim 2\left|\mathbf{u}_P\right| \frac{\left|\mathbf{v}_P\right|}{v_0^z}. \quad (6.146)$$

In the stationary limit, we have

$$\sigma_{eP} = \frac{\left|\mathbf{v}_0 - \mathbf{v}_P\right|}{2\left|\mathbf{u}_p\right|} = \frac{\left|\mathbf{v}_0\right|}{2\left|\mathbf{u}_P\right|} = \frac{v_0^z}{2\left|\mathbf{u}_P\right|} \Rightarrow$$
$$\frac{\left|\mathbf{v}_P\right|}{\sigma_{eP}} = 2\left|\mathbf{u}_P\right| \frac{\left|\mathbf{v}_P\right|}{v_0^z}. \quad (6.147)$$

Hence, by comparing (6.146) with (6.147) we obtain that

$$\sqrt{1/s_2} \lesssim \frac{|\mathbf{v}_P|}{\sigma_{eP}} \Rightarrow \sqrt{1/s_2} \lesssim S_P, \tag{6.148}$$

since (see (6.96))

$$S_P \gtrsim \frac{|\mathbf{v}_P|}{\sigma_{eP}}. \tag{6.149}$$

With respect to this fact that $s_2 \leqslant s_1$, we must also have

$$\sqrt{1/s_1} \lesssim S_P. \tag{6.150}$$

From (6.148) and (6.150), it is concluded that in the stationary limit, the diameters of the ellipse along $L^{x'}$ and $L^{y'}$ must be smaller than $S_{P,D}$. This means that if it is assumed that $S_{P,D}$ are very small, then the components $L^{x'}$ and $L^{y'}$ must be negligible with respect to the diameter along $L^{z'}$ (or $\mathbf{v}_0$) which goes to infinity. As a result of this discussion, the phase reads

$$\phi_{ij}\left(T_0\right) = \frac{\Delta m_{ij}^2}{2\left|\mathbf{p}_0\right|}\left(|\mathbf{L}| + \mathcal{O}\left(S_{P,D}\right)\right), \tag{6.151}$$

which is again equal to the standard oscillation phase if $S_{P,D} \ll |\mathbf{L}| \approxeq L^{z'}$.

Now, with the help of (6.102) and (6.134) we obtain

$$\frac{\left(E_i - E_j\right)^2}{4F''\left(T_0\right)} = \left(\tilde{\rho}\tilde{\sigma}_{x\,\mathrm{eff}}\frac{\delta m_{ij}^2}{2\sqrt{2}\left|\mathbf{p}_0\right|}\right)^2. \tag{6.152}$$

Finally regarding (6.140) and (6.32), substituting $F\left(T_0\right)$ with (6.133), substituting $\phi_{ij}\left(T_0\right)$ with (6.143) or (6.151) (no difference) and applying (6.152), the probability transition for flavor mixing reads

$$\begin{aligned}\mathcal{P}_{\alpha\to\beta}\left(\mathbf{L}\right) \sim N_{\tilde{g}} \exp\left[-\frac{\left(\tilde{\mathbf{v}}_0 \times \tilde{\mathbf{L}}\right)^2}{2\sigma_x^2\tilde{\mathbf{v}}_0^2}\right] \sum_{i,j=1}^{3} U_{i\alpha}U_{j\beta}U_{\beta i}^*U_{\alpha j}^* \\ \times \exp\left[-2\pi i\frac{L}{L_{ij}^{\mathrm{osc}}} - \frac{\left(\delta m_i^2\right)^2 + \left(\delta m_j^2\right)^2}{16\tilde{\sigma}_m^2E_0^2} - 2\pi^2\left(\frac{\tilde{\rho}\tilde{\sigma}_{x\,\mathrm{eff}}}{L_{ij}^{\mathrm{osc}}}\right)^2\right],\end{aligned} \tag{6.153}$$

where

$$L_{ij}^{\mathrm{osc}} = \frac{4\pi\left|\mathbf{p}_0\right|}{\delta m_{ij}^2}. \tag{6.154}$$

The factor $N_{\tilde{g}}$ is a normalization constant which can be easily found as

$$\sum_{\beta}\int L^2\mathrm{d}\Omega\mathcal{P}_{\alpha\to\beta}\left(\mathbf{L}\right) = 1, \tag{6.155}$$

where Ω is the solid angle.

6.7 Transversal-dispersion regime

6.7.1 Definition

As time increases and violates (6.121), the story becomes different. This means that $\phi_j(\mathbf{p})$ becomes big and consequently, the function $\exp(-i\phi_j(\mathbf{p}))$ oscillates rapidly over most of the range of integration. Here, by rapidly we mean that in comparison to the overlap function, i.e., $\exp(-f_j(E,\mathbf{p}))$, $\exp(-i\phi_j(\mathbf{p}))$ varies more rapidly. Therefore from (6.82), we must have

$$\frac{|\mathbf{p}-\mathbf{p}_0|}{2\sigma_{pP}} \lesssim 1. \tag{6.156}$$

Since $\sigma_{pP} \ll 1$, then $\mathbf{p}$ and $\mathbf{p}_0$ must be nearly parallel. In other words, if $\mathbf{p}$ and $\mathbf{p}_0$ are nearly parallel, the oscillation amplitude becomes significantly different from zero.

Due to the fast oscillations, the oscillation amplitude is approximately zero except in a neighborhood around $\mathbf{p}_{\mathrm{cl},j}$ where this point, called *stationary point*, is the solution of

$$\nabla_{\mathbf{p}}\phi_j(\mathbf{p}) = 0. \tag{6.157}$$

As discussed in section (6.5), in this case integral (6.88) is solved with the stationary phase method in which both $f_j(\mathbf{p})$ and $\phi_j(\mathbf{p})$ must be Taylor expanded around $\mathbf{p}_{\mathrm{cl},j}$ to the second order. This is because the other (big) terms become very distant from $\mathbf{p}_{\mathrm{cl},j}$ and then their contributions practically vanish.

Before continuing, let us discuss a little more the dispersion. In the stationary limit, i.e., conditions (6.114), we can write (6.126) as

$$F_j(T) = \frac{\left(v_j^x T - L^x\right)^2}{4\sigma_x^2 + iT/2E_0} + \frac{\left(v_j^y T - L^y\right)^2}{4\sigma_{x-}^2 + iT/2E_0} + \frac{\left(v_j^z T - L^z\right)^2}{4\sigma_{x+}^2 + iT/2E_0\left(1-|\mathbf{v}_0|^2\right)}. \tag{6.158}$$

The above equation shows that in the direction of $\mathbf{p}_0$, i.e., the z axis, there is the relativistic time contraction $1-|\mathbf{v}_0|^2$. So in this direction, the temporal term effect appears later, in comparison with the directions transverse to $\mathbf{p}_0$.

Furthermore, since σ_{x+}^2 is much bigger than σ_x^2 (note: $\sigma_x \to \sigma_{x-}$), again in the direction of $\mathbf{p}_0$, the effect of σ_{x+}^2 is considerable. Hence, it is understood that the method of stationary phase is preferred in the transversal directions (i.e., the x and y axes) rather than the longitudinal direction (i.e., the z axis).

If the stationary boundary limit is not valid, we should diagonalize the matrix R in (6.126), i.e.,

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1-|\mathbf{v}_0|^2 \end{pmatrix}. \tag{6.159}$$

Thus, in the direction of $\mathbf{p}_0$ there is the relativistic contraction $1-|\mathbf{v}_0|^2$, whereas in the transversal directions we do not have this contraction. So, it can be said that in the directions transverse to $\mathbf{p}_0$, we should apply the stationary phase method to solve integral (6.88) rather than in the longitudinal direction.

When the stationary phase method has been chosen for the transversal directions whereas Laplace's method has been used for the longitudinal direction, we are in the *transversal-dispersion regime*.

But if time increases and crosses the transversal-dispersion regime, i.e., in spite of the time contraction or very big σ_{x+}, the temporal term effect is not negligible anymore, we must integrate in all directions, transversal or longitudinal, by the stationary phase method. In this case we are in the *longitudinal-dispersion regime* which is going to be discussed in section (6.8).

6.7.2 The transition amplitude

Now let us come back to the transversal-dispersion regime. Regarding (6.157) we have

$$\phi_j(\mathbf{p}) = \left(\sqrt{|\mathbf{p}|^2 + m_j^2}\right) T - \mathbf{p}\cdot\mathbf{L} \Rightarrow \nabla_{\mathbf{p}}\phi_j(\mathbf{p}) = \left(\frac{\mathbf{p}_{\mathrm{cl},j}}{E_{\mathrm{cl},j}} T - \mathbf{L}\right) = 0, \quad (6.160)$$

where $E_{\mathrm{cl},j} = \sqrt{\left|\mathbf{p}_{\mathrm{cl},j}\right|^2 + m_j^2}$. Due to (6.160) we see that

$$\mathbf{p}_{\mathrm{cl},j} \parallel \mathbf{L}. \quad (6.161)$$

From (6.156), we understand that $\mathbf{p}$ and $\mathbf{p}_0$ must be nearly parallel. In (6.161), it is seen that momentum at $\mathbf{p}_{\mathrm{cl},j}$ is parallel with $\mathbf{L}$. Therefore at $\mathbf{p}_{\mathrm{cl},j}$, $\mathbf{p}_0$ and $\mathbf{L}$ are nearly parallel. Since we do not go very far away from point $\mathbf{p}_{\mathrm{cl},j}$, then it can be said that $\mathbf{p}_0$ and $\mathbf{L}$ are nearly parallel for an amplitude significantly different from zero.

