Lorentz-Violating Models of Neutrino Oscillations

Kevin R. Labe
Swarthmore College

March 13, 2011

Contents

Acknowledgments

Thanks to Professor Matthew Mewes for advising me, without whose assistance this project would never have been completed, and to my parents for their many years of encouragement and support. Thanks also to Swarthmore College for funding the research displayed herein from January - August 2010, in part through the Pat Tarble Summer Research Fund.

Chapter 1

Introduction

After the 20^{th}-century triumphs of unification, the two major theories of physics are the Standard Model and General Relativity. The Standard Model describes the interactions (via the strong, weak, and electromagnetic forces) of all subatomic particles, while General Relativity describes the extremely long-range effects of gravity. It has long been understood that these models are incompatible at high energies, leading scientists to believe that these theories are the low energy limits of a more fundamental theory. Furthermore, certain known behaviors in particle physics may already shed light on this problem. One of the few observed effects which cannot be explained by the conventional Standard Model is the phenomenon of neutrino oscillations, in which a beam of neutrinos changes in flavor as it propagates. Neutrinos come in three flavors, and were thought to be massless until flavor-oscillations were discovered. A mass model was then developed and experimentally tested. Most experiments agree about the size of the masses; however, there are a few outliers that cannot be consistently explained. Therefore a new explanation may be in order, and such an explanation would necessarily go beyond the Standard Model. In this thesis, neutrino oscillation is analyzed to determine what it can tell us about physics beyond the Standard Model.

Among the possible candidate theories describing physics beyond the Standard Model is the Standard Model Extension (SME), which considers all possible Lorentz symmetry violating effects consistent with the standard particles. Lorentz symmetry is the frame-invariance of physical law which underlies relativity. We develop a model for neutrino oscillations in the SME framework which, although completely consistent with the conventional mass model over the range so far experimentally probed, nevertheless relies on Lorentz violations instead of mass to produce its effects. This is especially interesting since Lorentz violation effects usually appear only at high energy, whereas in this model the effect would appear at low energy.

Recent observations on neutrinos indicate that antineutrinos may interact differently than neutrinos,

which could be evidence of CPT violation. CPT symmetry (which stands for Charge-Parity-Time) is one of the presumed fundamental symmetries that allows all interactions to occur in reverse if the particles are switched to their antiparticles and everything runs in reversed handedness. CPT symmetry violations entail Lorentz violations, as will later be discussed, so this work is further motivated at present by these experimental hints of antipaticle asymmetry.

Chapter 2

The Standard Model

2.1 Standard Model Overview

In this section we discuss some basic facts about particles and their interactions. For a more detailed overview of this subject, see [19]. The Standard Model is supposed to describe all interactions of matter with the exception of the gravitational force, which is so weak at the scales canonically considered that it can be ignored. The Standard Model identifies 60 fundamental particles which comprise all matter, and sorts these into three categories: leptons, quarks, and mediators. These particles interact via three forces: the strong, weak and electromagnetic force. Although the electromagnetic force is the weakest at an atomic or subatomic level, it is responsible for most of the forces we feel in ordinary interactions (friction, normal forces, light, etc.). The strong and weak forces are responsible for the cohesion and decay of atomic nuclei.

There are six leptons, the most familiar of which is the electron. There are other particles, the muon and the tau, which are identical to the electron in all properties except that they are more massive. These three particles are said to form three "generations" of matter, of increasing mass, a pattern repeated elsewhere in the model. They are represented by the symbols e, μ, τ. There are also three neutrinos, the electron neutrino, muon neutrino and tau neutrino (represented ν_e, ν_μ, ν_τ), which are also identical except in lepton generation (and possibly in mass). All these leptons also have antiparticle partners. Leptons have spin 1/2.

Quarks come in 6 flavors (up, down, top, bottom, charm and strange) and three colors, which, together with their antiparticles come to 36 kinds. Quarks also have spin 1/2, and they combine to form most other particles, including the proton and the neutron. The quarks and leptons together form all massive matter.

The remaining particles mediate the three forces. These are the photon (which mediates the electromagnetic force), eight gluons (which mediate the strong force), and the Z and $W^{\pm}$ bosons (which mediate the

weak force). The 12 mediators have spin 1. Furthermore, there is expected to be at least one kind of Higgs particle, which would give mass to the other particles. There may also be a graviton, though this would stand outside the Standard Model. Although these many particles can be structured in group theoretic terms which lends the theory some unity, 18 parameter values must be selected before the Standard Model can make predictions, a fact many find unacceptably arbitrary in a final theory.

There are three fundamental interactions in the Standard Model, one for each force. These can be represented schematically by their Feynman diagrams as vertices between three particles. There is quark scattering via a gluon, electron scattering via a photon, and fermionic scattering via a Z or W boson. In addition, all interactions must respect the following conservation laws: charge, color, baryon (quark) number, and lepton number must all be conserved. Furthermore, with the exception of neutrino oscillations, the three lepton generations (e, μ, and τ) must be individually conserved.

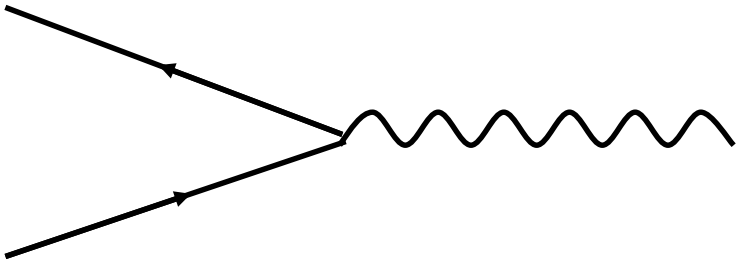

Figure 2.1: Fundamental vertex of a Feynman diagram. Straight lines represent fermions and squiggly lines represent bosons. This could represent the annihilation of an electron and a positron to form a photon. Here time is shown on the horizontal axis and position is shown on the vertical axis. This can be thought of as the two-dimensional projection of the collision as it occurs in four-dimensional space-time, or can be seen as simply schematic.

It is recognized that the Standard Model is inconsistent with General Relativity, and must therefore fail at very high energies. This energy scale is commonly called the Planck scale, and is believed to be on the order of 10^{19} GeV. In spite of this known limitation, the Standard Model is one of the most successful physical theories ever developed, both in the breadth of its applicability and in the accuracy of its predictions.

2.2 Neutrinos

Neutrinos are extremely low-mass subatomic particles. They may be – and were long believed to be – massless, but their mass, if nonzero, is on the order of eV (10^{-36} kg) or less, at least a million times lighter than even the electron. They are considered to be structureless point particles - some of the universal

building blocks of the Standard Model. Six kinds of neutrinos have been observed: the electron neutrino ν_e, the muon neutrino, ν_μ, and the tau neutrino, ν_τ. Three corresponding generations of antineutrinos also exist. Neutrinos are leptons (spin 1/2), and are therefore governed by the same Quantum Mechanical laws as electrons.

Neutrinos were first theorized by Pauli in 1930 to balance nuclear reactions, and were first observed in 1956. The muon neutrino was discovered in 1962 and the tau neutrino was first detected in 2000. Whether more flavors of neutrinos exist, particularly sterile forms that do not interact with ordinary matter, remains an open question. It should be noted, however, that the number of kinds of neutrinos and their masses are bounded by astrophysical considerations [5]. Neutrino oscillations were first theorized in the 1950s and first observed in the 1960s in the Homestake Mine Experiment, although definitive experimental evidence was not developed until about 2000.

Neutrinos interact only very rarely with other matter, and only through the weak force. Neutrinos have chirality, and in the Standard Model only left-handed neutrinos can interact with other matter. Right-handed neutrinos, if they exist, would interact only with other neutrinos, and are therefore called sterile. The three kinds of neutrinos are called flavors or generations. It is possible that these flavor eigenstates, the form in which neutrinos interact with other matter (and which are therefore the eigenstates of observation) are not the same as the mass (or, more generally, the propagation) eigenstates, and that this gives rise to flavor-mixing during propagation, observed as flavor-oscillation. This is the basic principle behind both the mass model and the new models to be introduced later.

There is some question as to whether neutrinos are their own antiparticles or not. Majorana proposed in 1937 that some particles could be their own antiparticles. However, only neutral particles can have this property without violating charge conservation, so of all the fermions in the Standard Model, neutrinos are the only candidates. Most particles are distinct from their antiparticles, in which case they are called Dirac particles. For example, electrons, protons and neutrons all fall into this category. A Dirac anti-neutrino would have opposite helicity (so that each flavor has 4 possible states). The two possibilities are not exclusive; in general, neutrinos could have both Dirac and Majorana masses. At this time, it is unknown whether either component of the mass is zero. If there were a non-zero Majorana mass, the phenomenon of "neutrinoless double beta decay" could occur, in which a nucleus emits two electrons and becomes positively charged without emitting neutrinos. This has not been observed, but experimental searches for it continue. However, Majorana masses are theoretically favored because of the role they play in explaining the tiny mass of the active neutrinos (this will be explained below in the section on the seesaw mechanism).

2.3 Relativistic Quantum Mechanics: the Dirac Equation

Because neutrinos travel at or near the speed of light, they must be handled by the relativistic version of quantum mechanics. The evolution of quantum particles in such a regime is described, for spin-1/2 particles, by the Dirac Equation:

$$i\gamma^{\mu}\partial_{\mu}\psi - m\psi = 0 \tag{2.1}$$

where γ represents the gamma matrices:

$$\gamma^{i} = \begin{bmatrix} 0 & \sigma^{i} \\ -\sigma^{i} & 0 \end{bmatrix} \text{, and } \gamma^{0} = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \tag{2.2}$$

(where σ are the usual Pauli polarization matrices). Here and throughout this thesis we will use natural units, in which $\hbar = c = 1$. We have also used Einstein summation notation, in which summation is implied by repeated indices (here μ). This will be used later without comment. The Dirac equation can be obtained by factoring the relativistic energy-momentum relation

$$p^{\mu}p_{\mu} - m^{2} = (\gamma^{\kappa}p_{\kappa} + m)(\gamma^{\lambda}p_{\lambda} - m) = 0 \tag{2.3}$$

One of these factors is then taken to be zero (which, if satisfied, will clearly satisfy the original equation), and substituting the quantum operator $i\partial_{\mu}$ for the momentum, we obtain the Dirac equation.

Observe that because the gamma matrices are four dimensional, the particle fields are also four dimensional vectors, called Dirac spinors (note that these are not four-vectors, however). Two dimensions correspond to the dynamics of one particle, while the other two describe the behavior of its antiparticle.

Chapter 3

Neutrino Oscillations

Neutrinos are famously difficult to detect. In fact, they pass without hinderance through huge bodies of matter, easily piercing the Earth at nearly the speed of light. Billions have passed through your hand in the time it took to read this sentence [19]. To study neutrino oscillations, a known source of neutrinos is measured at a given distance (the 'baseline' of the experiment) from the source and then the variance from the expected number of each type is measured. If neutrinos are missing or excess neutrinos were measured, it is concluded they oscillated into or out of the flavor being studied.

3.1 The Solar Neutrino Problem

In 1938, Hans Bethe developed a model for the way that the sun produces energy, now called the Standard Solar Model. He proposed that protons fuse to form Helium and emit electron neutrinos and photons in the process. In an attempt to test this model, Ray Davis developed the Homestake Experiment in 1968 to try to measure the flux of electron neutrinos coming from the sun. Because neutrinos pass unaltered through large amounts of mass that shield out other cosmic rays, the detector (a container of 615 metric tons of cleaning fluid) was built in a mine shaft deep underground. The solar neutrinos interacted with the fluid in the detector to occasionally produce Argon atoms, which could then be extracted from the material and counted to determine the number of neutrino interactions. 33 counts were observed after several months, which fell short of the expected value by a third. This discrepancy became known as the Solar Neutrino Problem

Solutions to the problem initially pointed to flaws in the experiment or in the Solar Model used to predict the neutrino flux, but further experimental evidence led some to take the problem more seriously. In 1968

Pontecorvo suggested for the first time that if neutrinos could oscillate into another state to which the detector was insensitive, the discrepancy could be explained.

Decisive experimental evidence that neutrino oscillation was responsible for the Solar Neutrino Problem was not obtained until 2001, when Super-Kamiokande [24] and the Sudbury Neutrino Observatory (SNO) [23] independently reported that, beside the unexpectedly small electron neutrino flux from the sun, there was also a component of muon or tau neutrinos in the solar radiation that were also unexpected based on the nuclear reactions known to occur in the Sun. Furthermore, it was found that the total neutrino flux was the amount originally predicted by the Solar Model. This led to the conclusion that some neutrinos, produced as ν_e on the Sun, nevertheless were detected as ν_μ or ν_τ on Earth, and that therefore some neutrinos change in flavor as they propagate.

3.2 Mass Model

Many results of Linear Algebra will be used in this thesis. An attempt has been made to define terms the first time they appear, but they have also been collected at the beginning of Appendix A. For more complete information, consult a text such as [31].

We detect the presence of neutrinos only when they interact with other matter - interaction mediated by the weak force. Therefore, when neutrinos are measured, they are measured in their flavor eigenbasis. Neutrino oscillation can only be made sense of if the eigenstates of propagation are incommensurable with the flavor eigenstates. Perhaps the simplest mechanism to invoke to explain these propagation states is to assume that there is a mass eigenbasis, which are mixtures of the flavor states. We could write the relationship between these two bases as a unitary[1] transformation of the form

$$|\nu_\alpha\rangle = \sum_i U_{\alpha i} |\nu_i\rangle \tag{3.1}$$

where greek-subscripted states ($\alpha = e, \mu, \tau$) represent flavor eigenstates, and latin-subscripted states ($i = 1, 2, 3$) are mass eigenstates. In other words, we are ascribing a mass matrix M to the neutrinos, which is not assumed to be diagonal in the flavor basis. This mass matrix can be diagonalized by the same unitary transformation U, and we will call its eigenvalues $\{m_1, m_2, m_3\}$. The matrix U then takes state vectors from the mass to the flavor basis. The mixture of flavor states in the mass eigenstates is shown schematically in Figure 3.1.

We now explain how neutrino masses give rise to flavor mixing. Let us assume we begin with a neutrino

[1]Recall that a unitary matrix U is one whose conjugate transpose is its inverse, ie, one for which $UU^\dagger = I$.

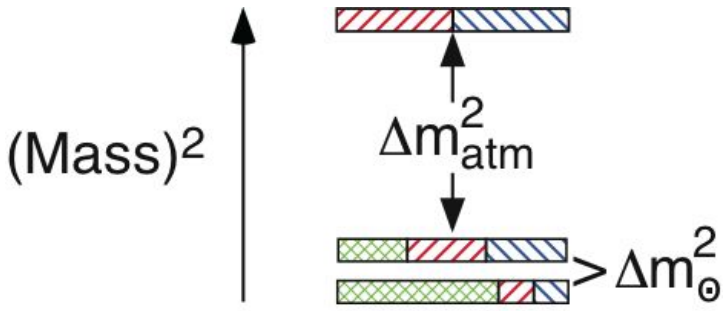

Figure 3.1: The flavor composition of the mass eigenstates. ν_e is indicated in green; ν_μ in blue and ν_τ in green. Taken from [5].

beam of a single flavor-type, i.e., $|\Psi(0)\rangle = |\nu_\alpha\rangle$. We could use the propagation basis instead, and write

$$|\Psi(0)\rangle = \sum_i U_{\alpha i} |\nu_i\rangle \tag{3.2}$$

We know that the time evolution operator is formed from the Hamiltonian, $e^{it\hat{H}}$, and that the space transformation operator is formed from linear momentum, e^{-ixp}.

The energy of a neutrino is given by the relativistic energy equation

$$E_i = \sqrt{p_\nu^2 + m_i^2} \approx p_\nu + \frac{m_i^2}{2p_\nu}, \tag{3.3}$$

where the approximation is good in the limit $p_\nu \gg m$, which is reasonable since it is known that m is on the order of eV at most. Then

$$|\Psi(x,t)\rangle = \sum_i U_{\alpha i} e^{-ixp} e^{itE_i} |\nu_i\rangle . \tag{3.4}$$

Note that the Hamiltonian has become the energy because the propagation states are the energy eigenstates. Since the neutrinos are traveling almost at the speed of light, $x \approx t$, and substituting in from (3.3), we have

$$|\Psi(x,t)\rangle = \sum_i U_{\alpha i} e^{im_i^2 t/2p} |\nu_i\rangle . \tag{3.5}$$

Now, when the beam is measured at a position a distance L from the source, we have $x = L$ but also $t \approx L$ because the neutrinos are travelling at almost the speed of light. Thus the final state of the beam is given by

$$|\Psi(L,L)\rangle = \sum_i U_{\alpha i} e^{im_i^2 L/2p} |\nu_i\rangle , \tag{3.6}$$

and the probability that we measure a state β is therefore

$$P_{\alpha\beta} = |\langle\nu_\beta|\Psi(L,L)\rangle|^2 = \left|\sum_i U^\dagger_{i\beta}\langle\nu_i|\sum_j U_{\alpha j}e^{im_j^2L/2p}|\nu_j\rangle\right|^2 \tag{3.7}$$

$$= \left|\sum_{i,j} U^\dagger_{i\beta}U_{\alpha j}e^{im_j^2L/2p}\langle\nu_i|\nu_j\rangle\right|^2$$

$$= \left|\sum_{i,j} U^\dagger_{i\beta}U_{\alpha j}e^{im_j^2L/2p}\delta_{ij}\right|^2$$

$$= \left|\sum_i U^\dagger_{i\beta}U_{\alpha i}e^{im_i^2L/2p}\right|^2. \tag{3.8}$$

The final form of the probability depends on the parameterization selected for the unitary matrix. The traditional representation for the two-generation case is simply the rotation matrix

$$U = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \tag{3.9}$$

and for the three generation case, the fully general expression is

$$U = \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{-i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{-i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{-i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{-i\delta} & c_{23}c_{13} \end{bmatrix} \times \begin{bmatrix} e^{-i\alpha_1/2} & 0 & 0 \\ 0 & e^{-i\alpha_2/2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{3.10}$$

where $c_{12} = \cos\theta_{12}$ and so on. There are six parameters which characterize U: three mixing angles (θ_{12}, θ_{13} and θ_{23}), a phase angle δ and two more overall phases α_1 and α_2 (which can be ignored when calculating probabilities). From experimental evidence, it is believed that $\theta_{13} \approx 0$ and that the effect of δ can be ignored [5]. This produces considerable simplification, so that the U usually used for calculations has the form (taking $\theta_{12} = \theta$, $\theta_{23} = \phi$)

$$U = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta\cos\phi & \cos\theta\cos\phi & \sin\phi \\ \sin\theta\sin\phi & -\cos\theta\sin\phi & \cos\phi \end{bmatrix}. \tag{3.11}$$

The transition probabilities between the different flavor eigenstates can be calculated from (3.8). For

example,

$$\begin{aligned}
P_{ee} &= \left|\sum_i U_{ie}^T U_{ei} e^{im_i^2 L/2p}\right|^2 \\
&= \left|\cos^2\theta e^{im_1^2 L/2p} + \sin^2\theta e^{im_2 L/2p}\right|^2 \\
&= (\cos^2\theta e^{-im_1^2 L/2p} + \sin^2\theta e^{-im_2^2 L/2p})(\cos^2\theta e^{im_1^2 L/2p} + \sin^2\theta e^{im_2^2 l/2p}) \\
&= \cos^4\theta + \sin^4\theta + \sin^2\theta\cos^2\theta(e^{iL(m_1^2-m_2^2)/2p} + e^{iL(m_2^2-m_1^2)/2p}) \\
&= \cos^4\theta + \sin^4\theta + \sin^2\theta\cos^2\theta(e^{i\Delta m_{12}^2 L/2p} + e^{-i\Delta m_{12}^2 L/2p}) \\
&= \cos^4\theta + \sin^4\theta + 2\sin^2\theta\cos^2\theta\cos(\Delta m_{12}^2 L/2p) \\
&= \cos^4\theta + \sin^4\theta + 2\sin^2\theta\cos^2\theta(1 - 2\sin^2(\Delta m_{12}^2 L/4p)) \\
&= 1 - 4\sin^2\theta\cos^2\theta\sin^2(\Delta m_{12}^2 L/4E). \qquad (3.12)
\end{aligned}$$

Here $\Delta m_{12}^2 = m_1^2 - m_2^2$; this is one of the mass-squared differences. The other Δm_{ij} are defined analogously. It is these values to which the standard neutrino oscillation experiments are sensitive. In the last step we have used the fact that $p \approx E$, which follows from (3.3).

