Imperial College London

Department of Theoretical Physics

Neutrino Oscillations and Non-Hermitian Quantum Mechanics

Jonathan White

September 23, 2010

Supervised by Dr Hugh Jones

Submitted in part fulfilment of the requirements for the degree of Master of Science in Quantum Fields and Fundamental Forces of Imperial College London

Abstract

In this dissertation we firstly present a brief review of the Standard Model of particle physics, focussing on aspects of the model that prove most relevant in the context of explaining the observed oscillation of neutrinos. The phenomenology of neutrino oscillations is then introduced, and different possible extensions to the Standard Model that would allow for non-zero neutrino masses, thereby allowing for the observed neutrino oscillations, are discussed. After introducing the $\mathcal{PT}$-symmetric formulation of quantum mechanics, we then turn to the so-called Model 8 solution of the $\mathcal{PT}$-symmetric Dirac equation found by Jones-Smith *et al.*, and discuss its potential relevance to the neutrino oscillation problem.

Contents

1 Outline

The Standard Model of particle physics is incredibly successful in explaining almost all experimental observations made at the energy scales currently available to us. The observation of neutrino oscillations, however, proves to be a major stumbling block. As we shall see, the fact that neutrinos can oscillate between different flavour states is taken as an indication that they in fact possess a non-zero mass. Under the constraints imposed on terms permissible within the Standard Model Lagrangian, namely that they be gauge invariant and renormalisable, generating this non-zero neutrino mass is not possible.

As such, we are forced to make extensions to the Standard Model, and, as will be shown, this broadly amounts to either supplementing the assumed particle content with additional light, as-yet unobserved particles, or conceding that the Standard Model is not a complete theory for energy scales greater than that of the electro-weak sector. With regard to this second possibility, the high energy nature of the new physics makes it very difficult to distinguish between alternative theories through experiment. Many questions still remain, therefore, as to the exact mechanism behind observed neutrino oscillations.

In the following two chapters of this dissertation we hope to present some of the key features of the Standard Model and the problems faced when trying to incorporate neutrino oscillations. We also discuss the phenomenology of neutrino oscillations and the implication that neutrinos appear to have non-zero mass. Finally we discuss some possible extensions to the Standard Model that could be responsible for generating these masses and their respective merits and downfalls.

One recent result in the field of non-Hermitian quantum mechanics, presented in [1] and [2], offers a potentially very interesting new possibility in the context of neutrino oscillations. The field of non-Hermitian quantum mechanics is concerned with finding Hamiltonians that, despite being non-Hermitian, still display the desired properties of having a real energy spectrum and unitary time-evolution. One particular subset of these non-Hermitian Hamiltonians consists of those displaying symmetry under the combined action of parity and time-reversal. This is the field of $\mathcal{PT}$-symmetric quantum mechanics, and has seen many developments over the last decade, [3]. In [1] and [2], Jones-Smith *et al.* consider solutions to a $\mathcal{PT}$-symmetric Dirac equation. The 4-dimensional solution, dubbed Model 4, is found to be exactly equivalent to that of the Hermitian Dirac equation. Unlike the 8-dimensional solution to the Hermitian Dirac equation, however, which simply decouples into two 4-dimensional solutions, the 8-dimensional solution of the $\mathcal{PT}$-symmetric Dirac equation, dubbed Model 8, is found to describe what looks like a new type of particle. Namely, under certain conditions, the solution would appear to describe two irreducible *massless* particles, despite having a non-zero mass matrix. As we shall see, this could potentially lead us to the possibility of oscillations between the two *massless* particles, which evidently is very appealing in the context of neutrino oscillations.

In chapter 4 of this dissertation we will review the principles behind a

$\mathcal{PT}$-symmetric formulation of quantum mechanics before providing explicit calculations for the Model 8 solution. We also explore the possibility of extending the model to describe three flavours of massless particle by attempting to make the natural extension to a 12-dimensional solution.

In our discussion of $\mathcal{PT}$-symmetric quantum mechanics, we shall see that it is always possible to find a mapping between the $\mathcal{PT}$-symmetric formulation and an equivalent Hermitian one. With the potentially striking implications of the new Model 8 solution, we will therefore finish by stressing the importance of determining this map. Doing so would allow us to confirm whether or not the solution really does correspond to a new type of particle, rather than being the $\mathcal{PT}$-symmetric equivalent to one that is already known.

2 The Standard Model

Here we review key elements of the Standard Model that will help us to understand the issues faced when trying to extend the model to include neutrino masses.

2.1 The gauge group and particle content

The Standard Model of particle physics is a gauge theory with gauge group $SU(3) \times SU(2) \times U(1)$. The $SU(3)$ element is responsible for interactions involving the strong force and, being an eight-dimensional group[1], has eight associated spin-one gauge fields, called gluons, which we will label G_μ^α, with $\alpha = 1, ..., 8$ and μ being the Lorentz index. The remaining $SU(2) \times U(1)$ is the electro-weak sector and gives us an additional three spin-one gauge fields associated with $SU(2)$ and one associated with $U(1)$. We denote these W_μ^a, with $a = 1, ..., 3$, and B_μ respectively. It is important to note that these latter four gauge fields do not correspond directly with the $W_\mu^\pm$ and Z_μ^0 responsible for mediating the weak force and the photon, A_μ, responsible for mediating the electromagnetic force. We will see, however, that these more familiar objects are indeed linear combinations of W_μ^a and B_μ.

The elementary particles of the Standard Model fall into representation of

[1]The group $SU(n)$ has dimension $n^2 - 1$, which follows from imposing the conditions $X^\dagger = -X$ and $tr(X) = 0$ on the group generators, which in turn follows from expressing $M \in SU(n)$ as $M = e^X$ and requiring $M^\dagger M = 1$ and $det(M) = 1$.

the $SU(3)$, $SU(2)$ and $U(1)$ elements of the gauge group and carry quantum numbers dubbed *colour*, *weak isospin* and *weak hypercharge* respectively. Guided by experimental results, and in the absence of neutrino masses, the particle content of the Standard Model is taken to be that given in Table 2.1.

Particles	$SU(3)$ rep.	$SU(2)$ rep.	$U(1)$ charge
$\begin{pmatrix} \nu_e \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau_L \end{pmatrix}$	$\mathbf{1}$	$\mathbf{2}$	$-\frac{1}{2}$
e_R, μ_R, τ_R	$\mathbf{1}$	$\mathbf{1}$	-1
$\begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \begin{pmatrix} t_L \\ b_L \end{pmatrix}$	$\mathbf{3}$	$\mathbf{2}$	$\frac{1}{6}$
u_R, c_R, t_R	$\mathbf{3}$	$\mathbf{1}$	$\frac{2}{3}$
d_R, s_R, b_R	$\mathbf{3}$	$\mathbf{1}$	$-\frac{1}{3}$

Table 2.1: Particle content of the Standard Model, with lepton and quark labels following the usual convention and subscripts L and R labelling the handedness of the particles. The neutrinos are taken to be left-handed (with no right-handed components, as is experimentally the case for massless neutrinos) and so the handedness label is omitted. The emboldened numbers also follow the conventional notation whereby, for example, the **2** in the first row is a statement that the lepton and its corresponding neutrino form a doublet under the action of the $SU(2)$ group. Here we have explicitly grouped leptons and neutrinos into their $SU(2)$ doublets and likewise for the quarks.

The particles listed in Table 2.1 are two-component Weyl spinors, and were introduced as such in order to appeal to the possibly more familiar notion of the Standard Model's particle content. We can then choose to present the left- and right-handed components of a particle in a single Dirac spinor or with two Majorana spinors. So, taking the electron as an example, we have[2]:

[2]Here σ_2 is the Pauli matrix $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, and Weyl spinor indices are suppressed to avoid over-complication of expressions.

$$\psi^e_{Dirac} = \begin{pmatrix} e_L \\ e_R \end{pmatrix} = e \tag{2.1}$$

$$\text{or} \quad \psi^{e_L}_{Majorana} = \begin{pmatrix} e_L \\ i\sigma_2 e_L^* \end{pmatrix} \quad \text{and} \quad \psi^{e_R}_{Majorana} = \begin{pmatrix} -i\sigma_2 e_R^* \\ e_R \end{pmatrix}. \tag{2.2}$$

Following the notation of [4], we choose to use the Majorana notation and also make the notation more compact so that:

$$\psi^{l^m_L}_{Majorana} = \mathcal{E}_m \quad \text{and} \quad \psi^{l^m_R}_{Majorana} = E_m, \tag{2.3}$$

where l^m labels e, μ and τ for $m = 1, 2, 3$ respectively. Similarly, labelling u, c and t with u^m, and d, s and b with d^m, where $m = 1, .., 3$ in both cases, we have:

$$\psi^{u^m_L}_{Majorana} = \mathcal{U}_m \quad \text{and} \quad \psi^{u^m_R}_{Majorana} = U_m \tag{2.4}$$

$$\psi^{d^m_L}_{Majorana} = \mathcal{D}_m \quad \text{and} \quad \psi^{d^m_R}_{Majorana} = D_m. \tag{2.5}$$

In the case of the neutrinos, we only have left-handed Weyl spinors and thus only require a single Majorana spinor:

$$\psi^{\nu^m}_{Majorana} = \mathcal{V}_m, \tag{2.6}$$

where ν^m corresponds to ν_e, ν_μ and ν_τ for $m = 1, 2, 3$ respectively. With this new notation, we can now re-express the particle content of the Standard Model as shown in Table 2.2.

Finally, with a total Majorana spinor containing both left and right-handed components as given in (2.2), it will be useful for us to know how the corresponding, opposite-handed components of our Majorana spinors

Particles	$SU(3)$ rep.	$SU(2)$ rep.	$U(1)$ charge
$P_L\mathbf{L}_m = \begin{pmatrix} P_L\mathcal{V}_m \\ P_L\mathcal{E}_m \end{pmatrix}$	$\mathbf{1}$	$\mathbf{2}$	$-\frac{1}{2}$
$P_R E_m$	$\mathbf{1}$	$\mathbf{1}$	-1
$P_L\mathbf{Q}_m = \begin{pmatrix} P_L\mathcal{U}_m \\ P_L\mathcal{D}_m \end{pmatrix}$	$\mathbf{3}$	$\mathbf{2}$	$\frac{1}{6}$
$P_R U_m$	$\mathbf{3}$	$\mathbf{1}$	$\frac{2}{3}$
$P_R D_m$	$\mathbf{3}$	$\mathbf{1}$	$-\frac{1}{3}$

Table 2.2: Particle content of the Standard Model using Majorana spinor notation. P_L and P_R are the projection operators that pick out the left- and right-handed components of a spinor respectively. The emboldened $\mathbf{L}_m$ and $\mathbf{Q}_m$ represent the explicit grouping of the particles into $SU(2)$ doublets.

will transform, i.e. $P_R\mathbf{L}_m$, $P_L E_m$, $P_R\mathbf{Q}_m$, $P_L U_m$ and $P_L D_m$. Since these objects are related to their partners by complex conjugation, they transform under the complex conjugate representations. See Table 2.3.

Particles	$SU(3)$ rep.	$SU(2)$ rep.	$U(1)$ charge
$P_R\mathbf{L}_m = \begin{pmatrix} P_R\mathcal{V}_m \\ P_R\mathcal{E}_m \end{pmatrix}$	$\mathbf{1}$	$\overline{\mathbf{2}}$	$\frac{1}{2}$
$P_L E_m$	$\mathbf{1}$	$\mathbf{1}$	1
$P_R\mathbf{Q}_m = \begin{pmatrix} P_R\mathcal{U}_m \\ P_R\mathcal{D}_m \end{pmatrix}$	$\overline{\mathbf{3}}$	$\overline{\mathbf{2}}$	$-\frac{1}{6}$
$P_L U_m$	$\overline{\mathbf{3}}$	$\mathbf{1}$	$-\frac{2}{3}$
$P_L D_m$	$\overline{\mathbf{3}}$	$\mathbf{1}$	$\frac{1}{3}$

Table 2.3: Transformation properties of the corresponding, opposite-handed components of the Majorana spinors set out in Table 2.2. The notation $\overline{\mathbf{3}}$ indicates that the object transforms in the complex conjugate of the triplet representation.

2.2 The Lagrangian

Given the ingredients of our gauge group and desired particle content, we can now go about constructing the Lagrangian. In doing so, we need to ensure that all terms are gauge invariant, Lorentz invariant and renormalisable. Once again following the notation of [4], it turns out that the most general such Lagrangian is of the form:

$$\begin{aligned}\mathcal{L}_0 =& -\frac{1}{2}\overline{\mathbf{L}}_m \not{D}\mathbf{L}_m - \frac{1}{2}\overline{E}_m \not{D} E_m \\ &- \frac{1}{2}\overline{\mathbf{Q}}_m \not{D}\mathbf{Q}_m - \frac{1}{2}\overline{U}_m \not{D} U_m - \frac{1}{2}\overline{D}_m \not{D} D_m \\ &- \frac{1}{4}G^{\alpha}_{\mu\nu}G^{\alpha\mu\nu} - \frac{1}{4}W^{a}_{\mu\nu}W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \\ &- \frac{g_3^2\Theta_3}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}G^{\alpha\mu\nu}G^{\alpha\lambda\rho} - \frac{g_2^2\Theta_2}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}W^{a\mu\nu}W^{a\lambda\rho} - \frac{g_1^2\Theta_1}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}B^{\mu\nu}B^{\lambda\rho}.\end{aligned} \tag{2.7}$$

The first two lines use the notation $\overline{E}_m = E_m^\dagger i\gamma^0$ and $\not{D} = \gamma^\mu D_\mu$, where γ^μ, with $\mu = 0, ..., 3$, are the Dirac γ matrices[3] and D_μ is the covariant derivative[4]. The exact form of the covariant derivative is specific to the field on which it is acting. Taking the generators of $SU(3)$ for the **3** representation to be $\frac{1}{2}\lambda_\alpha$ and the generators of $SU(2)$ for the **2** representation to be $\frac{1}{2}\sigma_a$, where λ_α with $\alpha = 1, ..., 8$ are the eight 3×3 Gell-Mann matrices and σ_a with $a = 1, .., 3$ are the three 2×2 Pauli matrices, the covariant derivatives are given by:

[3]Following the conventions of [4] we are using the metric diag$(-1, 1, 1, 1)$. This results in our γ^0 being anti-Hermitian, and it is related to the Hermitian γ_0 associated with the metric diag$(1, -1, -1, -1)$ by $\gamma^0_{+---} = i\gamma^0_{-+++}$. Thus, substituting into our more familiar expression $\overline{E}_m = E_m^\dagger\gamma^0_{+---}$, we pick up a factor of i.

[4]Not to be confused with the quark fields D_m.

$$D_\mu \mathbf{L}_m = \partial_\mu \mathbf{L}_m + \left[\frac{i}{2} g_1 B_\mu - \frac{i}{2} g_2 W_\mu^a \sigma_a\right] P_L \mathbf{L}_m + \left[-\frac{i}{2} g_1 B_\mu + \frac{i}{2} g_2 W_\mu^a \sigma_a^*\right] P_R \mathbf{L}_m, \tag{2.8}$$

$$D_\mu E_m = \partial_\mu E_m - i g_1 B_\mu (P_L E_m) + i g_1 B_\mu (P_R E_m), \tag{2.9}$$

$$D_\mu \mathbf{Q}_m = \partial_\mu \mathbf{Q}_m + \left[-\frac{i}{6} g_1 B_\mu - \frac{i}{2} g_2 W_\mu^a \sigma_a - \frac{i}{2} g_3 G_\mu^\alpha \lambda_\alpha\right] P_L \mathbf{Q}_m + \left[\frac{i}{6} g_1 B_\mu + \frac{i}{2} g_2 W_\mu^a \sigma_a^* + \frac{i}{2} g_3 G_\mu^\alpha \lambda_\alpha^*\right] P_R \mathbf{Q}_m, \tag{2.10}$$

$$D_\mu U_m = \partial_\mu U_m + \left[\frac{2i}{3} g_1 B_\mu + \frac{i}{2} g_3 G_\mu^\alpha \lambda_\alpha^*\right] P_L U_m + \left[-\frac{2i}{3} g_1 B_\mu - \frac{i}{2} g_3 G_\mu^\alpha \lambda_\alpha\right] P_R U_m \quad \text{and} \tag{2.11}$$

$$D_\mu D_m = \partial_\mu D_m + \left[-\frac{i}{3} g_1 B_\mu + \frac{i}{2} g_3 G_\mu^\alpha \lambda_\alpha^*\right] P_L D_m + \left[\frac{i}{3} g_1 B_\mu - \frac{i}{2} g_3 G_\mu^\alpha \lambda_\alpha\right] P_R D_m, \tag{2.12}$$

where, for example, $D_\mu \mathbf{L}_m = D_\mu(P_L \mathbf{L}_m + P_R \mathbf{L}_m)$, and g_1, g_2 and g_3 are the coupling strengths for the $U(1)$, $SU(2)$ and $SU(3)$ elements of the gauge group respectively.

Covariant derivatives take the place of ordinary derivatives in the first five terms of Lagrangian (2.7) in order that invariance can be preserved under *local* transformations. For this to be the case, we further require that the gauge fields themselves transform in a specific way under the action of the gauge group elements. Let us take $M_1 \in U(1)$, $M_2 \in SU(2)$ and

$M_3 \in SU(3)$ to be of the form:

$$M_1 = e^{ih\theta_1(x)}, \tag{2.13}$$

$$M_2 = e^{i\theta_2^a(x)\frac{\sigma_a}{2}} \quad a = 1, ..., 3 \tag{2.14}$$

$$\text{and} \quad M_3 = e^{i\theta_3^\alpha(x)\frac{\lambda_\alpha}{2}} \quad \alpha = 1, ..., 8, \tag{2.15}$$

where h is the weak hypercharge of the field on which M_1 is acting and $\theta_g^i(x)$ are the spatially dependent coefficients of the generators that make this a *local* gauge theory. Then the required infinitesimal transformations for our gauge fields will be:

$$\delta B_\mu = \frac{1}{g_1}\partial_\mu\theta_1(x), \tag{2.16}$$

$$\delta W_\mu^a = \frac{1}{g_2}\partial_\mu\theta_2^a(x) - \epsilon^{abc}\theta_2^b(x)W_\mu^c, \tag{2.17}$$

$$\text{and} \quad \delta G_\mu^\alpha = \frac{1}{g_3}\partial_\mu\theta_3^\alpha(x) - f_{\beta\gamma}^\alpha\theta_3^\beta(x)G_\mu^\gamma, \tag{2.18}$$

where ϵ^{abc} and $f_{\beta\gamma}^\alpha$ are the $SU(2)$ and $SU(3)$ Lie Algebra structure constants respectively.

The terms in line 3 of (2.7) are the so-called Yang-Mills field strengths associated with the gauge fields, and take the form:

$$G_{\mu\nu}^\alpha = \partial_\mu G_\nu^\alpha - \partial_\nu G_\mu^\alpha + g_3 f_{\beta\gamma}^\alpha G_\mu^\beta G_\nu^\gamma, \tag{2.19}$$

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g_2\epsilon_{abc}W_\mu^b W_\nu^c, \tag{2.20}$$

$$\text{and} \quad B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{2.21}$$

2.2.1 Lorentz invariance

Desiring a relativistic theory, we require all the terms in our Lagrangian to be Lorentz invariant. Dirac spinors carry a representation $\Lambda_{\frac{1}{2}}$ of the group $Sp(3,1) \cong SL(2,\mathbb{C})$[5], which is the double cover of $SO(3,1)$, so that under a Lorentz transformation $\psi \to \psi'$ as:

$$\psi'(x) = \Lambda_{\frac{1}{2}} \psi(\Lambda^{-1} x'), \tag{2.22}$$

where Λ is the Lorentz transformation for contravariant 4-vectors such that $x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu}$. The 4×4 representation $\Lambda_{\frac{1}{2}}$ is reducible into two 2×2 representations of $SL(2,\mathbb{C})$. This is made clear in the Weyl representation, where the field consists of a left- and right-handed component, each transforming differently under $SL(2,\mathbb{C})$. Using this reducibility we are able to derive the transformation properties of the objects contained in our Lagrangian.

Taking $\sigma_{\mu} = (\mathbb{1}, \sigma_1, \sigma_2, \sigma_3)$, where σ_i are the Pauli matrices, we can construct the bi-spinor $\mathbf{X}$:

$$\mathbf{X} = x^{\mu} \sigma_{\mu} = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix}. \tag{2.23}$$

Being a bi-spinor, $\mathbf{X}$ transforms under $SL(2,\mathbb{C})$ as $\mathbf{X}' = \mathbf{A}\mathbf{X}\mathbf{A}^{\dagger}$, where $\mathbf{A} \in SL(2,\mathbb{C})$. Thus, by taking $\mathbf{X}' = x'^{\mu}\sigma_{\mu} = \Lambda^{\mu}_{\nu} x^{\nu} \sigma_{\mu}$ we are able to find the matrix $\mathbf{A}$ equivalent to the Lorentz transformation Λ:

$$\mathbf{X}' = x'^{\mu}\sigma_{\mu} = \Lambda^{\mu}_{\nu} x^{\nu} \sigma_{\mu} = \mathbf{A}\mathbf{X}\mathbf{A}^{\dagger} = \mathbf{A} x^{\nu} \sigma_{\nu} \mathbf{A}^{\dagger} \tag{2.24}$$

$$\implies \mathbf{A}\sigma_{\nu}\mathbf{A}^{\dagger} = \Lambda^{\mu}_{\nu}\sigma_{\mu}. \tag{2.25}$$

[5] Where $\cong$ means there is an isomorphism between the two groups.

We have found that σ_μ transforms as a covariant vector under the Lorentz transformations. Similarly, taking $\tilde{\sigma}_\mu = (\mathbb{1}, -\sigma_1, -\sigma_2, -\sigma_3)$, we can construct the bi-spinor $\mathcal{X} = x^\mu \tilde{\sigma}_\mu$ transforming as $\mathcal{X}' = \mathbf{B}\mathcal{X}\mathbf{B}^\dagger$, which leads to the relation $\mathbf{B}\tilde{\sigma}_\nu \mathbf{B}^\dagger = \Lambda^\mu{}_\nu \tilde{\sigma}_\mu$. As we might expect, $\mathbf{A}$ and $\mathbf{B}$ are related, [5], and we find that for any Lorentz transformation $\mathbf{A}\mathbf{B}^\dagger = \mathbb{1}$. This allows us to show also that $\mathbf{A}^\dagger \tilde{\sigma}_\nu \mathbf{A} = \Lambda^\mu{}_\nu \tilde{\sigma}_\mu$ and $\mathbf{B}^\dagger \sigma_\nu \mathbf{B} = \Lambda^\mu{}_\nu \sigma_\mu$.