Suppose that $\mathbf{L}$ is along the z axis. Then, from (6.161) it is understood that the stationary points $p^x_{\mathrm{cl},j}$ and $p^y_{\mathrm{cl},j}$ are zero. So, to integrate along the axes x and y by stationary phase method, we should expand $\phi_j(\mathbf{p})$ and $f_j(\mathbf{p})$ around $p^x_{\mathrm{cl},j} = p^y_{\mathrm{cl},j} = 0$, to the second order as follows

$$\phi_j(p_x,p_y) \cong \phi_j(0,0) + \frac{1}{2}\left(p_x^2\phi_j^{xx}(0,0) + p_x^2\phi_j^{yy}(0,0)\right) + p_xp_y\phi_j^{xy}(0,0), \quad (6.162)$$

$$f_j(p_x,p_y) \cong p_xf_j(0,0) + p_xf_j^x(0,0) + p_yf_j^y(0,0) + \frac{1}{2}\left(p_x^2f_j^{xx}(0,0) + p_x^2f_j^{yy}(0,0)\right) + p_xp_yf_j^{xy}(0,0), \quad (6.163)$$

where the superscript x (xx) refers to the first (second) p_x-derivative of $f_j(p_x,p_y)$ or $\phi_j(p_x,p_y)$ at $p^x_{\mathrm{cl},j} = p^y_{\mathrm{cl},j} = 0$. There is the same story for the superscript y. Now, we should just put the expanded functions (6.162) and (6.163) into (6.88), and so the integral over p_x and p_y will be Gaussian. Hence from (6.58), (6.88)

becomes (note: $p^z \equiv p,\ \ p_0 \equiv p_0^z,\ \ v_p \equiv v_p^z$)

$$\mathcal{A}_j = \frac{N g(\mathbf{l})}{T - i\mu} \int \mathrm{d}p \exp\left(-i\phi_j(p) - f_j(p)\right), \tag{6.164}$$

where N contains the constant (independent of distance). In (6.164), $f_j(p) = f_{jP}(p) + f_{jD}(p)$ with

$$f_{jp}(p) = \frac{(p-p_0)^2}{4\sigma_{pP}^2} + \frac{\left(\sqrt{p^2+m_j^2} - E_0 - (p-p_0)v_p\right)^2}{4\sigma_{eP}^2}, \tag{6.165}$$

where E_0 absorbs $p_0^x v_p^x + p_0^y v_p^y$ into itself. Moreover,

$$g(\mathbf{l}) = \exp\left(-\frac{(\mathbf{p}_0 \times \mathbf{l})^2}{4\sigma_p^2}\right) \qquad \text{with} \qquad \mathbf{l} = \mathbf{L}/L, \tag{6.166}$$

and

$$\phi_j(p) = \sqrt{p^2+m_j^2}T - pL. \tag{6.167}$$

The constant μ in the denominator of (6.164), which is equal to $E_0/2\sigma_p^2$, guarantees that time must be bigger than μ, i.e., the no-dispersion regime has finished. In the limit where $T \gg \mu$, then the denominator becomes approximately T.

To have a significantly different amplitude from zero we must have

$$\frac{|\mathbf{p}_0 \times \mathbf{l}|^2}{4\sigma_p^2} = \frac{|\mathbf{p}_0|^2}{4\sigma_p^2}\sin^2(\theta) \lesssim 1, \tag{6.168}$$

where θ is the angle between $\mathbf{L}$ and $\mathbf{p}_0$. In fact, as it was discussed in the previous section, in the no-dispersion regime the transition probability is significantly different from zero within a cylinder of axis $\mathbf{p}_0$ (remember that $\mathbf{p}_0 \parallel \mathbf{L}$ in the no-dispersion regime). Now, in the transversal-dispersion regime, an angle θ appears between $\mathbf{p}_0$ and $\mathbf{L}$. Thus, we are left with a cone with the axis along $\mathbf{p}_0$ (Fig. (6.3)). However, as already explained $\mathbf{p}_0$ and $\mathbf{L}$ must be nearly parallel, i.e., the angle θ must be very small. From (6.168), it is seen that only if

$$\theta \lesssim \arcsin\left(\frac{\sigma_p}{|\mathbf{p}_0|}\right),$$

the transition amplitude is significantly different from zero.

Now, we are left with just one integral over p (recall that $p \equiv p^z$). To calculate this integral it is needed to know whether dispersion along $\mathbf{p}_0$ has started or not (the longitudinal dispersion is postponed with respect to the transversal ones). If dispersion has not yet started, then Laplace's method will be the best approach. In this case we should follow the steps such as those in the no-dispersion case.

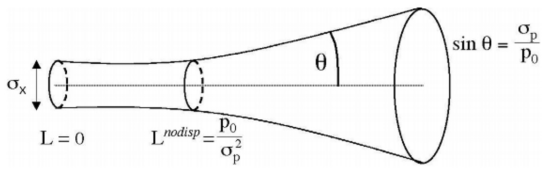


Figure 6.3: Probability becomes significantly different from zero within a cylinder in the no-dispersion regime and within a cone in the transversal-dispersion regime. The figure has been taken from Ref. [13].

This means that both function (6.165) and phase (6.167) should be Taylor expanded around the point p_j, for which $f_j(p)$ is extremal, to order $(p-p_j)^2$. Moreover, the phase must be slowly varying[13].

For finding p_j (to $\mathcal{O}(\epsilon)$), the first derivative of (6.165), $f_j'(p)$, should be set to zero, i.e.,

$$f_j'(p) = 0 \Rightarrow p_j = p_0 + (\rho - 1)\frac{\delta m_j^2}{2p_0}, \tag{6.169}$$

where

$$\rho = \sigma_{p\,\text{eff}}^2 \left(\frac{1}{\sigma_p^2} - \frac{v_p(v_0 - v_p)}{\sigma_{eP}^2} - \frac{v_D(v_0 - v_D)}{\sigma_{eD}^2} \right), \tag{6.170}$$

with the *effective momentum width*

$$\frac{1}{\sigma_{p\,\text{eff}}^2} = \frac{1}{\sigma_{pP}^2} + \frac{1}{\sigma_{pD}^2} + \frac{(v_0 - v_p)^2}{\sigma_{eP}^2} + \frac{(v_0 - v_D)^2}{\sigma_{eD}^2}. \tag{6.171}$$

Then, $f_j(p_j)$ to $\mathcal{O}(\epsilon^2)$ reads

$$f_j(p_j) = \frac{1}{4\sigma_m^2}\left(\frac{\delta m_j^2}{2E_0}\right)^2 + \mathcal{O}(\epsilon^3), \tag{6.172}$$

with

$$\frac{1}{\sigma_m^2} = \sigma_{p\,\text{eff}}^2 \left[\frac{1}{\sigma_p^2}\left(\frac{1}{\sigma_{eP}^2} + \frac{1}{\sigma_{eD}^2} \right) + \frac{(v_p - v_D)^2}{\sigma_{eP}^2 \sigma_{eD}^2} \right]. \tag{6.173}$$

Since $\sigma_{eP,D} \leqslant \sigma_{pP,D}$ (see (6.84)), then $\sigma_{p\,\text{eff}}$ is dominated by the smallest between σ_{eP} and σ_{eD}, say σ_{eP}. Moreover, there is the same story for σ_m. So, we

[13]Recall that we find p_j to $\mathcal{O}(\epsilon)$, $f_j(p_j)$ to $\mathcal{O}(\epsilon^2)$. The first derivative of the phase to $\mathcal{O}(\epsilon)$ and the second derivatives to $\mathcal{O}(\epsilon^0)$.

have $\sigma_m \sim v_0 \sigma_p$.

As the next step, $f_j''(p_j)$, i.e., the second derivative of $f_j(p)$ at p_j should be derived (to $\mathcal{O}(\epsilon^0)$). So,

$$\frac{1}{2} f_j''(p_j) = \frac{1}{4\sigma_{p\,\text{eff}}^2}. \tag{6.174}$$

Now, from (6.172) and (6.174), one can write the expansion of $f_j(p)$ as follows

$$f_j(p) = f_j(p_j) + \frac{(p-p_j)^2}{4\sigma_{p\,\text{eff}}^2}. \tag{6.175}$$

Since in Laplace's method the overlap function varies faster than the phase, so the momentum width of the oscillation process is the width of the overlap function. This means that with respect to (6.175), the physical meaning of $\sigma_{p\,\text{eff}}$ is the width of the oscillation process. Same as what we had in (6.110), $\sigma_{p\,\text{eff}}$ in (6.175) gives a constraint on the range of momentum, i.e.,

$$|p - p_j| \lesssim 2\sigma_{p\,\text{eff}}. \tag{6.176}$$

Furthermore, we define the effective spatial width $\sigma_{x\,\text{eff}}$ and according to the Heisenberg uncertainty we have $\sigma_{x\,\text{eff}}\sigma_{p\,\text{eff}} = 1/2$.