The other probabilities are calculated similarly:

$$P_{e\mu} = 4\sin^2\theta\cos^2\theta\cos^2\phi\sin^2\left(\frac{\Delta m_{21}^2 L}{4E}\right), \qquad (3.13)$$

$$P_{e\tau} = 4\sin^2\theta\cos^2\theta\sin^2\phi\sin^2\left(\frac{\Delta m_{21}^2 L}{4E}\right), \qquad (3.14)$$

$$\begin{aligned}
P_{\mu\mu} = 1 &- 4\sin^2\theta\cos^2\theta\cos^4\phi\sin^2\left(\frac{\Delta m_{21}^2 L}{4E}\right) - 4\sin^2\theta\sin^2\phi\cos^2\phi\sin^2\left(\frac{\Delta m_{31}^2 L}{4E}\right) \qquad (3.15) \\
&-4\cos^2\theta\sin^2\phi\cos^2\phi\sin^2\left(\frac{\Delta m_{32}^2 L}{4E}\right),
\end{aligned}$$

$$\begin{aligned}
P_{\mu\tau} = &-4\sin^2\theta\cos^2\theta\sin^2\phi\cos^2\phi\sin^2\left(\frac{\Delta m_{21}^2 L}{4E}\right) + 4\sin^2\theta\sin^2\phi\cos^2\phi\sin^2\left(\frac{\Delta m_{31}^2 L}{4E}\right) \qquad (3.16) \\
&+4\cos^2\theta\sin^2\phi\cos^2\phi\sin^2\left(\frac{\Delta m_{32}^2 L}{4E}\right),
\end{aligned}$$

$$\begin{aligned}
P_{\tau\tau} = 1 &- 4\sin^2\theta\cos^2\theta\sin^4\phi\sin^2\left(\frac{\Delta m_{21}^2 L}{4E}\right) - 4\sin^2\theta\sin^2\phi\cos^2\phi\sin^2\left(\frac{\Delta m_{31}^2 L}{4E}\right) \qquad (3.17) \\
&-4\cos^2\theta\sin^2\phi\cos^2\phi\sin^2\left(\frac{\Delta m_{32}^2 L}{4E}\right).
\end{aligned}$$

From this we see some of the characteristic features of the mass model predictions. The oscillations all depend on L/E; in particular, there is an inverse dependence on energy. Furthermore, there are two

independent mass-squared differences which control the absolute length scale on which oscillations occur. The mixing angles determine the amplitude of the oscillations.

The experimental results are best fit by the parameters

$$\begin{aligned} \theta &\approx 34^\circ \\ \phi &\approx 45^\circ \\ \Delta m^2_\odot = \Delta m^2_{12} &\approx 7.6 \times 10^{-5}\ \text{eV}^2 \\ \Delta m^2_{atm} = \Delta m^2_{23} &\approx .0024\ \text{eV}^2. \end{aligned} \tag{3.18}$$

Whether θ_{13} is identically zero remains an open question. The most robust answer on this subject is the result that $\sin^2\theta_{13} < .04$, determined by the Chooz collaboration [24].

We plot some of these oscillation probabilities in Figures 3.2 - 3.5. Because $\phi = 45^\circ$, ν_μ and ν_τ are mixed in equal quantities in the mass states, and so are interchangeable. You can see that the ν_e survival probability is almost the complement of the ν_e to ν_μ transition. However, the amplitude of the ν_e to ν_μ transition is only half the amplitude of the survival probability because the other half of the disappearing ν_e oscillate into ν_τ. The ν_μ behavior, shown in 3.3 and 3.4, is somewhat more complicated. There are two characteristic oscillation lengths: the short one showing strong (amplitude unity) oscillations between ν_μ and ν_τ and the long-wavelength weaker signal corresponding to the oscillation into ν_e shown in the previous figures. These interact to produce the complex signal graphed.

The probabilities can also be left in terms of the matrix U, which will be necessary if it is impossible to easily parameterize U, as will be the case in the general models to be discussed. We then have

$$P_{\alpha\beta} = \delta_{\alpha\beta} - 4\sum_{i>j} U_{\alpha i}U_{\beta i}U_{\alpha j}U_{\beta j}\sin^2\left(\frac{\Delta\lambda_{ij}L}{2}\right) \tag{3.19}$$

where $\{\lambda_i\}$ are the eigenvalues of the Hamiltonian diagonalized by U. In the case of the mass model, these are simply $\frac{m_i^2}{2E}$.

Note that there are ways to detect neutrino masses directly [25] instead of inferring mass squared differences, which is predicated on the accuracy of the mass model. Measurement of the arrival time of neutrinos from a supernova gave an upper limit on neutrino mass of 5.7 eV. However, such events are rare and unsuitable for repeated measurement. Another method is to look at the spectrum of β decay, in which a neutrino takes part, and calculate its mass directly from the relativistic kinematics involved. The best result from β decay experiments is $m_{\nu_e} < 2.3$ eV, attained by the Mainz Neutrino Mass Experiment [29]. A new version of this experiment, called KATRIN [30], is under development, and is expected to begin collecting data in 2012.

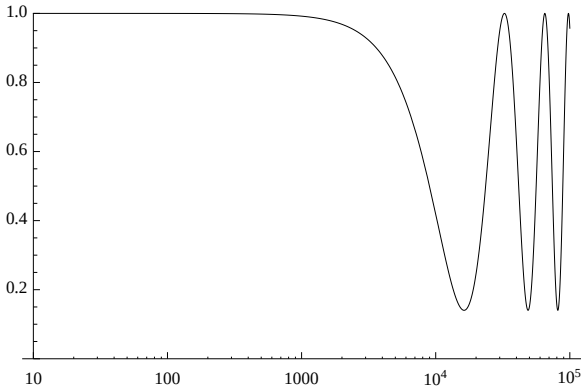

Figure 3.2: Survival Probability for the electron neutrino using the experimentally determined parameters. Horizontal axis is L/E in units [m/MeV].

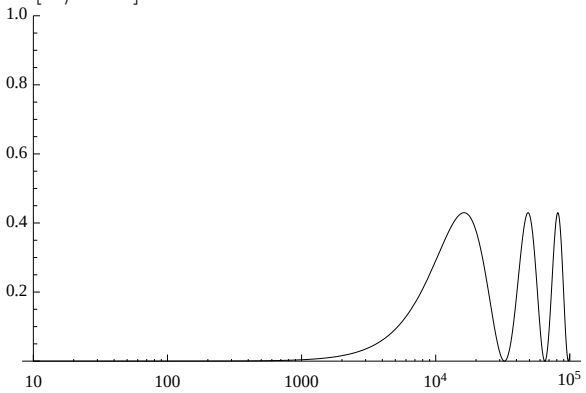

Figure 3.3: Oscillation Probability for $\nu_e \leftrightarrow \nu_\mu$ and $\nu_e \leftrightarrow \nu_\tau$. Horizontal axis is L/E in units [m/MeV].

There is a two-flavor limit to the mass model oscillations, which is frequently used in analyzing the experiments. Because the scales on which the different oscillation channels act are so different, experiments can focus on the effect of a single channel without really paying attention to the other oscillations which are (negligibly) going on. This gives rise to the nomenclature of atmospheric and solar mass squared differences, since experiments of these types are sensitive to only one of the two oscillation channels. In this limit the two oscillation probabilities become

$$P_{\mu\tau} = \sin^2(2\phi)\sin^2\left(\frac{\Delta m^2_{\text{atm}} L}{4E}\right) \tag{3.20}$$

$$P_{e\mu} = \sin^2(2\theta)\sin^2\left(\frac{\Delta m^2_{\odot} L}{4E}\right). \tag{3.21}$$

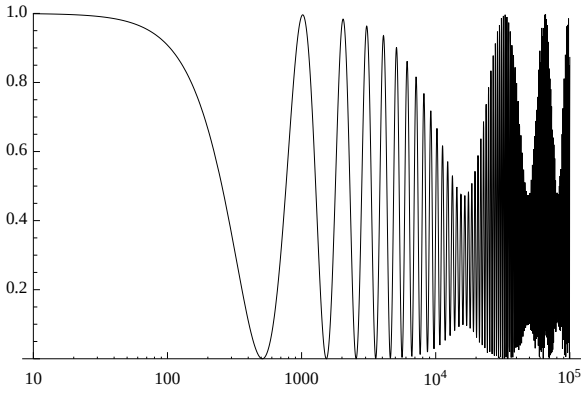

Figure 3.4: Survival Probability for the muon neutrino. Horizontal axis is L/E in units [m/MeV].

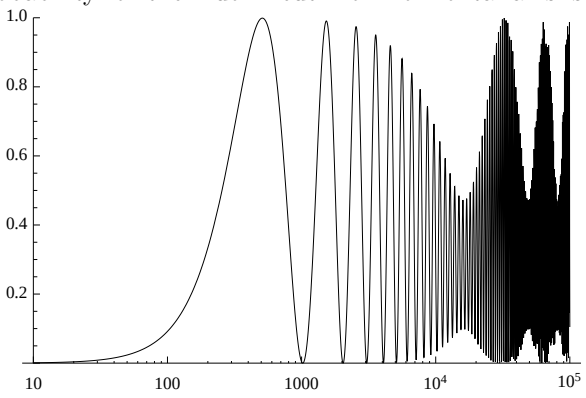

Figure 3.5: Oscillation Probability for $\nu_\mu \leftrightarrow \nu_\tau$. Horizontal axis is L/E in units [m/MeV].

3.3 Experiments

Neutrino oscillation experiments are generally classified by the neutrino source studied (atmospheric, reactor, solar, or accelerator), and each of these will be considered in more detail below. However, in all experiments, the same fundamental procedure is followed.

The flavor content of the neutrino beam at the source must be known - either from theoretical understanding of their production or from direct measurement. The flux of neutrinos must also be known, which requires a measurement, either direct or indirect (for example, of a particle produced in the same reaction that produces the neutrinos).

There are two general classes of experiments. Either a detector sensitive to the source flavor can be used to see whether all the neutrinos are still present at some distant location (a disappearance experiment) or a different flavor can be detected for, to see whether any neutrinos of a different flavor have suddenly appeared (an appearance experiment).

Table 3.1: Summary of results of key experiments about neutrino oscillation.

Name (Date)	Experiment	Distance	Energy	Result	Δm^2 (eV^2)	$\sin^2 2\theta$
LSND (2001) [14]	$\bar{\nu}_\mu \to \bar{\nu}_e$ appearance	30 m	60 MeV	present	4	.004
KARMEN (1998)	$\nu_\mu \to \nu_e$ appearance	18 m	30 MeV	absent		
	$\bar{\nu}_\mu \to \bar{\nu}_e$ appearance			absent		
	ν_e disappearance			absent		
MiniBooNE (2010+) [18]	$\nu_\mu \to \nu_e$ appearance	500 m	200-475 MeV	present	3	.002
			>475 MeV	absent		
	$\bar{\nu}_\mu \to \bar{\nu}_e$ appearance		200-475 MeV	absent		
			>475 MeV	absent		
MINOS (2010+) [10]	$\nu_\mu \to \nu_\tau$ disappearance	1 km, 735 km	120 GeV	present	2×10^{-3}	>.9
KamLAND (2008) [11]	$\bar{\nu}_e$ disappearance	180 km (avg)		present	7×10^{-5}	.9
SNO (2008) [12]	ν_e disappearance	-		present	8×10^{-5}	.87
K2K (2008)	$\nu_\mu \to \nu_\tau$	250 km	1GeV	present	3×10^{-3}	1
SuperK (1998) [24]	$\nu_\mu \to \nu_\tau$			present	1×10^{-3}	>.82

Many such experiments have been carried out since the 1990s. Some of the more important or relevant results for this study are summarized in Table 3.1.

3.3.1 Accelerator Neutrino Experiments

Particle accelerators produce many energetic particles which frequently decay into neutrinos. Because the source is controlled by the experimentors, such experiments uniquely facilitate direct measurements both near the site of production and at some longer distance, allowing for precise comparisons of the flavor composition of the beam as it propagates.

LSND (the Liquid Scintillation Neutrino Detector) collected data from 1993 until 1998, and published their surprising results in 2001 [14]. Because LSND used an unusually short baseline (only 30m), no oscillations were expected; however, they were measured. The primary mode of neutrino generation was through the reaction $\pi^+ \to \mu^+ + \nu_\mu$ and $\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu$. The source used was the Los Alamos Neutrino Science Center (LANSCE), a particle accelerator which produces a large number of pions. Because the source produces very few $\bar{\nu}_e$, the oscillation mode from $\bar{\nu}_\mu \to \bar{\nu}_e$ could be sensitively detected for through the likely

reaction $\bar{\nu}_e + p \rightarrow e^+ + n$. Their result was consistent with a neutrino mass squared difference of about $0.2 - 10$ eV^2.

The MiniBooNE collaboration was developed to try to confirm or reject the result observed at LSND. It uses a high-energy proton source to react with a Beryllium target to produce mesons which decay into neutrinos or antineutrinos [7]. The beam can be controlled to select for either neutrinos or antineutrinos. MiniBooNE initially ran in neutrino mode (because the statistics are better), and found results incompatible with LSND. However, beginning in 2006, MiniBooNE has been taking data in antineutrino mode, and the results are so far consistent with LSND and may therefore show signs of a neutrino - anti-neutrino asymmetry, which could indicate possible CPT violation.

In most neutrino oscillation experiments, a certain background of apparent oscillations must be dealt with. In the case of MiniBooNE, the detector can pick up electron neutrinos from cosmic radiation, there is a tiny contamination of electron neutrinos in the source, and some other particle reactions have signatures which can be confused with electron neutrinos. These effects must all be carefully quantified so that the expected number of detections - assuming no oscillations - can be predicted. The background must then be subtracted from the data to determine the number of "excess events" observed. In the case of MiniBooNE in antineutrino mode, after running for three years, this figure was still only 25 events. This excess can then be translated into an oscillation probability. This process is shown in Figures 3.6 and 3.7.

Another ongoing accelerator experiment, the Main Injector Neutrino Oscillation Search (MINOS) also measures the results of pion decays [10]. Its neutrino source is at Fermilab, and oscillations are measured by detecting the beam 735 km away in Minnesota. MINOS has so far released only preliminary results which seem to be consistent with the mass model, but may also show signs of a neutrino anti-neutrino asymmetry.

There are also many other accelerator experiments, which are consistent with the mass model, including CHORUS and NOMAD.

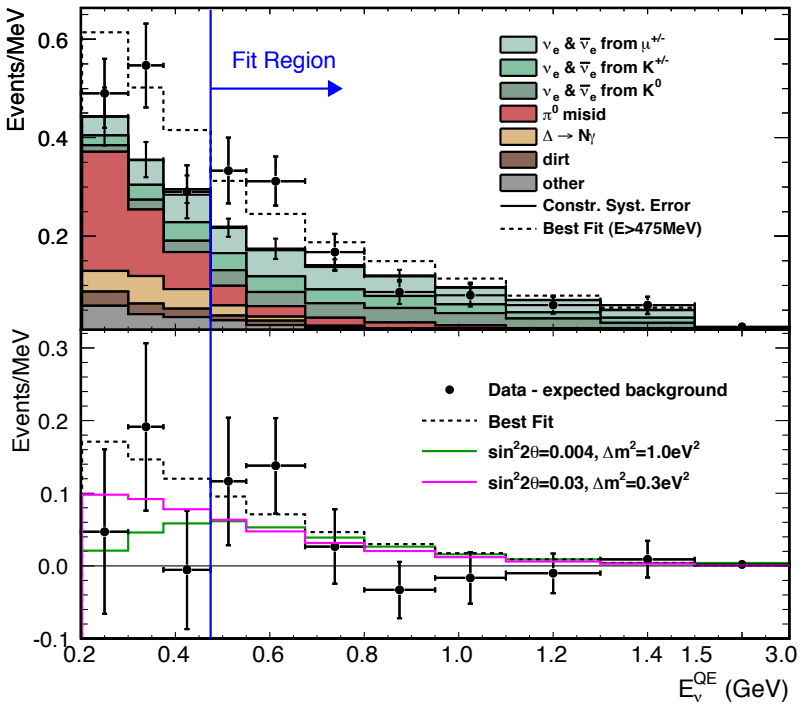

Figure 3.6: Data from MiniBooNE. Top: the colored columns indicate the expected contribution to the background signal from different sources; the dots indicate the actual data points. Bottom: the data after the background has been subtracted. Taken from [18].