In the Weyl representation we have γ matrices satisfying the Dirac algebra $\{\gamma_\mu, \gamma_\nu\} = \eta_{\mu\nu}\mathbb{1}$ given by[6]:

$$\gamma_\mu = \begin{pmatrix} 0 & i\sigma_\mu \\ i\tilde{\sigma}_\mu & 0 \end{pmatrix}, \tag{2.26}$$

and our left- and right-handed spinor components of ψ, ϕ_L and ϕ_R respectively, transform as $\phi_L \to \mathbf{A}\phi_L$ and $\phi_R \to \mathbf{B}\phi_R$ under Lorentz transformations. Thus, considering the object $\overline{\psi}\psi = \psi^\dagger i\gamma^0\psi$:

$$\overline{\psi}\psi = \begin{pmatrix} \phi_L^\dagger & \phi_R^\dagger \end{pmatrix} i \begin{pmatrix} 0 & -i\mathbb{1} \\ -i\mathbb{1} & 0 \end{pmatrix} \begin{pmatrix} \phi_L \\ \phi_R \end{pmatrix} \tag{2.27}$$

$$= \phi_L^\dagger \phi_R + \phi_R^\dagger \phi_L, \tag{2.28}$$

which, under a Lorentz transformation

$$\longrightarrow \quad \phi_L^\dagger \mathbf{A}^\dagger \mathbf{B} \phi_R + \phi_R^\dagger \mathbf{B}^\dagger \mathbf{A} \phi_L \tag{2.29}$$

$$= \phi_L^\dagger \phi_R + \phi_R^\dagger \phi_L, \tag{2.30}$$

we see that we have a Lorentz scalar. Next consider the object $\overline{\psi}\gamma_\mu\psi =$

[6] Recall that we are still using the metric diag$(-1, 1, 1, 1)$

$\psi^\dagger i\gamma^0\gamma_\mu\psi$:

$$\overline{\psi}\gamma_\mu\psi = \begin{pmatrix} \phi_L^\dagger & \phi_R^\dagger \end{pmatrix} i \begin{pmatrix} 0 & -i\mathbb{I} \\ -i\mathbb{I} & 0 \end{pmatrix} \begin{pmatrix} 0 & i\sigma_\mu \\ i\tilde{\sigma}_\mu & 0 \end{pmatrix} \begin{pmatrix} \phi_L \\ \phi_R \end{pmatrix} \tag{2.31}$$

$$= \begin{pmatrix} \phi_L^\dagger & \phi_R^\dagger \end{pmatrix} \begin{pmatrix} i\tilde{\sigma}_\mu & 0 \\ 0 & i\sigma_\mu \end{pmatrix} \begin{pmatrix} \phi_L \\ \phi_R \end{pmatrix} \tag{2.32}$$

$$= \phi_L^\dagger i\tilde{\sigma}_\mu\phi_R + \phi_R^\dagger i\sigma_\mu\phi_L, \tag{2.33}$$

which, under a Lorentz transformation

$$\longrightarrow \quad \phi_L^\dagger \mathbf{A}^\dagger i\tilde{\sigma}_\mu \mathbf{A}\phi_L + \phi_R^\dagger \mathbf{B}^\dagger i\sigma_\mu \mathbf{B}\phi_R \tag{2.34}$$

$$= \phi_L^\dagger \Lambda^\nu{}_\mu i\tilde{\sigma}_\nu\phi_L + \phi_R^\dagger \Lambda^\nu{}_\mu i\sigma_\nu\phi_R \tag{2.35}$$

$$= \Lambda^\nu{}_\mu \overline{\psi}\gamma_\nu\psi, \tag{2.36}$$

i.e. $\overline{\psi}\gamma_\mu\psi$ transforms as a covariant vector. Using the metric to raise the index we find that $\overline{\psi}\gamma^\mu\psi$ transforms as a contravariant vector $\overline{\psi}\gamma^\mu\psi \rightarrow \Lambda^\mu{}_\nu\overline{\psi}\gamma^\nu\psi$.

Given that our Lagrangian also contains terms involving the projection operators P_L and P_R (a general state $\psi = P_L\psi + P_R\psi$), it will also be useful to know how these behave under Lorentz transformations. P_L and P_R can be constructed in terms of a fifth γ matrix $\gamma^5 = -i\gamma^0\gamma^1\gamma^2\gamma^3$ as[7] :

$$P_L = \frac{1+\gamma^5}{2} \qquad P_R = \frac{1-\gamma^5}{2}. \tag{2.37}$$

As such, we need to know how the objects $\overline{\psi}\gamma^5\psi$ and $\overline{\psi}\gamma^\mu\gamma^5\psi$ transform. Following similar methods to those used above, it can be shown that these

[7] Once again following the conventions of [4], we choose to define γ^5 and the projection operators in such a way as to give left-handed particles as the chirality +1 states and right-handed particles the chirality -1 states.

transform as a scalar and contravariant vector respectively.

The first five terms of Lagrangian (2.7) are of the form $\overline{\psi}\gamma^\mu D_\mu\psi$, so, given the result $\overline{\psi}\gamma^\mu\psi \to \Lambda^\mu_{\ \nu}\overline{\psi}\gamma^\nu\psi$, we can see that the contraction between the contravariant Lorentz index of $\overline{\psi}\gamma^\mu\psi$ and the covariant Lorentz index of D_μ will indeed leave these terms Lorentz invariant.

Looking at the remaining six terms, and using the fact that the spin-one gauge fields transform as:

$$G'^{\alpha}_{\mu}(x) = \Lambda^{\nu}{}_{\mu} G^{\alpha}_{\nu}(\Lambda^{-1}x'), \tag{2.38}$$

$$W'^{a}_{\mu}(x) = \Lambda^{\nu}{}_{\mu} W^{a}_{\nu}(\Lambda^{-1}x'), \tag{2.39}$$

$$\text{and} \quad B'_{\mu}(x) = \Lambda^{\nu}{}_{\mu} B_{\nu}(\Lambda^{-1}x'), \tag{2.40}$$

we can once again see that the contraction over all Lorentz indices will leave terms of the form $B_{\mu\nu}B^{\mu\nu}$ invariant.

In summary, due to the nice transformation properties of the objects $\overline{\psi}\psi$, $\overline{\psi}\gamma^\mu\psi$, $\overline{\psi}\gamma^5\psi$ and $\overline{\psi}\gamma^\mu\gamma^5\psi$, ensuring Lorentz invariance amounts to requiring all Lorentz indices to be contracted over, leaving no 'free' Lorentz indices.

Aside on index notation

In the preceding discussion we have not explicitly included the $SL(2,\mathbb{C})$ indices. If we were to do so, then we would have two types of index - undotted and dotted - corresponding to the fundamental and complex conjugate representations. As such, in the case of $SL(2,\mathbb{C})$, complex conjugation adds or removes dots on indices. Each type of index can further be raised or lowered. Listing objects carrying the different types of index and their transformation properties under $SL(2,\mathbb{C})$, we therefore have

$$\lambda_\alpha \to A_\alpha{}^\beta \lambda_\beta, \tag{2.41}$$

$$\chi^\alpha \to \chi^\beta A^{-1}{}_\beta{}^\alpha, \tag{2.42}$$

$$(\lambda_\alpha)^* \equiv \overline{\lambda}_{\dot\alpha} \to \left(A_\alpha{}^\beta\right)^* \overline{\lambda}_{\dot\beta} = A^*{}_{\dot\alpha}{}^{\dot\beta} \overline{\lambda}_{\dot\beta} = \overline{\lambda}_{\dot\beta} A^{\dagger\dot\beta}{}_{\dot\alpha} \text{ and} \tag{2.43}$$

$$(\chi^\alpha)^* \equiv \overline{\chi}^{\dot\alpha} \to \overline{\chi}^{\dot\beta} \left(A^{-1}{}_\beta{}^\alpha\right)^* = \overline{\chi}^{\dot\beta} \left(A^{-1}\right)^*_{\dot\beta}{}^{\dot\alpha} = \left(A^{-1}\right)^{\dagger\,\dot\alpha}{}_{\dot\beta} \overline{\chi}^{\dot\beta}, \tag{2.44}$$

where $A_\alpha{}^\beta \in SL(2,\mathbb{C})$, and $\alpha, \beta, \dot\alpha, \dot\beta = 1, 2$. These transformation rules ensure that objects such as $\chi^\alpha \lambda_\alpha$ and $\chi^{\dot\alpha} \lambda_{\dot\alpha}$ are invariant under the action of $SL(2,\mathbb{C})$. Explicitly including indices for the objects discussed in the preceding section we would have

$$\gamma_\mu = \begin{pmatrix} 0 & i(\sigma_\mu)_{\alpha\dot\beta} \\ i(\tilde\sigma_\mu)^{\dot\gamma\delta} & 0 \end{pmatrix} \tag{2.45}$$

$$\text{and } \psi = \begin{pmatrix} (\phi_L)_\alpha \\ (\phi_R)^{\dot\beta} \end{pmatrix}, \tag{2.46}$$

$$\text{with } (\sigma_\nu)_{\alpha\dot\beta} \to \mathbf{A}_\alpha{}^\gamma (\sigma_\nu)_{\gamma\dot\delta} \mathbf{A}^{\dagger\dot\delta}{}_{\dot\beta}, \tag{2.47}$$

$$(\tilde\sigma_\nu)^{\dot\alpha\beta} \to \mathbf{B}^{\dot\alpha}{}_{\dot\gamma} (\sigma_\nu)^{\dot\gamma\delta} \mathbf{B}^{\dagger\beta}_\delta, \tag{2.48}$$

$$(\phi_L)_\alpha \to \mathbf{A}_\alpha{}^\beta (\phi_L)_\beta \tag{2.49}$$

$$\text{and } (\phi_R)^{\dot\alpha} \to \mathbf{B}^{\dot\alpha}{}_{\dot\beta} (\phi_R)^{\dot\beta}, \tag{2.50}$$

where $\mathbf{B} = (\mathbf{A}^{-1})^\dagger \to \mathbf{A}\mathbf{B}^\dagger = \mathbb{1}$.

2.2.2 Gauge invariance

In addition to Lorentz invariance, we also require that all the terms in our Lagrangian be invariant under the action of the gauge group. As previously mentioned, it is the construction of our covariant derivatives and the variation of our gauge fields under the group action that is key in ensuring this. Specifically, they are constructed in order that, for some spinor field ψ and some general element $\mathbf{M}$ of the gauge group acting on ψ, we have:

$$D_\mu \psi \longrightarrow \mathbf{M} D_\mu \psi. \tag{2.51}$$

Let us show for the most simple case that the forms of D_μ and gauge field transformations stated earlier satisfy this. Taking the field E_m, we have:

$$\begin{aligned}
D_\mu E_m =& D_\mu P_L E_m + D_\mu P_R E_m \qquad (2.52)\\
=& \partial_\mu E_m - i g_1 B_\mu P_L E_m + i g_1 B_\mu P_R E_m\\
=& \partial_\mu P_L E_m + \partial_\mu P_R E_m - i g_1 B_\mu P_L E_m + i g_1 B_\mu P_R E_m\\
\longrightarrow \quad & \partial_\mu \left(e^{i\theta_1(x)} P_L E_m\right) + \partial_\mu \left(e^{-i\theta_1(x)} P_R E_m\right)\\
& - i g_1 \left[B_\mu + \frac{1}{g_1}\partial_\mu \theta_1(x)\right] e^{i\theta_1(x)} P_L E_m\\
& + i g_1 \left[B_\mu + \frac{1}{g_1}\partial_\mu \theta_1(x)\right] e^{-i\theta_1(x)} P_R E_m \qquad (2.53)\\
=& i\partial_\mu \theta_1(x) e^{i\theta_1(x)} P_L E_m + e^{i\theta_1(x)} \partial_\mu P_L E_m\\
& - i\partial_\mu \theta_1(x) e^{-i\theta_1(x)} P_R E_m + e^{-i\theta_1(x)} \partial_\mu P_R E_m\\
& - i g_1 B_\mu e^{i\theta_1(x)} P_L E_m - i\partial_\mu \theta_1(x) e^{i\theta_1(x)} P_L E_m\\
& + i g_1 B_\mu e^{-i\theta_1(x)} P_R E_m + i\partial_\mu \theta_1(x) e^{-i\theta_1(x)} P_R E_m \qquad (2.54)\\
=& e^{i\theta_1(x)} \partial_\mu P_L E_m + e^{-i\theta_1(x)} \partial_\mu P_R E_m\\
& - i g_1 B_\mu e^{i\theta_1(x)} P_L E_m + i g_1 B_\mu e^{-i\theta_1(x)} P_R E_m\\
=& e^{i\theta_1(x)} D_\mu P_L E_m + e^{-i\theta_1(x)} D_\mu P_R E_m. \qquad (2.55)
\end{aligned}$$

So, as required, we have the results:

$$D_\mu P_L E_m \to e^{i\theta_1(x)} D_\mu P_L E_m \quad (2.56)$$

$$\text{and} \quad D_\mu P_R E_m \to e^{-i\theta_1(x)} D_\mu P_R E_m. \quad (2.57)$$

Note that we have had to consider the covariant derivative acting on $P_L E_m$ and $P_R E_m$ separately as they transform differently under the gauge group. Although we will not show it explicitly here, the same behaviour $D_\mu P_L \psi \longrightarrow \mathbf{M}_L D_\mu P_L \psi$ and $D_\mu P_R \psi \longrightarrow \mathbf{M}_R D_\mu P_R \psi$ holds for all the other fields and their covariant derivatives, where $\mathbf{M}_L$ and $\mathbf{M}_R$ are the forms of gauge transformation associated with the left- and right-handed components of the field ψ respectively.

We will now look at how these results give us gauge invariance. In doing so, we will make use of the following properties of the γ matrices and the projection operators:

$$P_L^2 = P_R^2 = 1 \quad (2.58)$$

$$P_L P_R = P_R P_L = 0 \quad (2.59)$$

$$\gamma^{5\dagger} = \gamma^5 \to P_L^\dagger = P_L \ , \ \ P_R^\dagger = P_R \quad (2.60)$$

$$\left\{\gamma^5, \gamma^\mu\right\} = 0 \to P_L \gamma^\mu = \gamma^\mu P_R \ , \ \ P_R \gamma^\mu = \gamma^\mu P_L. \quad (2.61)$$

Let us consider some general Majorana field ψ_M:

$$\overline{\psi_M} \gamma^\mu D_\mu \psi_M = \left(\overline{P_L \psi_M} + \overline{P_R \psi_M}\right) \gamma^\mu D_\mu \psi_M \quad (2.62)$$

$$= \left(\overline{P_L \psi_M} P_R + \overline{P_R \psi_M} P_L\right) \gamma^\mu D_\mu \psi_M \quad (2.63)$$

$$= \overline{P_L \psi_M} \gamma^\mu P_L D_\mu \psi_M + \overline{P_R \psi_M} \gamma^\mu P_R D_\mu \psi_M \quad (2.64)$$

$$= \overline{P_L \psi_M} \gamma^\mu D_\mu P_L \psi_M + \overline{P_R \psi_M} \gamma^\mu D_\mu P_R \psi_M. \quad (2.65)$$

In the case of E_m, with our results (2.56), (2.57) and the fact that $\overline{P_L E_m} \rightarrow e^{-i\theta_1(x)}\overline{P_L E_m}$ and $\overline{P_R E_m} \rightarrow e^{i\theta_1(x)}\overline{P_R E_m}$ under the action of the gauge group, it is clear that the $U(1)$ factors cancel to leave the expression (2.65) invariant under a gauge transformation. Next consider the field $P_L\mathbf{L}_m$ and explicitly include the $SU(2)$ indices. By convention, $P_L\mathbf{L}_m$ carries a lower index $(P_L\mathbf{L}_m)_a$, with $a = 1, 2$, and the act of taking the complex conjugate raises or lowers an index so that $\overline{P_L\mathbf{L}_m}$ carries a raised index $\left(\overline{P_L\mathbf{L}_m}\right)^a$. In particular,

$$\left(\overline{P_L\mathbf{L}_m}\right)^a = \left((P_L\mathbf{L}_m)^\dagger\right)^a \gamma^0 \tag{2.66}$$

$$= ((P_L\mathbf{L}_m)_a)^\dagger \gamma^0, \tag{2.67}$$

so, taking $(P_L\mathbf{L}_m)_a \longrightarrow (M_2)_a{}^b (P_L\mathbf{L}_m)_b$ for $M_2 \in SU(2)$, we have:

$$((P_L\mathbf{L}_m)_a)^\dagger \gamma^0 \longrightarrow \left((M_2)_a{}^b (P_L\mathbf{L}_m)_b\right)^\dagger \gamma^0 \tag{2.68}$$

$$= \left((P_L\mathbf{L}_m)^\dagger\right)^b \left((M_2)^\dagger\right)_b{}^a \gamma^0 \tag{2.69}$$

$$= \left(\overline{P_L\mathbf{L}_m}\right)^b \left((M_2)^\dagger\right)_b{}^a. \tag{2.70}$$

Thus, including the action of $U(1)$ as well ($\mathbf{L}_m$ being a singlet under the $SU(3)$ element of the gauge group), we have $\left(\overline{P_L\mathbf{L}_m}\right)^a \gamma^\mu D_\mu (P_L\mathbf{L}_m)_a$ transforming as:

$$\left(\overline{P_L\mathbf{L}_m}\right)^a \gamma^\mu D_\mu (P_L\mathbf{L}_m)_a \tag{2.71}$$

$$\longrightarrow \left(\overline{P_L\mathbf{L}_m}\right)^b e^{\frac{i}{2}\theta_1(x)}((M_2)^\dagger)_b{}^a \gamma^\mu (M_2)_a{}^c e^{-\frac{i}{2}\theta_1(x)} D_\mu (P_L\mathbf{L}_m)_c \tag{2.72}$$

$$= \left(\overline{P_L\mathbf{L}_m}\right)^b \gamma^\mu D_\mu (P_L\mathbf{L}_m)_b\,, \tag{2.73}$$

where we have used the fact that $M_2 \in SU(2)$ is unitary and so

$((M_2)^\dagger)_b{}^a\,(M_2)_a{}^c = \delta_b{}^c$. A similar result holds for the term $\overline{P_R\mathbf{L}_m}\gamma^\mu D_\mu P_R\mathbf{L}_m$, where $P_R\mathbf{L}_m$ will carry a raised index $(P_R\mathbf{L}_m)^a$ as it is related to $P_L\mathbf{L}_m$ by complex conjugation. Following exactly the same arguments for the fields $\mathbf{Q}$, U and D we can therefore see that the first two lines of (2.7) are indeed invariant under the action of the gauge group.

Finally we consider the field-strength terms. In doing so, let us take the case of the non-Abelian field strength $W^a_{\mu\nu}$, from which we will be able to infer the corresponding results for $G^\alpha_{\mu\nu}$ and $B_{\mu\nu}$. With the definitions

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g_2\epsilon^{abc}W^b_\mu W^c_\nu, \tag{2.74}$$

$$\delta W^a_\mu = \frac{1}{g_2}\partial_\mu\theta^a_2(x) - \epsilon^{abc}\theta^b_2(x)W^c_\mu \quad \text{and} \tag{2.75}$$

$$\delta W^a_{\mu\nu} = \partial_\mu\delta W^a_\nu - \partial_\nu\delta W^a_\mu + g_2\epsilon^{abc}\delta W^b_\mu W^c_\nu + g_2\epsilon^{abc}W^b_\mu\delta W^c_\nu, \tag{2.76}$$

we can then use the Jacobi identity for the structure constants and the fact that partial derivatives commute to show that

$$\delta W^a_{\mu\nu} = \theta^b_2(x)\epsilon^{bac}W^c_{\mu\nu} = i\theta^b_2(x)(T^b_{adj})^{ac}W^c_{\mu\nu}, \tag{2.77}$$

where $(T^b_{adj})^{ac} = -i\epsilon^{bac}$ are the three $SU(2)$ generators in the 3×3 adjoint representation, i.e. we have found that $W^a_{\mu\nu}$ transforms under the adjoint representation. If we then consider the term $-\frac{1}{4}W^a_{\mu\nu}W^{a\mu\nu}$ and use $\delta(-\frac{1}{4}W^a_{\mu\nu}W^{a\mu\nu}) = -\frac{1}{2}\delta W^a_{\mu\nu}W^{a\mu\nu} = -\frac{1}{2}\theta^b_2(x)\epsilon^{bac}W^c_{\mu\nu}W^{a\mu\nu}$, we see that because ϵ^{bac} is antisymmetric under $a\leftrightarrow c$, whilst $W^c_{\mu\nu}W^{a\mu\nu}$ is symmetric, this term is indeed invariant: $\delta(-\frac{1}{4}W^a_{\mu\nu}W^{a\mu\nu}) = 0$. A similar result holds for the term $-\frac{g_2^2\Theta_2}{64\pi^2}\epsilon_{\mu\nu\lambda\rho}W^{a\mu\nu}W^{a\lambda\rho}$ due to the fact that $\epsilon_{\mu\nu\lambda\rho} = \epsilon_{\lambda\rho\mu\nu}$. For the equivalent terms involving $G^\alpha_{\mu\nu}$ the same derivation applies, where we can use the antisymmetry of $f^\alpha{}_{\beta\gamma}$ under $\beta\leftrightarrow\gamma$. In the case of $B_{\mu\nu}$ the invariance is easier to see. As the group $U(1)$ is Abelian, the structure constants vanish, and so $\delta B_{\mu\nu} = 0$, immediately giving us the required invariance.

In short, if in each term of our Lagrangian we were to explicitly write all indices carried by the various objects under the $SU(2)$ and $SU(3)$ elements of the gauge group, then for gauge invariance we would require that all indices be contracted over, leaving no 'free' indices. We would further require that the charges under the $U(1)$ symmetry of all objects within each term sum to zero.