The width σ_m is called *mass width*. We know that $f_j(p_j)$ must be less than one, so from (6.172) it is seen that

$$\frac{|\delta m_j^2|}{2E_0} \lesssim \sigma_m \Rightarrow \left(\frac{|\delta m_i^2|}{2E_0}\right)^2 + \left(\frac{|\delta m_j^2|}{2E_0}\right)^2 \lesssim 2\sigma_m^2 \Rightarrow \left(\frac{\delta m_{ij}^2}{2E_0}\right)^2 \lesssim 2\sigma_m^2, \tag{6.177}$$

where the above result stems from the fact that

$$\left(\frac{|\delta m_i^2|}{2E_0}\right)^2 + \left(\frac{|\delta m_j^2|}{2E_0}\right)^2 \geqslant \left(\frac{\delta m_{ij}^2}{2E_0}\right)^2.$$

From (6.177), it is easily seen that σ_m puts a constraint on the mass difference. This is why σ_m is called the mass width.

To complete Laplace's method, we should Taylor expand the phase (6.167) around p_j (following (6.115) with respect to the fact that the non-zero momentum range is along z axis). So,

$$\phi_j(p) \cong \phi_j(p_j) + (v_j T - L)(p - p_j) + \frac{m_j^2 T}{2E_0^3}(p - p_j)^2, \tag{6.178}$$

where (6.93) has been used, i.e., $R^{33} = 1 - v_0^2$, and $v_0 = p_0/E_0$. Regarding (6.176), Laplace's method can be applied if

$$|v_j T - L|\, 2\sigma_{p\,\mathrm{eff}} \lesssim 1; \tag{6.179}$$

$$\frac{m_j^2 T}{2E_0^3} 4\sigma_{p\,\mathrm{eff}}^2 \lesssim 1. \tag{6.180}$$

As already shown, the first condition is automatically satisfied if the second one is met. Thus from (6.180), it is understood that if

$$T \lesssim \frac{E_0^3}{2m_j^2 \sigma_{p\,\mathrm{eff}}^2}, \tag{6.181}$$

then Laplace's method should be considered. We define *dispersion time* T_j^{disp} as

$$T_j^{\mathrm{disp}} = \frac{E_0^3}{2m_j^2 \sigma_{p\,\mathrm{eff}}^2}, \tag{6.182}$$

and consequently, *dispersion length* L_j^{disp} as $L_j^{\mathrm{disp}} = v_0 T_j^{\mathrm{disp}}$. Therefore, the transversal dispersion happens in this distance range:

$$\frac{p_0}{\sigma_p^2} \lesssim L \lesssim L_j^{\mathrm{disp}}. \tag{6.183}$$

The above range is called the *transversal-dispersion regime*. Now, it is enough to put (6.175) and (6.178) into (6.164) and then solve, by applying (6.58), the Gaussian integral. As a result,

$$\begin{aligned}\mathcal{A}_j =& \frac{N g\,(\mathbf{l})\,\sigma_{p\,\mathrm{eff}}}{T\sqrt{1 + iT/T_j^{\mathrm{disp}}}} \times \\ & \exp\left[-iE_j T + ip_j L - \frac{1}{4\sigma_m^2}\left(\frac{\delta m_j^2}{2E_0}\right)^2 - \frac{1}{1 + iT/T_j^{\mathrm{disp}}} \frac{(v_j T - L)^2}{4\sigma_{x\,\mathrm{eff}}^2}\right].\end{aligned} \tag{6.184}$$

For an amplitude significantly different from zero, the terms inside the exponential must be smaller than one. As for the fourth term, we see that condition (6.179) must be satisfied if condition (6.180) is met.

6.7.3 The transition probability

To find the oscillation probability, one should time integrate the amplitude squared. We know that $\delta m_{ij}^2/2E_0 \ll 1$. This condition is satisfied if

1. masses m_i and m_j are roughly degenerate, $|m_i - m_j| \ll m_i, m_j$.
2. The masses are so small (with respect to E_0), in other words, m_i and m_j must be relativistic.

Nearly degenerate masses. As for the first option, we can use the approximation $m_i \approx m_j \approx \tilde{m}_0$ ($\tilde{m}_0$ is the mass in the degenerate limit) in the denominator of (6.182). So, it is expected to have nearly one dispersion time for each mass eigenstate, i.e.,

$$\begin{aligned} T_i^{\text{disp}} \approxeq & T_j^{\text{disp}} = T^{\text{disp}} = \frac{E_0^3}{2\tilde{m}_0^2 \sigma_{p\,\text{eff}}^2} \Rightarrow \\ \mathcal{A}_i \mathcal{A}_j^* = & \frac{N^2 g^2 \left(\mathbf{1}\right) \sigma_{p\,\text{eff}}^2}{T^2 \sqrt{1 + \left(T/T^{\text{disp}}\right)^2}} \times \\ & \exp\left[-i\phi_{ij}\left(T, L\right) - \frac{\left(\delta m_i^2\right)^2 + \left(\delta m_j^2\right)^2}{16\sigma_m^2 E_0^2} - f_{ij}\left(T, L\right)\right], \end{aligned} \tag{6.185}$$

where

$$\begin{aligned} \phi_{ij}\left(T, L\right) = & \left(E_i - E_j\right) T - \left(p_i - p_j\right) L - \\ & \frac{T}{T^{\text{disp}}} \frac{1}{1 + \left(T/T^{\text{disp}}\right)^2} \frac{\left(v_i T - L\right)^2 - \left(v_j T - L\right)^2}{4\sigma_{x\,\text{eff}}^2}, \end{aligned} \tag{6.186}$$

and

$$f_{ij}\left(T, L\right) = \frac{1}{1 + \left(T/T^{\text{disp}}\right)^2} \frac{\left(v_i T - L\right)^2 + \left(v_j T - L\right)^2}{4\sigma_{x\,\text{eff}}^2}. \tag{6.187}$$

Now, we should calculate

$$\int \mathrm{d}T \mathrm{e}^{-i\phi_{ij}(T,L)} \mathrm{e}^{-f_{ij}(T,L)}. \tag{6.188}$$

This integral can be solved by Laplace's method. To do so, we should expand $f_{ij}\left(T, L\right)$ and $\phi_{ij}\left(T, L\right)$ around the point T_{ij}, where $f_{ij}\left(T, L\right)$ is minimum. Same as before, T_{ij} is derived to order $\mathcal{O}\left(\epsilon\right)$, $f_{ij}\left(T, L\right)$ to order $\mathcal{O}\left(\epsilon^2\right)$, the first derivative to order $\mathcal{O}\left(\epsilon\right)$ and the second derivatives to order $\mathcal{O}\left(\epsilon^0\right)$. Let us find T_{ij},

$$\begin{aligned} & \frac{\mathrm{d}f_{ij}\left(T, L\right)}{\mathrm{d}T} = 0 \Rightarrow \\ & T_{ij} = \frac{L}{v_0}\left(1 - \frac{v_i + v_j - 2v_0}{2v_0}\right) + \mathcal{O}\left(\epsilon^2\right) \Rightarrow \end{aligned} \tag{6.189}$$

$$f_{ij}\left(T_{ij}, L\right) = \frac{L^2}{1 + l^2} \frac{\left(v_i - v_j\right)^2}{8v_0^2 \sigma_{x\,\text{eff}}^2}, \tag{6.190}$$

where

$$l = \frac{L}{v_0 T^{\text{disp}}}. \tag{6.191}$$

In (6.169), p_j to $\mathcal{O}(\epsilon)$ was found. Now we would like to find E_j and v_j to $\mathcal{O}(\epsilon)$,

$$E_j = \sqrt{p_j^2 + m_j^2} = E_0 + \rho \frac{\delta m_j^2}{2E_0}, \tag{6.192}$$

$$v_j = \frac{p_j}{E_j} = v_0 + \left(\rho\left(1 - v_0^2\right) - 1\right) \frac{\delta m_j^2}{2p_0 E_0}. \tag{6.193}$$

where the parameter ρ was defined in (6.170). In view of (6.193), (6.190) becomes

$$\begin{aligned} f_{ij}(T_{ij}, L) &= \frac{L^2}{1+l^2} \frac{(v_i - v_j)^2}{8v_0^2 \sigma_{x\,\mathrm{eff}}^2} \\ &= \frac{E_0^4}{8\sigma_{x\,\mathrm{eff}}^2 \tilde{m}_0^4} \frac{l^2}{1+l^2} \left(\rho \frac{\tilde{m}_0^2}{E_0^2} - 1\right)^2 \left(\frac{\delta m_{ij}^2}{2p_0}\right)^2 + \mathcal{O}\left(\epsilon^3\right), \end{aligned} \tag{6.194}$$

where we used the approximation

$$1 - v_0^2 \approxeq \tilde{m}_0^2 / E_0^2. \tag{6.195}$$

The second derivative of $f_{ij}(T, L)$ at T_{ij} is

$$\frac{1}{2} \frac{\mathrm{d}^2 f_{ij}}{\mathrm{d}T^2}(T_{ij}, L) = \frac{2v_0^2 \sigma_{x\,\mathrm{eff}}^2}{1+l^2} + \mathcal{O}(\epsilon). \tag{6.196}$$