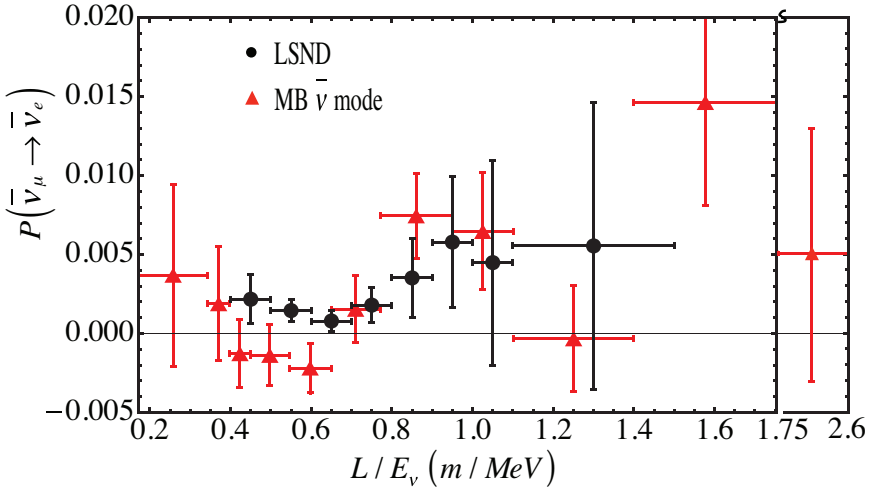

Figure 3.7: The raw data can be translated into probabilities, which are the theoretically preferred quantity. Here the red triangles indicate data from MiniBooNE in antineutrino mode while the black dots indicate data from LSND. Observe the general agreement. Also taken from [18].

3.3.2 Solar Neutrino Experiments

Solar Neutrino experiments are those which use the Sun as the neutrino source. The Sun produces electron neutrinos almost exclusively via the interaction $p + p \rightarrow d + e^- + \nu_e$ and other similar reactions in the fusion chain used by the Sun. In fact, the neutrinos usually detected are the more energetic products of the reaction ${}^8B \rightarrow {}^8 Be^* + e^+ + \nu_e$ even though the reaction is much less frequent. Because the sun's reactions are well understood, it is straightforward to accurately predict the Sun's neutrino flux at the Earth.

The Sudbury Neutrino Observatory (SNO) measured the flavor components of the solar neutrino flux by three different channels: the charged-current interaction $\nu_e + d \rightarrow p + p + e^-$, elastic scattering $\nu + d \rightarrow p + n + \nu$, and the neutral-current $\nu + e^- \rightarrow \nu + e^-$ [23]. Because some of these channels are specially activated only by electron neutrinos, but others are equally sensitive to all flavors of neutrinos, the experiment could measure both the electron neutrino and total neutrino flux coming from the Sun. Their data, collected between 2001 and 2003, provided the first definitive evidence that the solar neutrino problem could be solved by invoking neutrino oscillations. Their results through 2008 indicate that $\Delta m^2 = 7.59 \times 10^{-5}$ eV^2 and $\theta = 34.4^\circ$ [12]. There have been many other Solar neutrino experiment, including the famous Homestake experiment mentioned above.

3.3.3 Reactor Neutrino Experiments

Reactor experiments use nuclear power reactors as their neutrino source, despite the fact that these sources are often variable and their precise composition is unknown.

The Kamioka Liquid scintillator Anti-Neutrino Detector (KamLAND) collected data from 55 Japanese nuclear reactors, which produce $\bar{\nu}_e$, between 2002 and 2007 [11]. The detection reaction used was $\bar{\nu} + e + p \rightarrow e^+ + n$. They found that $\Delta m^2 = 7.58 \times 10^{-5}$ eV^2, and $\tan^2\theta = 0.56$. Notice that this result is in remarkably good agreement with the result from SNO listed above. There were many other reactor experiments, including Bugey, Gosgen, and CHOOZ; however, no others had a long enough baseline to detect any oscillations.

3.3.4 Atmospheric Neutrino Experiments

Many neutrinos are produced in the upper atmosphere through interactions with cosmic rays. The Super-Kamiokande experiment measured the oscillatory dependence on distance of survival probability of these atmospheric neutrinos. They measured for ν_μ travelling through the Earth after being created in the atmosphere. Since the distance to the atmosphere on the other side of the Earth depends on the angle at which you are looking into the Earth, they could easily change the baseline of their experiment. The results were consistent with $\Delta m^2 = 2.4 \times 10^{-3}$ eV^2 [26].

3.4 LSND and other Anomalies

Assuming there are only three kinds of neutrinos, the mass model has room for only two independent mass-squared differences. However, three such differences have been experimentally observed: one of order 10^{-5}, one of order 10^{-3}, and one of order 1 (eV^2). This last large mass difference has been measured in only two anomalous experiments: LSND and MiniBooNE (which was developed to confirm or reject the LSND result). Some of the explanations suggested to explain the LSND anomaly are that the experiment made some kind of error in computing the probability of oscillation; that the result indicates there are more than three kinds of neutrinos (either sterile neutrinos or a fourth generation of matter), which could possibly allow for a more general mass model explaining all the results; and that the mass model is simply wrong.

Other theoretically problematic results of neutrino oscillation experiments include the Gallium anomaly [15], in which extremely short-baseline oscillations were observed, and hints of a neutrino - antineutrino asymmetry which could represent a CPT violation.

3.5 Seesaw Mechanism

The seesaw mechanism is one conventional explanation for the scale of the neutrino masses. The fact that the neutrinos (if massive) have masses many orders of magnitude smaller than any other known particle is one of the hierarchy problems, which ask why such imbalances exist. The seesaw mechanism can explain how such masses could arise from the standard mass scales. Furthermore, the same mathematical effect will lead to very important results in the new models to be proposed in Chapter 5.

Consider a general hermitian[2] 2 x 2 matrix

$$\begin{bmatrix} a & b \\ b^* & c \end{bmatrix}. \tag{3.22}$$

The eigenvalues of this matrix are easy to compute from the definition. We find

$$\lambda = \frac{1}{2}(a + c \pm \sqrt{(a-c)^2 + 4b^2}). \tag{3.23}$$

Note that when the matrix becomes unbalanced, that is, when one of the terms dominates the others, the exact eigenvalues can be accurately approximated by a Taylor expansion. For example, when c is very large we obtain

$$\lambda = c + \frac{b^2}{c}\ ,\ a - \frac{b^2}{c} \tag{3.24}$$

[2]Remember that a *hermitian* matrix M is one whose conjugate transpose is itself, ie, for which $M = M^\dagger$. All hermitian matrices are diagonalizable by the Spectral Theorem.

Thus the matrix has one very large eigenvalue (c) and one very small eigenvalue (a). Since the diagonal entries are approximately the eigenvalues, we see that the matrix is almost diagonal. Thus as c grows, the behavior of the eigenvalues changes significantly in form - not simply in proportion to the change in c. This observation that the form of the eigenvalues changes as a matrix becomes unbalanced is called the seesaw mechanism.

The seesaw mechanism may explain why the mass of the neutrino is so much smaller than the mass of the other fundamental particles. A single generation of neutrino has a mass matrix of the form

$$\begin{bmatrix} m_L & m_D \\ m_D & m_R \end{bmatrix} \tag{3.25}$$

where m_R and m_L are the Majorana mass coupling of the right- and left- handed neutrinos, respectively, and m_D is the Dirac mass. We assign $m_L = 0$ since this is theoretically desirable (a Majorana mass term for ν_L violates electro-weak gauge invariance [9]), $m_D = 100$ GeV, and $m_R = 10^{15}$ GeV. These correspond to the energy scales of electroweak symmetry and the GUT scale, respectively, and are therefore "natural" mass scales. GUT stands for Grand Unified Theory, and the GUT scale is the energy range at which the strong, weak and electromagnetic forces are theorized to unify. Using these values, the mass matrix has a single very large diagonal entry, so its eigenvalues are given approximately by (3.24) as

$$m_R + \frac{m_D^2}{m_R} \text{ and } \frac{m_D^2}{m_R} \tag{3.26}$$

These masses are of order 10^{24} and 10^{-2} eV. The light mass is attributed to the left-handed active neutrino, while the heavy mass is attributed to the sterile, right-handed neutrino. This provides some explanation for why the active neutrinos have a positive mass so much smaller than any of the other known particles. The very large mass of the right-handed neutrino would make it very difficult to produce, and would thus also explain why right-handed neutrinos have yet to be observed.

We will later employ the seesaw mechanism to attempt to explain neutrino oscillations without appealing to masses, by considering a Hamiltonian which depends on energy in such a way that the behavior differs fundamentally at different energy scales.

Chapter 4

Lorentz Violation

4.1 Symmetry

Symmetry is probably one of the most important ideas in modern physics. Every conservation law represents an underlying symmetry of nature, an idea known as *Noether's Theorem*. For example, conservation of momentum is a consequence of the invariance of physical laws under spatial translations. Symmetries generally fall into one of two categories: *discrete* if only a finite set of symmetric states exist, and *continuous* if infinitely many equivalent states exist (such as with spatial translations). Again, more detailed treatment of this material can be found in [19].

4.1.1 Discrete Symmetries

Parity is mirror (or inversion) symmetry, the idea that the universe does not show a handed-ness preference. Violation of parity was first observed in 1956 in a study of cobalt made by Wu. Neutrinos badly (maximally) violate parity because active neutrinos are exclusively left-handed, while active anti-neutrinos are exclusively right-handed. If neutrinos are actually massive, they cannot travel at the speed of light, and there must be a frame in which the observed helicities are reversed; however, this has never been observed.

Charge conjugation or symmetry is really an operator that turns a particle into its antiparticle: $\hat{C}\,|p\rangle = |\bar{p}\rangle$, and is so-called because the antiparticle of a charged particle has the opposite charge. When parity symmetry was observed to be violated, some hoped that CP symmetry would still be maintained - and it is a very good symmetry, but not a perfect one. In 1964 Cronin and Fitch found evidence of CP violation in a neutral kaon system. CP violation creates an inequality between matter and antimatter and may explain why the universe is matter-dominated.

Time is another discrete symmetry, reflecting the idea that the laws of physics should work equally well in reverse. It has been proven that CPT is a perfect symmetry in any Lorentz-invariant Quantum Field Theory, so if such a theory describes the real world, T symmetry must also be broken (to counterbalance the broken CP symmetry).

4.1.2 Lorentz Symmetry

Lorentz symmetry is the symmetry underlying Special Relativity. Simply put, it means that there is no preferred frame in the universe. More technically, it says that all physical laws can be written in terms of Lorentz covariant variables, which remain fixed under Lorentz transformations (rotations in space and velocity boosts). This principle has been experimentally verified, but whether the symmetry is perfect is not knowable *a priori*. It has been proven that any CPT violation necessarily entails Lorentz violation, though the converse is not necessarily true [17].

4.2 Standard Model Extension

Because of the acknowledged shortcoming of the Standard Model, much work has been undertaken to extend its reach.

The Standard Model Extension (SME) [16] is a coherent framework in which to organize any work involving Lorentz violations. It makes only two assumptions: the laws of physics are coordinate system independent (so that the same results obtain whether cartesian coordinates, cylindrical coordinates, etc. are used), and that the Standard Model is the limiting case of a more general theory, in which Lorentz violations are permitted. In the Minimal SME, only renormalizable terms (those of mass dimension 3 or 4) are considered, but higher order terms have also been considered [27]. Any such theory can be formulated in terms of spontaneously formed background tensor fields, described by a set of coefficients.

As an example of how this works, we will discuss the Standard Model Extension of the Dirac equation for fermions, which include neutrinos. Recall (2.1) that the standard Dirac equation is written

$$i\gamma^\mu\partial_\mu\psi - m\psi = (i\gamma^\mu\partial_\mu - m)\psi = 0. \tag{4.1}$$

In the SME, we incorporate the effects of several Lorentz symmetry violating fields, which are given conventional names (a, b, c, and so on). The analogous equation [9] can be written

$$(i\Gamma^\mu\partial_\mu - M)\psi = 0 \tag{4.2}$$

where we define these new variables as follow:

$$\Gamma^{\nu} = \gamma^{\nu} + c^{\mu\nu}\gamma_{\mu} + d^{\mu\nu}\gamma_5\gamma_{\mu} + e^{\nu} + if^{\nu}\gamma_5 + \frac{1}{2}g^{\lambda\mu\nu}\sigma_{\lambda\mu} \tag{4.3}$$

and

$$M = m + a^{\mu}\gamma_{\mu} + b^{\mu}\gamma_5\gamma_{\mu} + \frac{1}{2}H^{\mu\nu}\sigma_{\mu\nu}. \tag{4.4}$$

where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$, and $\sigma_{\mu\nu} = \frac{i}{2}(\gamma_{\mu}\gamma_{\nu} - \gamma_{\nu}\gamma_{\mu})$. Note here that the first term in each sum corresponds to the Standard Model term, without Lorentz violations. Each of the other terms $a, b, c \ldots$ corresponds to a different Lorentz violation. For instance, e is a vector in space-time which gives some kind of structure or direction to space itself. These coefficients are assumed to be constant or nearly constant. Equation (4.2) describes the interactions of a single generation of particle, for example, ν_e by itself. If we want to consider several flavors at the same time, we can generalize (4.2) to a matrix equation.

There is a large experimental program ongoing to test Lorentz Symmetry and put upper bounds on the violation coefficients. In some sectors these have become very small, but most of the neutrino coefficients remain untested [28], primarily because of the difficulty of measuring neutrinos.

There are three primary effects of Lorentz violation. The first is that if rotational symmetry were broken, some physics could be observed to be direction-dependent. Many Lorentz-violating effects would then become apparent as daily or annual oscillations in measurements, as the Earth rotated through any possible Lorentz-violating field. The second possible signature of Lorentz violation would be velocity-dependent effects, which would arise if reference frame boost symmetry were broken. The last effect is an asymmetry between particles and antiparticles, which could appear if CPT symmetry were broken. Most signals are small and appear at high energies. However, the effects to be discussed in Chapter 5 are somewhat atypical in that the disagreements with the mass model predictions occur at low energy and are not necessarily characterized by annual or sidereal variations.

Some reference frame must be selected to discuss the projections of the SME coefficients, and the standard choice is a Sun-centered frame, in which the equatorial plane of the Earth corresponds to the x-y plane, and the axis of rotation of the Earth corresponds to the z-axis.

4.2.1 Neutrinos in the Standard Model Extension

The theory of Lorentz violations in neutrinos is discussed in detail in [9]. Of particular importance are the four Lorentz-violating field coefficients which appear in the effective Hamiltonian. These are a, c, g and H. a and g are CPT-odd and c and H are CPT-even. c and g are dimensionless, while a and H have dimensions of mass [28].

The fully general description of neutrinos includes three species of active, left-handed neutrinos, and three species of sterile, right-handed neutrinos, together with their antiparticles (note that anti- active neutrinos are right-handed while anti- sterile neutrinos are left-handed). We assumed small general Lorentz violations; because mass terms enter only at second order, they are neglected at this point. If we work in the basis 12-dimensional basis $\{\nu_L, \bar{\nu}_R, \nu_R, \bar{\nu}_L\}$ (each of these is a set of three flavors), the Hamiltonian describing the system is, to first order

$$H = \frac{1}{E}\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}. \tag{4.5}$$

This is a 12×12 matrix, which has been broken down in 6×6 blocks H_{ij}. These are defined

$$H_{11} = \begin{bmatrix} -c_L^{\mu\nu}p_\mu p_\nu + a_L^\mu p_\mu & -c_M^{\mu\nu}p_\mu p_\nu + a_M^\mu p_\mu \\ -c_M^{\dagger\mu\nu}p_\mu p_\nu + a_M^{\dagger\mu}p_\mu & -c_R^{T\mu\nu}p_\mu p_\nu - a_R^{T\mu}p_\mu \end{bmatrix} \tag{4.6}$$

$$H_{12} = \begin{bmatrix} -ig_D^{\lambda\mu\nu}p_\lambda q_\mu p_\nu + iH_D^{\lambda\mu}p_\lambda q_\mu & -ig_M^{\lambda\mu\nu}p_\lambda q_\mu p_\nu + iH_M^{\lambda\mu}p_\lambda q_\mu \\ -ig_M^{\dagger\lambda\mu\nu}p_\lambda q_\mu p_\nu + iH_M^{\dagger\lambda\mu}p_\lambda q_\mu & -ig_D^{T\lambda\mu\nu}p_\lambda q_\mu p_\nu - iH_D^{T\lambda\mu}p_\lambda q_\mu \end{bmatrix} \tag{4.7}$$

$$H_{21} = \begin{bmatrix} ig_D^{\lambda\mu\nu}p_\lambda q_\mu^* p_\nu - iH_D^{\lambda\mu}p_\lambda q_\mu^* & ig_M^{\lambda\mu\nu}p_\lambda q_\mu^* p_\nu - iH_M^{\lambda\mu}p_\lambda q_\mu^* \\ ig_M^{\dagger\lambda\mu\nu}p_\lambda q_\mu^* p_\nu - iH_M^{\dagger\lambda\mu}p_\lambda q_\mu^* & ig_D^{T\lambda\mu\nu}p_\lambda q_\mu^* p_\nu + iH_D^{T\lambda\mu}p_\lambda q_\mu^* \end{bmatrix} \tag{4.8}$$

$$H_{22} = \begin{bmatrix} -c_R^{\mu\nu}p_\mu p_\nu + a_R^\mu p_\mu & -c_M^{T\mu\nu}p_\mu p_\nu - a_M^{T\mu}p_\mu \\ -c_M^{*\mu\nu}p_\mu p_\nu - a_M^{*\mu}p_\mu & -c_L^{T\mu\nu}p_\mu p_\nu - a_L^{T\mu}p_\mu \end{bmatrix} \tag{4.9}$$

Each of these blocks H_{ij} is itself a 6×6 matrix which has been further decomposed into four 3×3 blocks. Each of these blocks implicitly contains the three neutrino flavors $\{e, \mu, \tau\}$. Thus the Lorentz-violating coefficients (eg, c_L) are now 3×3 matrices in flavor-space and tensors[1] in space-time. Here p_μ is the energy-momentum fourvector for the neutrino, and q_μ is a helicity state. A suitable choice for q is $q^\nu = \frac{1}{\sqrt{2}}(0; \hat{q_1} + i\hat{q_2})$ where $\{\hat{p}, \hat{q_1}, \hat{q_2}\}$ forms a right-handed orthonormal set. There are, in total, almost 1,000 variables describing the possible Lorentz-violations of neutrinos. There are three sets of c coefficients (c_L, c_M, c_R). Each of these act on three generations of neutrinos, and can therefore take different values for each of 9 oscillation channels. For each channel, the coefficient is a symmetric, two-dimensional hermitian tensor, which therefore has 10 degrees of freedom. So for the c's we have $3 \times 9 \times 10 = 270$ variables. There are also three a coefficients, which are only one-dimensional tensors, and therefore contain 108 variables. The two H's are antihermitian real two-dimensional tensors, and each contain 54 degrees of freedom. Lastly, the two g's are antihermitian real three-dimensional tensors, and each therefore contain 216 degrees of freedom. This makes a total of 918 independent coefficients. These coefficients are linear combinations of the Lorentz-violating fields discussed

[1] A tensor is a multidimensional matrix which behaves in a particular way under transformations.

in the previous section. For example, $c_L = c + d$. The other coefficients are similarly defined (details can be found in [9]).