It is important to note that the requirement for gauge invariance plays a key role in relation to particle masses. In the Lagrangian (2.7) we have no mass terms for any of our fields, let alone the neutrinos. By simply considering the $U(1)$ properties of a generic mass term of the form $\frac{1}{2}m^2\overline{\psi_M}\psi_M$[8], we can see that it is not gauge invariant. Using the properties (2.58)-(2.61) we have:

$$\frac{1}{2}m^2\overline{\psi_M}\psi_M = \frac{1}{2}m^2\left(\overline{P_L\psi_M} + \overline{P_R\psi_M}\right)\left(P_L\psi_M + P_R\psi_M\right) \tag{2.78}$$
$$= \frac{1}{2}m^2\left(\overline{P_L\psi_M}P_R + \overline{P_R\psi_M}P_L\right)\left(P_L\psi_M + P_R\psi_M\right) \tag{2.79}$$
$$= \frac{1}{2}m^2\left(\overline{P_L\psi_M}P_R\psi_M + \overline{P_R\psi_M}P_L\psi_M\right), \tag{2.80}$$

and given that $\overline{P_L\psi_M}$ and $P_R\psi_M$ both carry the same charge $-h$ under $U(1)$ and $\overline{P_R\psi_M}$ and $P_L\psi_M$ the same charge $+h$, this transforms as:

$$\frac{1}{2}m^2\left(\overline{P_L\psi_M}P_R\psi_M + \overline{P_R\psi_M}P_L\psi_M\right) \tag{2.81}$$
$$\longrightarrow \frac{1}{2}m^2\left(e^{-2ih\theta_1(x)}\overline{P_L\psi_M}P_R\psi_M + e^{2ih\theta_1(x)}\overline{P_R\psi_M}P_L\psi_M\right), \tag{2.82}$$

[8]Where the subscript M simply indicates that these are Majorana fields.

i.e. it is not a gauge invariant term and so cannot appear in our Lagrangian. This prohibition of explicit mass terms in our Lagrangian leads us to the Higgs mechanism as a source for our field masses.

2.2.3 Renormalisability

The last request that we make of a term in our Lagrangian is that it be renormalisable. For some given term in our Lagrangian, calculations of the corresponding physical interactions are performed perturbatively in its coupling strength, or vertex factor if we are thinking in terms of Feynman diagrams. For any diagram containing a loop there is an associated integral over the loop momentum, and this integral is often infinite if no upper bound is placed on the loop momentum. As we consider higher-order diagrams in our calculations, with more and more vertices and loops, we might expect to encounter many infinite terms.

Let us consider a term in our Lagrangian of the form $\lambda\psi_1\psi_2....\psi_m$, where λ is the coupling strength and $\psi_1\psi_2....\psi_m$ is some Lorentz and gauge invariant combination of fields. Working in natural units, if we were to put a cut-off on our loop integrals, at some value Λ say, then we would expect that for a diagram of order N our integral would behave as $\lambda^N\Lambda^D$, where D is the superficial degree of divergence. If $D > 0$ then the diagram is superficially divergent, as it will diverge in the limit $\Lambda \to \infty$[9]. By dimensional arguments, it is possible to show, [6], that for a diagram of order N with E external legs in d dimensions, the superficial degree of divergence is given by:

$$D = d - [\psi_1\psi_2....\psi_E] - N[\lambda], \tag{2.83}$$

[9] Note that the actual behaviour of a certain interaction calculation may differ from that predicted by D. Diagrams deemed not to be divergent by (2.83), may contain divergent sub-diagrams and therefore be divergent. Equally, divergent terms in the expansion may act so as to cancel.

where [ ] denotes the dimension of the enclosed object in natural units and $\psi_1\psi_2....\psi_E$ is the set of fields corresponding to the E external legs. From the form of (2.83) we can see that there are three possibilities[10]:

(1)	$[\lambda] > 0$	Degree of divergence decreases at higher orders, giving a finite number of divergent terms $\rightarrow$ *Super-renormalisable*
(2)	$[\lambda] = 0$	Degree of divergence independent of order $\rightarrow$ *Renormalisable*
(3)	$[\lambda] < 0$	Degree of divergence increases at higher orders, giving an infinite number of divergent terms $\rightarrow$ *Non-normalisable*

In order to determine $[\lambda]$, and thereby determine into which category we fall, we use the fact that the action containing our Lagrangian must be dimensionless:

$$[S] = 0 = [\int d^d x \mathcal{L}] = d[x] + [\mathcal{L}] \longrightarrow [\mathcal{L}] = d, \tag{2.84}$$

which in turn gives us:

$$[\mathcal{L}] = d = [\lambda\psi_1\psi_2....\psi_m] = [\lambda] + [\psi_1\psi_2....\psi_m] \tag{2.85}$$

$$\longrightarrow [\lambda] = d - [\psi_1\psi_2....\psi_m]. \tag{2.86}$$

In 4 dimensions, we therefore require that $[\psi_1\psi_2....\psi_m] \leq 4$ in order for our theory to be renormalisable.

But what does it mean for our theory to be renormalisable? We see that even in the case that $[\lambda] \geq 0$ we have some infinite terms in our expansion.

[10] Assuming $[\psi_1\psi_2....\psi_m] > 0$

What we find is that when performing a calculation of a certain interaction with a cut-off in place, our final expression will be a function of our coupling strength λ, our cut-off Λ and the coupling strengths of the other relevant terms in our Lagrangian. The coupling strengths in our theoretical Lagrangian, however, are not the coupling strengths that we can measure, but the two will be related. If, after re-expressing our final expression in terms of the coupling strengths we can measure, we find that the Λ dependence drops out, then we say our theory is renormalisable.

The Λ cut-off approach to renormalisation explicitly breaks gauge invariance. We are therefore forced to turn to other approaches, such as that of Dimensional Regularisation, if we desire a gauge invariant formulation. Despite this issue of gauge invariance, however, the inclusion of a cut-off in our calculations is nicely intuitive. It is the statement that we only expect our theory to be valid up to the energy scale Λ, beyond which there is new physics yet to be discovered. The effect of the new physics is to give us this discrepancy between the theoretical coupling strengths of our Lagrangian and the ones that we can measure, but once we account for this, in a renormalisable theory our calculations of interactions become independent of Λ, i.e. become independent of the high energy physics. This is a good indication that our theory is valid up to the energy scale Λ.

We will see later that the requirement for our theory to be renormalisable comes into question when considering extensions to the Standard Model that would allow for neutrino oscillations.

2.3 The Higgs Mechanism and particle masses

If we take our Lagrangian (2.7) and set aside all the interaction terms, i.e. only consider the terms at zeroth order in coupling factors g_1, g_2 and g_3, then we are left with the Lagrangian:

$$\begin{aligned}\mathcal{L}_{free} = &-\frac{1}{2}\overline{\mathbf{L}}_m \partial\!\!\!/ \mathbf{L}_m - \frac{1}{2}\overline{E}_m \partial\!\!\!/ E_m \\ &- \frac{1}{2}\overline{\mathbf{Q}}_m \partial\!\!\!/ \mathbf{Q}_m - \frac{1}{2}\overline{U}_m \partial\!\!\!/ U_m - \frac{1}{2}\overline{D}_m \partial\!\!\!/ D_m \\ &- \frac{1}{2}\left(\partial_\mu G^\alpha_\nu \partial^\mu G^{\alpha\nu} - \partial_\mu G^\alpha_\nu \partial^\nu G^{\alpha\mu}\right) \\ &- \frac{1}{2}\left(\partial_\mu W^a_\nu \partial^\mu W^{a\nu} - \partial_\mu W^a_\nu \partial^\nu W^{a\mu}\right) \\ &- \frac{1}{2}\left(\partial_\mu B_\nu \partial^\mu B^\nu - \partial_\mu B_\nu \partial^\nu B^\mu\right).\end{aligned} \tag{2.87}$$

We recognise this as the Lagrangian for fifteen *massless* spinor fields and twelve *massless* vector fields. As pointed out in section 2.2.2, explicit mass terms are prohibited by the requirement for gauge invariance.

In order to generate mass terms for our fields, we must therefore introduce a new field, the Higgs field, whose vacuum expectation value breaks the gauge symmetry such that:

$$SU(3) \times SU(2) \times U(1) \longrightarrow SU(3) \times U_\gamma(1). \tag{2.88}$$

Our electroweak symmetry $SU(2) \times U(1)$ is broken down to a single $U_\gamma(1)$ symmetry, which corresponds to electromagnetism. With the number of generators associated with the electroweak sector being reduced from four to one, we thereby imply that three have been broken. We can then infer that the three gauge fields associated with these three broken generators will have gained a mass, whilst the gauge field associated with the remaining $U_\gamma(1)$

symmetry remains massless. It is these broken and unbroken generators that correspond to $W^{\pm}_{\mu}$, Z^0_{μ} and A_{μ}.

The simplest object that we can introduce for breaking the symmetry is a doublet of complex scalar fields transforming under $SU(3) \times SU(2) \times U(1)$ as $(\mathbf{1}, \mathbf{2}, \frac{1}{2})$. Using the notation of [4], we label the new Higgs field ϕ with:

$$\phi = \begin{pmatrix} \phi^{+} \\ \phi^{0} \end{pmatrix}. \tag{2.89}$$

Taking the complex conjugate of this field, and for a moment including the $SU(2)$ indices, we get the object:

$$\phi^{*a} = (\phi_a)^* = \begin{pmatrix} \phi^{+*} \\ \phi^{0*} \end{pmatrix}^{a}, \tag{2.90}$$

which transforms under $SU(3) \times SU(2) \times U(1)$ as $(\mathbf{1}, \overline{\mathbf{2}}, -\frac{1}{2})$. But, under $SU(2)$ there is an invariant tensor ϵ_{ab}:

$$\epsilon_{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}_{ab}, \tag{2.91}$$

which we can use to lower the index on ϕ^* to give the object:

$$\tilde{\phi}_a = \epsilon_{ab}\phi^{*b} = \begin{pmatrix} \phi^{0*} \\ -\phi^{+*} \end{pmatrix}_{a} \tag{2.92}$$

transforming under $SU(3) \times SU(2) \times U(1)$ as $(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$. With these new objects, we are now able to include the following additional gauge invariant terms in our Lagrangian:

$$\begin{aligned}\mathcal{L}_H = &- (D_\mu \phi^\dagger)(D^\mu \phi) - V(\phi^\dagger \phi) \\ &- (f_{mn} \overline{P_L \mathbf{L}}_m P_R E_n \phi + f^\dagger_{nm} \phi^\dagger \overline{P_R E}_n P_L \mathbf{L}_m) \\ &- (h_{mn} \overline{P_L \mathbf{Q}}_m P_R D_n \phi + h^\dagger_{nm} \phi^\dagger \overline{P_R D}_n P_L \mathbf{Q}_m) \\ &- (g_{mn} \overline{P_L \mathbf{Q}}_m P_R U_n \tilde{\phi} + g^\dagger_{nm} \tilde{\phi}^\dagger \overline{P_R U}_n P_L \mathbf{Q}_m).\end{aligned} \tag{2.93}$$

Given the transformation properties of ϕ, its covariant derivative is given by:

$$D_\mu \phi = \partial_\mu \phi - \frac{i}{2} g_2 W^a_\mu \sigma_a \phi - \frac{i}{2} g_1 B_\mu \phi. \tag{2.94}$$

V must be a function of $\phi^\dagger \phi$ in order to satisfy gauge invariance. Furthermore, in order to satisfy renormalisability, we can only have terms of up to order $(\phi^\dagger \phi)^2$. The actual form of potential taken is:

$$V(\phi^\dagger \phi) = \lambda \left[\phi^\dagger \phi - \frac{\mu^2}{2\lambda} \right]^2 \tag{2.95}$$

$$= \lambda (\phi^\dagger \phi)^2 - \mu^2 \phi^\dagger \phi + \frac{\mu^4}{4\lambda}, \tag{2.96}$$

where λ and μ are constants. In order to ensure the reality of our Lagrangian we require λ and μ^2 to be real, for the potential to be bound from below we require that λ be positive, and in order that the potential have a vacuum expectation value that breaks the $SU(2) \times U(1)$ symmetry we require μ^2 to be positive.

The potential is minimised by the vacuum expectation value $\phi^\dagger \phi = \frac{\mu^2}{2\lambda}$. Without loss of generality, we may take our vacuum field ϕ_v to 'lie in a given direction', i.e. we are able to choose how to distribute this vacuum expectation value amongst the four components $\mathrm{Re}(\phi^+)$, $\mathrm{Im}(\phi^+)$, $\mathrm{Re}(\phi^0)$

and $\text{Im}(\phi^0)$. We make the choice which turns out to make the spectrum of our theory most transparent[11]:

$$\phi_v = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}, \quad \text{where } v \text{ is real and satisfies } v^2 = \frac{\mu^2}{\lambda}. \tag{2.97}$$

As hinted at previously, we find that there is only a single linear combination of our $SU(2) \times U(1)$ generators that leaves the vacuum state invariant. Let's call this linear combination ζ_I, which satisfies $\zeta_I \phi_v = 0$. There are a further three linear combinations of the $SU(2) \times U(1)$ generators, which we will label ζ_1, ζ_2 and ζ_3, that do not leave ϕ_v invariant, i.e. $\zeta_i \phi_v \neq 0$. It can be shown, [7], that the three objects $\zeta_i \phi_v$ span a three dimensional subspace of ϕ, which allows us to re-express our field ϕ as an expansion around its vacuum state ϕ_v using the following parameterisation:

$$\phi = e^{i\theta_i(x)\zeta_i} \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}\left(v + H(x)\right) \end{pmatrix} \quad i = 1, 2, 3, \tag{2.98}$$

where $H(x)$ is a real scalar field. Next, exploiting the gauge invariance of our Lagrangian, we can use a gauge transformation to take us into the unitary gauge, leaving us with:

$$\phi = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}\left(v + H(x)\right) \end{pmatrix}. \tag{2.99}$$

Substituting this into our Lagrangian and writing out the $SU(2)$ doublet components explicitly, we find a form for $\mathcal{L}_H$ from which we can read off our field masses:

[11] If a different 'direction' were chosen, we could always rotate it back to this form via a gauge transformation, exploiting the gauge invariance of our Lagrangian.

$$\begin{aligned}\mathcal{L}_H =& -\frac{1}{2}\partial_\mu H\partial^\mu H - \lambda v^2 H^2 - \lambda v H^3 - \frac{\lambda}{4}H^4 \\ & - \frac{1}{8}g_2^2(v+H)^2(W_\mu^1 + iW\mu^2)(W^{1\mu} - iW^{2\mu}) \\ & - \frac{1}{8}(v+H)^2(-g_2 W\mu^3 + g_1 B_\mu)(-g_2 W^{3\mu} + g_1 B^\mu) \\ & - \frac{1}{\sqrt{2}}(v+H)(f_{mn}\overline{P_L\mathcal{E}}_m P_R E_n + f^\dagger_{nm}\overline{P_R E}_n P_L \mathcal{E}_m) \\ & - \frac{1}{\sqrt{2}}(v+H)(h_{mn}\overline{P_L\mathcal{D}}_m P_R D_n + h^\dagger_{nm}\overline{P_R D}_n P_L \mathcal{D}_m) \\ & - \frac{1}{\sqrt{2}}(v+H)(g_{mn}\overline{P_L\mathcal{U}}_m P_R U_n + g^\dagger_{nm}\overline{P_R U}_n P_L \mathcal{U}_m).\end{aligned} \tag{2.100}$$

Comparing the term $-\lambda v^2 H^2$ to the standard mass term for a spin-zero scalar field, which takes the form $-\frac{1}{2}m_H^2 H^2$, we determine that, to lowest order, $m_H^2 = 2\lambda v^2 = 2\mu^2$.

The standard mass term for a spin-one field is $-\frac{1}{2}m^2 W_\mu W^\mu$. Thus, from the term $-\frac{1}{8}g_2^2(v+H)^2(W_\mu^1 W^{1\mu} + W\mu^2 W^{2\mu})$ we can read off the masses $m_{W^1}^2 = m_{W^2}^2 = \frac{1}{4}g_2^2 v^2$.

Next we make the substitution[12]:

$$Z_\mu^0 = \frac{-g_1 B_\mu + g_2 W_\mu^3}{\sqrt{g_1^2 + g_2^2}}, \tag{2.101}$$

and identify $\cos\theta_W = \frac{g_2}{\sqrt{g_1^2+g_2^2}}$ and $\sin\theta_W = \frac{g_1}{\sqrt{g_1^2+g_2^2}}$ so that

$$Z_\mu^0 = \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu, \tag{2.102}$$

where θ_W is the weak-mixing or Weinberg angle. This gives us the term $-\frac{1}{8}v^2(g_1^2+g_2^2)Z_\mu^0 Z^{0\mu}$, from which we read off the mass $m_Z^2 = \frac{1}{4}v^2(g_1^2+g_2^2)$.

[12] The $\sqrt{g_1^2+g_2^2}$ factor gives us the correct normalisation for the field strength terms we acquire for Z_μ^0 and A_μ on making the relevant substitutions for W_μ^3 and B_μ in our Lagrangian (2.87).

The massless combination of W_μ^3 and B_μ orthogonal to Z_μ^0, labelled A_μ, is given by:

$$A_\mu = \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu. \tag{2.103}$$

Being massless, A_μ is the gauge field associated with the remaining unbroken symmetry $U_\gamma(1)$, i.e. the photon, and the generators for this symmetry will be some linear combination of the four generators of the initial $SU(2) \times U(1)$ symmetry. Let us try to find this combination. Transforming as $(\mathbf{1}, \mathbf{2}, \frac{1}{2})$ under $SU(3) \times SU(2) \times U(1)$, our ϕ_v will transform under a constant gauge transformation as:

$$\begin{aligned}
\delta\phi_v = \frac{1}{\sqrt{2}}\delta \begin{pmatrix} 0 \\ v \end{pmatrix} &= \frac{i}{2}\theta_2^a \sigma_a \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} + \frac{i}{2}\theta_1 \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \\
&= \frac{i}{2}\frac{v}{\sqrt{2}} \begin{pmatrix} (\theta_2^1 - i\theta_2^2) \\ (\theta_1 - \theta_2^3) \end{pmatrix}.
\end{aligned} \tag{2.104}$$

Thus, in order that our ϕ_v be invariant, i.e. $\delta\phi_v = 0$, we require $\theta_2^1 = \theta_2^2 = 0$ and $\theta_1 = \theta_2^3 \equiv \theta_\gamma$. This tells us that the combination $Q = \frac{1}{2}\sigma_3 + h$ is the generator for our unbroken symmetry. The associated conserved quantity is the electric charge.

With our knowledge of this unbroken symmetry, we can now recover $W_\mu^\pm$ from the fields W_μ^1 and W_μ^2. If we substitute our particular transformation parameters in to equation (2.17), remembering that we are dealing with a rigid gauge transformation, then we have, in matrix form:

$$\delta \begin{pmatrix} W_\mu^1 \\ W_\mu^2 \\ W_\mu^3 \end{pmatrix} = \theta_\gamma \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} W_\mu^1 \\ W_\mu^2 \\ W_\mu^3 \end{pmatrix} = \begin{pmatrix} W_\mu^2 \\ -W_\mu^1 \\ 0 \end{pmatrix}, \tag{2.105}$$

from which we can see that W_μ^1 and W_μ^2 transform amongst themselves. Defining

$$W_\mu^\pm \equiv \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \tag{2.106}$$

we see that these fields will transform as $\delta W_\mu^\pm = \pm i\theta_\gamma W_\mu^\pm$, i.e. they carry charge ± 1 under the $U_\gamma(1)$ symmetry. Their masses are given by $m_{W^\pm} = m_{W^1} = m_{W^2}$. Also note that because the fields W_μ^3 and B_μ do not transform under the $U_\gamma(1)$ symmetry, Z_μ^0 and A_μ carry zero charge.

Thus we have recovered the familiar gauge bosons $W_\mu^\pm$, Z_μ^0 and A_μ, and all except the photon have acquired a mass through the Higgs Mechanism.

We now move on to the fermions. The relevant terms here are those in the last three lines of (2.100). As it stands, the matrices f_{mn}, g_{mn} and h_{mn} are not diagonal and so we are unable to read off the fermion masses immedately. However, by redefining our fermion fields as

$$\begin{aligned} P_R E_m &= M^e_{mn} P_R E^{'}_n, & P_L \mathcal{E}_m &= M^{e*}_{mn} P_L \mathcal{E}^{'}_n, \\ P_R U_m &= M^u_{mn} P_R U^{'}_n, & P_L \mathcal{U}_m &= M^{u*}_{mn} P_L \mathcal{U}^{'}_n, \\ P_R D_m &= M^d_{mn} P_R D^{'}_n \ \text{and} & P_L \mathcal{D}_m &= M^{d*}_{mn} P_L \mathcal{D}^{'}_n \end{aligned} \tag{2.107}$$

and making appropriate choices for M^e_{mn}, M^u_{mn} and M^d_{mn} we are able to bring these terms into the diagonal form[13]

$$
\begin{aligned}
&-\frac{1}{\sqrt{2}}(v+H)f_{\hat{m}}(\overline{P_L\mathcal{E}}_m P_R E_m + \overline{P_R E}_m P_L \mathcal{E}_m) \\
&-\frac{1}{\sqrt{2}}(v+H)h_{\hat{m}}(\overline{P_L\mathcal{D}}_m P_R D_m + \overline{P_R D}_m P_L \mathcal{D}_m) \\
&-\frac{1}{\sqrt{2}}(v+H)g_{\hat{m}}(\overline{P_L\mathcal{U}}_m P_R U_m + \overline{P_R U}_m P_L \mathcal{U}_m),
\end{aligned} \tag{2.108}
$$

where we have now dropped the primes and the hat on the third m index in each term indicates that it does not invoke Einstein summation. The $f_{\hat{m}}$, $g_{\hat{m}}$ and $h_{\hat{m}}$ are all real and positive. Also note that in order for the kinetic terms of our Lagrangian to remain in the standard form, we require M^e_{mn}, M^u_{mn} and M^d_{mn} to be unitary matrices. Using the result $P_R P_L = P_L P_R = 0$, we can now write (2.108) as

$$
\begin{aligned}
&-\frac{1}{\sqrt{2}}(v+H)f_{\hat{m}}\left((\overline{P_L\mathcal{E}}_m + \overline{P_R E}_m)(P_L\mathcal{E}_m + P_R E_m)\right) \\
&-\frac{1}{\sqrt{2}}(v+H)h_{\hat{m}}\left((\overline{P_L\mathcal{D}}_m + \overline{P_R D}_m)(P_L\mathcal{D}_m + P_R D_m)\right) \\
&-\frac{1}{\sqrt{2}}(v+H)g_{\hat{m}}\left((\overline{P_L\mathcal{U}}_m + \overline{P_R U}_m)(P_L\mathcal{U}_m + P_R U_m)\right).
\end{aligned} \tag{2.109}
$$

Finally, making the identifications $P_L\mathcal{E}_m + P_R E_m = e_m$ etc, where e_m are the Dirac fields for the electron, muon and taon for $m = 1, 2$ and 3 respectively, we have

$$
-\frac{1}{\sqrt{2}}(v+H)\left(f_{\hat{m}}\overline{e}_m e_m + h_{\hat{m}}\overline{d}_m d_m + g_{\hat{m}}\overline{u}_m u_m\right). \tag{2.110}
$$

From this we are now able to read off the masses of the Dirac fields e_m, d_m and u_m as $\frac{1}{\sqrt{2}}f_{\hat{m}}v$, $\frac{1}{\sqrt{2}}h_{\hat{m}}v$ and $\frac{1}{\sqrt{2}}g_{\hat{m}}v$ respectively.