Using (6.169), (6.189) and (6.192), (6.186) reads

$$\phi_{ij}(T_{ij}, L) = \frac{\delta m_{ij}^2}{2p_0} L + \mathcal{O}\left(\epsilon^2\right), \tag{6.197}$$

and the first derivative becomes

$$\begin{aligned} \frac{\mathrm{d}\phi_{ij}}{\mathrm{d}T}(T_{ij}, L) &= E_i - E_j - \frac{2l}{1+l^2} \sigma_{p\,\mathrm{eff}}^2 (v_i - v_j) L \\ &= \frac{\rho\left(\tilde{m}_0^2 / E_0^2\right) + l^2}{1+l^2} \frac{E_0 \delta m_{ij}^2}{2\tilde{m}_0^2} + \mathcal{O}\left(\epsilon^2\right), \end{aligned} \tag{6.198}$$

where in the second line, (6.192), (6.193) and (6.195) have been used. The second derivative of the phase does not contribute to the transition amplitude because it is of order ϵ. By putting (6.194), (6.196), (6.197) and (6.198) into (6.188), we obtain a Gaussian integral. By applying (6.58), it is derived that

$$\begin{aligned} &\int \mathrm{d}T \mathcal{A}_i \mathcal{A}_i^* = v_0 N^2 \sigma_{p\,\mathrm{eff}}^2 \frac{g^2(\mathbf{l})}{L^2} \times \\ &\exp\left(-2\pi i \frac{L}{L_{ij}^{\mathrm{osc}}} - \frac{\left(\delta m_i^2\right)^2 + \left(\delta m_j^2\right)^2}{16\sigma_m^2 E_0^2} - 2\pi^2 \left(\frac{\rho \sigma_{x\,\mathrm{eff}}^2}{L_{ij}^{\mathrm{osc}}}\right)^2 - \left(\frac{L}{L_{ij}^{\mathrm{coh}}}\right)^2\right), \end{aligned} \tag{6.199}$$

with oscillation length $L_{ij}^{\rm osc}$

$$L_{ij}^{\rm osc} = \frac{4\pi p_0}{\delta m_{ij}^2}, \tag{6.200}$$

and coherence length $L_{ij}^{\rm coh}$

$$L_{ij}^{\rm coh} = \frac{1}{\sqrt{2}\pi}\frac{p_0}{\sigma_{p\,\rm eff}} L_{ij}^{\rm osc}. \tag{6.201}$$

Note. Same as the exponential factors, we should expand the factor

$$\frac{1}{T^2\sqrt{1+\left(T/T^{\rm disp}\right)^2}}$$

appearing in (6.185) around T_{ij}. Since in this regime time is bigger than E_0/σ_p^2, then we just keep the term $1/T_{ij}^2$ and ignore the others. Because T_{ij} is in the denominator, from (6.189) we write

$$\frac{1}{T_{ij}} = \frac{v_0}{L}. \tag{6.202}$$

This is the reason of $\left(1/L^2\right)$ in (6.199). This prefactor comes from the spreading of the wave packets. As for the probability in the no-dispersion regime (i.e., (6.153)), there is not this prefactor. As a result, going further from the no-dispersion regime brings us smaller amplitudes. So, we have bigger amplitudes when

$$l = \frac{L}{L^{dis}} \ll 1. \tag{6.203}$$

Furthermore, it should be considered that in computing (6.199), the approximation $\Delta T \to \infty$ is good if ΔT is larger than the width of the overlap function[14], i.e.,

$$v_0 \Delta T \gtrsim \sigma_{x\,\rm eff}. \tag{6.204}$$

Relativistic masses. Now after discussing the nearly degenerate masses, let us move to the second option, i.e., relativistic neutrinos with very different masses[15]. Suppose

$$m_i \gg m_j \Rightarrow T_i^{\rm disp} \ll T_j^{\rm disp}. \tag{6.205}$$

[14] Same as what was completely discussed in (6.138).

[15] If the particles are relativistic, but the masses are nearly degenerate, the result will be (6.199).

In this case, decoherence starts before the dispersion length $L_i^{\rm disp}$, because from (6.182), (6.200) and (6.201) (taking $p_0 \approx E_0$ in the relativistic case) we have

$$\frac{L_{ij}^{\rm coh}}{L_i^{\rm dis}} \cong 4\sqrt{2}\frac{m_i^2\sigma_{p\,\rm eff}}{\Delta m_{ij}^2 E_0} \Rightarrow \quad \text{if } L_{ij}^{\rm coh} \lesssim L_i^{\rm dis} \Rightarrow$$
$$\frac{\delta m_{ij}^2}{m_i^2} \gtrsim \frac{\sigma_{p\,\rm eff}}{E_0}. \tag{6.206}$$

Since $m_i \gg m_j$, then the above condition will be satisfied because $\delta m_{ij}^2/m_i^2 \cong 1$. Hence, decoherence begins before $L_i^{\rm disp}$ and oscillations vanish before the smallest dispersion length. The fact which means that in the relativistic masses case, we are left with (6.199) too. So, in the transversal-dispersion regime we have

$$\frac{p_0}{\sigma_p^2} \lesssim L \lesssim \min\left(L_i^{\rm disp}, L_j^{\rm disp}\right). \tag{6.207}$$

With respect to (6.32), the flavor-mixing transition probability in the transversal-dispersion regime becomes

$$\mathcal{P}_{\alpha\to\beta}\left(\mathbf{L}\right) \sim v_0 N^2 \sigma_{p\,\rm eff}\frac{g^2\left(\mathbf{l}\right)}{L^2}\sum_{i,j=1}^{3} U_{i\alpha}U_{j\beta}U_{\beta i}^* U_{\alpha j}^*$$
$$\times \exp\left(-2\pi i\frac{L}{L_{ij}^{\rm osc}} - \frac{\left(\delta m_i^2\right)^2 + \left(\delta m_j^2\right)^2}{16\sigma_m^2 E_0^2} - 2\pi^2\left(\frac{\rho\sigma_{x\,\rm eff}^2}{L_{ij}^{\rm osc}}\right)^2 - \left(\frac{L}{L_{ij}^{\rm coh}}\right)^2\right), \tag{6.208}$$

where N is a normalization constant which can be obtained in the same way as (6.155). From the above probability, it is seen that

$$\rho\frac{\sigma_{x\,\rm eff}}{L_{ij}^{\rm ocs}} \lesssim 1, \qquad \text{and} \qquad \frac{L}{L_{ij}^{\rm coh}} \lesssim 1. \tag{6.209}$$

We shall speak about the above relations more.

6.7.4 Exponential terms of the probability

Coherence length. The factor $\exp\left(\left(L/L_{ij}^{\rm coh}\right)^2\right)$ appearing in (6.208) is a damping factor. In other words, oscillations happen only if $L \lesssim L_{ij}^{\rm coh}$. This is the reason why $L_{ij}^{\rm coh}$ is called the coherence length. Let us elaborate more on it.

The fact that after some length oscillations vanish comes from the different group velocities of the mass eigenstates, i.e., v_i and v_j. By this we mean that if $v_i \neq v_j$, so it is expected that after some length these two wave packets cross each other and so the interferences between them vanish[16]. This shows that the

[16]This is exactly the case which was discussed in quantum mechanical oscillations too, see Fig. (5.1).

origin of this decoherence comes from function (6.190) (overlap function) and the first derivative of the phase, i.e., (6.198).

In this section we are in the transversal-dispersion regime, so we discuss[17] $L_{ij}^{\mathrm{coh}} < \min\left(L_i^{\mathrm{dis}}, L_j^{\mathrm{dis}}\right)$. If $l \ll 1$, the main origin of the decoherence comes from the overlap function rather than (6.198).

Here, the coherence length has been discussed in configuration space. Now, let us speak about this length from another point of view, i.e., in energy-momentum space. If we do a coherent measurement in time at detector or in other words when $\sigma_{eD} \to 0$, then (from (6.171)) $\sigma_{p\,\mathrm{eff}} \to 0$.

This means that $L_{ij}^{\mathrm{coh}} \to \infty$ (see (6.201)). This result is obtained while two mass eigenstates could progressively separate spatially. What does it mean? It means that

a long coherent measurement in time can revive oscillations[18] [38].

Oscillation phase. It is worth stating that under special conditions, the transition probability (6.208) reduces to the standard equation (5.34) with the exception of the $1/L^2$ prefactor.

To show it, we examine each exponential term of (6.208). Let us start the proof from the phase (6.186) in the limit $T \ll T^{\mathrm{dis}}$. As a result, (6.186) becomes

$$\phi_{ij}\left(T, L\right) \cong \left(E_i - E_j\right) T - \left(p_i - p_j\right) L, \tag{6.210}$$

and by using (6.169) and (6.192), (6.210) reads

$$\phi_{ij}\left(T, L\right) \cong 2\pi \frac{L}{L_{ij}^{\mathrm{osc}}} + 2\pi\rho \frac{v_0 T - L}{L_{ij}^{\mathrm{osc}}}. \tag{6.211}$$

By looking at the fourth term inside the exponential (6.184), it is understood that for a highly non-zero amplitude we should have $v_j T - L \lesssim \sigma_{x\,\mathrm{eff}}$. Consequently, $v_0 T - L \lesssim \sigma_{x\,\mathrm{eff}}$. So, for the second term in (6.211) we have

$$\rho \frac{v_0 T - L}{L_{ij}^{\mathrm{osc}}} \lesssim \rho \frac{\sigma_{x\,\mathrm{eff}}}{L_{ij}^{\mathrm{osc}}}. \tag{6.212}$$

If (see (6.209))

$$\rho \frac{\sigma_{x\,\mathrm{eff}}}{L_{ij}^{\mathrm{ocs}}} \ll 1, \tag{6.213}$$

then we have

$$\rho \frac{v_0 T - L}{L_{ij}^{\mathrm{osc}}} \ll 1 \Rightarrow \phi_{ij}\left(T, L\right) \cong 2\pi \frac{L}{L_{ij}^{\mathrm{osc}}}, \tag{6.214}$$

which is the standard phase (5.31). Phase (6.214) results in oscillations.