These are only the variables affecting neutrinos. Each particle field (top quarks, photons, electrons, and so on) would be affected by its own set of parameters, of which there would be a similar number. Many of these parameters are well-restricted by experimental evidence, but, for example, with neutrinos, only about 20 of the thousand parameters have calculated bounds [28]. Obviously, only a small subset of these will be considered here in any detail.

4.3 Review of Lorentz-violating Models

The models introduced in the next chapter are not the first models of neutrino oscillations to incorporate Lorentz violations. At least two similar models have been proposed already: the Bicycle Model [1] and the Tandem Model [2]. Both models use a generalized Hamiltonian of the form

$$\hat{H}_{ab} = E\delta_{ab} + \frac{1}{2E}M^2_{ab} + (A_L)_{ab} - \frac{4}{3}(C_L)_{ab}E \tag{4.10}$$

This is a generalization of the Mass Model Hamiltonian to include Lorentz-violating terms. This Hamiltonian can be obtained from the general SME Lagrangian for neutrinos, which is related to the Dirac equation (4.2). The details of this calculation are given in [9]. These models are especially relevant to consider because both use the seesaw mechanism to predict different behavior over different energy ranges. This mechanism will also be used by the new models introduced below.

4.3.1 Bicycle Model

The Bicycle Model was introduced by Kostelecký and Mewes in 2004 [1]. It assumes a Hamiltonian of the form

$$\hat{H}_{BM} = \begin{bmatrix} -2cE & -\frac{a}{\sqrt{2}}\cos\Theta & -\frac{a}{\sqrt{2}}\cos\Theta \\ -\frac{a}{\sqrt{2}}\cos\Theta & 0 & 0 \\ -\frac{a}{\sqrt{2}}\cos\Theta & 0 & 0 \end{bmatrix} \tag{4.11}$$

in the basis $\{\nu_e, \nu_\mu, \nu_\tau\}$, which is the standard basis that will be used throughout. This corresponds to a choice of the Lorentz-violating parameters $(a_L)^Z_{e\mu} = (a_L)^Z_{e\tau} = \frac{a}{\sqrt{2}}$, $\frac{4}{3}(c_L)^{TT}_{ee} = 2c$, with all other terms being zero. The Θ refers to the angle between the neutrino momentum and the (Lorentz-violating) field vector a_L, which has been arbitrarily chosen to point in the $\hat{z}$ direction, along the axis of rotation of the Earth. The variation in Θ due to the motion of the Earth would give rise to sidereal oscillations typical of the SME. In

the new models introduced later, only the isotropic components of the tensors will be taken to be nonzero, so no sidereal variations would be predicted. Here only the isotropic component of c_L, (c_L^{TT}, is selected to be nonzero. This would correspond to an isotropic background which violates the symmetry of boosts but not rotations. The bicycle model is particularly simple: it has only two free parameters, fewer than even the traditional mass model it seeks to mimic. It still predicts quite complex behavior, however.

We want to find the unitary matrix U that will diagonalize H because we know that the matrix of transition amplitudes is given in terms of the time development operator by $S = e^{-itH}$, or

$$S = Ue^{-itU^\dagger HU}U^\dagger \tag{4.12}$$

and taking the absolute square of each of the components of this matrix S will give the transition probabilities.

We begin by finding the eigenvalues of the matrix:

$$\lambda = \{0, -cE \pm \sqrt{(cE)^2 + (\frac{a\cos\Theta}{\sqrt{2}})^2}\} \tag{4.13}$$

From this we see that $(0, \frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}})$ is an eigenvector of the Hamiltonian, and is therefore one of the columns of the matrix in question[2]. Because the Orthogonal Group in three dimensions is itself three dimensional, and since we have, in effect, selected two directions in the form of the first eigenvalue, we know we can parameterize the diagonalizing matrix in terms of one more direction, and we write

$$U = \begin{bmatrix} 0 & \cos\theta & -\sin\theta \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\sin\theta & \frac{1}{\sqrt{2}}\cos\theta \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\sin\theta & \frac{1}{\sqrt{2}}\cos\theta \end{bmatrix} \tag{4.14}$$

(noting that θ is unrelated to Θ).

If we multiply out $U^\dagger HU$ we find

$$H' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2cE\sin^2\theta - 2a\cos\Theta\sin\theta\cos\theta & 2cE\sin\theta\cos\theta + a\cos\Theta(\cos^2\theta - \sin^2\theta) \\ 0 & 2cE\sin\theta\cos\theta + a\cos\Theta(\cos^2\theta - \sin^2\theta) & -2cE\cos^2\theta + 2a\cos\Theta\sin\theta\cos\theta \end{bmatrix} \tag{4.15}$$

and, imposing the fact that this matrix is diagonal with the eigenvalues as entries, we know

$$2cE\sin\theta\cos\theta - a\cos\Theta(\cos^2\theta - \sin^2\theta) = 0 \tag{4.16}$$

This equation can be solved by isolating the $\sin\theta$, squaring both sides, and using the pythagorean identity. We find

$$\cos^2\theta = \frac{1}{2}[1 - \frac{cE}{\sqrt{(cE)^2 + (a\cos\Theta)^2}}] \tag{4.17}$$

[2]Recall that the matrix of normalized eigenvectors diagonalizes a matrix.

Now that U and H' are known completely, the transition probabilities can be calculated easily from (4.9). I find

$$P_{ee} = 1 - 4\sin^2\theta\cos^2\theta\sin^2(\Delta_{31}L/2) \tag{4.18}$$

$$P_{e\mu} = P_{e\tau} = 2\sin^2\theta\cos^2\theta\sin^2(\Delta_{31}L/2) \tag{4.19}$$

$$P_{\mu\mu} = P_{\tau\tau} = 1 - \sin^2\theta\cos^2\theta\sin^2(\Delta_{31}L/2) - \sin^2\theta\sin^2(\Delta_{21}L/2) - \cos^2\theta\sin^2(\Delta_{32}L/2) \tag{4.20}$$

$$P_{\mu\tau} = \sin^2\theta\sin^2(\Delta_{21}L/2) + \cos^2\theta\sin^2(\Delta_{32}L/2) - \sin^2\theta\cos^2\theta\sin^2(\Delta_{31}L/2) \tag{4.21}$$

where I have defined the Δ's in terms of the eigenvalue differences:

$$\begin{aligned} \Delta_{21} &= \sqrt{(cE)^2 + (a\cos\Theta)^2} - cE \\ \Delta_{31} &= 2\sqrt{(cE)^2 + (a\cos\Theta)^2} \\ \Delta_{32} &= \sqrt{(cE)^2 + (a\cos\Theta)^2} + cE \end{aligned} \tag{4.22}$$

These are exactly the results in [1]. The bicycle model also includes an energy-dependent seesaw effect, because at high energies, $\sin\theta$ goes to 0, so that electron neutrinos cannot oscillate, while the muon and tau neutrinos have a simple two-flavor oscillation. However, at low energies, the behavior is more complicated. We can simplify the expressions significantly if we take either the high- or low-energy limit. To first order at high energies, E $\gg$ a/c

$$P_{ee} = 1 \tag{4.23}$$

$$P_{e\mu} = P_{e\tau} = 0 \tag{4.24}$$

$$P_{\mu\mu} = P_{\tau\tau} = 1 - \sin^2(a^2L/4cE) \tag{4.25}$$

$$P_{\mu\tau} = \sin^2(a^2L/4cE) \tag{4.26}$$

We see that this mimics the usual energy dependence derived for the mass model (compare 3.12, 3.13). Thus in the high-energy limit, an effective mass-squared-difference $\Delta m^2 = a^2/c$ appears even though no masses were used in the model. In the low-energy limit that E $\ll$ a/c

$$P_{ee} = 1 - \sin^2(aL) \tag{4.27}$$

$$P_{e\mu} = P_{e\tau} = \frac{1}{2}\sin^2(aL) \tag{4.28}$$

$$P_{\mu\mu} = P_{\tau\tau} = 1 - \frac{1}{4}\sin^2(aL) - \sin^2(aL/2) \tag{4.29}$$

$$P_{\mu\tau} = \sin^2(aL/2) - \frac{1}{4}\sin^2(aL) \tag{4.30}$$

we see that the oscillation lengths are energy-independent.

The important feature to notice here is that a seesaw mechanism turns on at certain energies above which the eigenvalues depend on E^{-1}, allowing the model to reproduce the predictions of the mass model even though no masses are involved. The bicycle model is attractive for its simplicity; however, in a 2007 paper of Barger, Marfatia and Whisnant [8], it was shown that the bicycle model is incompatible with the combined results of solar, reactor, and accelerator experiments. In fact, the paper even considered a generalized version of the Hamiltonian (4.12) in which the direction of a_L could be specified, and in which direction-independent (isotropic) terms were included by replacing the entry $a\cos\Theta$ with $a(\cos\rho + \sin\rho\cos\Theta)$. This is sometimes called the BMW Model. If we want to reproduce the data, more complicated models must therefore be considered.

4.3.2 Tandem Model

The Tandem Model was developed by Kostelecký, Katori and Tayloe in 2006 [2]. It assumes a Hamiltonian of the form

$$\hat{H}_{TM} = \begin{bmatrix} cE & a & a \\ a & 0 & a \\ a & a & \frac{m^2}{2E} \end{bmatrix} \tag{4.31}$$

Notice that this model includes a real mass for ν_τ but that oscillations are still caused by Lorentz-violating fields. The Tandem Model is extremely complicated to discuss in full generality, because it involves formally solving the cubic. More details can be found in [2]. I restrict myself here to the high and low energy limits, which are

$$H_{highE} = \begin{bmatrix} cE & a & a \\ a & 0 & a \\ a & a & 0 \end{bmatrix} \tag{4.32}$$

and

$$H_{lowE} = \begin{bmatrix} 0 & a & a \\ a & 0 & a \\ a & a & \frac{m^2}{2E} \end{bmatrix} \tag{4.33}$$

The low- and high-energy eigenvalues are

$$\lambda_{lowE} = -a, \frac{2aE + m^2}{4E} \pm \sqrt{\frac{m^4}{4E^2} - \frac{m^2 a}{E} + 9a^2} \tag{4.34}$$

$$\lambda_{highE} = -a, \frac{a + cE}{2} \pm \sqrt{(cE)^2 - 2cEa + 9a^2}. \tag{4.35}$$

From this it is clear that as the system passes from the low to the high energy limit, cE comes to play the same role as $m^2/2E$ (which is also apparent from the form of the Hamiltonian). However, this means that the energy-dependence of the model becomes inverted as the energy increases.

The Hamiltonian in the high energy limit can be diagonalized by the matrix

$$U = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ \frac{1}{\sqrt{2}}\sin\theta & \frac{-1}{\sqrt{2}}\cos\theta & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}}\sin\theta & \frac{-1}{\sqrt{2}}\cos\theta & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{4.36}$$

where $\sin^2\theta = 1 - \frac{4a^2}{D^2+(a-cE)D}$ where $D = \sqrt{(cE)^2 - 2cEa + 9a^2}$. We can use this matrix to calculate the transition probabilities as in the case of the bicycle model. In fact, the probabilities now have exactly the same form, except that θ has been redefined and now

$$\begin{aligned} \Delta_{31} &= \frac{1}{2}(cE + 3a + D) \\ \Delta_{21} &= \frac{1}{2}(cE + 3a - D) \\ \Delta_{32} &= D \end{aligned} \tag{4.37}$$

Again taking strong Taylor approximations in the high energy limit we can get simplified transition probabilities:

$$P_{ee} = 1 \tag{4.38}$$

$$P_{e\mu} = P_{e\tau} = 0 \tag{4.39}$$

$$P_{\mu\mu} = P_{\tau\tau} = 1 - \sin^2(aL) \tag{4.40}$$

$$P_{\mu\tau} = \sin^2(aL) \tag{4.41}$$

For the Low Energy limit the diagonalizing matrix is

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\cos\theta & \frac{-1}{\sqrt{2}}\sin\theta \\ \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\cos\theta & \frac{-1}{\sqrt{2}}\sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \tag{4.42}$$

where we now define $\sin^2\theta = 1 - \frac{4a^2}{D^2+(\frac{m^2}{2E}-a)D}$ where $D = \sqrt{\frac{m^4}{4E^2} - \frac{m^2 a}{E} + 9a^2}$. This yields transition probabilities

$$P_{ee} = P_{\mu\mu} = 1 - \sin^2\theta\cos^2\theta\sin^2(\frac{DL}{2}) - \sin^2\theta\sin^2(\frac{\frac{m^2}{2E} + 3a + D}{4}L) - \cos^2\theta\sin^2(\frac{\frac{m^2}{2E} + 3a - D}{4}L) \tag{4.43}$$

$$P_{e\mu} = \sin^2\theta\sin^2(\frac{\frac{m^2}{2E} + 3a + D}{4}L) + \cos^2\theta\sin^2(\frac{\frac{m^2}{2E} + 3a - D}{4}L) - \sin^2\theta\cos^2\theta\sin^2(\frac{D}{2}L) \tag{4.44}$$

$$P_{e\tau} = P_{\mu\tau} = 2\sin^2\theta\cos^2\theta\sin^2(\frac{D}{2}L) \tag{4.45}$$

$$P_{\tau\tau} = 1 - 4\sin^2\theta\cos^2\theta\sin^2(\frac{D}{2}L) \tag{4.46}$$

Note that this is simply a flavor-permutation of the previous (high-energy) probabilities, where $e \leftrightarrow \tau$, once Δ_{ij} has been appropriately redefined. We can again calculate the limiting results in the low energy limit $E \ll \frac{m}{\sqrt{2}c}$, $E \ll \frac{m^2}{2a}$. We have:

$$P_{ee} = P_{\mu\mu} = 1 - \sin^2(\frac{m^2L}{2E}) - \frac{a^2E}{2m^4}\sin^2(aL) \approx 1 - \sin^2(\frac{m^2L}{2E}) \tag{4.47}$$

$$P_{e\mu} = (1 - \frac{a^2E}{m^4})\sin^2(\frac{m^2L}{2E}) + \frac{a^2E}{2m^4}\sin^2(aL) \approx \sin^2(\frac{m^2L}{2E}) \tag{4.48}$$

$$P_{e\tau} = P_{\mu\tau} = \frac{a^2E}{m^4}\sin^2(\frac{mL}{2}) \approx 0 \tag{4.49}$$

$$P_{\tau\tau} = 1 - \frac{2a^2E}{m^4}\sin^2(\frac{mL}{2}) \approx 1 \tag{4.50}$$

We see from all of these results that having a single on-diagonal term much larger than any of the other terms in the Hamiltonian effectively stops oscillations for that flavor.

Both these models employ Lorentz Violations to cause (in whole or in part) neutrino oscillations; furthermore, they both use a seesaw mechanism to predict different behavior in different energy ranges. Both of these features will be repeated in the new models about to be discussed; however, the new models will be more general in form than the two considered here.

Chapter 5

New Models

In this section, a series of original Lorentz-violating models are introduced and their behavior is explored. The early models are designed to give a general impression of the possible results in a simple (fully solvable) setting before approximation methods must be employed.

We know that a mass model can explain a large portion of the experimental data, and so any theory of neutrino oscillations must have mass-like behavior at least in the regions where this behavior has been observed. In particular, when Standard Model Extension terms are introduced, the characterizing feature is the energy dependence - thus, we hope to find E^{-1} energy dependence over a considerable region in order to mimic mass. Producing such behavior will be the motivation behind the models proposed below.

The full Hamiltonian used in these models was given above (4.1), but we here make the simplifying assumption that the Lorentz-violating parameters $g = H = 0$, so we can focus on just one quadrant of the Hamiltonian, of the form

$$H = \frac{1}{E}\begin{bmatrix} -c_L^{\mu\nu}p_\mu p_\nu + a_L^\mu p_\mu & -c_M^{\mu\nu}p_\mu p_\nu + a_M^\mu p_\mu \\ -c_M^{\dagger\mu\nu}p_\mu p_\nu + a_M^{\dagger\mu}p_\mu & -c_R^{T\mu\nu}p_\mu p_\nu - a_R^{T\mu}p_\mu \end{bmatrix}. \tag{5.1}$$

This choice is physically equivalent to preventing neutrino-antineutrino oscillations. Neutrinos then oscillate only into other flavors and other helicities. The behavior of the antiparticles will be closely related. Note that the basis used for expressing this Hamiltonian is $\{\nu_L, \bar{\nu}_R\}$. ν_L represents the three generations $\{\nu_e, \nu_\mu, \nu_\tau\}$ of the standard active neutrinos, while $\bar{\nu_R}$ represents the three corresponding sterile anti-neutrinos. Thus c_L and a_L control mixing among active neutrinos, c_M and a_M control mixing between active and sterile neutrinos, and c_R and a_R control mixing among sterile neutrinos, though these last effects are not physical. Also, observe that this quadrant (5.1) is actually a 6×6 matrix, since three generations are

considered.