The key point here with regard to neutrino masses is that because there is

[13] See [4] for proof that this is always possible.

no right-handed neutrino in the Standard Model, we are unable to construct a gauge invariant Yukawa term involving the Higgs field that would generate a neutrino mass in the same way as shown above for the other fermions.

2.4 Accidental symmetries and additive conserved charges

Whilst we will not explicitly go through all the interaction terms of the Standard Model Lagrangian here, what one finds is that in addition to being invariant under the action of the gauge group there are other global symmetries, namely

$$U_e(1) \times U_\mu(1) \times U_\tau(1) \times U_B(1). \tag{2.111}$$

These symmetries were not requirements of our initial Lagrangian, but are by-products of requirements we have imposed on it. As such, we call them accidental symmetries. For the leptons we have the three symmetries[14]

$$P_L \mathbf{L}_m \to e^{i\alpha_m} P_L \mathbf{L}_m, \tag{2.112}$$

$$P_R E_m \to e^{i\alpha_m} P_R E_m \tag{2.113}$$

and correspondingly

$$P_R \mathbf{L}_m \to e^{-i\alpha_m} P_R \mathbf{L}_m \text{ and} \tag{2.114}$$

$$P_L E_m \to e^{-i\alpha_m} P_L E_m. \tag{2.115}$$

[14]Note that the repeated index is not summed over here.

In terms of Dirac fields, this is equivalent to $e_m \to e^{i\alpha_m} e_m$ and $v_m \to e^{i\alpha_m} v_m$. By Noether's theorem we know that associated with any such symmetry there is a conserved current

$$J^\mu = \frac{\delta \mathcal{L}}{\delta(\partial_\mu \psi_a)} \delta\psi_a, \tag{2.116}$$

where ψ_a represent the fields contained in the Lagrangian, and a conserved charge

$$Q = \int d^3x J^0 \tag{2.117}$$

In our case, expressing the relevant kinetic terms as

$$\begin{aligned} -\frac{1}{2}\overline{\mathbf{L}}_m \gamma^\mu \partial_\mu \mathbf{L}_m - \frac{1}{2}\overline{E}_m \gamma^\mu \partial_\mu E_m = & -\frac{1}{2}\overline{P_L \mathbf{L}}_m \gamma^\mu \partial_\mu P_L \mathbf{L}_m \\ & -\frac{1}{2}\overline{P_R \mathbf{L}}_m \gamma^\mu \partial_\mu P_R \mathbf{L}_m \\ & -\frac{1}{2}\overline{P_L E}_m \gamma^\mu \partial_\mu P_L E_m \\ & -\frac{1}{2}\overline{P_R E}_m \gamma^\mu \partial_\mu P_R E_m \end{aligned} \tag{2.118}$$

we have

$$\begin{aligned} J^\mu = & -\frac{1}{2}\overline{P_L \mathbf{L}}_m \gamma^\mu i\alpha_m P_L \mathbf{L}_m + \frac{1}{2}\overline{P_R \mathbf{L}}_m \gamma^\mu i\alpha_m P_R \mathbf{L}_m \\ & + \frac{1}{2}\overline{P_L E}_m \gamma^\mu i\alpha_m P_L E_m - \frac{1}{2}\overline{P_R E}_m \gamma^\mu i\alpha_m P_R E_m, \end{aligned} \tag{2.119}$$

which in turn gives us[15]

$$
\begin{aligned}
Q =& \frac{1}{2}\alpha_m \int d^3x \Big[-(P_L\mathbf{L}_m)^\dagger P_L\mathbf{L}_m + (P_R\mathbf{L}_m)^\dagger P_R\mathbf{L}_m \qquad (2.120)\\
&+ (P_LE_m)^\dagger P_LE_m - (P_RE_m)^\dagger P_RE_m \Big] \\
=& \frac{1}{2}\alpha_m \int d^3x \Big[- e^\dagger_{Lm} e_{Lm} - v^\dagger_{Lm} v_{Lm} \qquad (2.121)\\
&- e^T_{Lm} e^*_{Lm} - v^T_{Lm} v^*_{Lm} - e^T_{Rm} e^*_{Rm} + e^\dagger_{Rm} e_{Rm} \Big] \\
=& -\alpha_m \int d^3x \Big[e^\dagger_m e_m + v^\dagger_m v_m \Big], \qquad (2.122)
\end{aligned}
$$

where

$$
e_m = \begin{pmatrix} e_{Lm} \\ e_{Rm} \end{pmatrix} \text{ and } v_m = \begin{pmatrix} v_{Lm} \\ 0 \end{pmatrix} \qquad (2.123)
$$

are the Dirac fields for the three flavours of lepton and neutrino respectively. In the canonical second quantisation we have Dirac fields given by

$$
e_m(x) = \int \frac{d^3\mathbf{p}}{(2\pi)^3\sqrt{2E_\mathbf{p}}} \sum_s \left[u^m(\mathbf{p},s)\mathbf{a}^m_{\mathbf{p},s} e^{ip\cdot x} + v^m(\mathbf{p},s)\mathbf{b}^{m\dagger}_{\mathbf{p},s} e^{-ip\cdot x} \right], \qquad (2.124)
$$

where $\mathbf{a}^m_{\mathbf{p},s}$ and $\mathbf{b}^{m\dagger}_{\mathbf{p},s}$ can be viewed as the electron annihilation operator and anti-electron creation operator respectively, $u^m(\mathbf{p},s)$ and $v^m(\mathbf{p},s)$ are four-component vectors and $\mathbf{p}$ and s label the momentum and spin state of the particles respectively. Moving into momentum space we thus find

$$
Q = -\alpha_m \int \frac{d^3\mathbf{p}}{(2\pi)^3} \sum_s \left[\mathbf{a}^{m\dagger}_{\mathbf{p},s}\mathbf{a}^m_{\mathbf{p},s} - \mathbf{b}^{m\dagger}_{\mathbf{p},s}\mathbf{b}^m_{\mathbf{p},s} + \mathbf{c}^{m\dagger}_{\mathbf{p},s}\mathbf{c}^m_{\mathbf{p},s} - \mathbf{d}^{m\dagger}_{\mathbf{p},s}\mathbf{d}^m_{\mathbf{p},s} \right], \qquad (2.125)
$$

where $\mathbf{c}^{m\dagger}_{\mathbf{p},s}/\mathbf{c}^m_{\mathbf{p},s}$ and $\mathbf{d}^{m\dagger}_{\mathbf{p},s}/\mathbf{d}^m_{\mathbf{p},s}$ are the creation/annihilation operators for the neutrinos and anti-neutrinos respectively.

[15] Remembering that as we are using the metric $\eta^{\mu\nu} = \text{diag}(-1,1,1,1)$, $\overline{\psi} = \psi^\dagger i\gamma^0$ and $(\gamma^0)^2 = -\mathbb{1}$.

As we know from the formalism of creation and annihilation operators, objects of the form $\mathbf{a}^\dagger\mathbf{a}$ correspond to number operators. Thus, as we are integrating over all particle momenta and summing over all spin states, this conserved charge corresponds to a conservation of particle numbers. Denoting the number of m-flavour leptons N_m^+, the number of m-flavour anti-leptons N_m^-, the number of m-flavour neutrinos N_m^v and the number of m-flavour anti-neutrinos $N_m^{\overline{v}}$ we have the conserved quantities

$$\text{individual Lepton number} \quad L_m = N_m^+ + N_m^v - N_m^- - N_m^{\overline{v}} \tag{2.126}$$

$$\text{and total Lepton number} \quad L = \sum_m L_m. \tag{2.127}$$

In the case of the quarks, it turns out that whilst we do have conservation of the total quark number, we do not have conservation of individual quark generation numbers. This is a result of the fact that when we consider the charged-current interactions of the electro-weak sector[16] and express fields in terms of their mass eigenstates, we pick up terms of the form

$$V_{mn}\overline{u}_m\gamma^\mu P_L d_n + V_{mn}^\dagger \overline{d}_m\gamma^\mu P_L u_n, \tag{2.128}$$

where V_{mn} is the Kobayashi-Maskawa matrix. Being non-diagonal, these terms don't allow us the individual $U_m(1)$ symmetries $u_m, d_m \to e^{i\alpha_m}u_m$, but do allow us the symmetry $u_m, d_m \to e^{i\alpha}u_m$ for all m. This is the $U_B(1)$ symmetry, where B is the Baryon number and the fields u_m and d_m are in fact chosen to carry charge $\frac{1}{3}$ so that $u_m, d_m \to e^{i\frac{\alpha}{3}}u_m$. This choice of fractional charge is made in order that the baryons/anti-baryons, which are states of 3 quarks/antiquarks, have charge $\pm$ under the symmetry.

Briefly returning to the leptons, it is worth mentioning that when we

[16]i.e. those involving $W^\pm$

consider their charged-current interactions we also find terms of the form

$$U_{mn}\overline{e}_m\gamma^\mu P_L v_n + U^\dagger_{mn}\overline{v}_m\gamma^\mu P_L e_n, \tag{2.129}$$

where U_{mn} is a unitary matrix. However, because there are no mass terms for the neutrinos and the rest of the Lagrangian is invariant under the redefinition of our fields $v_m \to U^\dagger_{mn}v_n$, we are able to exploit $U^\dagger U = \mathbb{1}$ and bring the terms into diagonal form.

The accidental symmetries just considered are well confirmed by experiments. However, we will see in the following section that in introducing a neutrino mass we are potentially forced to relinquish Lepton number conservation.

3 Neutrino oscillations and extensions to the Standard Model

Neutrinos within the Standard Model are taken to be left handed and massless. However, there is now much experimental evidence indicating that free neutrinos oscillate between flavours. For example, as first detected by Raymond Davis' experiment in the late 1960's, the flux of solar neutrinos detected on earth is a third of what we expect, suggesting that electron-neutrinos oscillate into other states whilst in transit from the sun to the earth. Similarly, experiments such as Kamiokande have detected a deficit in atmospheric muon-neutrinos, and this deficit is seen to increase with the distance between the point of creation and detection.

As originally set out by Gribov and Pontecorvo, [8], the observation of neutrino oscillations is taken as evidence that they are in fact not massless. As such, any theory of particle physics is required to accommodate non-zero neutrino masses. For the Standard Model, this proves to be a major stumbling block as, with the assumed particle content, gauge invariance and renormalisability, there is no way in which a non-zero neutrino mass can be generated. The Standard Model must therefore be extended in order to accommodate neutrino masses, and in doing so we are forced to allow the conditions of assumed particle content and renormalisability to be relaxed. In the following section we hope to outline the reasons for this connection between neutrino oscillations and their having mass. We will also discuss some of the most popular extensions to the Standard model that accommodate these masses and what it means to relax our initial criteria in making these extensions. For a review of neutrino oscillation phenomena and experimental results see, for example, [9] and [10].

3.1 Oscillation phenomenon and particle masses

The basic principle behind neutrino oscillations is fairly simple. If we allow neutrinos to have a mass, then, as with the quarks, the neutrino *mass* eigenstates and *weak interaction* or *flavour* eigenstates will in general not be the same. In any experiment where neutrinos are measured, the measurement is made via the neutrinos weak interaction with the detector. As such, our measurements tell us the flavour of the neutrino. Equally, the neutrinos being detected will originally have been created via a weak interaction, and thus start off in some particular flavour eigenstate. Between the point of creation and the point of detection, however, the states will evolve according to the free Hamiltonian, whose eigenstates are the mass eigenstates[1]. Thus, if we expand our initial flavour eigenstate in terms of mass eigenstates and each of these components evolve differently according to the free Hamiltonian, then our neutrino will not remain in the initial flavour eigenstate. Instead, whatever the final state is in terms of the mass eigenstates, we must then re-express this as a sum of flavour eigenstates in order to determine a prediction for the neutrino flavour we will measure.

[1] We restrict ourselves to considering oscillations in vacuo here.

Let us take our flavour eigenstates $|\nu_\alpha\rangle$ and the mass eigenstates $|\nu_i\rangle$ to be related by the unitary mixing matrix U as

$$|\nu_\alpha\rangle = U_{\alpha i}|\nu_i\rangle \;\text{ and }\; |\nu_i\rangle = U^\dagger_{i\alpha}|\nu_\alpha\rangle, \tag{3.1}$$

where we take the general case of n mass and flavour eigenstates so that $i, \alpha = 1, ..., n$. If our state is initially a flavour eigenstate and we take it to be produced at $\mathbf{x} = 0$ and $t = 0$, then we have

$$|\nu(0,0)\rangle = |\nu_\alpha\rangle = U_{\alpha i}|\nu_i\rangle. \tag{3.2}$$

After time t and at position $\mathbf{x}$, this state will have evolved to

$$|\nu(\mathbf{x},t)\rangle = U_{\alpha i}e^{-iE_i t + i\mathbf{p}_i\cdot\mathbf{x}}|\nu_i\rangle = U_{\alpha i}e^{-iE_i t + i\mathbf{p}_i\cdot\mathbf{x}}U^\dagger_{i\beta}|\nu_\beta\rangle. \tag{3.3}$$

If we are then interested in the probability of measuring the neutrino to be in some flavour eigenstate $|\nu_\gamma\rangle$, then we must calculate the amplitude $\langle\nu_\gamma|\nu(\mathbf{x},t)\rangle$:

$$\langle\nu_\gamma|\nu(\mathbf{x},t)\rangle = e^{-iE_i t + i\mathbf{p}_i\cdot\mathbf{x}}U_{\alpha i}U^\dagger_{i\beta}\langle\nu_\gamma|\nu_\beta\rangle = e^{-iE_i t + i\mathbf{p}_i\cdot\mathbf{x}}U_{\alpha i}U^\dagger_{i\gamma}. \tag{3.4}$$

The relativistic dispersion relation for a particle of mass m_i and momentum $\mathbf{p}_i$ is given by

$$E_i = \sqrt{|\mathbf{p}_i|^2 + m_i^2}. \tag{3.5}$$

If we assume that the energy of our neutrino is known accurately enough to allow us to say $E_i = E$ for all i but that the finite size of the source and detector allow for a range of possible neutrino momenta, then we can write

$$|\mathbf{p}_i| = \sqrt{E^2 - m_i^2} = E\sqrt{1 - \frac{m_i^2}{E^2}} \simeq E\left(1 - \frac{m_i^2}{2E^2}\right), \tag{3.6}$$

where we have used $E \gg m_i$. Substituting this result into (3.4) and using the approximation $x \approx t$ for small m_i we get[2]

$$\langle \nu_\gamma | \nu(\mathbf{x}, t) \rangle = e^{-i\frac{m_i^2 x}{2E}} U_{\alpha i} U_{i\gamma}^\dagger. \tag{3.7}$$

To determine the probability of measuring neutrino flavour state $|\nu_\gamma\rangle$ we require $|\langle \nu_\gamma | \nu(\mathbf{x}, t) \rangle|^2$:

$$\begin{aligned} |\langle \nu_\gamma | \nu(\mathbf{x}, t) \rangle|^2 &= e^{i\frac{m_j^2 x}{2E}} U_{\alpha j}^* U_{j\gamma}^T e^{-i\frac{m_i^2 x}{2E}} U_{\alpha i} U_{i\gamma}^\dagger & (3.8) \\ &= e^{i\frac{(m_j^2 - m_i^2)x}{2E}} U_{\alpha i} U_{i\gamma}^\dagger U_{\alpha j}^* U_{j\gamma}^T & (3.9) \\ &= e^{i\frac{\Delta m_{ji}^2 x}{2E}} U_{j\alpha}^\dagger U_{\alpha i} U_{i\gamma}^\dagger U_{\gamma j}, & (3.10) \end{aligned}$$

where $\Delta m_{ji}^2 = m_j^2 - m_i^2$. As such, we see that our probability for measuring a given flavour state has acquired a phase that depends on x, E and Δm_{ji}^2, and it is this phase factor that can account for the observed oscillation phenomenon. If all our neutrinos were massless, then this phase factor would not have appeared and we would not expect to see oscillations. Another important point to note is that our measurements are only sensitive to mass-squared differences. This means that we would only require one neutrino to be massive in order to observe oscillations, and also means that if all neutrino masses were non-zero but degenerate then we would not observe the oscillation effect.

[2] Note that without loss of generality we have also made the simplification of taking $\mathbf{p}$ and $\mathbf{x}$ to lie along the same direction.

The arguments outlined above are somewhat simplistic, but determining a full Quantum Field Theory formulation of the oscillation phenomenon is something that has proven to be difficult. For a discussion of the simplifications made in the above formulation and a review of the numerous other approaches that have been taken and their associated problems see [11]. Surprisingly, the final predictions of the simple analysis included here and the various, more complete QFT formulations do, in general, all agree. However, it has been suggested in [12] that a full QFT formulation does in fact give corrections to these standard oscillation equations.

3.2 Mechanisms for massive neutrinos

As already mentioned, with the assumed particle content and conditions of invariance under the action of the gauge group and renormalisability, there is no way in which a neutrino mass can be generated within the Standard Model. In light of the experimental evidence for neutrinos having mass, we are therefore forced to consider relaxing one or more of the constraints of the Standard Model. In the following section we consider some such possibilities and find that, as well as having to relax the condition of renormalisability and alter the assumed particle content, the conservation of lepton numbers is also brought into question.

3.2.1 Relinquishing particle content

Perhaps the most natural extension to propose would, in analogy with E_m, U_m and D_m, be the introduction of right-handed neutrino fields V_m transforming as $(\mathbf{1}, \mathbf{1}, 0)$ under the action of the gauge group. Being singlets under the $SU(3)$ and $SU(2)$ elements of the gauge group and carrying zero charge under the $U(1)$ element, this type of neutrino is said to be *sterile*. With the introduction of these fields we are then able construct the Dirac neutrino field

$$\nu_m = P_L \mathcal{V}_m + P_R V_m = \begin{pmatrix} \nu_{Lm} \\ \nu_{Rm} \end{pmatrix} \tag{3.11}$$

just as we were able to for the other fermions. Furthermore, we are also able to construct the Yukawa term

$$-(k_{mn} \overline{P_L \mathbf{L}}_m P_R V_n \tilde{\phi} + k^{\dagger}_{nm} \tilde{\phi}^{\dagger} \overline{P_R V}_n P_L \mathbf{L}_m), \tag{3.12}$$

which, after symmetry breaking, will give us

$$-\frac{1}{\sqrt{2}}(v + H)\,(k_{\hat{m}} \overline{\nu}_m \nu_m) \tag{3.13}$$

in exact analogy with the other leptons considered earlier. From this we are able to read off the neutrino masses $\frac{1}{\sqrt{2}} k_{\hat{m}} v$. Unlike the other leptons, however, we are faced with an additional complication due to the fact that the fields V_m transform as $(\mathbf{1}, \mathbf{1}, 0)$. Given their transformation properties, we are permitted to construct the gauge invariant Majorana mass term $\frac{1}{2} M_{mn} \overline{V}_m V_n$. In fact, as such, the Dirac neutrino we defined in (3.11) is actually ill-defined unless $M_{mn} = 0$. Furthermore, this Majorana mass term does not preserve individual or *total* Lepton numbers. Once again in analogy with the the other leptons, under the individual Lepton number symmetries we have $P_R V_m \to e^{i\alpha_m} P_R V_m$, which gives us

$$\begin{aligned}\frac{1}{2} M_{mn} \overline{V}_m V_n &= \frac{1}{2} M_{mn} (\overline{P_L V}_m P_R V_n + \overline{P_R V}_m P_L V_n) \\ &\to \frac{1}{2} M_{mn} (e^{i(\alpha_n + \alpha_m)} \overline{P_L V}_m P_R V_n + e^{-i(\alpha_n + \alpha_m)} \overline{P_R V}_m P_L V_n),\end{aligned} \tag{3.14}$$

and thus demonstrates that $\frac{1}{2} M_{mn} \overline{V}_m V_n$ is not invariant under the individual transformations or under the total Lepton symmetry where $P_R V_m \to e^{i\alpha} P_R V_m$ for all m. In order to preserve even the total Lepton number we must therefore take $M_{mn} = 0$, which in turn allows us to define the Dirac spinor as in (3.11).

Even after taking $M_{mn} = 0$ there are some important issues remaining. Firstly, in order to produce the observed tiny neutrino masses, we would require the matrix elements of k_{mn} to be many orders of magnitude smaller than those of f_{mn}, g_{mn} and h_{mn}. Whilst this is not disallowed, and we concede that there is already an element of arbitrariness in the values of other fundamental constants in the Standard Model and in setting $M_{mn} = 0$, such a fine tuning of these matrix elements close to zero does not seem very desirable.

Secondly, whilst total Lepton number is still preserved, we see that introducing the Yukawa term involving V_m does cause the individual Lepton number symmetries to be broken. As discussed in section 2.4, the reason that individual Lepton numbers were conserved whereas only the total quark number was conserved was due to the fact that in the case of the leptons we were free to redefine our fields $\mathcal{V}_m$ as we liked in order to leave the charged-current interaction terms diagonal. With the introduction of the new Yukawa term, however, this is no longer the case and, as with the quarks, we pick up a CKM-like matrix called the *Pontecorvo Maki Nak-*

agawa Sakata (PMNS) matrix in our charged-current interactions. The resulting non-diagonal form then prevents individual Lepton numbers from being conserved.

Another natural question to ask is why do we not observe these additional right-handed neutrino states? Despite the fact that our new fields are sterile, we would expect the existence of these sterile neutrinos (and in general we might consider the possibility that there are more than three additional sterile neutrinos) to have some effect on the observed rates of interactions involving the observed neutrinos, as the known neutrinos can oscillate into the sterile states. We are able to use the lack of evidence for such effects to constrain the mixing between the sterile neutrinos and those observed.