[17] In section (6.8) the case $L_{ij}^{\mathrm{coh}} > L^{\mathrm{disp}}$ will be regarded.

[18] For more information about increasing the coherence length by accurate energy measurements at the detector, see Ref. [38].

Localization. According to what was said in section 5.4, if the oscillation length becomes smaller than the spatial uncertainty, oscillations vanish. Remark that spatial uncertainty is $\sigma_{x\,\mathrm{eff}}$ (the width of the overlap function), so we must have

$$L_{ij}^{\mathrm{osc}} \gtrsim \sigma_{x\,\mathrm{eff}}. \tag{6.215}$$

By using (6.200), we see that (6.215) becomes

$$\frac{\delta m_{ij}^2}{p_0} \lesssim \sigma_{p\,\mathrm{eff}}, \tag{6.216}$$

which says that if momentum measurements are so exact, then there will be no interferences and so oscillations disappear (see (5.60)). This is the reason why (6.215) or (6.216) is an observability constraint.

All of the terms in (6.208) which impose constraint (6.215) are called *localization terms*. To find the first localization term in (6.208), let us write the second exponential term in (6.208) as

$$\frac{\left(\delta m_i^2\right)^2 + \left(\delta m_j^2\right)^2}{16\sigma_m^2 E_0^2} = \frac{\left(\delta m_{ij}^2\right)^2}{32\sigma_m^2 E_0^2} + \frac{\left(\delta m_i^2 + \delta m_j^2\right)^2}{32\sigma_m^2 E_0^2}, \tag{6.217}$$

where the following identity has been used

$$-2\left(\delta m_i^2\right)\left(\delta m_i^2\right) = \left(\delta m_{ij}^2\right)^2 - \left(\delta m_i^2\right)^2 - \left(\delta m_j^2\right)^2.$$

Each term of (6.217) must be smaller than one. By plugging $\sigma_m \sim v_0\sigma_p$ (see the explanation under (6.173)) into the first term of (6.217), a localization term as follows

$$L_{ij}^{\mathrm{osc}} \gtrsim \sigma_x, \tag{6.218}$$

is obtained. The above constraint satisfies (6.215). Because from (6.171), it is seen that $\sigma_{x\,\mathrm{eff}} \geqslant \sigma_x$. The second term of (6.217) does not give any localization term. It is just an energy-momentum conservation term[19].

The second localization term comes from the third term in (6.208), i.e.,

$$L_{ij}^{\mathrm{osc}} \gtrsim |\rho|\,\sigma_{x\,\mathrm{eff}}. \tag{6.219}$$

An important question is whether (6.219) is stronger than (6.218) or not. In other words, is $|\rho|\,\sigma_{x\,\mathrm{eff}} \gg \sigma_x$? From (6.171), it is seen that in the limit of zero energy uncertainty at detection ($\sigma_{eD} \to 0$), we obtain $\sigma_{p\,\mathrm{eff}} \to 0$, and consequently $\sigma_{x\,\mathrm{eff}} \to \infty$. Hence, condition $\sigma_{x\,\mathrm{eff}} \gg \sigma_x$ is satisfied. Maybe $|\rho|\,\sigma_{x\,\mathrm{eff}} \gg \sigma_x$ is also valid, let us check it. From (6.170), we have

$$\lim_{\sigma_{eD}\to 0} |\rho|\,\sigma_{x\,\mathrm{eff}} = \frac{|v_D|}{\sigma_{eD}} \lesssim S_D, \tag{6.220}$$

[19]Here, we are not going to discuss the reason, see Ref. [13].

where S_D is the size of the macroscopic detection region. So in view of (6.219), we have

$$L_{ij}^{\mathrm{osc}} \gtrsim S_D. \tag{6.221}$$

In the limit of zero energy uncertainty at detection, $\sigma_{p\,\mathrm{eff}} \to 0$. Thus from (6.201), we see that L_{ij}^{coh} increases. This means that by accurate energy measurement or in other words long coherent measurement in time, the coherence length can be increased. This is exactly the important fact which was already said. Furthermore, if approximation (6.213) is taken into consideration, then (6.221) becomes

$$L_{ij}^{\mathrm{osc}} \gg S_D. \tag{6.222}$$

Here, it has been shown that under condition (6.213) (or (6.222)), the probability transition (6.208) becomes the standard formula derived by the quantum mechanical approach.

6.8 Longitudinal-dispersion regime

6.8.1 The transition amplitude

When time increases and crosses the dispersion time, integral (6.164) cannot be solved by Laplace's method anymore and instead, the stationary phase method must be applied. As already said, in this approach we should find the point $p_{\mathrm{cl},j}$ where the phase $\phi_j(p)$ is extremal, i.e.,

$$\frac{\mathrm{d}\phi_j(p)}{\mathrm{d}p} = 0 \Rightarrow p_{\mathrm{cl},j} = m_j \frac{v_{\mathrm{cl}}}{\sqrt{1-v_{\mathrm{cl}}^2}}, \tag{6.223}$$

where $v_{\mathrm{cl}} = L/T$. Then, after expanding the phase and function $f_j(p)$ to the second order around $p_{\mathrm{cl},j}$, we are left with a Gaussian integral whose solution is

$$\begin{aligned}\mathcal{A}_j =& \frac{Ng(\mathbf{1})\,\sigma_{p\,\mathrm{eff}}}{T\sqrt{1+iT/T_j^{\mathrm{disp}}}} \times \\ & \exp\left[-im_j\sqrt{T^2-L^2} - f_j(p_{\mathrm{cl},j}) + \sigma_{p\,\mathrm{eff}}^2 \frac{\left(f_j'(p_{\mathrm{cl},j})\right)^2}{1+iT/T_j^{\mathrm{disp}}}\right].\end{aligned} \tag{6.224}$$

The function $f_{jP,D}(p_{\mathrm{cl},j})$ is (6.165) at the point $p_{\mathrm{cl},j}$ and $f'_{jP,D}(p_{\mathrm{cl},j})$ is its first momentum derivative, i.e.,

$$f_{jP,D}(p_{\mathrm{cl},j}) = \frac{(p_{\mathrm{cl},j}-p_0)^2}{4\sigma_{pP,D}^2} + \frac{(E_{\mathrm{cl},j}-E_0-(p_{\mathrm{cl},j}-p_0)\,v_{p,D})^2}{4\sigma_{eP,D}^2}, \tag{6.225}$$

$$f'_{jP,D}(p_{\mathrm{cl},j}) = \frac{p_{\mathrm{cl},j}-p_0}{2\sigma_{pP,D}^2} + (v_{\mathrm{cl}}-v_{P,D})\,\frac{E_{\mathrm{cl},j}-E_0-(p_{\mathrm{cl},j}-p_0)\,v_{p,D}}{2\sigma_{eP,D}^2}, \tag{6.226}$$

with

$$E_{\mathrm{cl},j} = \sqrt{p_{\mathrm{cl},j}^2 + m_j^2} = \frac{m_j T}{\sqrt{T^2 - L^2}}. \tag{6.227}$$

In the stationary phase method, the overlap function varies much more slowly than the phase, i.e., we must have $|p_{\mathrm{cl},j} - p_0| \lesssim \sigma_{pP,D}$ (see (6.225)). Similarly for the mass eigenstate m_i, we will have $|p_{\mathrm{cl},i} - p_0| \lesssim \sigma_{pP,D}$. Then

$$\left.\begin{array}{l} |p_{\mathrm{cl},j} - p_0| \lesssim \sigma_{pP,D} \\ |p_{\mathrm{cl},i} - p_0| \lesssim \sigma_{pP,D} \end{array}\right\} \Rightarrow |p_{\mathrm{cl},i} - p_{\mathrm{cl},j}| \lesssim \sigma_{pP,D}. \tag{6.228}$$

As a result of the above condition, with respect to (6.223), we have

$$\frac{|\delta m_{ij}|}{m} \lesssim \frac{\sigma_{pP,D}}{p},$$

where m and p are defined by $p = mL\left(T^2 - L^2\right)^{-1/2}$. Therefore in view of the above relation, if masses are nearly degenerate, then the interference $\mathcal{A}_i \mathcal{A}_j^*$ becomes non-zero. So, we can define only one dispersion time T^{disp}, i.e., $T_i^{\mathrm{disp}} \approxeq T_i^{\mathrm{disp}} \approxeq T^{\mathrm{disp}}$. This result was predictable, since in the previous section it was seen that if masses were not degenerate, decoherence would begin before L_i^{disp}, i.e., when $L \gtrsim L_j^{\mathrm{disp}}$, due to this decoherence we would have zero amplitude.