5.1 Special Case (Tricycle Model)

In this section we will develop a simple model which has the benefit of being exactly solvable. In general, we will focus on Hamiltonians of the form

$$H = \begin{bmatrix} 0 & a_M \\ a_M^\dagger & c_R E \end{bmatrix}. \tag{5.2}$$

This is a special case of the general Hamiltonian (5.1) in which we have selected $a_L = a_R = c_L = c_M = 0$, and we are looking only at the isotropic components of the tensors. This relatively simple class of Hamiltonians will have all the properties in which we are interested. For instance, at high energy, the right-handed component of the matrix will dominate, which will cut off oscillations between left- and right-handed neutrinos. At low energy, the matrix will be off-diagonal, so there will be mixing between active and sterile neutrinos. Because sterile neutrinos cannot be observed, such oscillations would appear experimentally as spontaneous neutrino creation or annihilation. Since this has not been observed in the experimental energy domains, it is important that these oscillations are suppressed at high energy.

To be more quantitative about the behavior predicted by this model, we must find the eigenvalues of (5.2). Although it is straightforward to write down an equation for the eigenvalues of a 2×2 matrix, finding the eigenvalues of larger matrices quickly becomes difficult because it involves solving general polynomials of increasing degree, which becomes impractical for a 3×3 matrix, and analytically impossible for a 5×5 matrix. Therefore a different approach is required for finding the eigenvalues of large matrices. This is block diagonalization. Because this is somewhat technical and complicated, we have put the work on this subject in an appendix.

We can see (from A.13) that in the high energy limit, the Hamiltonian in question (5.2) will be block diagonalized to

$$H' = \begin{bmatrix} -a_M (c_R E)^{-1} a_M & 0 \\ 0 & c_R E \end{bmatrix}. \tag{5.3}$$

Thus we see that the conventional E^{-1} energy dependence will arise for the active neutrinos in the high-energy limit. Treating even this relatively simple Hamiltonian in full generality is rather complicated, so we leave this problem for section 5.3 and here begin by building intuition with an exactly solvable version, in which we select a_M to be real symmetric and c_R to be any matrix which commutes with a_M. At this point we drop the subscripts since there is no room for confusion. Thus we have

$$H_T = \begin{bmatrix} 0 & A \\ A & CE \end{bmatrix} \tag{5.4}$$

where both A and C are 3×3 matrices. In order to illustrate what is going on, we simplify even further and treat the case of one flavor of neutrino, where we have only a 2×2 matrix

$$\begin{bmatrix} 0 & a \\ a & cE \end{bmatrix}. \tag{5.5}$$

It is now straightforward to calculate the eigenvalues directly from the definition. We find $\lambda = \{\frac{cE \pm \sqrt{(cE)^2+4a^2}}{2}\}$. We also know that (5.5) can be diagonalized by a rotation matrix of the form (3.9). Using these facts, one can easily find the value of θ as in the models explored in Chapter 4. I find

$$\sin^2\theta = \frac{1}{2}[1 - \frac{1}{\sqrt{1+(\frac{2a}{cE})^2}}] \tag{5.6}$$

Notice that for $cE \gg a$, we get $\sin^2\theta \approx 0$, or $\theta \approx 0$. Furthermore, if $cE \ll a$, then $\sin^2\theta \approx \frac{1}{2}$. These are the high- and low-energy limits of the model. Thus for the one-generation model, we have a disappearance probability of $P = 4\sin^2\theta\cos^2\theta\sin^2(cEL)$. We see that $P \to 0$ in the high-energy limit. Thus we see that mixing is decreased as one of the components of the Hamiltonian dominates.

This can now be easily generalized to the three flavor case. We diagonalize this matrix in two steps: we first diagonalize A and C to bring the Hamiltonian into the form

$$\begin{bmatrix} 0 & 0 & 0 & a_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_3 \\ a_1 & 0 & 0 & c_1 & 0 & 0 \\ 0 & a_2 & 0 & 0 & c_2 & 0 \\ 0 & 0 & a_3 & 0 & 0 & c_3 \end{bmatrix} \tag{5.7}$$

where $\{a_i\}$ are the eigenvalues of A and $\{c_i\}$ are the eigenvalues of C. In this form, the Hamiltonian is, in effect, three intermeshed copies of the 2×2 matrix of the form (5.5) analyzed above. We can therefore fully diagonalize by a second transformation made up of three intermeshed copies of (3.9).

To perform the first step, we make use of the following idea: we see that because the sterile neutrinos are not measurable we can always choose a basis for $\{\nu_R\}$ in which C is diagonal. Furthermore, recall that commuting matrices can be simultaneously diagonalized. Then the same transformation, applied to the left- and right-handed bases at the same time, will diagonalize both A and C. We assume that because our results

must be consistent with the experimental results, the change of basis needed to diagonalize C will have the form

$$U = \begin{bmatrix} U_s^\dagger & 0 \\ 0 & U_s^\dagger \end{bmatrix} \tag{5.8}$$

where U_s is the standard diagonalizing matrix as defined in (3.11). In essence, this partially reparameterizes the model in terms of the conventional mixing angles, making it easier to compare the results to experimental data. We can now perform the second step and fully diagonalize the matrix by conjugating by

$$V = \begin{bmatrix} \mathrm{Cos}\Theta & \mathrm{Sin}\Theta \\ -\mathrm{Sin}\Theta & \mathrm{Cos}\Theta \end{bmatrix} = \begin{bmatrix} \cos\alpha_1 & 0 & 0 & \sin\alpha_1 & 0 & 0 \\ 0 & \cos\alpha_2 & 0 & 0 & \sin\alpha_2 & 0 \\ 0 & 0 & \cos\alpha_3 & 0 & 0 & \sin\alpha_3 \\ -\sin\alpha_1 & 0 & 0 & \cos\alpha_1 & 0 & 0 \\ 0 & -\sin\alpha_2 & 0 & 0 & \cos\alpha_2 & 0 \\ 0 & 0 & -\sin\alpha_3 & 0 & 0 & \cos\alpha_3 \end{bmatrix}. \tag{5.9}$$

Thus we finally have

$$V^\dagger U^\dagger H U V = H' \tag{5.10}$$

where H' is the matrix of eigenvalues

$$\lambda_{i^\pm} = \frac{c_i E \pm \sqrt{(c_i E)^2 + 4a_i^2}}{2} \tag{5.11}$$

The mixing angles are defined in analogy with the one dimensional case (5.6):

$$\sin^2\alpha_i = \frac{1}{2}[1 - \frac{1}{\sqrt{1 + (\frac{2a_i}{c_i E})^2}}]. \tag{5.12}$$

The transition amplitude matrix S is then given by

$$S = UVe^{-itH'}V^\dagger U^\dagger = \begin{bmatrix} S_{LL} & S_{LR} \\ S_{RL} & S_{RR} \end{bmatrix} \tag{5.13}$$

where each block is a 3×3 matrix describing oscillations among one of the kinds of neutrinos. For instance, S_{LL} is the transition amplitudes among the active neutrinos, while S_{LR} describes oscillations from active to sterile neutrinos. If we define

$$\lambda_+ = \begin{bmatrix} \lambda_{1+} & 0 & 0 \\ 0 & \lambda_{2+} & 0 \\ 0 & 0 & \lambda_{3+} \end{bmatrix}, \tag{5.14}$$

and λ_- in analogy, then we can write the transition matrix among the left-handed neutrinos in the form

$$S_{LL} = U_s^\dagger \mathrm{Sin}\Theta e^{-it\lambda_+}\mathrm{Sin}\Theta U_s + U_s^\dagger \mathrm{Cos}\Theta e^{-it\lambda_-}\mathrm{Cos}\Theta U_s \tag{5.15}$$

and numbering the eigenvalues one through six in the order $\{\lambda_{1^+}, \lambda_{2^+}, \lambda_{3^+}, \lambda_{1^-}, \lambda_{2^-}, \lambda_{3^-}\}$ we can write down the transition probabilities in terms of the various parameters by taking the absolute squares of the entries of the transition matrix S_{LL}, as for the models in Chapter 4:

$$\begin{aligned} P_{ee} = 1 &- 4\cos^4\theta\sin^2\alpha_1\cos^2\alpha_1\sin^2\left(\frac{\Delta\lambda_{41}L}{2}\right) - 4\sin^4\theta\sin^2\alpha_2\cos^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{52}L}{2}\right) \\ &-4\sin^2\theta\cos^2\theta\left\{\sin^2\alpha_1\sin^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{21}L}{2}\right) + \sin^2\alpha_1\cos^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{51}L}{2}\right)\right. \\ &\left. + \cos^2\alpha_1\sin^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{24}L}{2}\right) + \cos^2\alpha_1\cos^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{54}L}{2}\right)\right\} \end{aligned} \tag{5.16}$$

$$\begin{aligned} P_{e\mu} = 4\sin^2\theta\cos^2\theta\cos^2\phi\Bigg\{&-\sin^2\alpha_1\cos^2\alpha_1\sin^2\left(\frac{\Delta\lambda_{41}L}{2}\right) - \sin^2\alpha_2\cos^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{52}L}{2}\right) \\ &+\sin^2\alpha_1\sin^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{21}L}{2}\right) + \sin^2\alpha_1\cos^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{51}L}{2}\right) \\ &+\cos^2\alpha_1\sin^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{24}L}{2}\right) + \cos^2\alpha_1\cos^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{54}L}{2}\right)\Bigg\} \end{aligned} \tag{5.17}$$

$$\begin{aligned} P_{e\tau} = 4\sin^2\theta\cos^2\theta\sin^2\phi\Bigg\{&-\sin^2\alpha_1\cos^2\alpha_1\sin^2\left(\frac{\Delta\lambda_{41}L}{2}\right) - \sin^2\alpha_2\cos^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{52}L}{2}\right) \\ &+\sin^2\alpha_1\sin^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{21}L}{2}\right) + \sin^2\alpha_1\cos^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{51}L}{2}\right) \\ &+\cos^2\alpha_1\sin^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{24}L}{2}\right) + \cos^2\alpha_1\cos^2\alpha_2\sin^2\left(\frac{\Delta\lambda_{54}L}{2}\right)\Bigg\} \end{aligned} \tag{5.18}$$

Notice that α_3 does not appear in any of these expressions. This is because $\theta_{13} = 0$, which has simplified these particular oscillation probabilities. α_3 does appear in the other probabilities, which can be similarly computed; however, these extra terms make them quite lengthy expressions, so I have omitted them here.

These expressions can be simplified in the high energy limit. If we assume $E \gg \frac{2a}{c}$ we know that α and β approach 0 (this follows from their definitions), and we can also apply a Taylor expansion to the eigenvalues to find that, for example,

$$P_{e\mu} = 4\sin^2\theta\cos^2\theta\cos^2\phi\sin^2\left(\frac{(\frac{a_1^2}{c_1E} - \frac{a_2^2}{c_2E})L}{2}\right) \tag{5.19}$$

and from comparison to (3.13) we can see that, to be consistent with with mass model, we must have

$$\frac{a_1^2}{c_1} - \frac{a_2^2}{c_2} = \frac{\Delta m_{21}^2}{2} \tag{5.20}$$

From this we see that the models are consistent in the high energy limit if we choose

$$\frac{a_i^2}{c_i} = \frac{1}{2}m_i^2 \tag{5.21}$$

In the low energy limit, we see from the discussion in (A.1.2) that the a terms dominate, and the oscillation lengths become energy-independent. We define a set of critical energies

$$E_{c_i} = \frac{a_i}{c_i}. \tag{5.22}$$

If $E \gg E_c$, the model will look asymptotically similar to the mass model. If $E \ll E_c$, the oscillations become energy-independent and typically occur only at lengths longer than those predicted by the mass model. When $E \approx E_c$, the model can predict enhanced oscillations at short distances compared to the mass model (see Figure 5.2). There are three different energy scales, one for each neutrino flavor, which can be individually tuned.

Considered from these phenomenological parameters, we have four choices to make for the model: the energies at which the seesaw mechanism is activated (the E_c's), and an overall mass term which allows us to determine the relative widths of the intermediate zones. The mixing angles θ and ϕ in U are fixed by experiment. These four parameters can all be varied without altering the mass-model behavior in the high-energy limit.

We note that $P_{ee} + P_{e\mu} + P_{e\tau} < 1$ if $\alpha_i \neq 0$. In these Lorentz-violating cases, $P_{LR} \neq 0$. These oscillations are given by

$$S_{LR} = U_s^\dagger [Cos\Theta e^{-it\lambda_-} Sin\Theta - Sin\Theta e^{-it\lambda_+} Cos\Theta] U_s \tag{5.23}$$

We calculate as an example

$$\begin{aligned} P_{e\bar{e}} = {} & 4\cos^4\theta \sin^2\alpha \cos^2\alpha \sin^2(\frac{\Delta\lambda_{41}L}{2}) + 4\sin^4\theta \sin^2\beta \cos^2\beta \sin^2(\frac{\Delta\lambda_{52}L}{2}) \\ & + 4\sin^2\theta \cos^2\theta \sin\alpha \cos\alpha \sin\beta \cos\beta [\sin^2(\frac{\Delta\lambda_{42}L}{2}) + \sin^2(\frac{\Delta\lambda_{51}L}{2}) - \sin^2(\frac{\Delta\lambda_{54}L}{2}) - \sin^2(\frac{\Delta\lambda_{21}L}{2})] \end{aligned} \tag{5.24}$$

The other probabilities can be similarly computed. The left-to-right transition probabilities go to zero at high energies, but can be quite significant at low energies.

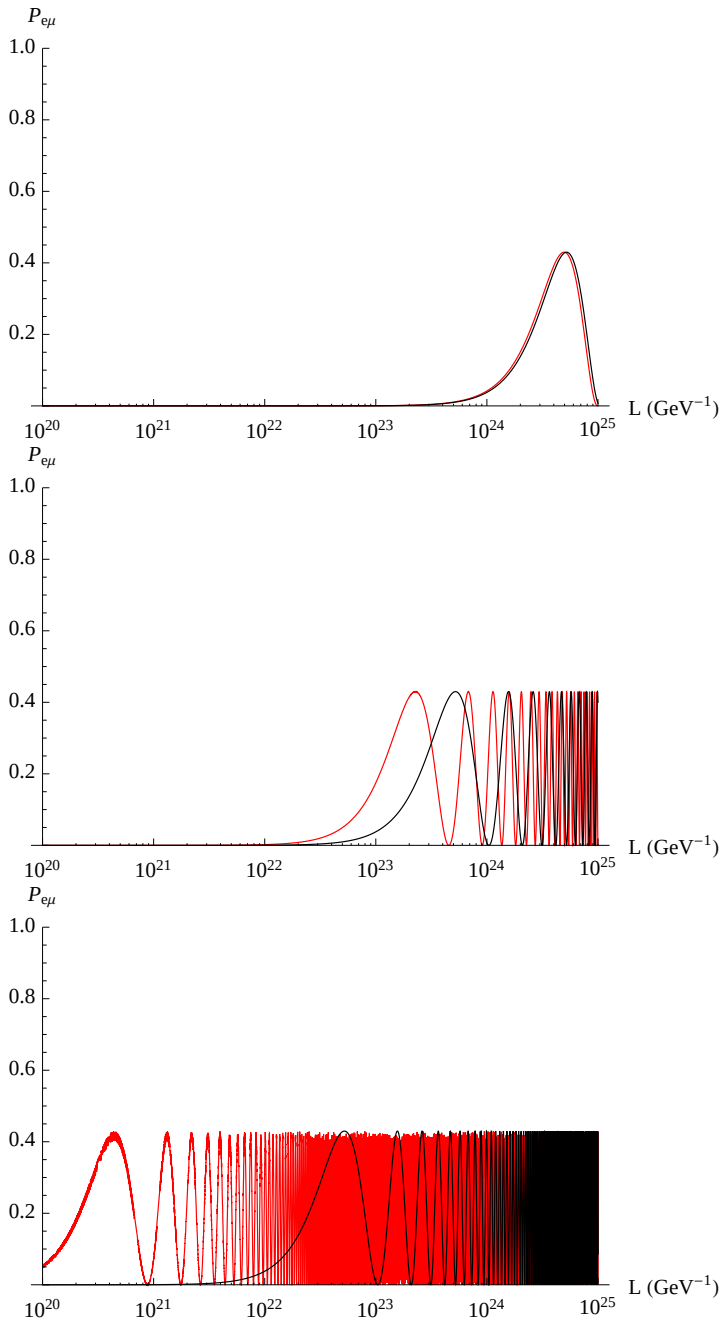

Figure 5.1: Oscillation Probability for the electron to muon neutrino in the mass model (black) and the model discussed here (red). The three plots are shown for decreasing energies (E = 63 GeV, E = 6.3 GeV, E = 0.63 GeV). The Horizontal axis is distance, in units of GeV^{-1}. Notice how at high energies, the model discussed here is indistinguishable from the mass model - only as the energy decreases do the behaviors begin to diverge and then become unrelated. The parameter values used here are $c_1 = 6.03 \times 10^{-17}, c_2 = 1.66 \times 10^{-17}, c_3 = 10^{-17}, a_1 = 8.92 \times 10^{-18}$ GeV, $a_2 = 4.68 \times 10^{-18}$ GeV, $a_3 = 3.63 \times 10^{-18}$ GeV (see 5.30).