If, for a moment, we were to allow for the non-conservation of Lepton number, then after symmetry breaking our mass terms could be expressed in matrix form as

$$-\frac{1}{2}\begin{pmatrix} (\overrightarrow{P_L\mathcal{V}})^T & (\overrightarrow{P_LV})^T \end{pmatrix}\begin{pmatrix} 0 & \frac{v}{\sqrt{2}}k \\ \frac{v}{\sqrt{2}}k^T & M \end{pmatrix}\begin{pmatrix} \overrightarrow{P_R\mathcal{V}} \\ \overrightarrow{P_RV} \end{pmatrix} + h.c., \tag{3.15}$$

where, for example, $\overrightarrow{P_RV}$ denotes a 3-component vector composed of the elements P_RV_m for $m = 1, 2, 3$, and *h.c.* refers to the Hermitian conjugate. Diagonalising this mass matrix we find that there are three mass eigenstates whose masses are eigenvalues of the matrix M and three whose masses are eigenvalues of the matrix $\mu M^{-1}\mu^T$. Thus, as the magnitude of the elements of M are increased, three of the states masses also increase, whilst the remaining three become inversely smaller. This is known as the *seesaw* mechanism, and is particularly relevant if we imagine there to be as yet unseen particles at mass scales much greater than that of the electro-weak sector. It would appear to offer a more natural explanation of the small neutrino masses observed than requiring the elements k_{mn} to be very small. Furthermore, if new particles and their corresponding physics are present beyond the energy scales of the Standard Model, then we may well find that Lepton number conservation does break down at these scales, in which case the non-zero M_{mn} becomes acceptable.

3.2.2 Relinquishing renormalisability

As discussed in section 2.2.3, enforcing renormalisabilty amounts to the statement that we consider our theory to be correct for the energy scales in question, and we therefore restrict ourselves to interaction terms with $[\psi_1, ..., \psi_m] \leq 4$. This is equivalent to restricting the dimension of the interaction term coupling constants, let us denote this by $[\lambda]$, to $[\lambda] \geq 0$. In allowing non-renormalisable terms with $[\psi_1, ..., \psi_m] > 4$, their coupling constants correspondingly satisfy $[\lambda] = -d$, where $d > 0$. If we take the energy scale at which our theory breaks down to be Λ, then we can take $\lambda \propto \Lambda^{-d}$ and, as such, the non-renormalisable interaction terms are suppressed by a factor of Λ^{-d}. The greater the degree of non-renormalisability, the more the corresponding interaction term will be suppressed. This is an indication of the limited effect we expect the high-energy physics to have on our lower energy calculations. If Λ is sufficiently greater than the energy scales in which we are interested, then we might justifiably restrict ourselves to considering 'first-order' non-renormalisable terms with $d = 1$, or equivalently $[\psi_1, ..., \psi_m] = 5$.

In the Standard Model we have the set of fermion fields $\mathbf{L}_m, E_m, \mathbf{Q}_m, U_m$ and D_m. Let us label these fermion fields generically with F_i, where $i = 1, ..., 15$. We also have the gauge bosons G^α_μ, W^a_μ and B_μ. Let us label these generically as X_j with $j = 1, ..., 12$. Finally we have the Higgs field ϕ. Also note that the covariant derivative D_μ has $[D_\mu] = 1$. The kinetic terms for the fermion fields are all of the form $\overline{F}_i \not{\partial} F_i$. From the requirement that the action be dimensionless and in the case of working in 4 dimensions, we can therefore deduce that $[F_i] = \frac{3}{2}$. Similarly, from the Yang-Mills field strength terms we can deduce that $[X_j] = 1$, and from the Higgs field kinetic term we also have $[\phi] = 1$. In trying to find an interaction term with dimension 5, we are therefore limited in the possible combinations of the various fields and covariant derivatives. Requiring the dimension 5 term to be Lorentz and gauge invariant further restricts the possible combinations. It turns out, [4], that the only permitted dimension 5 term is

$$-q_{mn}\tilde{\phi}_a(\overline{P_L\mathbf{L}}^a_m P_R\mathbf{L}^b_n)\tilde{\phi}_b - q^\dagger_{nm}\tilde{\phi}^{*b}(\overline{P_R\mathbf{L}}_{nb}P_L\mathbf{L}_{ma})\tilde{\phi}^{*b}, \tag{3.16}$$

where we have explicitly included the $SU(2)$ indices and $q_{mn} \propto \Lambda^{-1}$. Because of the success of the Standard Model at describing most of particle physics at the energy scales we have probed, we expect Λ, i.e. the new physics, to be at a scale much greater than that of the electro-weak sector. As such, we can assume that it is a good approximation to restrict ourselves to considering only the dimension-5 interaction term and to neglect higher dimension interaction terms which will be suppressed by factors of Λ^{-2} and greater.

After symmetry breaking, and in the unitary gauge, the term (3.16) becomes

$$-\frac{1}{2}(v+H)^2(q_{mn}\overline{P_L\mathcal{V}}_m P_R\mathcal{V}_n + q^\dagger_{nm}\overline{P_R\mathcal{V}}_n P_L\mathcal{V}_m), \tag{3.17}$$

which we can recognise as giving the neutrino field a Majorana mass term and additional neutrino-Higgs interaction terms. As such, the only effect the dimension-5 term produces on the energy scales that we have been able to probe thus far is the introduction of a neutrino mass of order $v^2 q_{mn}$. Given the Λ^{-1} dependence of q_{mn}, the size of this neutrino mass naturally comes out small, which is perhaps more desirable than the scenario considered earlier where we were forced to choose tiny k_{mn} elements.

The high-energy physics

Having established that the existence of new physics at energy scales far greater than that of the Standard Model provides an appealing explanation for the observed neutrino oscillations and associated tiny neutrino masses, this naturally begs the question - what is the new physics? Whilst there are many proposed ideas as to what lies beyond the Standard Model, here we return to one possible extension as was discussed in section 3.2.1. If we introduce heavy, sterile, right-handed neutrino fields labelled V_m, then our Standard Model Lagrangian can be supplemented by

$$\mathcal{L}_V = -\frac{1}{2}\overline{V}_m \partial\!\!\!/ V_m - \frac{1}{2}M_{mn}\overline{V}_m V_n - (k_{mn}\overline{P_L \mathbf{L}}_m P_R V_n \tilde{\phi} + k^{\dagger}_{nm}\tilde{\phi}^{\dagger}\overline{P_R V}_n P_L \mathbf{L}_m). \tag{3.18}$$

As we saw in section 3.2.1, taking M_{mn} to be large would generate the observed small neutrino masses via the seesaw mechanism. Because the sterile neutrinos are taken to be very massive, we would not expect to see them at the energy scales we are capable of probing, but we might expect to see indications of their existence through non-renormalisable terms in our Lagrangian. Following the argument set out in [4], we are able to see that a possible source of the dimension-5 non-renormalisable term mentioned above is the interaction shown in figure 3.1, where a virtual sterile neutrino is exchanged between two Higgs fields and two leptonic fields. The matrix

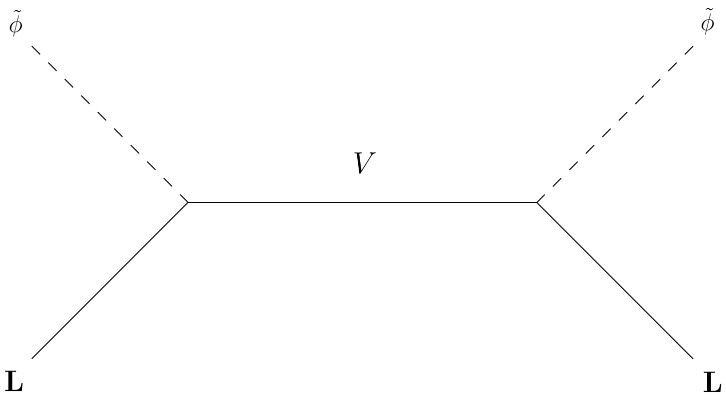

Figure 3.1: Feynman diagram for an interaction between two Higgs fields $\tilde{\phi}$ and two lepton fields **L** via the exchange of a virtual sterile neutrino.

element for this interaction is given by

$$k_{mo}k^T_{pn}\tilde{\phi}_a \left(\overline{P_L \mathbf{L}}^a_m \left[\frac{M - i\not{p}}{M^2 + p^2} \right]_{op} P_R \mathbf{L}^b_n \right) \tilde{\phi}_b + h.c., \tag{3.19}$$

where $h.c.$ refers to the Hermitian conjugate. In the case where we are taking M to be large, so that $M \gg p$, this amplitude approximates to the form

$$\begin{aligned} & k_{mo}k^T_{pn}\tilde{\phi}_a \left(\overline{P_L \mathbf{L}}^a_m M^{-1}_{op} P_R \mathbf{L}^b_n \right) \tilde{\phi}_b + h.c. \qquad (3.20) \\ = & \, (kM^{-1}k^T)_{mn}\tilde{\phi}_a \left(\overline{P_L \mathbf{L}}^a_m P_R \mathbf{L}^b_n \right) \tilde{\phi}_b + (kM^{-1}k^T)_{nm}\tilde{\phi}^{*b} \left(\overline{P_R \mathbf{L}}_{nb} P_L \mathbf{L}_{ma} \right) \tilde{\phi}^{*a}. \end{aligned}$$

This is exactly the dimension-5 interaction term discussed above with $q_{mn} = (kM^{-1}k^T)_{mn}$, and after symmetry breaking we will get mass terms with masses of order $v^2(kM^{-1}k^T)_{mn}$. We can see that the more massive we take the scale of the new physics to be, i.e. the larger we take $M(\leftrightarrow \Lambda)$ to be, the lighter our standard neutrinos will be. This is an example of the seesaw mechanism embedded within the particular scenario of beyond-the-

Standard-Model physics where very massive neutrino states are postulated to exist.

In summary, we have seen that the observed oscillation of neutrinos can be explained by allowing neutrinos to have a small but non-zero mass. As it stands, these non-zero masses cannot be generated within the Standard Model, but there are numerous ways in which the model can be extended to do so. In doing so, however, we are forced to either alter the particle content of the Standard Model or to concede that our theory is not a complete one if we consider energy scales beyond that of the electro-weak sector.

The lack of evidence for the additional light particles required in the first type of extension would perhaps point us towards preferring the second, and at the very least it places limits on the mixing that can take place between these new particles and those already known to exist. Furthermore, the first type of extension requires certain constants within the model to be tuned very close to zero, which, although not forbidden, does not seem very desirable.

In the second type of extension, the suggestion that there exists new physics at energy scales beyond those already probed would seem perfectly reasonable. Due to the very nature of this high energy physics, however, we can only hope to observe clues of its existence through the suppressed, non-renormalisable interaction terms in our Lagrangian. As such, it is not possible to determine the exact nature of the new physics, and at present there are numerous propositions that could account for the oscillation of neutrinos. One important consequence of the particular extension we considered is the appearance of a term which breaks Lepton number conservation. Observation of this non-conservation could therefore play a crucial role in elucidating the nature of the mechanism whereby neutrino masses are generated.

One particular process in which we might expect to see this breaking of Lepton number conservation is double Beta decay. This is the process whereby certain nuclei stable against single Beta decay can decay via the very rare process of *two* neutrons simultaneously decaying to protons. Within the Standard Model this takes place via two decays of the form $n \to p + e^- + \overline{\nu}_e$. If Lepton number conservation is violated, however, we can also have $n + n \to p + p + e^- + e^-$. The two different processes can be distinguished by the energy spectra of the resulting electrons, and as such it is possible to place limits on the size of coupling constants for Lepton non-conserving interaction terms in our Lagrangian.

4 Non-Hermitian Quantum Mechanics

In Quantum Mechanics, one of the fundamental requirements is that the Hamiltonian be Hermitian. Imposing $H^\dagger = H$ ensures that the eigenvalue spectrum is real and that the time evolution operator $U(t) = \exp(-iHt)$ is unitary, which in turn ensures that quantities such as $\langle \psi \mid A \mid \phi \rangle$ are time independent, where A is an operator corresponding to some time-independent observable and therefore commutes with H. Whilst the consequences of taking $H^\dagger = H$ are desirable, the condition itself is purely mathematical, with no physical motivation. As such, one might consider whether this requirement is too restrictive. Perhaps there is a wider spectrum of Hamiltonians that would still give us the desired real eigenvalues and unitary time evolution. Furthermore, perhaps this more general set of Hamiltonians could be dictated by some underlying physical reasoning.

The mathematical field of *pseudo-Hermiticity* is concerned exactly with the relaxing of this requirement for Hermiticity. A linear operator A is said to be pseudo-Hermitian if there exists an Hermitian operator η such that $A^\dagger = \eta A \eta^{-1}$, and it has been shown that every Hamiltonian with a real spectrum is pseudo-Hermitian, [13]. Note that in the case of Hermitian operators, $\eta = \mathbb{1}$.

Hamiltonians displaying symmetry under the combined action of parity, $\mathcal{P}$, and time-reversal, $\mathcal{T}$, form a subset of these pseudo-Hermitian Hamiltonians, and have received particular attention over the past decade. It is this subset that we will now discuss.

4.1 $\mathcal{PT}$ symmetric Hamiltonians

Using the notation $H^{\mathcal{PT}} = \mathcal{PT}H(\mathcal{PT})^{-1}$, the requirement for our Hamiltonian to display $\mathcal{PT}$ symmetry gives us

$$H^{\mathcal{PT}} = H \implies \mathcal{PT}H(\mathcal{PT})^{-1} = H \implies [H, \mathcal{PT}] = 0. \tag{4.1}$$

Under a certain set of conditions, such a Hamiltonian can be used to construct a $\mathcal{PT}$-formulation of quantum mechanics. Let us first remind ourselves of what we have in the standard formulation of quantum mechanics where $H^\dagger = H$:

(i) The set of states $|\psi_i\rangle$ satisfying $H|\psi_i\rangle = E_i|\psi_i\rangle$ form a basis for the Hilbert space of the system, where E_i are the positive, real energy eigenvalues of H and $|\psi_i\rangle$ the corresponding eigenstates[1].

(ii) With respect to the standard inner product on the Hilbert space, defined as $(\psi, \phi) = \langle\psi|\phi\rangle$, where $\langle\psi| = |\psi\rangle^\dagger$, these states are orthogonal.

(iii) Also with respect to the standard inner product, the states have positive-definite norm. For suitably normalised states, these last two requirements can be summarised as orthonormality, requiring $\langle\psi_i|\psi_j\rangle = \delta_{ij}$.

(iv) The states $|\psi_i\rangle$ obey unitary time-evolution, which by the definition of

[1] We do not consider degenerate eigenstates here.

the standard inner product ensures the time-independence of the inner product and of the expectation values of time-independent observables.

In our $\mathcal{PT}$-formulation, we would therefore like to reproduce this set of properties. In trying to do so, there are two scenarios that we must consider: that where $\mathcal{T}^2 = 1$, which is appropriate to bosons, and that where $\mathcal{T}^2 = -1$, which is appropriate to fermions. To see that these two possible scenarios exist we must consider the nature of the time-reversal operator $\mathcal{T}$. If we wish for energies to remain positive under time-reversal then we require that it be an anti-linear operator, [14]. As such, $\mathcal{T}$ acts on a state $|\psi\rangle$ as $\mathcal{T}|\psi\rangle = Z|\psi\rangle^*$, where Z is a linear operator. Requiring that $\mathcal{T}^2$ leave a state unchanged up to a phase factor we have

$$\mathcal{T}^2|\psi\rangle = ZZ^*|\psi\rangle = \lambda|\psi\rangle \implies ZZ^* = \lambda \implies Z^* = \lambda Z^{-1}, \tag{4.2}$$

where $|\lambda|^2 = 1$. Taking the complex conjugate of this we get $Z = \lambda^* Z^{-1*}$, which, acting from the right with Z^* gives us $ZZ^* = \lambda^*$. This in turn gives us the result $(ZZ^*)^2 = |\lambda|^2 = 1$, but we also have $(ZZ^*)^2 = \lambda^2$, from which we deduce $\lambda = \pm 1$. Let us consider each of these in turn.

4.1.1 Even $\mathcal{T}$ symmetry: $\mathcal{T}^2 = 1$

In this scenario, let us assume that the eigenstates of a $\mathcal{PT}$-symmetric Hamiltonian are also eigenstates of the operator $\mathcal{PT}$. Note that if $\mathcal{PT}$ were a linear operator then this would follow directly from the statement $[H, \mathcal{PT}] = 0$, but with $\mathcal{PT}$ being anti-linear, this is no longer the case. Taking $\mathcal{PT}|\psi\rangle = \lambda|\psi\rangle$ we act from the left with $\mathcal{PT}$ and use the properties

$[\mathcal{P}, \mathcal{T}] = 0$, $\mathcal{P}^2 = 1$, $\mathcal{T}^2 = 1$ and thus $(\mathcal{PT})^2 = 1$, which gives us

$$(\mathcal{PT})^2|\psi\rangle = |\psi\rangle = \mathcal{PT}\lambda|\psi\rangle = \mathcal{PT}\lambda(\mathcal{PT})^2|\psi\rangle = \lambda^*\lambda|\psi\rangle \tag{4.3}$$

$$\implies |\lambda|^2 = 1. \tag{4.4}$$

We now take the expression $H|\psi\rangle = E|\psi\rangle$, act from the left with $\mathcal{PT}$, use the property $[H, \mathcal{PT}] = 0$ and insert the identity to give

$$\mathcal{PT}H|\psi\rangle = \mathcal{PT}E|\psi\rangle = \mathcal{PT}E(\mathcal{PT})^2|\psi\rangle = E^*\lambda|\psi\rangle \tag{4.5}$$

$$= H\mathcal{PT}|\psi\rangle = H\lambda|\psi\rangle = E\lambda|\psi\rangle \tag{4.6}$$

$$\implies E = E^*. \tag{4.7}$$

We have therefore shown that if an energy eigenstate $|\psi\rangle$ of H is simultaneously an eigenstate of $\mathcal{PT}$ then the energy eigenvalues are real. The regime in which this requirement holds is referred to as that of 'unbroken' $\mathcal{PT}$-symmetry, and within this regime we see that we satisfy the first requirement for our $\mathcal{PT}$-formulation of quantum mechanics. Proving that one is in the regime of 'unbroken' $\mathcal{PT}$-symmetry, however, can be very difficult.

The remaining properties required of the states, namely that they be orthonormal and evolve in time in such a way as to preserve the inner product are, however, not satisfied if we consider these properties with respect to the standard inner product. As such, we are led to consider an alternative inner product on our Hilbert space.

Alternative inner products

An inner product is a way to combine two vectors in a vector space to give a complex number. Taking a vector space V and vectors $v_i \in V$, the inner product between two vectors, denoted (v_i, v_j), has the following properties:

- $(v_i, v_j) = (v_j, v_i)^*$
- $(v_i, v_j + v_k) = (v_i, v_j) + (v_i, v_k)$ and $(v_i + v_j, v_k) = (v_i, v_k) + (v_j, v_k)$
- $(v_i, cv_j) = c(v_i, v_j)$ and $(cv_i, v_j) = c^*(v_i, v_j)$ for $c \in \mathbb{C}$

We can further require that the inner product be positive-definite, that is $(v_i, v_i) \geq 0$ and $(v_i, v_i) = 0$ only for $v_i = 0$.

If we consider a set of basis states for our Hilbert space, labelled e_i, then we are able to express the inner product of any two vectors in our vector space in terms of those between these basis states. For

$$v_1 = \sum_i v_1^i e_i \text{ and } v_2 = \sum_j v_2^j e_j, \tag{4.8}$$

and using the properties of the inner product, (v_1, v_2) is given by

$$(v_1, v_2) = (\sum_i v_1^i e_i, \sum_j v_2^j e_j) \tag{4.9}$$

$$= \sum_{i,j} (v_1^i e_i, v_2^j e_j) \tag{4.10}$$

$$= \sum_{i,j} v_2^j (v_1^i e_i, e_j) = \sum_{i,j} v_1^{i*} v_2^j (e_i, e_j). \tag{4.11}$$

The inner product (e_i, e_j) is called the kernel of the inner product. In the case of an orthonormal basis, the kernel is simply δ_{ij}, which leaves us with the familiar expression

$$(v_1, v_2) = \sum_i v_1^{i*}(1) v_2^i = v_1^\dagger v_2. \tag{4.12}$$

In Hermitian quantum mechanics we have $e_i = \psi_i$ and $(\psi_i, \psi_j) = \langle \psi_i | \psi_j \rangle =$

δ_{ij}, which means that we recover the familiar expression for the standard inner product (4.12). Furthermore, as $|\psi_i\rangle$ evolves in time as $|\psi_i(t)\rangle = \exp(-iHt)|\psi_i(0)\rangle$, $\langle\psi_i| = |\psi_i\rangle^\dagger$ and $H = H^\dagger$, we find that the inner product is time independent:

$$\langle\psi_i(t)|\psi_j(t)\rangle = \langle\psi_i(0)|e^{iH^\dagger t}e^{-iHt}|\psi_j(0)\rangle \quad (4.13)$$
$$= \langle\psi_i(0)|e^{iHt}e^{-iHt}|\psi_j(0)\rangle \quad (4.14)$$
$$= \langle\psi_i(0)|\psi_j(0)\rangle. \quad (4.15)$$

In $\mathcal{PT}$ quantum mechanics, given that $H^\dagger \neq H$, we lose this time-independence of the inner product, as well as the orthogonality of the different eigenstates of H, [15]. We must therefore look to define our inner product differently. In analogy with the Hermitian case, a sensible guess might be to define it as

$$(\psi_i, \psi_j)_{\mathcal{PT}} = (\mathcal{PT}|\psi_i\rangle)^T|\psi_j\rangle. \quad (4.16)$$

It turns out, [3], that such a definition can only get us so far. The orthogonality condition is satisfied and, due to the fact that $\mathcal{PT}$ commutes with H, this inner product is also time-independent[2]. However, the norm of a state with respect to this inner product is neither positive-definite nor correctly normalised, and can even be zero!