6.8.2 The transition probability

Now, to find the probability we should integrate the amplitude squared over time, i.e.,

$$\begin{aligned} &\int \mathrm{d}T \mathcal{A}_i \mathcal{A}_j^* = \\ &\int \mathrm{d}T \frac{N^2 g^2 \left(\mathbf{1}\right) \sigma_{p\,\mathrm{eff}}^2}{T^2 \sqrt{1 + \left(T/T^{\mathrm{disp}}\right)^2}} \exp\left(-i\tilde{\phi}_{ij}\left(T, L\right) - \tilde{f}_{ij}\left(T, L\right)\right), \end{aligned} \tag{6.229}$$

with the phase

$$\tilde{\phi}_{ij}\left(T, L\right) = \delta m_{ij} \sqrt{T^2 - L^2} + \sigma_{p\,\mathrm{eff}}^2 \frac{T}{T^{\mathrm{disp}}} \frac{\left(f_i'\left(p_{\mathrm{cl},i}\right)\right)^2 - \left(f_j'\left(p_{\mathrm{cl},j}\right)\right)^2}{1 + \left(T/T^{\mathrm{disp}}\right)^2}, \tag{6.230}$$

where $\delta m_{ij} = m_i - m_j$. The function $\tilde{f}_{ij}\left(T, L\right)$ is also defined as

$$\tilde{f}_{ij}\left(T, L\right) = f_i\left(p_{\mathrm{cl},i}\right) + f_j\left(p_{\mathrm{cl},j}\right) - \sigma_{p\,\mathrm{eff}}^2 \frac{\left(f_i'\left(p_{\mathrm{cl},i}\right)\right)^2 + \left(f_j'\left(p_{\mathrm{cl},j}\right)\right)^2}{1 + \left(T/T^{\mathrm{disp}}\right)^2}. \tag{6.231}$$

The above integral is going to be solved with Laplace's method, however, there is a big difference. Here, we have m_j and not m_j^2. As it was explained very

early, $\delta m_j = m_j - m_0$ is not necessarily very small (in the relativistic case). This is the reason why we have been working with the mass squared. Therefore, expansion around m_0 will not work anymore and a new expansion parameter, i.e., $\tilde{m}_0$, is defined as follows

$$\tilde{m}_0 = (m_i + m_j)/2, \tag{6.232}$$

and

$$\delta\tilde{m}_j = m_j - \tilde{m}_0. \tag{6.233}$$

The parameter $\delta\tilde{m}_j$ is of order ϵ. It is clear that due to (6.232), we can have

$$\delta m_{ij} = \frac{\delta m_{ij}^2}{2\tilde{m}_0} \Rightarrow \tag{6.234}$$

$$\tilde{m}_0^2 (\delta m_{ij})^2 = \frac{1}{4}\left(\delta m_i^2 - \delta m_j^2\right)^2, \tag{6.235}$$

and

$$\left(\delta\tilde{m}_0^2\right)^2 = \frac{1}{4}\left(\delta m_i^2 + \delta m_j^2\right)^2. \tag{6.236}$$

So from (6.235) and (6.236), it is obtained that

$$2\tilde{m}_0^2 (\delta m_{ij})^2 + 2\left(\delta\tilde{m}_0^2\right)^2 = \left(\delta m_i^2\right)^2 + \left(\delta m_j^2\right)^2. \tag{6.237}$$

We shall use from (6.234) and (6.237) later on.

Now same as before, first the point $\tilde{T}_0$, where $\tilde{f}_{ij}(T, L)$ becomes minimum should be found. This point cannot be calculated exactly. However, by taking the first time derivative from (6.231), we see if $m_i = m_j = \tilde{m}_0$, then the solution would be any T satisfying $p_{\mathrm{cl},j} = p_0$.

In the case of nearly degenerate masses, it is expected that the solution will be any T satisfying $p_{\mathrm{cl},j} \approxeq p_0$. As a result, $\tilde{T}_0$ will be equal to a term of order $(p_{\mathrm{cl},j} - p_0)^0$ (i.e., when $p_{\mathrm{cl},j} = p_0$), plus a term of order $(p_{\mathrm{cl},j} - p_0)$ and the other terms of higher orders are neglected. To find $\tilde{T}_0$ (to order ϵ), let us expand time in (6.227) around $p_{\mathrm{cl},j} = p_0$, and $m_j^2 = \tilde{m}_0^2$. Hence,

$$\begin{aligned} T &= L\frac{\sqrt{p_{\mathrm{cl},j}^2 + m_j^2}}{p_{\mathrm{cl},j}} \\ &\approxeq \frac{\tilde{E}_0 L}{p_0} - L\frac{\tilde{m}_0^2}{p_0^2 E_0}(p_{\mathrm{cl},j} - p_0) + L\frac{1}{2p_0 E_0}\left(m_j^2 - \tilde{m}_0^2\right), \end{aligned} \tag{6.238}$$

where $\tilde{E}_0 = \sqrt{p_0^2 + \tilde{m}_0^2}$. In the denominators, $\tilde{E}_0$ has been approximated by E_0. The first term in (6.238) satisfies $p_{\mathrm{cl},j} = p_0$, and the other terms are perturbations around p_0 and $\tilde{m}_0$. Accordingly, T in (6.238) is the point that we are looking for, i.e., $\tilde{T}_0$.

Since $\tilde{m}_0 = (m_i + m_j)/2$, in the case of the nearly degenerate masses, we may write

$$\tilde{m}_0^2 = \tilde{m}_0 \frac{m_i + m_j}{2} = \frac{\tilde{m}_0 m_i}{2} + \frac{\tilde{m}_0 m_j}{2} \approx \tilde{m}_0 m_j. \tag{6.239}$$

Then, let us write the term $m_j^2 - \tilde{m}_0^2$ in (6.238) as follows,

$$m_j^2 - \tilde{m}_0^2 = m_j^2 - 2\tilde{m}_0^2 + \tilde{m}_0^2 \approx \underbrace{(m_j - \tilde{m}_0)^2}_{(\delta \tilde{m}_j)^2}. \tag{6.240}$$

Therefore, to order ϵ, the last term in (6.238) vanishes and we are left with

$$\tilde{T}_0 \approxeq \frac{\tilde{E}_0 L}{p_0} - L \frac{\tilde{m}_0^2}{p_0^2 E_0} (p_{\mathrm{cl},j} - p_0). \tag{6.241}$$

We know that $|p_{\mathrm{cl},j} - p_0| \lesssim \sigma_{pP,D}$. Additionally, it was shown that $|p_j - p_0| \lesssim \sigma_{p\mathrm{eff}}$, where $\sigma_{p\mathrm{eff}} \leqslant \sigma_{pP,D}$. Regarding this, we can have $|p_{\mathrm{cl},j} - p_0| \sim |p_j - p_0|$, and so from (6.169) and (6.240), it is obtained that

$$\begin{aligned} |p_{\mathrm{cl},j} - p_0| &= (\rho - 1) \frac{m_j^2 - m_0^2}{2p_0} = (\rho - 1) \frac{m_j^2 - \tilde{m}_0^2 + \tilde{m}_0^2 - m_0^2}{2p_0} \\ &\approxeq (\rho - 1) \frac{\delta \tilde{m}_0^2}{2p_0} + \mathcal{O}\left(\epsilon^2\right), \end{aligned} \tag{6.242}$$

where $\delta \tilde{m}_0^2 = \tilde{m}_0^2 - m_0^2$. So, (6.241) becomes

$$\tilde{T}_0 \approxeq \frac{\tilde{E}_0 L}{p_0} - \frac{\tilde{m}_0^2 \delta \tilde{m}_0^2}{2p_0^3 E_0} (1 - \rho) L + \mathcal{O}\left(\epsilon^2\right). \tag{6.243}$$

Now the rest of the process is same as before, i.e., $\tilde{f}_{ij}\left(\tilde{T}_0, L\right)$ should be found to order ϵ^2:

$$\begin{aligned} \tilde{f}_{ij}\left(\tilde{T}_0, L\right) = &\frac{\tilde{m}_0^2 (\delta m_{ij})^2 + \left(\delta \tilde{m}_0^2\right)^2}{8\sigma_m^2 E_0^2} \\ &+ \frac{E_0^4}{8\tilde{m}_0^4 \sigma_{p\mathrm{eff}}^2} \frac{l^2}{1 + l^2} \left(\rho \frac{\tilde{m}_0^2}{E_0^2} - 1\right)^2 \left(\frac{\tilde{m}_0 \delta m_{ij}}{p_0}\right)^2 + \mathcal{O}\left(\epsilon^3\right), \end{aligned} \tag{6.244}$$

where l and σ_m are defined by (6.191) and (6.173), respectively. By applying (6.234) and (6.237), the above equation becomes

$$\begin{aligned} \tilde{f}_{ij}\left(\tilde{T}_0, L\right) = &\frac{\left(\delta m_i^2\right)^2 + \left(\delta m_j^2\right)^2}{16\sigma_m^2 E_0^2} \\ &+ \frac{E_0^4}{8\tilde{m}_0^4 \sigma_{p\mathrm{eff}}^2} \frac{l^2}{1 + l^2} \left(\rho \frac{\tilde{m}_0^2}{E_0^2} - 1\right)^2 \left(\frac{\delta m_{ij}^2}{2p_0}\right)^2 + \mathcal{O}\left(\epsilon^3\right). \end{aligned} \tag{6.245}$$

It is seen that the first term is $f_i(p_i)+f_j(p_j)$ (see (6.172)) and the second term is (6.194). The second $T-$derivative of $\tilde{f}_{ij}(T,L)$ at $\tilde{T}_0$ becomes

$$\frac{1}{2}\frac{\mathrm{d}^2\tilde{f}_{ij}}{\mathrm{d}T^2}\left(\tilde{T}_0,L\right)=\frac{2v_0^2\sigma_{p\,\mathrm{eff}}^2}{1+l^2}+\mathcal{O}(\epsilon), \tag{6.246}$$

which is exactly same as (6.196). The phase (6.230) at $\tilde{T}_0$, to order ϵ, is

$$\tilde{\phi}_{ij}\left(\tilde{T}_0,L\right)=\frac{\delta m_{ij}^2}{2p_0}L+\mathcal{O}\left(\epsilon^2\right), \tag{6.247}$$

which is (6.197). The first derivative of $\tilde{\phi}_{ij}(T,L)$ at $\tilde{T}_0$ is

$$\frac{\mathrm{d}\tilde{\phi}_{ij}}{\mathrm{d}T}\left(\tilde{T}_0,L\right)=\frac{\rho\left(\tilde{m}_0^2/E_0^2\right)+l^2}{1+l^2}\frac{E_0\delta m_{ij}}{\tilde{m}_0}+\mathcal{O}\left(\epsilon^2\right), \tag{6.248}$$

where by applying (6.234), the above relation becomes (6.198). Hence, the derived equations for the longitudinal-dispersion regime is exactly same as the transversal-dispersion regime. This means that the transition probability for $L\gtrsim L^{\mathrm{disp}}$ is also (6.208).