5.1.1 Eigenvalue Behavior

We begin by looking at a plot of the eigenvalues: observe that the there are three low-energy limits to the

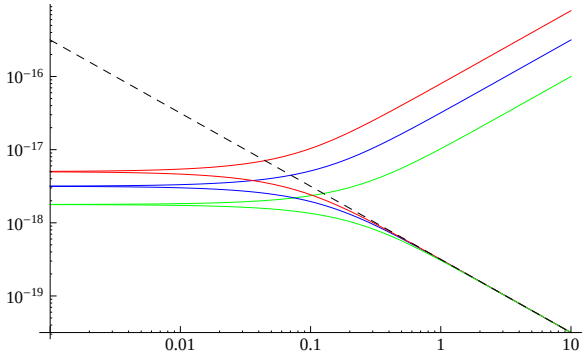

Figure 5.2: The behavior of the six eigenvalues of the special model, as a function of energy, for typical parameter values. The dotted line indicates a typical mass line.

eigenvalues, one for each flavor, and that from each of these values, two eigenvalues diverge at high energies, the one sloping down corresponding to active neutrinos, the one sloping up corresponding to sterile neutrinos. At high energies, the eigenvalues have the behavior of masses, which is shown in Figure 5.2 by the dotted line.

Because the characteristic oscillation length is given by the eigenvalue differences, the probabilities predicted by the model can be summarized by looking at how the eigenvalue differences behave at different energies. If we look in the E-L plane [9], we can schematically determine whether oscillations will be present under different parameters regimes purely on the basis of the locations of the eigenvalue difference lines.

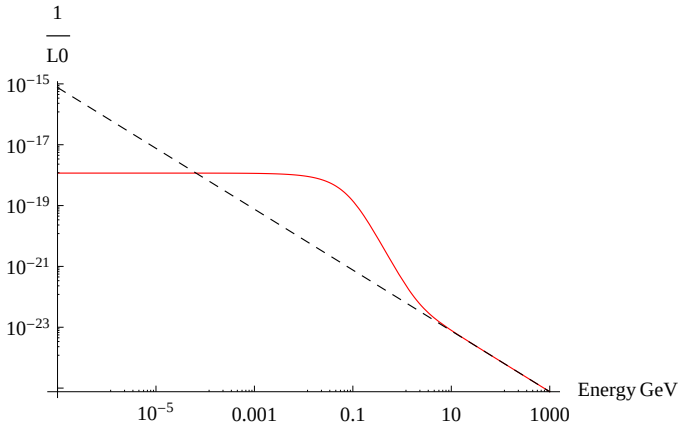

Figure 5.3: The dashed line represents the inverse oscillation length in the mass model. The solid line is the inverse oscillation length in the new model, using the same parameter values as above.

Shown in Figure 5.3 is the value of $\Delta\lambda_{54}$ as a function of energy, together with $\frac{\Delta m^2_{21}}{E}$. We see that the energy spectrum is partitioned into three regions: at the highest energy the two models agree; at low energies, the mass model predicts stronger oscillations; but in the middle region, the new model predicts stronger oscillation.

5.1.2 CPT Symmetry

CPT violations are of particular interest because some recent experiments (MINOS, MiniBooNE) have hinted that neutrino oscillations are different for neutrinos and antineutrinos, which might represent a violation of CPT symmetry. Furthermore, any CPT violation entails Lorentz violation, so if such an effect were decisively measured, it would be a clear sign that at least one of the Lorentz-violating parameters is non-zero. Additionally, such a measurement would show decisively that neutrino mass cannot be the only cause of oscillation, because neutrinos and antineutrinos must have the same mass in any local Quantum Field Theory [17].

The effects of the Special Model discussed above are invariant under CPT transformations. Such transformations introduce negative signs to the CPT-odd coefficients (in this case A). However, the oscillation probabilities do not depend on the sign of A. The probabilities are a function of the eigenvalues and the mixing angles α, β, and γ, and it can be seen from their definitions that they are all independent of the sign of A.

More generally speaking, it is possible to prove that any Hamiltonian of the form

$$H = \begin{bmatrix} 0 & A \\ A^\dagger & CE \end{bmatrix} \tag{5.25}$$

will be CPT-invariant. Let us represent the CPT-transformed Hamiltonian by H'. Then it is easy to check that H and H' are similar matrices[1]:

$$H = \begin{bmatrix} 0 & A \\ A^\dagger & CE \end{bmatrix} \sim \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} 0 & A \\ A^\dagger & CE \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} = \begin{bmatrix} 0 & -A \\ -A^\dagger & CE \end{bmatrix} = H' \tag{5.26}$$

This shows that in general, the transformed eigenvalues are unchanged. Furthermore, the two Hamiltonians share eigenvectors that differ only by sign, componentwise. This can be seen by considering that if x is an eigenvector of H, then

$$x' = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} x \tag{5.27}$$

[1]Recall that two matrices A and B are called *similar* if there exists a third matrix M such that $A = MBM^{-1}$. Similar matrices share eigenvalues and have related eigenvectors.

will be an eigenvector of H' corresponding to the same eigenvalue. Thus we have

$$S' = U'^{\dagger} e^{-itH'} U' = \begin{bmatrix} U_1^T \\ -U_2^T \end{bmatrix} e^{-itH} \begin{bmatrix} U_1 & -U_2 \end{bmatrix} = \begin{bmatrix} S_{LL} & -S_{LR} \\ -S_{RL} & S_{RR} \end{bmatrix} \tag{5.28}$$

from which it can be seen that the probabilities are unchanged by a CPT transformation, since the probabilities are the squares of the entries of S, and the signs in S are therefore unimportant. Any Hamiltonian of the form (5.17) will preserve CPT symmetry; it can only be broken by the mixing of A and C terms, such as would occur by rotating the model. It is this idea we are about to pursue.

5.1.3 Data fitting

We want to know whether the new model proposed here is able to explain all the data on neutrino oscillations, or at least do as well as the mass model. To do this we examine particular numerical parameter values and see how well the data is fit.

We first remark that the parameters in these models can be selected to be completely consistent with the mass model over the range of energies probed to date. As was discussed (5.17), we can select the energy at which the seesaw mechanism turns off. Thus if we take a to be large enough, we can push the cut-off energy below the level of experimental data. This is a sort of degenerate solution, since there is no need to posit the additional complications of Lorentz-violation if the mass model is simply reproduced, although it does show that these models cannot be immediately ruled out by the data.

It would be more interesting, however, if the model could look like the mass model but also explain some of the anomalous data. We have examined in some detail the predictions of the basic model made with the following parameters

$$\begin{aligned} a_1 &= 8.92 \times 10^{-18} \text{ GeV} \\ a_2 &= 4.68 \times 10^{-18} \text{ GeV} \\ a_3 &= 3.63 \times 10^{-18} \text{ GeV} \\ c_1 &= 6.03 \times 10^{-17} \\ c_2 &= 1.66 \times 10^{-17} \\ c_3 &= 1.00 \times 10^{-17} \end{aligned} \tag{5.29}$$

These were the parameter values used to produce the plots already shown (Figure 5.1). They were selected because they match reasonably well the experimental results seen at LSND and MiniBooNE. Although the fit is not perfect, this demonstrates that probabilities in the correct range can be predicted. This model also

reproduces the mass predictions at high energies. Since it is likely to agree with the high-energy experiments (those above the critical energies of about 100 MeV), we investigated whether the model agrees with the results of some low-energy experiments. It does match the results of Bugey and Gosgen, predicting only negligible oscillation probabilities in the experimental region tested ($P_{e\mu} \approx .05\%$). Further research is needed to determine whether this model can account for other experimental results.

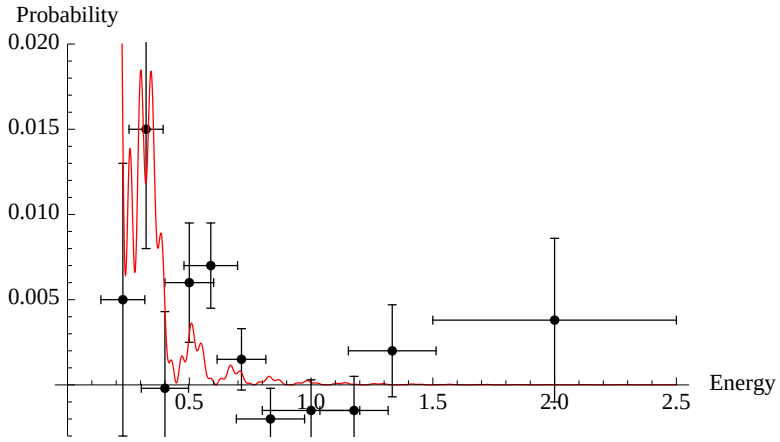

Figure 5.4: Data from MiniBooNE interpreted from [18] (see Figure 3.6), overlain with prediction of the model with the given parameters. Energy in GeV.

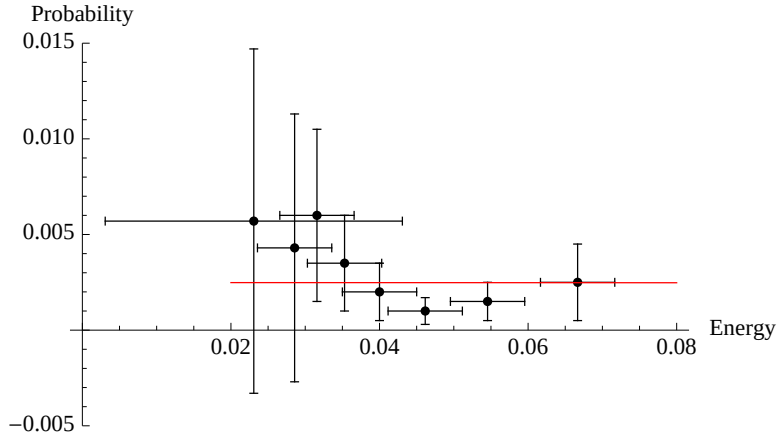

Figure 5.5: Data from LSND interpreted from [18], overlain with the model with the given parameters. Energy in GeV.

5.2 Rotated Model

There is a strong motivation to consider a model which, unlike the special case already discussed, allows for CPT violations, because experimental results (most notably from MiniBooNE) are beginning to hint at a CPT violation in the neutrino sector, in the form of different behavior for neutrinos and antineutrinos.

We therefore consider rotating the model discussed above through an angle ϵ, which will mix the A and C terms. This is equivalent to considering the Hamiltonian of the previous model to be written in a new basis of mixed left- and right-handed states. We can write it in the traditional basis by using a rotation matrix R (of the form (5.8) if $\alpha_i = \epsilon$ for all i). We introduce the following notation: Let the matrix M be considered as a many-dimensional function of the angle θ. Then M_ϵ will represent $M(\theta + \epsilon)$. This simplifies the following argument. Then the rotated model H_ϵ is given in terms of the old model H by $H_\epsilon = RHR^\dagger$, or

$$H_\epsilon = \begin{bmatrix} -\sin^2 \epsilon CE & A + \sin\epsilon\cos\epsilon CE \\ A - \sin\epsilon\cos\epsilon CE & \cos^2 \epsilon CE \end{bmatrix}. \tag{5.30}$$

The calculation of probabilities for the rotated model is straightforward given the probabilities of the unrotated model. Then the transition amplitude matrix of the rotated model, S, will be given by

$$\begin{aligned} S &= U_{total} e^{-itU_{total}^\dagger H_\epsilon U_{total}} U_{total}^\dagger \\ &= RUV e^{-itV^\dagger U^\dagger R^\dagger H_\epsilon RUV} V^\dagger U^\dagger R^\dagger \\ &= RUV e^{-itV^\dagger U^\dagger HUV} V^\dagger U^\dagger R^\dagger \\ &= RUV e^{-itH'} V^\dagger U^\dagger R^\dagger \end{aligned} \tag{5.31}$$

This simply says that S acts on the same eigenvalues as in the unrotated model, which is fairly obvious. But we see that there is an additional term contributing to the mixing angles. However, since U is block diagonal, it commutes with R; furthermore, because R and V are both rotation matrices, they multiply to form a single rotation matrix:

$$\begin{aligned} S &= URV e^{itH'} V^\dagger R^\dagger U^\dagger \\ &= UV_\epsilon e^{-itH'} V_\epsilon^\dagger U^\dagger \\ &= S_\epsilon \end{aligned} \tag{5.32}$$

This tells us that the transition probabilities for the rotated model can be written in terms of the probabilities

for the unrotated model by adding a small constant to the mixing angles. For example:

$$\begin{aligned} P_{e\mu} = 4\sin^2\theta\cos^2\theta\cos^2\phi[&-\sin^2(\alpha_1+\epsilon)\cos^2(\alpha_1+\epsilon)\sin^2\left(\frac{\Delta\lambda_{41}L}{2}\right) - \sin^2(\alpha_2+\epsilon)\cos^2(\alpha_2+\epsilon)\sin^2\left(\frac{\Delta\lambda_{52}L}{2}\right) \\ &+\sin^2(\alpha_1+\epsilon)\sin^2(\alpha_2+\epsilon)\sin^2\left(\frac{\Delta\lambda_{21}L}{2}\right) + \sin^2(\alpha_1+\epsilon)\cos^2(\alpha_2+\epsilon)\sin^2\left(\frac{\Delta\lambda_{51}L}{2}\right) \\ &+\cos^2(\alpha_1+\epsilon)\sin^2(\alpha_2+\epsilon)\sin^2\left(\frac{\Delta\lambda_{24}L}{2}\right) + \cos^2(\alpha_1+\epsilon)\cos^2(\alpha_2+\epsilon)\sin^2\left(\frac{\Delta\lambda_{54}L}{2}\right)]. \end{aligned} \tag{5.33}$$

5.2.1 CPT Symmetry

To see what happens to this model under CPT transformations, we begin by observing that

$$H'_\epsilon = \begin{bmatrix} -\sin^2\epsilon CE & -A+\sin\epsilon\cos\epsilon CE \\ -A-\sin\epsilon\cos\epsilon CE & \cos^2\epsilon CE \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} H_{-\epsilon} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \sim H_{-\epsilon} \tag{5.34}$$

Therefore, drawing on the analysis of 5.1.2, we see that S'_ϵ looks like $S_{-\epsilon}$ up to a sign, so $P'_\epsilon \to P_{-\epsilon}$. Thus the probabilities transform by replacing ϵ with $-\epsilon$. CPT symmetry is broken in this model because although $\sin^2\alpha$ is invariant under CPT transformations, $\sin^2(\alpha+\epsilon)$ is not. In general, the magnitude of the CPT violation will depend of the size of the parameter a.

5.3 General Massless Model

In the previous sections we wanted to solve for the probabilities exactly, and so we imposed a strong condition on our model: we assumed that A and C commute. It would be preferable to be able to consider all models of the general form (5.2), so that the entire class could be compared to experimental results. However, we will no longer be able to find the exact eigenvalues in the general case. We will need to use the block diagonalization when the submatrices do not commute.

We can always assume that C is diagonal without loss of generality, because we are free to choose any basis we like for the right-handed neutrinos, since they have not been observed. Furthermore, recall that we have seen in the special case that it is the high energy limit which approximates the mass model. Since we know the mass model describes most of the data well, we will consider only the high-energy limit, when the Lorentz violating model looks asymptotically similar to the mass model, and we deduce the correction terms to the mass model. Remember that at high energy

$$\begin{bmatrix} 0 & A \\ A^\dagger & CE \end{bmatrix} \sim \begin{bmatrix} -AC^{-1}A^\dagger/E & 0 \\ 0 & CE \end{bmatrix} \tag{5.35}$$

(to first order). Observe that if this is to be consistent with the mass model, this puts some restrictions on the form of A - exactly how A is restricted is not obvious, but a parameterization of A will be given later. It is the E^{-1} dependence of some of the eigenvalues that allows the model to mimic mass behavior at high energies. We assume that U_s still diagonalizes $-AC^{-1}A^\dagger$ into the mass squared matrix M^2, so that the regular mass predictions come out to first order. We will find the next-to-leading order terms in the eigenvalues and diagonalizing matrix, since these will correspond to the Lorentz-violating corrections to the mass model. We are interested, then, in what the block diagonalization looks like to higher order. The reader is again referred to the appendix for the details.

We remind the reader of the summation form of writing transition probabilities introduced in (3.19). We can write

$$P_{\alpha\beta} = \delta_{\alpha\beta} - 4\sum_{i>j} U_{\alpha i}U_{\beta i}U_{\alpha j}U_{\beta j}\sin^2(\frac{\Delta\lambda_{ij}L}{2}). \tag{5.36}$$

We need the corrections to both λ and U. We begin with λ. The argument proceeds in three steps. We first block diagonalize H to next-to-leading order (because we are assuming that the leading order term gives the mass prediction). We then approximately diagonalize the block eigenvalue with U_s since it is approximately the mass matrix. Finally, we use standard perturbation theory to find the small corrections to the eigenvalues.