We firstly turn to the issue of normalisation. In (4.4) we showed that in the 'unbroken' regime the eigenstates satisfy $\mathcal{PT}|\psi_i\rangle = \lambda_i|\psi_i\rangle$, where λ_i is a pure phase. Thus, if we define a new state $|\psi_i\rangle' = \lambda_i^{+\frac{1}{2}}|\psi_i\rangle$ we have the

[2]We are making the simplifying assumption here that H is symmetric. However, this is not a necessary condition. See [3] and references therein.

result

$$\mathcal{PT}|\psi_i\rangle' = \mathcal{PT}\lambda_i^{+\frac{1}{2}}(\mathcal{PT})^2|\psi_i\rangle \tag{4.17}$$

$$= \lambda_i^{-\frac{1}{2}}\lambda_i|\psi_i\rangle = \lambda_i^{+\frac{1}{2}}|\psi_i\rangle \tag{4.18}$$

$$= |\psi_i\rangle', \tag{4.19}$$

which gives us the correct normalisation. Note that $|\psi_i\rangle'$ is still an eigenstate of H, as required.

In order to tackle the issue of non-positive-definite norm we introduce a new linear operator $\mathcal{C}$, which acts on the states as follows[3]:

$$\mathcal{C}|\psi_i\rangle = s_i|\psi_i\rangle, \tag{4.20}$$

where s_i is the sign of $(\psi_i, \psi_i)_{\mathcal{PT}}$ and we have dropped the prime so that $|\psi_i\rangle$ now corresponds to the correctly normalised state. From this definition we see that $\mathcal{C}^2 = 1$. $\mathcal{C}$ also satisfies the conditions

$$[\mathcal{C}, H] = 0 \quad \text{and} \quad [\mathcal{C}, \mathcal{PT}] = 0, \tag{4.21}$$

which in turn gives us the result

$$[\mathcal{CPT}, H] = 0. \tag{4.22}$$

With this new linear operator we are now able to define the $\mathcal{CPT}$ inner product

$$(\psi_i, \psi_j)_{\mathcal{CPT}} = (\mathcal{CPT}|\psi_i\rangle)^T|\psi_j\rangle \tag{4.23}$$

that, in addition to satisfying the conditions already satisfied by the $\mathcal{PT}$

[3]Not to be confused with the charge conjugation operator.

inner product (4.16), also satisfies the condition of a positive-definite norm.

Given that $\mathcal{C}$ is defined explicitly in terms of the eigenstates of the Hamiltonian, (4.20), it will in turn depend on H itself. We therefore say that the $\mathcal{CPT}$ inner product is dynamically determined. In order to construct $\mathcal{C}$ and subsequently $(\psi_i, \psi_j)_{\mathcal{CPT}}$ we first require knowledge of all the eigenstates of H and the signs of their norms under the $\mathcal{PT}$ inner product. As such, this task is in general a very difficult one and often can only be carried out perturbatively.

Pseudo-Hermiticity and equivalent Hermitian Hamiltonians

Perhaps a more elegant way in which to recast some of the preceding discussion is in the context of pseudo-Hermiticity. Following the arguments of [16], the statement of pseudo-Hermiticity mentioned earlier, whereby H is pseudo-Hermitian if $H^\dagger = \eta H \eta^{-1}$, with η being some linear, Hermitian, invertible operator on the vector space spanned by the eigenstates of H, is equivalent to the requirement that H be Hermitian with respect to the inner product $\langle\langle\psi, \phi\rangle\rangle$ defined as $\langle\langle\psi, \phi\rangle\rangle = (\psi, \eta\phi) = \langle\psi|\eta|\phi\rangle$. That is to say $\langle\langle H\psi, \phi\rangle\rangle = \langle\langle\psi, H\phi\rangle\rangle$. This equivalence can be shown as follows:

$$\langle\langle H\psi, \phi\rangle\rangle = \langle\langle\psi, H\phi\rangle\rangle \tag{4.24}$$

$$\implies \langle\psi|H^\dagger\eta|\phi\rangle = \langle\psi|\eta H|\phi\rangle \tag{4.25}$$

$$\implies H^\dagger\eta = \eta H \tag{4.26}$$

$$\implies H^\dagger = \eta H \eta^{-1}. \tag{4.27}$$

It is shown in [17] that if we take the case where H is symmetric under an anti-linear symmetry such as $\mathcal{PT}$ and we restrict ourselves to the regime of unbroken $\mathcal{PT}$ symmetry, which ensures that the eigenvalue spectrum of H

is real, then we can always construct an inner product of the form $\langle\langle\psi, \phi\rangle\rangle$ that is positive-definite. As such, we can then take η to be a positive-definite operator, which in turn allows us to express η in terms of its positive square root, $\eta = \rho^2$. ρ will be linear, Hermitian[4] and invertible.

If we now define a new Hamiltonian h as

$$h = \rho H \rho^{-1}, \tag{4.28}$$

then we can see that it is Hermitian with respect to the standard inner product as follows:

$$h^\dagger = \rho^{-1\dagger} H^\dagger \rho^\dagger = \rho^{-1} \eta H \eta^{-1} \rho = \rho^{-1} \rho^2 H \rho^{-2} \rho = \rho H \rho^{-1} = h. \tag{4.29}$$

The Hermitian Hamiltonian h is equivalent to H in that it has the same eigenvalue spectrum. The eigenstates of h are given by $\rho|\psi_i\rangle$, where $|\psi_i\rangle$ are the eigenstates of H, so, if we consider the inner product of two such eigenstates of h, then we get $\langle\psi_i|\rho^\dagger\rho|\psi_j\rangle = \langle\psi_i|\eta|\psi_j\rangle$, which we know to be the correct inner product for this system. Thus, although the eigenvalue spectra of h and H are identical, relations between their eigenvectors will differ. This could potentially be very useful, as it suggests that by moving from one basis to the other, one might be able to take two states that are initially very 'close' and difficult to distinguish between and make them much easier to distinguish.

Another nice property of the above formulation of our positive-definite inner product is that its time-independence is clear to see[5]:

---

[4]Note that η and ρ are Hermitian with respect to the standard inner product rather than with respect to $\langle\langle\ ,\ \rangle\rangle$.

[5]Here we use the result $M^{-1} e^A M = e^{M^{-1} A M}$, where A and M are linear operators.

$$\langle\psi_i(t)|\eta|\psi_j(t)\rangle = \langle\psi_i(0)|e^{iH^\dagger t}\eta e^{-iHt}|\psi_j(0)\rangle \tag{4.30}$$

$$\text{inserting the identity} \quad = \langle\psi_i(0)|\eta\eta^{-1}e^{iH^\dagger t}\eta e^{-iHt}|\psi_j(0)\rangle \tag{4.31}$$

$$= \langle\psi_i(0)|\eta e^{i\eta^{-1}H^\dagger\eta t}e^{-iHt}|\psi_j(0)\rangle \tag{4.32}$$

$$= \langle\psi_i(0)|\eta e^{iHt}e^{-iHt}|\psi_j(0)\rangle \tag{4.33}$$

$$= \langle\psi_i(0)|\eta|\psi_j(0)\rangle. \tag{4.34}$$

In our specific case, we have shown that the $\mathcal{CPT}$ inner product is positive-definite and, as such, we find that $\eta^{-1} = \mathcal{CP}$, [16][18]. This makes it clear that, as with the construction of $\mathcal{C}$, the construction of η and subsequently ρ is in general not an easy task, as it requires knowledge of all the eigenstates of H and the sign of their norm with respect to the $\mathcal{PT}$ inner product.

One useful method for constructing $\mathcal{C}$ is to parametrise it as

$$\mathcal{C} = e^{\mathcal{Q}(\hat{x},\hat{p})}\mathcal{P}, \tag{4.35}$$

where $\mathcal{Q}$ is an Hermitian function of $\hat{x}$ and $\hat{p}$, [3]. We then try to solve for $\mathcal{Q}$ by imposing the conditions $[\mathcal{C},\mathcal{PT}] = 0$, $[\mathcal{C},H] = 0$ and $\mathcal{C}^2 = 1$. Given this parametrisation we see that $\eta^{-1} = e^{\mathcal{Q}}$ and so $\rho = e^{-\frac{\mathcal{Q}}{2}}$, which tells us that the equivalent Hermitian formulation of the system described by H will in general be non-local, as h is potentially a function of arbitrarily high powers of $\hat{p}$.

Taking this form of ρ we can now show explicitly for the case of a $\mathcal{PT}$

symmetric Hamiltonian that h as defined in (4.28) is indeed Hermitian, [3]:

$$h^\dagger = e^{\frac{\mathcal{Q}}{2}} H^\dagger e^{-\frac{\mathcal{Q}}{2}} \tag{4.36}$$

$$= e^{\frac{-\mathcal{Q}}{2}} e^{\mathcal{Q}} H^\dagger e^{-\mathcal{Q}} e^{\frac{\mathcal{Q}}{2}}. \tag{4.37}$$

Now note that $e^{\mathcal{Q}} = \eta^{-1} = \mathcal{CP}$ and $e^{-\mathcal{Q}} = \eta = \mathcal{PC}$, so we have

$$h^\dagger = e^{\frac{-\mathcal{Q}}{2}} \mathcal{CP} H^\dagger \mathcal{PC} e^{\frac{\mathcal{Q}}{2}}, \tag{4.38}$$

but as $H^\dagger = \eta H \eta^{-1} = \mathcal{PC}H\mathcal{CP}$, $[\mathcal{C}, H] = 0$ and $\mathcal{C}^2 = 1$, we have $H^\dagger = \mathcal{P}H\mathcal{P}$, leaving us with

$$h^\dagger = e^{\frac{-\mathcal{Q}}{2}} H e^{\frac{\mathcal{Q}}{2}} = h. \tag{4.39}$$

One final comment we make is relating to observables within this $\mathcal{PT}$ formulation of quantum mechanics. In the normal Hermitian formulation, an operator A must satisfy $A^\dagger = A$ in order to be associated with some observable, as this ensures its eigenvalues are real and the corresponding eigenvectors are orthonormal. This comes from the requirement that A be self-dual with respect to the standard inner product, i.e. $(A\psi, \phi) = (\psi, A\phi)$. Thus, the natural extension into non-Hermitian quantum mechanics is to require that the operator for an observable be self-dual with respect to the relevant inner product, i.e. $\langle\langle A\psi, \phi\rangle\rangle = \langle\langle \psi, A\phi\rangle\rangle$. In the case of a $\mathcal{PT}$-symmetric Hamiltonian we therefore require $(A\psi, \phi)_{\mathcal{CPT}} = (\psi, A\phi)_{\mathcal{CPT}}$. To see how this gives us our real eigenvalue spectrum and orthonormality of eigenvectors, we use the standard argument that exploits the properties of an inner product:

$$\langle\langle A\psi_i, \psi_j\rangle\rangle = \langle\langle \psi_i, A\psi_j\rangle\rangle \tag{4.40}$$

$$\implies \langle\langle \lambda_i\psi_i, \psi_j\rangle\rangle = \langle\langle \psi_i, \lambda_j\psi_j\rangle\rangle \tag{4.41}$$

$$\implies \lambda_i^*\langle\langle \psi_i, \psi_j\rangle\rangle = \lambda_j\langle\langle \psi_i, \psi_j\rangle\rangle \tag{4.42}$$

$$\implies (\lambda_i^* - \lambda_j)\langle\langle \psi_i, \psi_j\rangle\rangle = 0, \tag{4.43}$$

where ψ_i are eigenvectors of A with eigenvalue λ_i. Thus, considering the case $i = j$ and exploiting the fact that $\langle\langle\ ,\ \rangle\rangle$ is defined so as to give a non-zero, positive definite norm, we establish that the eigenvalues are real. And in the case $i \neq j$, assuming no degeneracy, we therefore establish $\langle\langle \psi_i, \psi_j\rangle\rangle = 0$.

4.1.2 Odd $\mathcal{T}$ symmetry: $\mathcal{T}^2 = -1$

The generalisation of the $\mathcal{PT}$ formulation of quantum mechanics to the case of $\mathcal{T}^2 = -1$ follows the same principles as outlined in the previous section, namely we look for a new inner product that gives us our desired properties for the eigenvectors and eigenvalues of H. It is made more tricky, however, by the fact that we are unable to find states which are invariant under the action of $\mathcal{PT}$. The details of how to construct the relevant inner product are given in [19] and [2], and we will not include them here. The relevant inner product is found to be

$$(\psi, \phi)_{\mathcal{CPT}} = (\mathcal{CPT}|\psi\rangle)^T Z|\phi\rangle, \tag{4.44}$$

where Z corresponds to the linear operator used to define the anti-linear operator $\mathcal{T}$ as $\mathcal{T}|\psi\rangle = Z|\psi\rangle^*$.

Armed with these new inner products, we are thus able to construct a formulation of quantum mechanics that allows us to replace the requirement for the Hamiltonian to be Hermitian with the requirement that it be $\mathcal{PT}$-symmetric. This condition of $\mathcal{PT}$-symmetry is perhaps more appealing than that of Hermiticity in the sense that, rather than being purely mathematical, it does have some physical motivation.

4.2 $\mathcal{PT}$-symmetric Dirac equation

Having introduced the principles behind a formulation of quantum mechanics that imposes $\mathcal{PT}$ symmetry on the Hamiltonian rather than Hermiticity, we now look at what consequences this has when applied to the Dirac equation. We follow the analysis set out in [1] and [2], where a new solution to the Dirac equation is discovered that would appear to describe two flavours of massless particle, despite having a non-zero mass matrix. This leads us to the possibility of having mass and flavour eigenstates that do not coincide, and in the case of massive neutrinos we saw that this in turn led to the possibility of flavour oscillations. We finish by attempting to extend the model to describe three flavours of massless particle and questioning how certain we can be that the findings of Jones-Smith *et al.* do indeed point to a new type of solution.

4.2.1 The Dirac Equation

Before the Dirac equation, negative energies and negative probability densities were two key problems associated with taking the wavefunction of a particle to satisfy the relativistic Klein-Gordon equation, [20]. It was observed that the second of these issues could be alleviated by replacing the Klein-Gordon equation with one that only contained first-order time derivatives. As such, desiring a relativistically covariant equation that still reproduced the correct relativistic energy-momentum relation for a massive particle, Dirac proposed[6]

$$E\psi = i\frac{\partial \psi}{\partial t} = -i\boldsymbol{\alpha}\cdot\nabla\psi + m\beta\psi = H\psi \tag{4.45}$$

and required that the α_i, with $i = 1, 2, 3$, and β satisfy the relations

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij}, \quad \{\alpha_i, \beta\} = 0 \quad \text{and} \quad \beta^2 = 1, \tag{4.46}$$

known as the Dirac algebra, [21]. In order to ensure that the energy eigenvalues of (4.45) be real, the α_i and β were all taken to be Hermitian. The requirements (4.46) are such that on squaring (4.45) we reproduce the relativistic energy-momentum relation , and we can check that this is indeed the case:

$$E^2\psi = (-i\boldsymbol{\alpha}\cdot\nabla + \beta m)^2\,\psi \tag{4.47}$$

$$= \left(-\alpha_i\alpha_j\nabla_i\nabla_j - im\alpha_i\beta\nabla_i - im\beta\alpha_i\nabla_i + m^2\beta^2\right)\psi \tag{4.48}$$

$$= \left(-\frac{1}{2}\left(\alpha_i\alpha_j\nabla_i\nabla_j + \alpha_j\alpha_i\nabla_j\nabla_i\right) - im\{\alpha_i, \beta\}\nabla_i + m^2\right)\psi \tag{4.49}$$

$$= \left(-\frac{1}{2}\{\alpha_i, \alpha_j\}\nabla_i\nabla_j + m^2\right)\psi \tag{4.50}$$

$$= \left(-\nabla_i\nabla_i + m^2\right)\psi \tag{4.51}$$

$$= \left(\mathbf{p}^2 + m^2\right)\psi \quad \text{as required,} \tag{4.52}$$

where we have used the fact that $\nabla_i\nabla_j = \nabla_j\nabla_i$ and $\mathbf{p} = -i\nabla$.

Evidently the α_i and β cannot be normal numbers if they are to satisfy the Dirac algebra, but we are able to find matrices that do so. From the

[6]Note that we are using natural units.

conditions (4.46), we can see that $\alpha_i^2 = 1$ and $\beta^2 = 1$, which tells us that the α_i and β have eigenvalues ± 1. We can further show that all the elements of the algebra are traceless. Taking $\alpha_i \beta + \beta \alpha_i = 0$ and using $\beta^2 = 1$ we deduce $\mathrm{Tr}(\alpha_i) = -\mathrm{Tr}(\beta \alpha_i \beta)$, and by the cyclicity of the trace and once again using $\beta^2 = 1$ we are left with $\mathrm{Tr}(\alpha_i) = -\mathrm{Tr}(\alpha_i)$, which can only hold if $\mathrm{Tr}(\alpha_i) = 0$. By a similar argument using the fact that $\alpha_i^2 = 1$ and the cyclicity of the trace, we also find $\mathrm{Tr}(\beta) = 0$. We now note that if we have a set of α_i and β that solve the algebra, then so too will the set $U \alpha_i U^{-1}$ and $U \beta U^{-1}$. In the case where we require α_i and β to be Hermitian, the U is a unitary transformation. By taking four different transformations that diagonalise the four elements α_i and β in turn, and using the fact that $\mathrm{Tr}(U \alpha_i U^{-1}) = \mathrm{Tr}(\alpha_i) = 0$, our knowledge that the diagonal elements are ± 1 leads us to the conclusion that the matrix forms of α_i and β must be of even dimension.

We start by considering a 2×2 representation. We know that the Pauli matrics

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{and} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.53}$$

satisfy the algebra $\{\sigma_i, \sigma_j\} = 2\delta_{ij}$, so we can make the association $\alpha_i \leftrightarrow \pm \sigma_i$. However, in trying to find a 2×2 matrix that anti-commutes with all the σ_i, it can be shown by considering some general 2×2 matrix that such a matrix does not exist. As such, we are unable to satisfy the algebra, except in the case where $m = 0$. It turns out, however, that the algebra can be satisfied by a 4×4 representation, and one particular form of this is

$$\alpha_i = \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}. \tag{4.54}$$

As such, the wavefunction ψ must be a four-component object, and Dirac showed that such a wavefunction could be interpreted as one describing a spin-half particle and its antiparticle, each having two spin states.

In this block-diagonal form, we see that the α_is are just the direct sum of the 2×2 representations $\pm\sigma_i$. We note here that $\pm\sigma_i$ are independent representations as if one could be mapped to the other by a similarity transformation, $U\sigma_i U^{-1} = -\sigma_i$, then this would imply $U\sigma_i + \sigma_i U = 0$, which we know to be impossible. If we therefore take our four-component wavefuntion to be of the form

$$\psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}, \tag{4.55}$$

where ψ_L and ψ_R both have two components, then the Dirac equation has two parts:

$$-i\sigma_i \nabla_i \psi_L + m\psi_R = i\frac{\partial \psi_L}{\partial t} \tag{4.56}$$

$$\text{and} \qquad i\sigma_i \nabla_i \psi_R + m\psi_L = i\frac{\partial \psi_R}{\partial t}. \tag{4.57}$$

So we see that we can consider our system in terms of two two-component objects coupled by a mass. As outlined in [2], a more general representation for α_i would be

$$\alpha_i = \begin{pmatrix} V\sigma_i V^{-1} & 0 \\ 0 & -W\sigma_i W^{-1} \end{pmatrix}, \tag{4.58}$$

where V and W represent similarity transformations. Any 4×4 representation of α_i and β can be shown to be unitarily equivalent to the α_i given in (4.58) and β given in (4.54).

Having found a 4×4 representation for the Dirac algebra, one might try to look for higher-dimensional ones. However, it can be shown, [2], that if one tries to find an 8×8 representation, it simply decouples into two independent 4×4 representations.

4.2.2 Useful properties of the Dirac algebra and Pauli matrices

Before moving on to the non-Hermitian case, we first point out some results that will be useful to us later on. Similar to the proof that there is no matrix that anti-commutes with all three Pauli matrices, by simply considering some general 2×2 matrix A, it can be shown that

(i) if $[A, \sigma_i] = 0$ for $i = 1, 2, 3$, then $A = a\sigma_0$, where $a \in \mathbb{C}$,

(ii) there is no A satisfying $A\sigma_i^* = \sigma_i A$ and

(iii) if $A\sigma_i^* = -\sigma_i A$ then $A = ai\sigma_2$, where $a \in \mathbb{C}$.

Another useful observation, as discovered by Dirac, is the relation between the α_i and the Lorentz group generators. Taking the generators of the boosts to be K_i, where $i = 1, 2, 3$ labels boosts in the x, y and z directions respectively, and the generators of rotations to be J_i, where $i = 1, 2, 3$ labels rotations about the x, y and z axes respectively, the Lie algebra of the Lorentz group can be summarised as

$$[J_i, J_j] = i\epsilon_{ijk}J_k, \quad [J_i, K_j] = i\epsilon_{ijk}K_k \quad \text{and} \quad [K_i, K_j] = -i\epsilon_{ijk}J_k. \quad (4.59)$$

If we make the associations $J_i \leftrightarrow -\frac{i}{2}\alpha_j\alpha_k$ and $K_i \leftrightarrow \frac{i}{2}\alpha_i$, where in the association for J_i the labels i, j and k are always an even permutation of 123, then, using the Dirac algebra, it can be shown that these objects do indeed satisfy the Lorentz algebra. The α_i can thus be used to form a representation of the Lorentz group, and this observation allows one to prove the Lorentz covariance of the Dirac equation, [2].

4.2.3 Swapping Hermiticity for $\mathcal{PT}$-symmetry

In moving to the case of a $\mathcal{PT}$-symmetric Dirac Hamiltonian, the method employed in [2] was to try and start with the most general possible forms for objects such as the α_i and β, and to let the specific forms be deduced by imposing the desired requirements such as that the Hamiltonian be $\mathcal{PT}$-symmetric.

A relativistic energy-momentum relation

Starting from the same proposed form for the Dirac Hamiltonian (4.45) and imposing that the correct relativistic energy-momentum relation be reproduced on squaring this, we recover the same requirements $\{\alpha_i, \alpha_j\} = 2\delta_{ij}$ and $\{\alpha_i, \beta\} = 0$. Hermiticity played no role in deriving these conditions originally, but rather was imposed independently, so we would expect to reproduce these same anti-commuation relations.