6.8.3 Coherence length and temporal ranges

Regarding (6.246), the spatial width of the oscillating particle becomes

$$\textit{spatial width}=\sigma_{x\,\mathrm{eff}}\sqrt{1+l^2}\cong l\sigma_{x\,\mathrm{eff}}=\frac{\sigma_{p\,\mathrm{eff}}}{p_0}\frac{\tilde{m}_0^2}{E_0^2}L, \tag{6.249}$$

where the approximation has been derived in the limit $l\gg 1$. The time interval ΔT on which we take the Gaussian integral must go to infinity. To satisfy this condition, it is supposed that (same as what was did in the previous sections) $v_0\Delta T$ must be bigger than the spatial width (6.249), i.e.,

$$v_0\Delta T\gtrsim l\sigma_{x\,\mathrm{eff}}. \tag{6.250}$$

Let spatial width be bigger than $L_{ij}^{\mathrm{osc}}\gamma^{-2}$ ($\gamma=E_0/\tilde{m}_0$ is the Lorentz factor), then from (6.249) we have

$$l\sigma_{x\,\mathrm{eff}}\gtrsim\frac{\tilde{m}_0^2}{E_0^2}L_{ij}^{\mathrm{osc}}\Rightarrow\frac{\sigma_{p\,\mathrm{eff}}}{p_0}\frac{L}{L_{ij}^{\mathrm{osc}}}\gtrsim 1\Rightarrow\frac{L}{L_{ij}^{\mathrm{coh}}}\gtrsim 1, \tag{6.251}$$

where we have used (6.201). The above result states that if the coherence length is larger than the dispersion length, decoherence starts if the spatial width becomes bigger than $\gamma^{-2}L_{ij}^{\mathrm{osc}}$.

It is important that contrary to the case where $L_{ij}^{\mathrm{coh}}\lesssim L^{disp}$ (already discussed), in the case $L_{ij}^{\mathrm{coh}}\gtrsim L^{disp}$ (for $l\gg 1$), we are not going to have any decoherence due to the different group velocities. This fact can be seen easily from (6.194) and (6.198) that in the case $l\gg 1$, the terms containing group

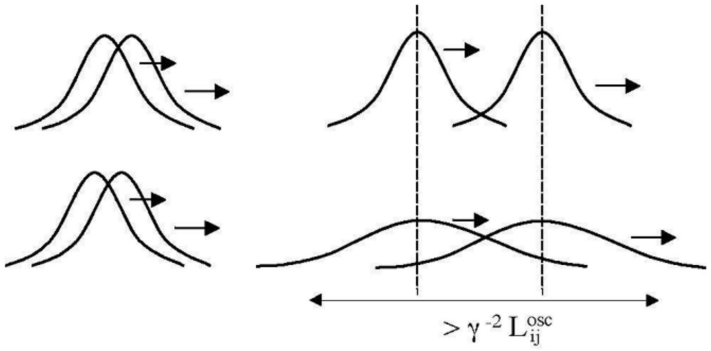


Figure 6.4: The first drawing refers to the decoherence that we spoke about in the transversal-dispersion regime or in quantum mechanical approach, i.e., different group velocities lead to the decoherence. In the longitudinal-dispersion regime, decoherence does not happen anymore due to the different group velocities, but, it happens when the spatial width becomes bigger than $\gamma^2 \mathrm{L}_{ij}^{\mathrm{osc}}$ (the second drawing). The drawings have been taken from Ref. [13].

velocities become very small[20] (see Fig. (6.4)).

Let us come back to (6.250). By using (6.249), we obtain

$$\Delta T \gtrsim \frac{\sigma_{p\,\mathrm{eff}}}{p_0} \frac{L}{L_{ij}^{\mathrm{osc}}} \gamma^{-2} T_{ij}^{\mathrm{osc}}, \tag{6.252}$$

where $v_0 T_{ij}^{\mathrm{osc}} = L_{ij}^{\mathrm{osc}}$. In (6.251), it was proved that if the spatial width becomes bigger than $L_{ij}^{\mathrm{osc}} \gamma^{-2}$, then $(\sigma_{p\,\mathrm{eff}}/p_0)\left(L/L_{ij}^{\mathrm{osc}}\right) \gtrsim 1$. In the case of very large distances, i.e., $L \gg L_{ij}^{\mathrm{coh}}$, although oscillations vanish, condition (6.252) or (6.250) violates. Therefore, instead of (6.250), it is enough to assume that

$$\Delta T \gtrsim \gamma^{-2} T^{\mathrm{osc}}. \tag{6.253}$$

Until now we have found three ranges for ΔT, i.e., no-dispersion regime (6.138), transversal-dispersion regime (6.204) and longitudinal-dispersion regime (6.253). Putting together these three conditions (with $\tilde{\sigma}_{x\,\mathrm{eff}} \sim \sigma_{x\,\mathrm{eff}}$) we see that

$$\Delta T \gtrsim \max\left(\frac{\sigma_{x\,\mathrm{eff}}}{v_0}, \frac{T_{ij}^{\mathrm{osc}}}{\gamma^2}\right). \tag{6.254}$$

6.8.4 Longitudinal-dispersion phase vs standard phase

Finally, let us show that in the limit $T \gg T^{\mathrm{dis}}$, phase (6.230) is the standard oscillation phase (5.29) or (5.31). From (6.230), we have

$$\tilde{\phi}_{ij}(T, L) \cong \delta m_{ij} \sqrt{T^2 - L^2} \qquad \text{for} \quad T \gg T^{\mathrm{dis}} \;\; (l \gg 1). \tag{6.255}$$

[20] Although (6.194) and (6.198) were derived in the transversal-dispersion regime, it was shown that they are still valid in the longitudinal-dispersion regime.

Now by expanding the phase around $\tilde{T}_0$ and using (6.247) and (6.248) (in the limit $l \gg 1$), we have

$$\tilde{\phi}_{ij}\left(T, L\right) \approxeq 2\pi \frac{L}{L_{ij}^{\text{osc}}} + \frac{E_0 \delta m_{ij}}{\tilde{m}_0}\left(T - \tilde{T}_0\right). \tag{6.256}$$

The time range $T - \tilde{T}_0$ is constrained by ΔT, i.e., $\left|T - \tilde{T}_0\right| \lesssim \Delta T$. So, with respect to (6.252) and using (6.200), (6.201) and (6.234), the second term of (6.256) becomes

$$\frac{E_0 \delta m_{ij}}{\tilde{m}_0}\left(T - \tilde{T}_0\right) \lesssim \frac{L}{L_{ij}^{\text{coh}}}. \tag{6.257}$$

If (see (6.209))

$$\frac{L}{L_{ij}^{\text{coh}}} \ll 1, \tag{6.258}$$

then the second term in (6.256) is negligible and the phase becomes the standard phase. In general, it can be said that if (6.258) and (6.222) are satisfied, then (6.208) becomes the standard transition probability (5.34) with the exception $1/L^2$.

We have not shown that the probability (6.153) derived in the no-dispersion regime reduces to the standard formula. However with the similar calculations, it can be shown that if $|\mathbf{v}_0 \times \mathbf{L}| \lesssim |\mathbf{v}_0|\, \sigma_x$ and (6.222) are satisfied [13], then (6.153) reduces to the standard formula. In conclusion, under especial conditions, the probabilities derived in QFT reduce to the standard probability.

Chapter 7

Conclusions

In this thesis we discussed neutrino oscillations. We started to review the physics of neutrino before and after the SM. It was explained that as a result of the SM, there is no constraint on the masses of fermions contrary to bosons which become massive after symmetry breaking and interacting with the Higgs boson. As for neutrinos, it was observed that they appear in three different flavors and according to the SM they must be massless.

On the other hand in view of our observations, there was a problem called the mystery of the missing neutrino. This problem stated that the electron neutrinos created in the Sun did not completely reach us on earth. With the help of the further experiments, it was proved that there are oscillations between different flavors of neutrino, the fact which was called neutrino oscillations phenomenon and leads to the massive neutrino.