The next-to-leading order correction to the block eigenvalues occurs at third order. This is discussed in the appendix. We see (A.24) that we can write

$$U\begin{bmatrix} 0 & A \\ A^\dagger & CE \end{bmatrix}U^\dagger = H' = \begin{bmatrix} H_1 & 0 \\ 0 & H_2 \end{bmatrix} \tag{5.37}$$

where

$$\begin{aligned} U &= (I+\Delta_2)(I+\Delta_1+\frac{1}{2}\Delta_1^2+\frac{1}{6}\Delta_1^3) \\ &= \begin{bmatrix} I-\frac{1}{2}\frac{AC^{-2}A}{E^2} & -\frac{AC^{-1}}{E}+\frac{1}{2}\frac{AC^{-2}A^\dagger AC^{-1}}{E^3}+\frac{AC^{-1}A^\dagger AC^{-2}}{E^3} \\ \frac{C^{-1}A^\dagger}{E}-\frac{1}{2}\frac{C^{-1}A^\dagger AC^{-2}A^\dagger}{E^3}-\frac{C^{-2}A^\dagger AC^{-1}A^\dagger}{E^3} & I-\frac{1}{2}\frac{C^{-1}A^\dagger AC^{-1}}{E^2} \end{bmatrix} \end{aligned} \tag{5.38}$$

and

$$H_1 = -\frac{AC^{-1}A^\dagger}{E}+\frac{1}{2}\frac{AC^{-2}A^\dagger AC^{-1}A^\dagger}{E^3}+\frac{1}{2}\frac{AC^{-1}A^\dagger AC^{-2}A^\dagger}{E^3} \tag{5.39}$$

$$H_2 = CE+\frac{1}{2}\frac{C^{-1}A^\dagger A}{E}+\frac{1}{2}\frac{A^\dagger AC^{-1}}{E}-\frac{3}{4}\frac{C^{-1}A^\dagger AC^{-1}A^\dagger AC^{-1}}{E^2}-\frac{1}{6}\frac{C^{-1}A^\dagger AC^{-2}A^\dagger A}{E^3}-\frac{1}{6}\frac{A^\dagger AC^{-2}A^\dagger A}{E^3} \tag{5.40}$$

We can then write (to first order in δU)

$$\begin{aligned} P'_{\alpha\beta} = \delta_{\alpha\beta} &- 4\sum_{i>j} U_{\alpha i}U_{\beta i}U_{\alpha j}U_{\beta j}\sin^2(\frac{\Delta\lambda'_{ij}L}{2}) \\ &- 4\sum_{i>j} U_{\alpha i}U_{\beta i}U_{\alpha j}\delta U_{\beta j}\sin^2(\frac{\Delta\lambda'_{ij}L}{2}) - 4\sum_{i>j} U_{\alpha i}U_{\beta i}\delta U_{\alpha j}U_{\beta j}\sin^2(\frac{\Delta\lambda'_{ij}L}{2}) \\ &- 4\sum_{i>j} U_{\alpha i}\delta U_{\beta i}U_{\alpha j}U_{\beta j}\sin^2(\frac{\Delta\lambda'_{ij}L}{2}) - 4\sum_{i>j} \delta U_{\alpha i}U_{\beta i}U_{\alpha j}U_{\beta j}\sin^2(\frac{\Delta\lambda'_{ij}L}{2}) \end{aligned} \tag{5.41}$$

The question then remains how to find the λ' and to write δU exactly. To answer the first question we turn to time-independent perturbation theory. We know that H has six eigenvalues - three approximately like the mass eigenvalues (H_1) and three that look approximately like cE terms (H_2). We re-introduce the notion of a diagonal mass matrix M. We are still assuming that the flavor-basis mass matrix is diagonalized by U and that the unperturbed solution of the model is consistent with the mass model. We can then express the left-handed block diagonal matrix

$$\begin{aligned} H'_1 &= U_s\left(\frac{-AC^{-1}A^\dagger}{E} + \frac{1}{2}\frac{AC^{-2}A^\dagger AC^{-1}A^\dagger}{E^3} + \frac{1}{2}\frac{AC^{-1}A^\dagger AC^{-2}A^\dagger}{E^3}\right)U_s^\dagger \\ &= \frac{M^2}{2E} + \frac{1}{2}\frac{UAC^{-2}A^\dagger U^\dagger M^2}{2E^3} + \frac{1}{2}\frac{M^2UAC^{-2}A^\dagger U^\dagger}{2E^3} \\ &= h_0 + \delta h \end{aligned} \tag{5.42}$$

Because we are in the high-energy limit, we know that $\delta h \ll h_0$, so we know that the eigenvalues of H_1 are approximately the mass eigenvalues with corrections we can obtain from standard perturbation theory. We know the uncorrected eigenvalues are simply

$$\lambda_i^{(0)} = \frac{m_i^2}{2E} \tag{5.43}$$

and that the corresponding uncorrected eigenvectors are simply the basis vectors $\{\hat{e_i}\}$. Then perturbation theory tells us that

$$\lambda_i^{(1)} = \delta h_{ii} \tag{5.44}$$

so that

$$\lambda'_i = \lambda_i^{(0)} + \lambda_i^{(1)} = \frac{m_i^2}{2E}(I + \frac{UAC^{-2}A^\dagger U^\dagger}{E^2})_{ii} \tag{5.45}$$

We define $\epsilon = AC^{-1}$ to be the fundamental perturbing quantity that appears everywhere. Note that ϵ has units of energy but is independent of the energy of the experiment. We can then write

$$\lambda'_i = \frac{m_i^2}{2E}(1 + \frac{(U\epsilon\epsilon^\dagger U^\dagger)_{ii}}{E^2}) \tag{5.46}$$

We will denote the eigenvalues of the right-handed block diagonal matrix $\{\lambda_{\tilde{i}}\}$. By analogy with the above we have

$$\lambda'_{\tilde{i}} = c_i E(1 + \frac{(\epsilon^\dagger \epsilon)_{ii}}{E^2}) \tag{5.47}$$

Now that we have computed the corrections to the eigenvalues, we turn our attention to the changes in the diagonalizing matrix U'. Recall that we have built up U' in three steps: a block diagonalization, an approximate diagonalization from the mass model, and finally the correction due to perturbation theory. Thus we have

$$U' = \begin{bmatrix} I+J & 0 \\ 0 & I+K \end{bmatrix} \begin{bmatrix} U_s & 0 \\ 0 & I \end{bmatrix} (I+\Delta_2)(I+\Delta_1+\frac{1}{2}\Delta_1^2+\frac{1}{6}\Delta_1^3) = U + \delta U \tag{5.48}$$

where J and K are small matrices obtained from perturbation theory. Then we see

$$\delta U = \begin{bmatrix} UJ - \frac{1}{2}\frac{AC^{-2}A^\dagger}{E^2}U(I+J) & \left(\frac{AC^{-1}A^\dagger AC^{-2}}{E^3} + \frac{1}{2}\frac{AC^{-2}A^\dagger AC^{-1}}{E^3} - \frac{AC^{-1}}{E}\right)(I+K) \\ \left(\frac{C^{-1}A^\dagger}{E} - \frac{1}{2}\frac{C^{-1}A^\dagger AC^{-2}A^\dagger}{E^3} - \frac{C^{-2}A^\dagger AC^{-1}A^\dagger}{E^3}\right)U(I+J) & K - \frac{1}{2}\frac{C^{-1}A^\dagger AC^{-1}}{E^2}(I+K) \end{bmatrix} \tag{5.49}$$

This is cumbersome, and we really only need δU to leading order, but to determine what is important to keep, we need to know the orders of J and K. This can be computed, again using time-independent perturbation theory, because we are really asking at what order the eigenvectors of the perturbed Hamiltonian deviate from the standard basis vectors. This leads to

$$J_{ij} = \begin{cases} \frac{\delta h_{ji}}{N_i \Delta\lambda_{ij}} & \text{for i} \neq \text{j} \\ 1 - \frac{1}{N_i} & \text{for i} = \text{j} \end{cases} \tag{5.50}$$

where N_i is just some normalizing factor. $N_i - 1$ is fourth order and so we take $N_i = 1$. Thus we see that J is of order δh, which is second order in ϵ. We can write

$$J_{ij} = \begin{cases} 0 & \text{for i} = \text{j} \\ \frac{1}{2}\frac{m_i^2+m_j^2}{m_i^2-m_j^2}\frac{[U\epsilon\epsilon^\dagger U^\dagger]_{ij}}{E^2} & \text{for i} \neq \text{j} \end{cases} \tag{5.51}$$

We can also explicitly calculate K, but since it will not be needed we skip it. In summary, the leading (less than third) order corrections are

$$\delta U \approx \begin{bmatrix} UJ - \frac{1}{2}\frac{\epsilon\epsilon^\dagger}{E^2}U & \frac{-\epsilon}{E} \\ \frac{\epsilon^\dagger}{E}U & K - \frac{1}{2}\frac{\epsilon^\dagger\epsilon}{E^2} \end{bmatrix} \tag{5.52}$$

Using the symmetries of the problem, we can further simplify the expressions for the transition probabilities. Here tildes represent sterile neutrinos. The corrections to the mass model are second order in ϵ. We summarize all the probabilities below in terms of ϵ and $\xi = \delta U^{LL} = UJ - \frac{1}{2}\frac{\epsilon\epsilon^\dagger}{E^2}U$:

$$P_{\alpha\alpha} = 1 - 2\sum_{i\neq j} U_{\alpha i}^2 U_{\alpha j}(U+4\xi)_{\alpha j}\sin^2\left(\frac{\Delta\lambda'_{ij}L}{2}\right) - 4\sum_{i,j} U_{\alpha i}^2\left(\frac{\epsilon_{\alpha j}}{E}\right)^2\sin^2\left(\frac{\Delta\lambda'_{i\tilde{j}}L}{2}\right) \tag{5.53}$$

$$P_{\alpha\beta} = -\sum_{i\neq j} U_{\alpha i}U_{\beta i}\left\{U_{\alpha j}(U+4\xi)_{\beta j} + (U+4\xi)_{\alpha j}U_{\beta j}\right\}\sin^2\left(\frac{\Delta\lambda'_{ij}L}{2}\right) - 4\sum_{i,j} U_{\alpha i}U_{\beta i}\frac{\epsilon_{\alpha j}\epsilon_{\beta j}}{E^2}\sin^2\left(\frac{\Delta\lambda'_{i\tilde{j}}L}{2}\right) \tag{5.54}$$

$$P_{\alpha\tilde{\beta}} = P_{\tilde{\beta}\alpha} = -2\sum_{i\neq j} U_{\alpha i}U_{\alpha j}\epsilon_{i\beta}\epsilon_{j\beta}\sin^2\left(\frac{\Delta\lambda'_{ij}L}{2}\right) + 4\sum_{i} U_{\alpha i}\frac{\epsilon_{i\alpha}\epsilon_{\alpha\beta}}{E^2}\sin^2\left(\frac{\Delta\lambda'_{i\tilde{\beta}}L}{2}\right) \tag{5.55}$$

$$P_{\alpha\tilde{\alpha}} = 1 - 2\sum_{i}\left(\frac{\epsilon_{i\alpha}}{E}\right)^2\sin^2\left(\frac{\Delta\lambda'_{i\tilde{\alpha}}L}{2}\right) \tag{5.56}$$

$$P_{\tilde{\alpha}\tilde{\beta}} = 0 \tag{5.57}$$

One of the unique predictions made by this model is a "disappearance" probability. Since the model permits the oscillation of active neutrinos into sterile neutrinos (and vice versa), and since sterile neutrinos would not be detectable, neutrinos could disappear from the flux. This is

$$P_{\alpha R} = 4\sum_{i,j} U_{\alpha i}\frac{\epsilon_{i\alpha}\epsilon_{\alpha j}}{E^2}\sin^2\left(\frac{\Delta\lambda'_{i\tilde{j}}L}{2}\right) - 2\sum_{i\neq j}\sum_{k} U_{\alpha i}U_{\alpha j}\frac{\epsilon_{ik}\epsilon_{jk}}{E^2}\sin^2\left(\frac{\Delta\lambda'_{ij}L}{2}\right) \tag{5.58}$$

5.3.1 Restrictions of the model

There is also the question of the restrictions on ϵ. We do not want to leave ϵ in terms of A because we know that consistency with the mass model imposes restrictions on A but we do not know how to parameterize the free variables. It would be better to write ϵ in terms of exclusively known or free parameters. If this model is to make predictions consistent with the mass model in the high energy limit, we must have

$$U^\dagger A C^{-1} A^\dagger U = M \tag{5.59}$$

or

$$(AC^{-1/2})(AC^{-1/2})^\dagger = UMU^\dagger \tag{5.60}$$

thus

$$\epsilon = RC^{-1/2} \tag{5.61}$$

where R is a solution to $RR^T = \tilde{M}$. The details of solving this equation can be found in the appendix, but the result tells us that

$$\epsilon = UM^{1/2}OC^{-1/2} \tag{5.62}$$

and that ϵ has seven degrees of freedom: one in the values of M (which is uncertain up to one neutrino mass), three in the general orthogonal matrix O, and the three choices of C.

As was said, the model is only a high-energy limit; in particular, both the block diagonalization and the perturbation theory used apply only when the energy is large. The particular limit which must obtain is

$$E^2 \gg \frac{|M|}{|C|} \tag{5.63}$$

5.3.2 Complex-valued parameters

We have so far assumed that the coefficients would be real, but in general they may be complex-valued. In that case, the probability expression becomes more complicated:

$$P_{\alpha\beta} = \delta_{\alpha\beta} - 4\sum_{i>j} \Re(U^*_{\alpha i}U_{\beta i}U_{\alpha j}U^*_{\beta j}) \sin^2\left(\frac{\Delta\lambda_{ij}L}{2}\right) + 2\sum_{i>j} \Im(U^*_{\alpha i}U_{\beta i}U_{\alpha j}U^*_{\beta j}) \sin^2\left(\frac{\Delta\lambda_{ij}L}{2}\right) \tag{5.64}$$

Permitting complex-valued coefficients broadens the possible values of the perturbing variable. We still assume that M and U are real (since this is a consequence of $\theta_{13} = 0$, which has been experimentally verified to be a reasonable approximation), but we now take $\epsilon = UM^{1/2}U'C^{-1/2}$ where U' is any unitary matrix. With this convention, the transition probabilities are

$$P_{\alpha\alpha} = 1 - 2\sum_{i\neq j} U^2_{\alpha i}U_{\alpha j}(U + 4\Re(\xi))_{\alpha j} \sin^2\left(\frac{\Delta\lambda'_{ij}L}{2}\right) - 4\sum_{i,j} U^2_{\alpha i}\left|\frac{\epsilon_{\alpha j}}{E}\right|^2 \sin^2\left(\frac{\Delta\lambda'_{ij}L}{2}\right) \tag{5.65}$$

$$\begin{aligned} P_{\alpha\beta} = &-2\sum_{i\neq j} U_{\alpha i}U_{\beta i}\left\{U_{\alpha j}(U + 2\Re(\xi))_{\beta j} + (U + 2\Re(\xi))_{\alpha j}U_{\beta j}\right\} \sin^2\left(\frac{\Delta\lambda'_{ij}L}{2}\right) \\ &-4\sum_{i,j} U_{\alpha i}U_{\beta i}\frac{\Re(\epsilon^*_{\alpha j}\epsilon_{\beta j})}{E^2} \sin^2\left(\frac{\Delta\lambda'_{ij}L}{2}\right) + 2\sum_{i,j} U_{\alpha i}U_{\beta i}\frac{\Im(\epsilon^*_{\alpha j}\epsilon_{\beta j})}{E^2}\sin(\Delta\lambda'_{ij}L) \end{aligned} \tag{5.66}$$

$$\begin{aligned} P_{\alpha\tilde{\beta}} = P_{\tilde{\beta}\alpha} = &-2\sum_{i\neq j} U_{\alpha i}U_{\alpha j}\frac{\Re(\epsilon_{\alpha\beta}\epsilon^*_{j\beta})}{E^2} \sin^2\left(\frac{\Delta\lambda'_{ij}L}{2}\right) + 4\sum_{i} U_{\alpha i}\frac{\Re(\epsilon^*_{\alpha\beta}\epsilon_{i\beta})}{E^2} \sin^2\left(\frac{\Delta\lambda'_{i\tilde{\alpha}}L}{2}\right) \\ &-2\sum_{i} U_{\alpha i}\frac{\Im(\epsilon^*_{\alpha\beta}\epsilon_{i\beta})}{E^2}\sin(\Delta\lambda'_{i\tilde{\alpha}}L) \end{aligned} \tag{5.67}$$

$$P_{\tilde{\alpha}\tilde{\alpha}} = 1 - 2\sum_{i}\left|\frac{\epsilon_{i\alpha}}{E}\right|^2 \sin^2\left(\frac{\Delta\lambda'_{i\tilde{\alpha}}L}{2}\right) \tag{5.68}$$

$$P_{\tilde{\alpha}\tilde{\beta}} = 0 \tag{5.69}$$

5.3.3 CPT Symmetry

Notice that the analysis is section 5.1.2 still applies to this model. The assumptions that A is real symmetric and that A and C commute were not used anywhere in the calculation. Therefore that work shows that even this very general model cannot predict CPT violations.

5.4 Generalization including Mass

Because the massless model is difficult to parameterize, we can actually state the space of possibilities much more simply if we consider neutrino masses. We assume both small Dirac and small Majorana masses.

Normally, it assumed that m_R is very large in comparison to the other masses, so that a seesaw mechanism can be invoked to explain the smallness of the left-handed neutrino masses, and to explain why oscillations into sterile neutrinos are unobserved. However, we will assume that these masses are all relatively small, and that the Lorentz-violating terms produce the desired effects instead. Thus we begin with the general Hamiltonian

$$H = \begin{bmatrix} \frac{m_L^2}{2E} & A + \frac{m_D^2}{2E} \\ A^\dagger + \frac{m_D^2}{2E} & cE + \frac{m_R^2}{2E} \end{bmatrix} \tag{5.70}$$

If we still assume that the term $cE + \frac{m_R^2}{2E}$ dominates the matrix at high energies, the block diagonalization introduced for the previous model is still valid, so the mass-like block eigenvalue will be given by

$$\Lambda = \frac{m_L^2}{2E} - (A + \frac{m_D^2}{2E})(cE + \frac{m_R^2}{2E})^{-1}(A^\dagger + \frac{m_D^2}{2E}) \tag{5.71}$$

This can be simplified by approximating the inverse of a perturbed a matrix. For example, let us assume X is a matrix small in comparison to M. Then

$$(M+X)^{-1} = (M(I+M^{-1}X))^{-1} = (I+M^{-1}X)^{-1}M^{-1} \approx (I-M^{-1}X)M^{-1} = M^{-1} - M^{-1}XM^{-1} \tag{5.72}$$

Applying this result with $M = cE$ and $X = \frac{m_R^2}{2E}$, we have

$$\Lambda = \frac{m_L^2}{2E} - \frac{Ac^{-1}A^\dagger}{E} - \frac{Ac^{-1}m_D^2}{2E^2} - \frac{m_D^2c^{-1}A^\dagger}{2E^2} + \cdots \tag{5.73}$$

from which we see that the observed neutrino mass would actually be a sum of a real mass and a Lorentz-violating apparent mass. The fact that oscillations do not appear to depend on E^{-2} could give a bound on the possible size of m_D (though this would of course be model-dependent). Observe that the second term is the one that appeared in the non-massive model. Since we still want the regular behavior at high energy (and assuming that the E^{-2} terms do not contribute significantly), we know

$$\tilde{M} = m_L^2 - AC^{-1}A^\dagger \tag{5.74}$$

Since each of these terms is Hermitian, this means we can choose any A and C we like, and then set m_L to satisfy the equation in terms of the known $\tilde{M}$. This means the massive model has 13 free parameters: one in $\tilde{M}$, three in C, and nine in A. This is the expedience of introducing mass - it allows the perturbation variable ϵ to be completely free. Thus we can reuse the results of the previous model, allowing ϵ to be any complex-valued matrix. This model then has 18 degrees of freedom.