Parity, Time-reversal and Lorentz Transformations

If we take the specific forms for $\mathcal{P}$ and $\mathcal{T}$ acting on some state $\psi(\mathbf{r})$ as $\mathcal{P}\psi(\mathbf{r}) = S\psi(-\mathbf{r})$ and $\mathcal{T}\psi(\mathbf{r}) = Z\psi^*(\mathbf{r})$, and recall that $\mathcal{P}^2 = 1$, $\mathcal{T}^2 = -1$

and $[\mathcal{P},\mathcal{T}]=0$, then we have the following results

$$\mathcal{P}^2 = 1 \;\Rightarrow\; \mathcal{P}^{-1} = \mathcal{P}, S^2 = 1 \;\Rightarrow\; S^{-1} = S \tag{4.60}$$

$$\mathcal{T}^2 = -1 \;\Rightarrow\; ZZ^* = -1 \;\Rightarrow\; Z^* = -Z^{-1} \;\Rightarrow\; \mathcal{T}^{-1}\psi = -Z\psi^* \tag{4.61}$$

$$[\mathcal{P},\mathcal{T}] = 1 \;\Rightarrow\; SZ = ZS^*. \tag{4.62}$$

The operators for rotations about axis i and boosts along axis i are given by

$$R_i(\theta) = e^{-i\theta J_i} \quad \text{and} \quad B_i(\zeta) = e^{-i\zeta K_i} \tag{4.63}$$

respectively, where θ is the angle of rotation and ζ the rapidity of the boost. Under parity and time-reversal transformations these transform as

$$\mathcal{P}R_i(\theta)\mathcal{P}^{-1} = R_i(\theta), \tag{4.64}$$

$$\mathcal{P}B_i(\zeta)\mathcal{P}^{-1} = B_i(-\zeta), \tag{4.65}$$

$$\mathcal{T}R_i(\theta)\mathcal{T}^{-1} = R_i(\theta) \tag{4.66}$$

$$\text{and} \quad \mathcal{T}B_i(\zeta)\mathcal{T}^{-1} = B_i(-\zeta). \tag{4.67}$$

Thus, by acting on some state $\psi(\mathbf{r})$ with these operators and using the results (4.60)-(4.62), we deduce

$$Se^{-\frac{1}{2}\theta\alpha_j\alpha_k}S\psi(\mathbf{r}) = e^{-\frac{1}{2}\theta\alpha_j\alpha_k}\psi(\mathbf{r}) \longrightarrow S\alpha_j\alpha_k S = \alpha_j\alpha_k \tag{4.68}$$

$$Se^{\frac{1}{2}\zeta\alpha_i}S\psi(\mathbf{r}) = e^{-\frac{1}{2}\zeta\alpha_i}\psi(\mathbf{r}) \longrightarrow S\alpha_i S = -\alpha_i \tag{4.69}$$

$$Ze^{-\frac{1}{2}\theta\alpha_j^*\alpha_k^*}(-Z^*)\psi(\mathbf{r}) = e^{-\frac{1}{2}\theta\alpha_j\alpha_k}\psi(\mathbf{r}) \longrightarrow Z\alpha_j^*\alpha_k^* Z^* = -\alpha_j\alpha_k \tag{4.70}$$

$$Ze^{\frac{1}{2}\zeta\alpha_i^*}(-Z^*)\psi(\mathbf{r}) = e^{-\frac{1}{2}\zeta\alpha_i}\psi(\mathbf{r}) \longrightarrow Z\alpha_i^* Z^* = \alpha_i. \tag{4.71}$$

It can be seen that if the second and fourth of these conditions hold, then the first and third follow, so we can summarise the results as

$$\{S, \alpha_i\} = 0 \quad \text{and} \quad Z\alpha_i^* = -\alpha_i Z. \tag{4.72}$$

A $\mathcal{PT}$-symmetric Hamiltonian

The requirement for the Hamiltonian to be $\mathcal{PT}$-symmetric can be expressed as $[H, \mathcal{PT}] = 0$. Thus, by acting on some state $\psi(\mathbf{p}, \mathbf{r})$, which we take to be a plane-wave solution, and taking the matrix forms of $\mathcal{P}$ and $\mathcal{T}$, we have

$$\begin{aligned} \alpha_i \mathbf{p}_i SZ\psi^*(\mathbf{p}, -\mathbf{r}) + \beta SZ\psi^*(\mathbf{p}, -\mathbf{r}) &= SZ\left(\alpha_i \mathbf{p}_i \psi(\mathbf{p}, -\mathbf{r})\right)^* + SZ\beta^*\psi^*(\mathbf{p}, -\mathbf{r}) \\ \implies \quad \alpha_i SZ = SZ\alpha_i^* \ \text{ and } \ \beta SZ &= SZ\beta^*. \end{aligned} \tag{4.73}$$

Self-duality of the Hamiltonian

As discussed previously, for a Hamiltonian displaying invariance under some given anti-linear symmetry, and in the regime of that symmetry being unbroken, the appropriate inner product for formulating a quantum theory is one with respect to which the Hamiltonian is self-dual. In the case of a $\mathcal{PT}$-symmetric hamiltonian we require the $\mathcal{CPT}$ inner product, defined as

$$(\psi, \phi)_{\mathcal{CPT}} = (\mathcal{CPT}|\psi\rangle)^T(-Z^\dagger)|\phi\rangle, \tag{4.74}$$

and so self-duality of the Hamiltonian gives us

$$(H\psi, \phi)_{\mathcal{CPT}} = (\psi, H\phi)_{\mathcal{CPT}} \tag{4.75}$$

$$\implies \quad (\mathcal{CPT}H|\psi\rangle)^T Z^\dagger|\phi\rangle = (\mathcal{CPT}|\psi\rangle)^T Z^\dagger H|\phi\rangle \tag{4.76}$$

$$= \ (\mathcal{PT}H\mathcal{C}|\psi\rangle)^T Z^\dagger|\phi\rangle = (\mathcal{PTC}|\psi\rangle)^T Z^\dagger H|\phi\rangle, \tag{4.77}$$

where we have used the fact that $[\mathcal{C}, \mathcal{PT}] = 0$ and $[\mathcal{C}, H] = 0$. If we take the explicit matrix forms for our $\mathcal{C}$, $\mathcal{P}$ and $\mathcal{T}$ operators[7] then this gives us

$$(SZH^*C^*|\psi\rangle^*)^T Z^\dagger|\phi\rangle = (SZC^*|\psi\rangle^*)^T Z^\dagger H|\phi\rangle \tag{4.78}$$

$$\langle\psi|C^\dagger H^\dagger Z^T S^T Z^\dagger|\phi\rangle = \langle\psi|C^\dagger Z^T S^T Z^\dagger H|\phi\rangle. \tag{4.79}$$

Combining these results with (4.73) we find

$$\beta = -Z^T\beta^T Z^\dagger \quad \text{and} \quad \alpha_i = Z^T\alpha_i^T Z^\dagger. \tag{4.80}$$

With this set of conditions, it is now possible to deduce the forms of α_i and β required to solve the $\mathcal{PT}$-symmetric Dirac equation.

4.3 Solutions to the $\mathcal{PT}$-symmetric Dirac equation

4.3.1 Model 4

Using the terminology of [2], Model 4 is the four-dimensional solution to the $\mathcal{PT}$-symmetric Dirac equation. They start by noting that, as with the Hermitian Dirac equation, any 2×2 representation for the α_i must be of the form $V\sigma_i V^{-1}$ or $-W\sigma_i W^{-1}$. In the Hermitian case, the similarity transformations V and W were required to be unitary, in order that Hermiticity of the α_i be preserved. Here, however, we no longer need to make that restriction. Having generalised the two possible types of 2×2 representations, the

[7] We have not had to use an explicit form for $\mathcal{C}$ as yet; here we will simply use $\mathcal{C}|\psi\rangle = C|\psi\rangle$.

Model 4 representation is taken to be the direct sum of these:

$$\begin{pmatrix} V\sigma_i V^{-1} & 0 \\ 0 & -W\sigma_i W^{-1} \end{pmatrix}. \tag{4.81}$$

With this as a starting point, the procedure is then to apply the various conditions set out in the previous section and thereby determine the exact forms of β, S and Z. All the details of this procedure for Model 4 are given in [2], and we will not include them here. The resulting α_i and β are given by

$$\alpha_i = U \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} U^\dagger \quad \text{and} \quad \beta = U \begin{pmatrix} 0 & m\sigma_o \\ m\sigma_0 & 0 \end{pmatrix} U^\dagger, \tag{4.82}$$

where U is a unitary matrix. This result is somewhat surprising. Despite the fact that Hermiticity was not imposed at the outset, the various conditions imposed on the system in requiring it to be $\mathcal{PT}$-symmetric conspire to give us Hermitian α_i and β unitarily equivalent to the solutions (4.54) we found for the Hermitian Dirac equation. As such, it would appear that in terms of a theory for free fermions, we could equally well replace our assumption that the Dirac Hamiltonian be Hermitian with the one that is $\mathcal{PT}$-symmetric.

4.3.2 Model 8

After establishing the equivalence of the Model 4 solution with that of the Hermitian Dirac equation, Jones-Smith *et al.* then construct an eight-dimensional representation for α_i and β. The procedure is exactly as for Model 4 and so is not explicitly given in [2]. We therefore use this as an opportunity to give an example of the procedure.

In analogy with Model 4, the starting point is to take α_i to be the direct sum of four 2×2 representations

$$\alpha_i = \begin{bmatrix} V\sigma_i V^{-1} & 0 & 0 & 0 \\ 0 & V\sigma_i V^{-1} & 0 & 0 \\ 0 & 0 & -W\sigma_i W^{-1} & 0 \\ 0 & 0 & 0 & -W\sigma_i W^{-1} \end{bmatrix}. \tag{4.83}$$

Next we take Z to be of the general form

$$Z = \begin{bmatrix} A & B & C & D \\ E & F & G & H \\ I & J & K & L \\ M & N & O & P \end{bmatrix}, \tag{4.84}$$

where each entry is a 2×2 matrix. By imposing the condition $Z\alpha_i^* = -\alpha_i Z$ and using the result that there is no A satisfying $A\sigma_i^* = \sigma_i A$ and that if $A\sigma_i^* = -\sigma_i A$ then $A = ai\sigma_2$, where $a \in \mathbb{C}$, we find that Z can take the form

$$Z = \begin{bmatrix} Vai\sigma_2 V^{-1*} & 0 & 0 & 0 \\ 0 & Vbi\sigma_2 V^{-1*} & 0 & 0 \\ 0 & 0 & Wci\sigma_2 W^{-1*} & 0 \\ 0 & 0 & 0 & Wdi\sigma_2 W^{-1*} \end{bmatrix},$$

where $a, b, c, d \in \mathbb{C}$. Requiring $ZZ^* = -1$ then gives us $|a|^2 = |b|^2 = |c|^2 =$

$|d|^2 = 1$, which means that they are all pure phases. We thus have

$$Z = \begin{bmatrix} Ve^{i\phi_a} i\sigma_2 V^{-1*} & 0 & 0 & 0 \\ 0 & Ve^{i\phi_b} i\sigma_2 V^{-1*} & 0 & 0 \\ 0 & 0 & We^{i\phi_c} i\sigma_2 W^{-1*} & 0 \\ 0 & 0 & 0 & We^{i\phi_d} i\sigma_2 W^{-1*} \end{bmatrix}.$$

By imposing the condition $\alpha_i = Z^T \alpha_i^T Z^\dagger$ and using the properties $(i\sigma_2)^\dagger = (i\sigma_2)^T = -i\sigma_2$ and $\sigma_2 \sigma_i^T \sigma_2 = -\sigma_i$ we find

$$\begin{bmatrix} V\sigma_i V^{-1} & 0 & 0 & 0 \\ 0 & V\sigma_i V^{-1} & 0 & 0 \\ 0 & 0 & -W\sigma_i W^{-1} & 0 \\ 0 & 0 & 0 & -W\sigma_i W^{-1} \end{bmatrix} = \tag{4.85}$$

$$\begin{bmatrix} V^{-1\dagger}\sigma_i V^\dagger & 0 & 0 & 0 \\ 0 & V^{-1\dagger}\sigma_i V^\dagger & 0 & 0 \\ 0 & 0 & -W^{-1\dagger}\sigma_i W^\dagger & 0 \\ 0 & 0 & 0 & -W^{-1\dagger}\sigma_i W^\dagger \end{bmatrix},$$

which gives us

$$V^\dagger V \sigma_i = \sigma_i V^\dagger V \quad \text{and} \quad W^\dagger W \sigma_i = \sigma_i W^\dagger W. \tag{4.86}$$

Next recall that if $[A, \sigma_i] = 0$ for $i = 1, 2, 3$, then $A = a\sigma_0$, where $a \in \mathbb{C}$. Furthermore, as the combinations $V^\dagger V$ and $W^\dagger W$ are Hermitian, we know that in their diagonal forms the entries will be real and positive. We can thus take $V^\dagger V = v^2 \sigma_0$ and $W^\dagger W = w^2 \sigma_0$, where $v, w \in \mathbb{R}$. From this we can also infer that $V^{-1} = \frac{1}{v^2} V^\dagger$ and $W^{-1} = \frac{1}{w^2} W^\dagger$, which in turn tells us that the matrices $\frac{1}{v} V$ and $\frac{1}{w} W$ are unitary. Plugging these results into our expressions for α_i and Z we have

$$\alpha_i = \begin{bmatrix} v^{-1}V\sigma_i v^{-1}V^\dagger & 0 & 0 & 0 \\ 0 & v^{-1}V\sigma_i v^{-1}V^\dagger & 0 & 0 \\ 0 & 0 & -w^{-1}W\sigma_i w^{-1}W^\dagger & 0 \\ 0 & 0 & 0 & -w^{-1}W\sigma_i w^{-1}W^\dagger \end{bmatrix},$$

$$Z = \begin{bmatrix} v^{-2}Ve^{i\phi_a}i\sigma_2 V^T & 0 & 0 & 0 \\ 0 & v^{-2}Ve^{i\phi_b}i\sigma_2 V^T & 0 & 0 \\ 0 & 0 & w^{-2}We^{i\phi_c}i\sigma_2 W^T & 0 \\ 0 & 0 & 0 & w^{-2}We^{i\phi_d}i\sigma_2 W^T \end{bmatrix}. \tag{4.87}$$

If we then define the unitary matrix U as

$$U = \begin{bmatrix} v^{-1}Ve^{\frac{1}{2}i\phi_a} & 0 & 0 & 0 \\ 0 & v^{-1}Ve^{\frac{1}{2}i\phi_b} & 0 & 0 \\ 0 & 0 & w^{-1}We^{\frac{1}{2}i\phi_c} & 0 \\ 0 & 0 & 0 & w^{-1}We^{\frac{1}{2}i\phi_d} \end{bmatrix}, \tag{4.88}$$

then we have

$$\alpha_i = U \begin{bmatrix} \sigma_i & 0 & 0 & 0 \\ 0 & \sigma_i & 0 & 0 \\ 0 & 0 & -\sigma_i & 0 \\ 0 & 0 & 0 & -\sigma_i \end{bmatrix} U^\dagger \tag{4.89}$$

$$\text{and} \quad Z = U \begin{bmatrix} i\sigma_2 & 0 & 0 & 0 \\ 0 & i\sigma_2 & 0 & 0 \\ 0 & 0 & i\sigma_2 & 0 \\ 0 & 0 & 0 & i\sigma_2 \end{bmatrix} U^T. \tag{4.90}$$

Given these forms for α_i and Z, we posit that S and β are of the form $S = U\Sigma U^\dagger$ and $\beta = U\mathcal{B}U^\dagger$. Now, taking the condition $\{S, \alpha_i\} = 0$, we have

$$\Sigma \begin{bmatrix} \sigma_i & 0 & 0 & 0 \\ 0 & \sigma_i & 0 & 0 \\ 0 & 0 & -\sigma_i & 0 \\ 0 & 0 & 0 & -\sigma_i \end{bmatrix} + \begin{bmatrix} \sigma_i & 0 & 0 & 0 \\ 0 & \sigma_i & 0 & 0 \\ 0 & 0 & -\sigma_i & 0 \\ 0 & 0 & 0 & -\sigma_i \end{bmatrix} \Sigma = 0. \tag{4.91}$$

Exploiting the results that there is no matrix that anti-commutes with all σ_i and that if $[A, \sigma_i] = 0$ for $i = 1, 2, 3$, then $A = a\sigma_0$, where $a \in \mathbb{C}$, we find that Σ can take the form

$$\Sigma = \begin{bmatrix} 0 & 0 & a\sigma_0 & 0 \\ 0 & 0 & 0 & b\sigma_0 \\ c\sigma_0 & 0 & 0 & 0 \\ 0 & d\sigma_0 & 0 & 0 \end{bmatrix}, \tag{4.92}$$

where $a, b, c, d \in \mathbb{C}$. The condition $S^2 = 1$ also gives us $\Sigma^2 = 1$, which in turn gives us $c = a^{-1}$ and $d = b^{-1}$. The condition $SZ = ZS^*$ is equivalent to

$$\Sigma \begin{bmatrix} i\sigma_2 & 0 & 0 & 0 \\ 0 & i\sigma_2 & 0 & 0 \\ 0 & 0 & i\sigma_2 & 0 \\ 0 & 0 & 0 & i\sigma_2 \end{bmatrix} = \begin{bmatrix} i\sigma_2 & 0 & 0 & 0 \\ 0 & i\sigma_2 & 0 & 0 \\ 0 & 0 & i\sigma_2 & 0 \\ 0 & 0 & 0 & i\sigma_2 \end{bmatrix} \Sigma^*, \tag{4.93}$$

from which we deduce that a and b are in fact real.

Now we consider the condition $\{\alpha_i, \beta\} = 0$. In terms of $\mathcal{B}$ this gives us

$$\begin{bmatrix} \sigma_i & 0 & 0 & 0 \\ 0 & \sigma_i & 0 & 0 \\ 0 & 0 & -\sigma_i & 0 \\ 0 & 0 & 0 & -\sigma_i \end{bmatrix} \mathcal{B} + \mathcal{B} \begin{bmatrix} \sigma_i & 0 & 0 & 0 \\ 0 & \sigma_i & 0 & 0 \\ 0 & 0 & -\sigma_i & 0 \\ 0 & 0 & 0 & -\sigma_i \end{bmatrix} = 0, \tag{4.94}$$

which, using the known properties of matrices that commute or anti-commute with all σ_i, gives the most general possible form of $\mathcal{B}$ as

$$\mathcal{B} = \begin{bmatrix} 0 & 0 & k\sigma_0 & l\sigma_0 \\ 0 & 0 & m\sigma_0 & n\sigma_0 \\ p\sigma_0 & q\sigma_0 & 0 & 0 \\ r\sigma_0 & h\sigma_0 & 0 & 0 \end{bmatrix}, \tag{4.95}$$

with $k, l, m, n, p, q, r, s \in \mathbb{C}$. Next, taking the condition $\beta = -Z^T \beta^T Z^\dagger$, which is equivalent to

$$\mathcal{B} = - \begin{bmatrix} i\sigma_2 & 0 & 0 & 0 \\ 0 & i\sigma_2 & 0 & 0 \\ 0 & 0 & i\sigma_2 & 0 \\ 0 & 0 & 0 & i\sigma_2 \end{bmatrix} \mathcal{B}^T \begin{bmatrix} i\sigma_2 & 0 & 0 & 0 \\ 0 & i\sigma_2 & 0 & 0 \\ 0 & 0 & i\sigma_2 & 0 \\ 0 & 0 & 0 & i\sigma_2 \end{bmatrix}, \tag{4.96}$$

we find

$$\begin{bmatrix} 0 & 0 & k\sigma_0 & l\sigma_0 \\ 0 & 0 & m\sigma_0 & n\sigma_0 \\ p\sigma_0 & q\sigma_0 & 0 & 0 \\ r\sigma_0 & h\sigma_0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & p\sigma_0 & r\sigma_0 \\ 0 & 0 & q\sigma_0 & s\sigma_0 \\ k\sigma_0 & m\sigma_0 & 0 & 0 \\ l\sigma_0 & n\sigma_0 & 0 & 0 \end{bmatrix} \tag{4.97}$$

$$\implies \quad \mathcal{B} = \begin{bmatrix} 0 & 0 & k\sigma_0 & l\sigma_0 \\ 0 & 0 & m\sigma_0 & n\sigma_0 \\ k\sigma_0 & m\sigma_0 & 0 & 0 \\ l\sigma_0 & n\sigma_0 & 0 & 0 \end{bmatrix}. \tag{4.98}$$

Finally, taking the condition $S\beta = \beta^\dagger S$, which amounts to $\Sigma\mathcal{B} = \mathcal{B}^\dagger\Sigma$, we have

$$\begin{bmatrix} ak\sigma_0 & am\sigma_0 & 0 & 0 \\ bl\sigma_0 & bn\sigma_0 & 0 & 0 \\ 0 & 0 & a^{-1}k\sigma_0 & a^{-1}l\sigma_0 \\ 0 & 0 & b^{-1}l\sigma_0 & b^{-1}n\sigma_0 \end{bmatrix} = \begin{bmatrix} k^*a^{-1}\sigma_0 & l^*b^{-1}\sigma_0 & 0 & 0 \\ m^*a^{-1}\sigma_0 & n^*b^{-1}\sigma_0 & 0 & 0 \\ 0 & 0 & k^*a\sigma_0 & m^*b\sigma_0 \\ 0 & 0 & l^*a\sigma_0 & n^*b\sigma_0 \end{bmatrix}. \tag{4.99}$$

From this we can read off the following conditions

$$a^2k = k^* \Rightarrow a = \pm 1 \text{ and } k \text{ is real or } a = 0 \tag{4.100}$$

$$b^2n = n^* \Rightarrow b = \pm 1 \text{ and } n \text{ is real or } b = 0 \tag{4.101}$$

$$abl = m^* \tag{4.102}$$

$$\text{and} \quad bam = l^*. \tag{4.103}$$

Taking the case $a = b = 1$, we have

$$\mathcal{B} = \begin{bmatrix} 0 & 0 & k\sigma_0 & l\sigma_0 \\ 0 & 0 & l^*\sigma_0 & n\sigma_0 \\ k\sigma_0 & l^*\sigma_0 & 0 & 0 \\ l\sigma_0 & n\sigma_0 & 0 & 0 \end{bmatrix}, \tag{4.104}$$

which, in terms of the real parameters m_0, m_1, m_2 and m_3, we can express as

$$\mathcal{B} = \begin{bmatrix} 0 & 0 & (m_0+m_3)\sigma_0 & (m_1-im_2)\sigma_0 \\ 0 & 0 & (m_1+im_2)\sigma_0 & (m_0-m_3)\sigma_0 \\ (m_0+m_3)\sigma_0 & (m_1+im_2)\sigma_0 & 0 & 0 \\ (m_1-im_2)\sigma_0 & (m_0-m_3)\sigma_0 & 0 & 0 \end{bmatrix}. \tag{4.105}$$

Summarising, we have

$$\alpha_i = U \begin{bmatrix} \sigma_i & 0 & 0 & 0 \\ 0 & \sigma_i & 0 & 0 \\ 0 & 0 & -\sigma_i & 0 \\ 0 & 0 & 0 & -\sigma_i \end{bmatrix} U^\dagger, \tag{4.106}$$

$$\beta = U \begin{bmatrix} 0 & M \\ M^* & 0 \end{bmatrix} U^\dagger, \tag{4.107}$$

$$S = U \begin{bmatrix} 0 & 0 & \sigma_0 & 0 \\ 0 & 0 & 0 & \sigma_0 \\ \sigma_0 & 0 & 0 & 0 \\ 0 & \sigma_0 & 0 & 0 \end{bmatrix} U^\dagger \quad \text{and} \tag{4.108}$$

$$Z = U \begin{bmatrix} i\sigma_2 & 0 & 0 & 0 \\ 0 & i\sigma_2 & 0 & 0 \\ 0 & 0 & i\sigma_2 & 0 \\ 0 & 0 & 0 & i\sigma_2 \end{bmatrix} U^T, \tag{4.109}$$

$$\text{where} \quad M = \begin{bmatrix} (m_0+m_3)\sigma_0 & (m_1-im_2)\sigma_0 \\ (m_1+im_2)\sigma_0 & (m_0-m_3)\sigma_0 \end{bmatrix}. \tag{4.110}$$

These are the results given in [1] and [2]. Note especially the non-Hermitian nature of β, which means that unlike Model 4, Model 8 does differ from its counterpart in the Hermitian formulation.