We saw that the Lagrangian which is responsible for the neutrino mass generation can be written in two different forms, i.e., the Dirac and Majorana Lagrangians. The difference of these two approaches is that in the Dirac case there are distinct neutrino and antineutrino while in the Majorana case both of them become one particle.

In order to explain theoretically neutrino oscillations, we supposed that the flavor eigenstates are superpositions of the mass eigenstates and relate these two states with the unitary mixing matrices U and V. At the end, these matrices were found in the case of the Dirac and Majorana neutrinos. It was seen that the creation (or annihilation) of neutrinos is described by interaction Lagrangian while propagation Lagrangian is responsible for the mass generation. The difference between these two Lagrangians leads to neutrino oscillations.

Then we started discussing quantum mechanical oscillations with the simplest model, i.e., the plane wave treatment and the oscillation phase and probability were found. However due to some problems such as delocalization of plane waves, the wave packet approach was applied to describe neutrino oscillations. As a result, we derived the coherence length after which oscillations vanish and the fact that there will be no oscillations if the oscillation length becomes smaller than the spatial uncertainty.

In spite of the success of the wave packet approach, there were still some problems. For instance, it is not possible to define the flavor eigenstates. Thus, oscillations were presented in quantum field theory in which neutrinos are created in interactions of some fields as propagators, and they are no more free particles. We solved the transition amplitude integral with Gaussian overlap function with Laplace's and stationary phase method. A time-dependent function, which may lead to dispersion, appeared. Accordingly, the solution of the integral was divided into three: no-, transversal- and longitudinal-dispersion regimes.

Consequently, in each case the probabilities were found and it became clear that the transition probabilities of the transversal- and longitudinal-dispersion regimes are exactly the same. Additionally, same as the QM wave packet approach, here, in QFT we obtained the coherence length too. Finally, it was seen that the oscillation formulas derived in QFT reduced to the standard oscillation relation.

Bibliography

[1] S. Bilenky, *Lecture Notes in Physics 817, Introduction to the Physics of Massive and Mixed Neutrinos*, Springer, Berlin Heidelberg 2010.

[2] A. Bettini, *Introduction to Elementary Particle Physics*, Cambridge University Press, 2008.

[3] D. Griffiths, *Introduction to Elementary Particles*, Wiley-VCH, 2008.

[4] B. R. Martin and G. Shaw, *Particle Physics*, Chichester: Wiley, 1992.

[5] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Addison-Wesley, 1995.

[6] M. K. Gaillard, P. D. Grannis and F. J. Sciulli, *The Standard Model of Particle Physics*, Rev. Mod. Phys. **71** (1999) S96-S111 [arXiv:hep-ph/9812285].

[7] S. F. King, *Neutrino mass*, [arXiv:0712.1750 [physics.pop-ph]].

[8] J. Bahcall, *Solving the Mystery of the Missing Neutrinos*, [arXiv:physics/0406040 [physics.pop-ph]].

[9] The ATLAS Collaboration, *Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC*, Phys. Lett. B **716**, (2012) 1-29 [arXiv:1207.7214 [hep-ex]].

[10] M. Novello, *A mechanism to generate mass: the case of fermions*, [arXiv:1003.5126[physics.gen-ph]].

[11] D. Haidt, *The Discovery of Weak Neutral Currents*, `http://www.desy.de/~haidt/nc30_text.pdf`, 2003.

[12] G. D'Ambrosio and G. Isidori, *CP violation in kaon decays*, Int. J. Mod. Phys. **A 13** (1998) [arXiv:hep-ph/9611284].

[13] M. Beuthe, *Oscillation of neutrinos and mesons in quantum field theory*, Phys. Rept. **375** (2003) 105-218 [arXiv:hep-ph/0109119].

[14] Y. Fukuda et al., [The Super-Kamiokande Collaboration], *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. 81 (1988) 1562-1567 [arXiv:hep-ex/9807003].

[15] B. T. Cleveland et al., *Measurement of the solar electron neutrino flux with Homestake chlorine detector*, Astrophys. J. 496 (1998) 505-526.

[16] K. S. Hirata et al., *Real-time, directional measurement of ^8B solar neutrinos in the Kamiokande II detector*, Phys. Rev. D **44** 2241 (1991).

[17] P. Anselmann et al., [GALLEX Collaboration], *GALLEX solar neutrino observations: complete results for GALLEX II*, Phys. Lett. B **357** (1995) 237-247.

[18] E. W. Otten and C. Weinheimer, *Neutrino mass limit from tritium $\beta-$decay*, Rept. Prog. Phys. **71** (2008) 086201 [arXiv:0909.2104 [hep-ex]].

[19] C. Kraus et al., *Final Results from phase II of the Mainz Neutrino Mass Search in Tritium β Decay*, Eur. Phys. J. C **40** (2005) 447-468 [arXiv:hep-ex/0412056].

[20] H. H. Chen et al., *Direct Approach to Resolve the Solar-Neutrino Problem*, Phys. Rev. Lett. **55**, (1985) 1534-1536.

[21] Q. R. Ahmad et al., *Measurement of the Rate of $\nu_e+d \rightarrow p+p+e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **87** (2001) 071301.

[22] R. Davis et al., *Search for Neutrinos from the Sun*, Phys. Rev. Lett. **20** (1968) 1205.

[23] J. N. Abdurashitov et al., *Measurement of the solar neutrino capture rate with gallium metal*, Phys. Rev. C **60** (1999) 055801 [arXiv:astro-ph/9907113].

[24] Q. R. Ahmad et al., *Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **89** (2002) 011301 [arXiv:nucl-ex/0204008].

[25] G. Danby et al., *Observation of High-energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos*, Phys. Rev. Lett. **9** (1962) 36.

[26] T. D. Lee and C. N. Yang, *Theoretical Discussions on Possible High-energy Neutrino Experiments*, Phys. Rev. Lett. **4** (1960) 307.

[27] S. Frankel et al., *New Limit on the $e + \gamma$ Decay Mode of the Muon*, Phys. Rev. Lett. **8** (1962) 123.

[28] K. Kodama et al. [DONUT Collaboration], *Observation of Tau Neutrino Interactions*, Phys. Lett. B **504** (2001) 218-224 [arXiv:hep-ex/0012035].

[29] C. S. Wu et al., *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105**, 1413 (1957).

[30] C. Giunti and C. W. Kim *Quantum Mechanics of Neutrino Oscillations*, Found. Phys. Lett. **14** (2001) 213-229 [arXiv:hep-ph/0011074].

[31] B. Kayser, *On the quantum mechanics of neutrino oscillation*, Phys. Rev. D**24** (1981) 110.

[32] C. Giunti et al., *Treatment of neutrino oscillations without resort to weak eigenstates*, Phys. Rev. D 48 (1993) 4310 [arXiv:hep-ph/9305276].

[33] C. Giunti and C. W. Kim, *Coherence of neutrino oscillations in the wave packet approach* Phys. Rev. D **58** (1998) 017301 [arXiv:hep-ph/9711363].

[34] C. Giunti et al., *Remarks on the weak states of neutrinos*, Phys. Rev. D**45** (1992) 2414.

[35] C. Giunti et al., *When do neutrinos really oscillate? Quantum mechanics of neutrino oscillations*, Phys. Rev. D **44** (1991) 3635.

[36] D. V. Ahluwalia and T. Goldman, *Interplay of Non-Relativistic and Relativistic Effects in Neutrino Oscillations*, Phys. Rev. D **56** (1997) 1698-1703 [arXiv:hep-ph/9702308].

[37] L. B. Okun and I. S. Tsukerman, *Comment on Equal Velocity Assumption for Neutrino Oscillations*, Mod. Phys. Lett. A **15** (2000) 1481-1482 [arXiv:hep-ph/0007262].

[38] K. Kiers, S. Nussinov and N. Weiss, *Coherence Effects in Neutrino Oscillations*, Phys. Rev. D **53** (1996) 537-547 [arXiv:hep-ph/9506271].

[39] G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nucl. Phys. B **44**, (1972) 189.

[40] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.

[41] Peter W. Higgs, *Broken Symmetry and the Masses of Gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[42] K. A. Olive et al. (Particle Data Group), Chin. Phys. C**38**, 090001 (2014) (URL:`http://pdg.lbl.gov`).

[43] H. J. Lipkin, *Quantum Mechanics of Neutrino Oscillations - Hand Waving for pedestrians*, [arXiv:hep-ph/9901399].

[44] R. G. Winter, *Neutrino Oscillation Kinematics*, Lett. Nuovo Cimento 30 (1981) 101.

[45] I. Yu. Kobzarev et al. Sum rules for neutrino oscillations, Sov. J. Nucl. Phys. **35** (1982) 708 [Yad. Fiz. **35** (1982) 1210].

[46] R. Jacob and R. G. Sachs, Mass and Lifetime of Unstable Particles, Phys. Rev. **121** (1961) 350.

[47] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, Sov. Phys. JETP 7 (1958) 172 [Zh. Eksp. Teor. Fiz. 34 (1957) 247].

[48] G. Feinberg, *Decays of the μ Meson in the Intermediate-Meson Theory*, Phys. Rev. **110** (1958) 1482.

[49] A. Kurup, *Muon to electron conversion: how to find an electron in a muon haystack*, Phil. Trans. R. Soc. A **368** (2010) 3645-3655.

[50] CMS Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30-61, [arXiv:1207.7235 [hep-ex]].