Chapter 6

Conclusions

Because the neutrino is the lightest known massive particle, it is opening the first window into physics beyond the Standard Model. The experimental result at LSND, partially confirmed at MiniBooNE, shows that a three-flavor, massive model of neutrino oscillations will not explain all the data. Furthermore, possible evidence of CPT-violation in the results of MiniBooNE and MINOS suggests that Lorentz-violation may be a natural choice for an alternative explanation.

Lorentz-violating models can predict a rich variety of behavior. On the basis of previous similar work, I have developed a sequence of models which are able to mimic mass behavior at high energy while retaining many degrees of freedom at low energies. This is a relatively unique characteristic, insofar as Lorentz violations are usually predicted to appear at high energy; whereas in this model, the disagreement appears at low energy. This is a consequence of the seesaw mechanism embedded in the model.

The models introduced here have been arranged in order of increasing complexity, and each has included the previous model as a subcase. However, the solutions become increasingly approximate. These models can describe an enormous range of behavior, and we have seen that they can do well describing the data.

These models rely on Lorentz violations to produce neutrino oscillations which precisely mimic the behavior of the mass model over certain energy ranges. These models have primarily been explored using the isotropic components of the Lorentz violating tensor fields. This gives rise to isotropic effects unlike the sidereal or annual effects frequently associated with Lorentz violation. Naturally, non-isotropic components could be added if experimental evidence demands it. Furthermore, because of the seesaw mechanism embedded in the models, the new prediction of these models that could be tested to differentiate them from the standard mass model occurs at low energies rather than the high energies usually associated with Lorentz violations. All the models proposed here can predict "disappearance", in which active neutrinos oscillate

into sterile neutrinos, or vice versa. This mixing generally occurs only at low energies. Furthermore, one of these models, the rotated model, is able to predict CPT violations, which may be necessary to explain the neutrino-antineutino asymmetry being experimentally observed.

These models were found to be capable of reproducing the mass model probabilities over certain energy ranges, and therefore cannot be immediately ruled out. Furthermore, their flexibility in predicting CPT violation and their relatively good fit to the experimental data, disagreeing only with the very low energy experiments, seems to indicate that models like these may be tenable.

Appendix A

Matrix Analysis

A.1 Glossary

Some basic facts and definitions from linear algebra are collected below for the reference of the reader. Further details can be found in [31] or similar texts.

- **Diagonalizable Matrices** of dimension n are those which have n eigenvalues (λ) and n independent eigenvectors (x). These satisfy the eigenvalue equation $Mx = \lambda x$. M is then diagonalized by conjugation by the matrix formed by adjoining the eigenvectors. Observe that any scalar multiple of an eigenvector is also an eigenvector; in particular, eigenvectors can always be normalized.

- **Symmetric Matrices** are matrices which are their own transpose, ie, $M = M^T$, or $m_{ij} = m_{ji}$ for all i, j.

- **Orthogonal Matrices** are matrices for which $MM^T = MM^{-1} = I$. The rows (and columns) of an orthogonal matrix are orthonormal.

- **Unitary Matrices** are the complex-valued analogue of orthogonal matrices; those for which $MM^\dagger = MM^{-1} = I$.

- **Hermitian Matrices** are those for which $M = M^\dagger$. All Hermitian matrices are diagonalizable (a consequence of the Spectral Theorem). An *antihermitian* matrix will satisfy $M = -M^\dagger$.

- A **is similar to** B if there exists an invertible matrix M with the property $A = MBM^{-1}$. Similar matrices share eigenvalues and have related eigenvectors. In particular, rotating a matrix does not change its eigenvalues.

We also remind the reader that the tensor is like a high-dimensional matrix which has particular transformational properties. In particular, if a coordinate transform $x \to x'$ is applied by the transformation $L^\nu_\mu x^\mu = x'^\nu$, then all tensors transform in this way, with one factor of L for each dimension.

A.2 Block Diagonalization

Finding the eigenvalues of general large-dimension matrices is intractable, so the method of block diagonalization is usually used instead. Because we will need this result frequently, we will derive it here in general and then apply it as necessary.

We study here block hermitian matrices, represented generally as

$$M = \begin{bmatrix} B & A \\ A^\dagger & C \end{bmatrix} \tag{A.1}$$

where B and C are themselves hermitian matrices. The matrices A, B, and C may be of any size. We will attempt to block-diagonalize the matrix, that is, transform it into a new basis in which it is of the form

$$M' = \begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix} \tag{A.2}$$

where M_1 and M_2 are written in terms of the original block components of the matrix. It is then assumed that it is easy to find the eigenvalues of the submatrices M_1 and M_2, which will together be the eigenvalues of the original matrix. In particular, we are interested in finding the approximate solution in the case where the matrix is unbalanced.

A.2.1 Approximation for large C

If C has a large norm in comparison to A and B, we know the matrix is already approximately diagonal, so it will be diagonalized by a matrix which is approximately the identity. We therefore conjugate M by P, defined

$$P = I + \Delta \tag{A.3}$$

where Δ is a small matrix. Then P has inverse

$$P^{-1} = I - \Delta \tag{A.4}$$

to first order. (In general, we can work to k^{th} order by letting $P = e^\Delta = I + \sum_{n=1}^{k} \frac{1}{n!}\Delta^n$ and $P^{-1} = e^{-\Delta} = I + \sum_{n=1}^{k} \frac{(-1)^n}{n!}\Delta^n$.)

We find

$$M' = PMP^{-1} = M - M\Delta + \Delta M - \Delta M\Delta \tag{A.5}$$

and we assume that, since we are working only to first order, the last term can be ignored. In addition, we know that Δ must be antihermitian[1] because P must be unitary (by the Spectral Theorem). In this case we have

$$\Delta = \begin{bmatrix} D_1 & D_2 \\ -D_2^\dagger & D_4 \end{bmatrix} \tag{A.6}$$

where D_1 and D_4 must also be antihermitian. Requiring that P block-diagonalizes M yields two equivalent equations, one of which is

$$A - BD_2 - AD_4 + D_1A + D_2C = 0 \tag{A.7}$$

To zeroth order, we approximate that since A, B, and D_i are all small in comparison to C, the products of two small terms can be ignored, and we are left with the simple equation

$$A + D_2C = 0 \tag{A.8}$$

which clearly has solution $D_2 = -AC^{-1}$. We can then complete Δ by assigning $D_1 = D_4 = 0$. This gives the diagonalization

$$M = \begin{bmatrix} B & A \\ A^\dagger & C \end{bmatrix} \sim \begin{bmatrix} B - 2AC^{-1}A^\dagger & BAC^{-1} \\ C^{-1}A^\dagger B & C + C^{-1}A^\dagger A - A^\dagger AC^{-1} \end{bmatrix} = M' \tag{A.9}$$

Since this is not a very good approximation, we hope we may do better by not throwing away all of the terms with two small factors. We still hope that we can take $D_1 = D_4 = 0$, in which case the equation to be solved is

$$DC - BD = -A \tag{A.10}$$

Solving this equation (known to mathematicians as *Sylvester's Equation*) is nontrivial. For matrices of small dimension, the matrix D can be explicitly computed by writing a system of linear equation which it satisfies. For example, for 2×2 matrices, we can write

$$D = -\begin{bmatrix} c_1 - b_1 & -b_3 & c_2 & 0 \\ -b_2 & c_1 - b_4 & 0 & c_2 \\ c_3 & 0 & c_4 - b_1 & b_3 \\ 0 & c_3 & b_2 & c_4 - b_4 \end{bmatrix}^{-1} A \tag{A.11}$$

[1]Recall that an antihermitian matrix M is one for which $M = -M^\dagger$.

where the matrices have been made into vectors for notational simplicity. Of course, even this expression involves finding the inverse of a 4×4 matrix, which is itself no simple problem. Mathematicians assure us that the equation has a unique solution as long as A and C do not share eigenvalues [22].

The equation can more usefully be solved by the method of successive approximations. Assuming that A and B are small and C is large, we have the solution

$$D = -AC^{-1} - BAC^{-2} - B^2AC^{-3} + \dots \tag{A.12}$$

At any rate, once D is known, it is relatively straightforward to compute that

$$\begin{bmatrix} B & A \\ A^\dagger & C \end{bmatrix} \sim \begin{bmatrix} B + AD^\dagger + DA^\dagger & 0 \\ 0 & C - A^\dagger D - D^\dagger A \end{bmatrix} \tag{A.13}$$

A.2.2 Approximation for large A

If we are interested instead in the case of A large, we need to perturb about a different diagonalizing matrix (I worked in the case of C large only because the matrix was already approximately diagonal). As our simple case we consider the matrix

$$M = \begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix} \tag{A.14}$$

which can be diagonalized by

$$U_0 = \frac{1}{\sqrt{2}} \begin{bmatrix} I & I \\ -I & I \end{bmatrix} \tag{A.15}$$

with

$$U_0^{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} I & -I \\ I & I \end{bmatrix} \tag{A.16}$$

we can then compute $U_0 M U_0^{-1}$ to see that

$$\begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix} \sim \begin{bmatrix} -A & 0 \\ 0 & A \end{bmatrix}. \tag{A.17}$$

To first order, then, the block eigenvalues of a matrix with large off-diagonal elements is

$$\begin{bmatrix} A + \frac{B+C}{2} & \frac{C-B}{2} \\ \frac{B+C}{2} & -A + \frac{B+C}{2} \end{bmatrix}. \tag{A.18}$$

A.3 Third-order Block diagonalization

Finding higher-order diagonalizations is difficult in general, be can be accomplished for matrices of the form (A.1) if we take $B = 0$. This is also the form relevant to the models discussed. Therefore we assume a matrix of the form

$$M = \begin{bmatrix} 0 & A \\ A^\dagger & C \end{bmatrix} \tag{A.19}$$

and find the next-to-leading order correction to its block diagonal entries. We can begin by keeping higher order terms in the diagonalization. For instance, to second order

$$(I+\Delta+\frac{1}{2}\Delta^2)M(I-\Delta+\frac{1}{2}\Delta^2) = \begin{bmatrix} -AC^{-1}A^\dagger & -AC^{-1}A^\dagger AC^{-1} - \frac{1}{2}AC^{-1}C^{-1}A^\dagger A \\ -C^{-1}A^\dagger AC^{-1}A^\dagger - \frac{1}{2}A^\dagger AC^{-1}C^{-1}A^\dagger & C + \frac{1}{2}C^{-1}A^\dagger A + \frac{1}{2}A^\dagger AC{-}1 \end{bmatrix} \tag{A.20}$$

where Δ is as defined in section (A.1) and for the case of this matrix $D = AC^{-1}$. This corrects the off-diagonal terms, but not the (block) eigenvalues, which is what we are interested in. We must therefore continue to the third order, computing

$$(I + \Delta + \frac{1}{2}\Delta^2 + \frac{1}{6}\Delta^3)M(I - \Delta + \frac{1}{2}\Delta^2 - \frac{1}{6}\Delta^3) \tag{A.21}$$

I find corrected block eigenvalues

$$\begin{aligned} M_1 &= -AC^{-1}A^\dagger + \frac{1}{2}AC^{-2}A^\dagger AC^{-1}A^\dagger + \frac{1}{2}AC^{-1}A^\dagger AC^{-2}A^\dagger \\ M_2 &= C + \frac{1}{2}C^{-1}A^\dagger A + \frac{1}{2}A^\dagger AC^{-1} - \frac{3}{4}C^{-1}A^\dagger AC^{-1}A^\dagger AC^{-1} - \frac{1}{6}C^{-1}A^\dagger AC^{-2}A^\dagger A - \frac{1}{6}A^\dagger AC^{-2}A^\dagger A \end{aligned} \tag{A.22}$$

But the off-diagonal terms are still

$$-C^{-1}A^\dagger AC^{-1}A^\dagger - \frac{1}{3}A^\dagger AC^{-2}A^\dagger. \tag{A.23}$$

Thus we now see that the original definition of Δ no longer suffices to diagonalize the matrix, since there are off-diagonal terms of second order, while we are interested in the third-order eigenvalues. We must therefore treat this as the first step in a diagonalization, and assume that this matrix is diagonalized by a new matrix $U = I + \Delta_2$. As before, we assume that Δ_2 is off-diagonal and anti-hermitian (ie, $\Delta_2 = \begin{bmatrix} 0 & D_2 \\ -D_2^\dagger & 0 \end{bmatrix}$) and solve

$$(I + \Delta_2)(I + \Delta_1 + \frac{1}{2}\Delta_1^2 + \frac{1}{6}\Delta_1^3)M(I - \Delta_1 + \frac{1}{2}\Delta_1^2 - \frac{1}{6}\Delta_1^3)(I - \Delta_2) = \begin{bmatrix} M_1' & 0 \\ 0 & M_2' \end{bmatrix} \tag{A.24}$$

for D_2 to third order. We find

$$D_2 = AC^{-1}A^\dagger AC^{-2} + \frac{1}{3}AC^{-2}A^\dagger AC^{-1} \tag{A.25}$$

Including this extra step in the diagonalization does not change the diagonal blocks (to third order), that is, $M_1' = M_1$ and $M_2' = M_2$.

A.4 Matrix Roots

We discuss here how to find square roots of diagonalizable matrices. These results are special cases of more general theorems discussed in [20]. Although finding one square root is not especially difficult, describing all possible roots is very difficult. We know, of course, that if Λ is the matrix of eigenvalues of M, then

$$X^2 = (S\Lambda^{1/2}S^{-1})^2 = M \tag{A.26}$$

But there are many other roots of M as well, in general. For instance,

$$\begin{bmatrix} 2 & -1 \\ 3 & 2 \end{bmatrix}^2 = I \tag{A.27}$$

In particular, 3 x 3 matrices with unique eigenvalues have square roots in exactly 8 similarity classes. Thus, if X is a root of M, and we know that M has eigenvalues $\{\lambda_i\}$ then

$$X = S \begin{bmatrix} \pm\sqrt{\lambda_1} & 0 & 0\dots & 0 \\ 0 & \pm\sqrt{\lambda_2} & 0\dots & 0 \\ & & \dots & \\ 0 & 0\dots & 0 & \pm\sqrt{\lambda_n} \end{bmatrix} S^{-1} \tag{A.28}$$

There is a similar class of matrices X for which $XX^T = M$. This is solved by the symmetric roots of M, and also possibly by other matrices which are not roots of M. Based on counting degrees of freedom, I believe the full set of solutions to this auxiliary equation is given by

$$X = U\Lambda^{1/2}U^\dagger O \tag{A.29}$$

where O is any orthogonal (or, in general, unitary) matrix.

References

[1] V. Alan Kostelecký and Matthew Mewes, Phys. Rev., **D70**, 031902 (2004).

[2] Teppei Katori, V. Alan Kostelecký and Rex Tayloe, Phys. Rev., **D74**, 105009 (2006).

[3] Fundamentals of Neutrino Physics and Astrophysics. Giunti and Kim, 2007.

[4] Physics of Massive Neutrinos. Boehm and Vogel, 1993.

[5] C. Amsler *et al.*, Phys. Lett. **B667**, 1 (2008).

[6] Jorge S. Diaz, V. Alan Kostelecký, and Matthew Mewes, Phys. Rev., **D80** 076007 (2009).

[7] A. A. Aguilar-Arevalo, *et al.* (MiniBooNE), Phys. Rev. Lett., **103**, 111801 (2009).

[8] V. Barger, D. Marfatia and K. Whisnant, Phys. Rev. Lett., **B653**, 267-277 (2007).

[9] V. Alan Kostelecký and Matthew Mewes, Phys. Rev. Lett., **D69**, 016005 (2004).

[10] Alec Habig, Mod. Phys. Lett. A **25**, 1219-1231 (2010).

[11] S. Abe *et al.* (KamLAND), Phys. Rev. Lett., **100**, 221803 (2008).

[12] B. Aharmin *et al.* (SNO), Phys. Rev. Lett., **101**, 111301 (2008).

[13] A. A. Aguilar-Arevalo *et al.* (MiniBooNE), Phys. Rev. Lett., **102**, 101802 (2009).

[14] A. Aguilar *et al.* (LSND), Phys. Rev., **D64**, 112007 (2001).

[15] Statistical Significance of the Gallium Anomaly. Guinti and Laveder, 2010.

[16] Don Colladay and V. Alan Kostelecký, Phys. Rev., **D58**, 116002 (1998).

[17] O. W. Greenberg, Phys. Rev. Lett. **89**, 231602 (2002).

[18] A. A. Aguilar-Arevalo *et al.*, Phys. Rev. Lett., **105**, 181801 (2010).

[19] Introduction to Particle Physics. David Griffiths, 2001.

[20] Topics in Matrix Analysis. Horn and Johnson, 1991.

[21] Hamann, *et al.* "Cosmology seeking friendship with sterile neutrinos," 2010.

[22] Eigenvalues of Matrices. Francoise Chateline, 1993.

[23] S. N. Ahmed *et al.* (SNO), Phys. Rev. Lett. **92** 181301 (2004).

[24] R. Wendell *et al.* (Super-Kamiokande), Phys. Rev. **D81** 092004 (2010).

[25] C. Weinheimer. "Direct Measurement of Neutrino Mass from Tritium Beta Decay." arXiv:0912.1619.

[26] Y. Ashie *et al.* (Super-K), Phys. Rev. Lett. **93** 101801 (2004).

[27] V. Alan Kostelecký and Matthew Mewes. Phys. Rev **D80** 015020 (2009).

[28] Alan Kostelecký and Neil Russell. "Data Tables for Lorentz and CPT Symmetry." arXiv:0801.0287v3.

[29] Ch. Kraus, *et al.* Eur. Phys. J. **C40**: 447 - 465 (2005).

[30] Marcus Beck. J. Phys.: Conf. Ser. 203 012097 (2010).

[31] Linear Algebra. Fraleigh and Beauregard, 1995.