Following the method of [2], we now look at solutions to a Dirac equation with α_i and β of the form above. If we consider plane-wave solutions of the form $\psi = u\exp(i(\mathbf{p}\cdot\mathbf{r} - Et))$, where u is an eight-component vector, our Dirac equation looks like

$$\begin{bmatrix} \boldsymbol{\sigma}\cdot\mathbf{p} & 0 & (m_0+m_3)\sigma_0 & (m_1-im_2)\sigma_0 \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{p} & (m_1+im_2)\sigma_0 & (m_0-m_3)\sigma_0 \\ (m_0+m_3)\sigma_0 & (m_1+im_2)\sigma_0 & -\boldsymbol{\sigma}\cdot\mathbf{p} & 0 \\ (m_1-im_2)\sigma_0 & (m_0-m_3)\sigma_0 & 0 & -\boldsymbol{\sigma}\cdot\mathbf{p} \end{bmatrix} u = Eu. \tag{4.111}$$

Now note that the diagonal elements of our Hamiltonian are proportional to the two-dimensional helicity operator. As such, if we consider solutions of the form

$$u = \begin{pmatrix} a\xi_\pm \\ b\xi_\pm \\ c\xi_\pm \\ d\xi_\pm \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \otimes \xi_\pm, \tag{4.112}$$

where $\xi_\pm$ are the positive and negative helicity eigenstates satisfying

$\boldsymbol{\sigma} \cdot \boldsymbol{p}\, \xi_\pm = \pm|\mathbf{p}|\xi_\pm = \pm p\xi_\pm$, then our problem reduces to

$$\begin{bmatrix} \pm p & 0 & (m_0+m_3) & (m_1-im_2) \\ 0 & \pm p & (m_1+im_2) & (m_0-m_3) \\ (m_0+m_3) & (m_1+im_2) & \mp p & 0 \\ (m_1-im_2) & (m_0-m_3) & 0 & \mp p \end{bmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = E \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}. \tag{4.113}$$

As in [2], we now restrict ourselves to the case where $m_1 = m_3 = 0$ and we suggestively write the remaining equation as

$$\begin{bmatrix} \pm p\sigma_0 & m_0\sigma_0 + m_2\sigma_2 \\ m_0\sigma_0 - m_2\sigma_2 & \mp p\sigma_0 \end{bmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = E \begin{pmatrix} A \\ B \end{pmatrix}, \tag{4.114}$$

where A and B are two-component vectors. This is named the restricted Model 8. From its association with the y-component of spin for a spin-half particle, we know σ_2 to have two eigenvectors with eigenvalues ± 1. If we denote these $\zeta_\pm$ and consider solutions of the form

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \alpha\zeta_\pm \\ \beta\zeta_\pm \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_\pm, \tag{4.115}$$

where $\alpha, \beta \in \mathbb{C}$, then we further reduce the problem to

$$\begin{bmatrix} \pm p & m_0 + m_2 \\ m_0 - m_2 & \mp p \end{bmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{for } \zeta_+ \tag{4.116}$$

$$\text{and} \quad \begin{bmatrix} \pm p & m_0 - m_2 \\ m_0 + m_2 & \mp p \end{bmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{for } \zeta_-. \tag{4.117}$$

Regardless of our choice of signs, i.e. our choice of $\xi_\pm$ and $\zeta_\pm$, the determinant of these matrices is $-p^2 - m_0^2 + m_2^2$. Further note that in all cases the matrix is traceless, which tells us that the eigenvalues are of equal magnitude and opposite sign. Combining these two observations gives us

$$E = \pm\sqrt{p^2 + m_0^2 - m_2^2}. \tag{4.118}$$

Thus there are eight solutions in total, which we list here for completeness[8]:

$$\begin{aligned}
\begin{bmatrix} p & m_0 + m_2 \\ m_0 - m_2 & -p \end{bmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_+ \otimes \xi_+ &= E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_+ \otimes \xi_+ \\
\begin{bmatrix} -p & m_0 + m_2 \\ m_0 - m_2 & p \end{bmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_+ \otimes \xi_- &= E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_+ \otimes \xi_- \\
\begin{bmatrix} p & m_0 - m_2 \\ m_0 + m_2 & -p \end{bmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_- \otimes \xi_+ &= E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_- \otimes \xi_+ \\
\begin{bmatrix} -p & m_0 - m_2 \\ m_0 + m_2 & p \end{bmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_- \otimes \xi_- &= E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_- \otimes \xi_-
\end{aligned} \tag{4.119}$$

Defining $m_{eff}^2 = m_0^2 - m_2^2$, the eight solutions can be associated with two spin-$\frac{1}{2}$ particles of mass m_{eff} and their antiparticles.

A case of particular interest is when $m_0 = m_2$, i.e. $m_{eff} = 0$. In this situation we have a Hamiltonian that describes two massless fermions but that also has a non-zero mass matrix. As we have seen in earlier discussions, if a Lagrangian has both mass and interaction terms for some set of fields, these terms are in general not simultaneously diagonal, i.e. mass and flavour eigenstates do not coincide. In the case of neutrinos, we saw that

[8] There are eight as each of the four scenarios has a positive and a negative energy solution.

this leads to the possibility of oscillations between the flavour eigenstates. As such, the observation of particle flavour oscillations is usually taken as an indication that those particles are massive. In the $\mathcal{PT}$ formulation of the Dirac equation, however, the properties of this new eight-dimensional solution would seem to suggest that perhaps we no longer need draw this conclusion, which is potentially very appealing in the context of the neutrino oscillation problem.

Given that there are three neutrino flavours, we now attempt to extend the findings of [1] to three-flavours of massless particle.

4.3.3 Model 12

In order to try and describe three particle flavours we move to a twelve-dimensional representation for α_i and β, and start with α_i of the form

$$\alpha_i = \begin{bmatrix} V\sigma_i V^{-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & V\sigma_i V^{-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & V\sigma_i V^{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & -W\sigma_i W^{-1} & 0 & 0 \\ 0 & 0 & 0 & 0 & -W\sigma_i W^{-1} & 0 \\ 0 & 0 & 0 & 0 & 0 & -W\sigma_i W^{-1} \end{bmatrix} \tag{4.120}$$

The procedure is then entirely equivalent to that shown in the previous section for Model 8, and we will not give the details here. After successively applying the various conditions on α_i, β, S and Z, we find the 12×12

representations to be the natural extension of those from Model 8. Namely

$$\alpha_i = U \begin{bmatrix} \sigma_i & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_i & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_i & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_i & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_i & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_i \end{bmatrix} U^\dagger, \tag{4.121}$$

$$\beta = U \begin{bmatrix} 0 & M \\ M^* & 0 \end{bmatrix} U^\dagger, \tag{4.122}$$

$$S = U \begin{bmatrix} 0 & 0 & 0 & \sigma_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_0 \\ \sigma_0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_0 & 0 & 0 & 0 \end{bmatrix} U^\dagger \tag{4.123}$$

$$\text{and } \; Z = U \begin{bmatrix} i\sigma_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & i\sigma_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & i\sigma_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & i\sigma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & i\sigma_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & i\sigma_2 \end{bmatrix} U^T, \tag{4.124}$$

where

$$U = \begin{bmatrix} v^{-1}Ve^{\frac{i}{2}\phi_a} & 0 & 0 & 0 & 0 & 0 \\ 0 & v^{-1}Ve^{\frac{i}{2}\phi_b} & 0 & 0 & 0 & 0 \\ 0 & 0 & v^{-1}Ve^{\frac{i}{2}\phi_c} & 0 & 0 & 0 \\ 0 & 0 & 0 & w^{-1}We^{\frac{i}{2}\phi_d} & 0 & 0 \\ 0 & 0 & 0 & 0 & w^{-1}We^{\frac{i}{2}\phi_e} & 0 \\ 0 & 0 & 0 & 0 & 0 & w^{-1}We^{\frac{i}{2}\phi_f} \end{bmatrix} \tag{4.125}$$

and

$$M = \begin{bmatrix} (m_0+m_7)\sigma_0 & (m_1-im_2)\sigma_0 & (m_3-im_4)\sigma_0 \\ (m_1+im_2)\sigma_0 & m_0\sigma_0 & (m_5-im_6)\sigma_0 \\ (m_3+im_4)\sigma_0 & (m_5+im_6)\sigma_0 & (m_0+m_7)\sigma_0 \end{bmatrix}. \tag{4.126}$$

Taking plane-wave solutions of the form $\psi = u\exp(i(\mathbf{p}\cdot\mathbf{r} - Et))$, where u is now a twelve-component vector, our Dirac equation looks like

$$\begin{bmatrix} \boldsymbol{\sigma}\cdot\mathbf{p} & 0 & 0 & & & \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{p} & 0 & & M & \\ 0 & 0 & \boldsymbol{\sigma}\cdot\mathbf{p} & & & \\ & & & -\boldsymbol{\sigma}\cdot\mathbf{p} & 0 & 0 \\ & M^* & & 0 & -\boldsymbol{\sigma}\cdot\mathbf{p} & 0 \\ & & & 0 & 0 & -\boldsymbol{\sigma}\cdot\mathbf{p} \end{bmatrix} u = Eu \tag{4.127}$$

As with M8, we then take solutions of the form

$$u = \begin{pmatrix} a\xi_\pm \\ b\xi_\pm \\ c\xi_\pm \\ d\xi_\pm \\ e\xi_\pm \\ f\xi_\pm \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} \otimes \xi_\pm. \tag{4.128}$$

Also notice that hidden within M is the operator for the y-component of spin in the spin-1 representation. We therefore choose to restrict ourselves to the case where $m_1 = m_3 = m_4 = m_5 = m_7 = 0$ and $m_2 = m_6$, in order that we can reduce our equation to

$$\begin{bmatrix} \pm p\mathbb{1} & m_0\mathbb{1} + m_2 J_y \\ m_0\mathbb{1} - m_2 J_y & \mp p\mathbb{1} \end{bmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = E \begin{pmatrix} A \\ B \end{pmatrix}, \tag{4.129}$$

where A and B are three-component vectors and

$$J_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \tag{4.130}$$

Note that we have absorbed a factor of $\sqrt{2}$ into m_2 in order to keep things tidy. We call this the restricted Model 12.

We know J_y to have three eigenvectors with eigenvalues ± 1 and 0. We label these eigenvectors $\zeta_{\pm,0}$ and consider solutions of the form

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \alpha\zeta_{\pm,0} \\ \beta\zeta_{\pm,0} \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_{\pm,0}, \tag{4.131}$$

where α, $\beta \in \mathbb{C}$. This now gives us three sets of solutions:

$$\begin{bmatrix} \pm p & m_0 + m_2 \\ m_0 - m_2 & \mp p \end{bmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{for } \zeta_+, \tag{4.132}$$

$$\begin{bmatrix} \pm p & m_0 \\ m_0 & \mp p \end{bmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{for } \zeta_0 \tag{4.133}$$

$$\text{and} \quad \begin{bmatrix} \pm p & m_0 - m_2 \\ m_0 + m_2 & \mp p \end{bmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = E \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{for } \zeta_-. \tag{4.134}$$

Listing all possible combinations of $\xi_\pm$ and $\zeta_{\pm,0}$ we have

$$\begin{bmatrix} p & m_0+m_2 \\ m_0-m_2 & -p \end{bmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\otimes\zeta_+\otimes\xi_+ = E\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\otimes\zeta_+\otimes\xi_+$$
$$\begin{bmatrix} -p & m_0+m_2 \\ m_0-m_2 & p \end{bmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\otimes\zeta_+\otimes\xi_- = E\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\otimes\zeta_+\otimes\xi_-$$
$$\begin{bmatrix} p & m_0 \\ m_0 & -p \end{bmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\otimes\zeta_0\otimes\xi_+ = E\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\otimes\zeta_0\otimes\xi_+$$
$$\begin{bmatrix} -p & m_0 \\ m_0 & p \end{bmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\otimes\zeta_0\otimes\xi_- = E\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\otimes\zeta_0\otimes\xi_-$$
$$\begin{bmatrix} p & m_0-m_2 \\ m_0+m_2 & -p \end{bmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\otimes\zeta_-\otimes\xi_+ = E\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\otimes\zeta_-\otimes\xi_+$$
$$\begin{bmatrix} -p & m_0-m_2 \\ m_0+m_2 & p \end{bmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\otimes\zeta_-\otimes\xi_- = E\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\otimes\zeta_-\otimes\xi_- \qquad (4.135)$$

In all six cases we see that the matrices are traceless, meaning that they each have two eigenvalues of equal magnitude but opposite sign. In the case of ζ_+ and ζ_-, we have the same situation as in Model 8, with energies given by $E = \pm\sqrt{p^2 + m_{eff}^2}$ and $m_{eff}^2 = m_0^2 - m_2^2$. In the case of ζ_0, however, the energies are given by $E = \pm\sqrt{p^2 + m_0^2}$.

Altogether we therefore have twelve solutions to our equation, corresponding to the three spin-$\frac{1}{2}$ particles. Two of these are potentially massless, if $m_0 = m_2$ holds, but the third has a non-zero mass m_0. As such, this first natural extension of the surprising Model 8 solution does not appear to give us another new massless solution describing three flavours of massless particle with a non-zero mass matrix. It consequently seems not to be a candidate for potentially allowing oscillations between three flavours of massless particle.

The appearance of a massive particle in this 12×12 representation was a result of the fact that in decomposing the solutions we made use of the spin-1 representations of spin operators and their eigenvectors. Specifically, the massive state corresponds to the fact that one of the eigenvectors has eigenvalue zero. It would therefore be interesting to look at higher dimension solutions, for example 16×16, where it might be possible to decompose the solutions using the spin-$\frac{3}{2}$ representation, which has no zero eigenvalue.

Without going through a complete analysis, one might expect to find that the $2N \times 2N$ representation of the $\mathcal{PT}$-symmetric Dirac equation generalises to

$$\begin{bmatrix} \boldsymbol{\sigma} \cdot \boldsymbol{p} \otimes \mathbb{1} & M \\ M^* & -\boldsymbol{\sigma} \cdot \boldsymbol{p} \otimes \mathbb{1} \end{bmatrix} u = Eu \tag{4.136}$$

for plane-wave solutions, where the form of M is the natural extension of (4.126) to an $N \times N$ matrix and u is a $2N$-component vector. If we then take solutions of the form $u = v \otimes \xi_{\pm}$, where v is an N-component vector, and further assume that the real parameters within M can be chosen such as to give

$$\begin{bmatrix} \pm p\mathbb{1} & m_0\mathbb{1} + m_2 S_y \\ m_0\mathbb{1} - m_2 S_y & \mp p\mathbb{1} \end{bmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = E \begin{pmatrix} A \\ B \end{pmatrix}, \tag{4.137}$$

where S_y is the $N/2$-dimensional representation for the y-component of spin and A and B are $N/2$-component vectors, then we can see that there will always be at most two particles that have the same m_{eff}. Being more explicit, if we take solutions of the form

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \alpha\zeta_n \\ \beta\zeta_n \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \otimes \zeta_n, \tag{4.138}$$

where $\alpha, \beta \in \mathbb{C}$ and ζ_n for $n = 1, ..., N/2$ are the $N/2$ eigenvectors of S_y with eigenvalues $s_y = -\frac{N/2-1}{2}, ..., \frac{N/2-1}{2}$, then we will have m_{eff} given by $m_{eff}^2 = m_0^2 - s_y^2 m_2^2$. Consequently, except for the case where $s_y = 0$, the two eigenvectors that have eigenvalues of the same magnitude but opposite sign will be associated with particles of the same mass m_{eff}. Thus, if we want to choose m_0 and m_2 in such a way as to give $m_{eff} = 0$, this will hold for at most two particles, and only one in the case that our particle corresponds to $s_y = 0$. This would suggest that within the framework of a restricted form of M, in going to higher dimensions we will not find any new massless solutions that could potentially allow for oscillations between more than two flavours of massless particle. Performing a more complete analysis where these restrictions on M are lifted would certainly be of interest.

4.3.4 Does the restricted Model 8 describe new physics?

As we have seen, the 8-dimensional solution to the $\mathcal{PT}$-symmetric Dirac equation displays the striking characteristic of allowing a non-zero mass matrix to describe two particles of zero mass. If considered alongside interaction terms in a more general Lagrangian, this would in general lead to a set of mass eigenstates and a set of flavour eigenstates that do not coincide. In the case of neutrinos, we saw that a similar discrepancy between mass and flavour eigenstates led us to the possibility of neutrino oscillations, and so we might expect the same to be true here. If this were to be the case, oscillations between flavours of massless particles would be a new phenomenon, and certainly of great interest in relation to the problem of observed neutrino oscillations. We must note, however, that the exact mechanism for the oscillations could not be as described earlier for the neutrinos. In the formulation for neutrinos, we relied on the fact that mass and flavour eigenstates did not coincide *and* the fact that the different mass eigenstates evolved differently between creation and detection as a result of their having different masses. Explicitly, the phase factor giving rise to the observed oscillation effect was a function of $\Delta m_{ij}^2 = m_i^2 - m_j^2$. Thus, in the case of the restricted Model 8 solution, where $m_i = m_j = 0$, the phase factor would vanish and we would not predict the observation of oscillations, despite the fact that the mass and flavour eigenstates may not coincide.

Even if an alternative mechanism for achieving oscillations does not exist, however, the restricted Model 8 solution still represents a remarkable new type of solution, and so it is natural to approach it with some caution. In the Hermitian case we find that the 8-dimensional representation of the Dirac equation can be decomposed into a direct sum of two 4-dimensional ones via a unitary transformation, and as such it does not describe a new type of particle. A first step in testing the validity of the new solution would therefore be to confirm that the same decomposition is not possible for the non-Hermitian case. If we were able to decompose the representation into two 4-dimensional ones, then, due to the known equivalence of the Hermitian and $\mathcal{PT}$-symmetric 4-dimensional representations, we would establish that the 8-dimensional solution did not represent a new type of particle.

In the context of pseudo-Hermiticity, finding the equivalent Hermitian Hamiltonian and corresponding eigenstates is also of great interest to us, as it would allow us to verify that the new solution is not simply an already known solution in disguise. As we saw previously, in going from a $\mathcal{PT}$

formulation to the equivalent Hermitian one, the solutions $|\psi\rangle$ get mapped to $\rho|\psi\rangle$, where ρ can be parameterised using $\mathcal{Q}(\hat{x},\hat{p})$, a Hermitian function of $\hat{x}$ and $\hat{p}$, as $\rho = e^{-\frac{\mathcal{Q}}{2}}$. It is argued by Jones-Smith *et al.* that due to the non-local nature of this mapping, the resulting solutions $\rho|\psi\rangle$ cannot be equivalent to the reducible 8-dimension solutions found by directly solving the 8-dimensional Hermitian Hamiltonian. As such, it is argued that these new solutions are indeed irreducible and represent a new type of particle.

Whilst this argument is certainly valid, the explicit form of ρ is yet to be calculated, and thus we suggest that one cannot be certain of the drawn conclusion. The reason for our doubting the certainty of this conclusion is based on the results of [22]. In this paper, the $\mathcal{PT}$ symmetric Hamiltonian

$$\mathcal{H}(x,t) = \overline{\psi}(x,t)(-i\nabla\!\!\!/ + m_1 + m_2\gamma^5)\psi(x,t), \tag{4.139}$$

where $m_2 \in \mathbb{R}$, is shown to describe a particle with physical mass $\mu^2 = m_1^2 - m_2^2$. This mass is therefore real in the scenario $m_1^2 \geq m_2^2$, and this corresponds to the region of unbroken $\mathcal{PT}$ symmetry. Bender *et al.* then proceed to calculate $\mathcal{Q}$ and indeed find that the Hermitian Hamiltonian, h, equivalent to $\mathcal{H}$ is that of a particle with mass $\mu^2 = m_1^2 - m_2^2$, namely

$$h = \overline{\psi}(x,t)(-i\nabla\!\!\!/ + \mu)\psi(x,t). \tag{4.140}$$

Furthermore, whilst the transformation ρ of ψ is in general non-local, in this case ψ is found to transform as

$$\psi \to e^{-\frac{1}{2}\eta\gamma^5}\psi \tag{4.141}$$

$$\text{and} \quad \overline{\psi} \to \overline{\psi}e^{-\frac{1}{2}\eta\gamma^5}, \tag{4.142}$$

where m_2 is parameterised as $m_2 = \epsilon m_1$ with $\epsilon \leq 1$ and $\epsilon = \tanh(\eta)$, which is in fact local. As such, we suggest that a similar result could be found for the transformations of the restricted Model 8 solutions.

One way in which to establish whether or not a similar result holds in the restricted Model 8 is to explicitly calculate ρ given our knowledge of the solutions set out in (4.119). Alternatively, it might prove easier to try and exploit the results of [22]. In the formulation we have followed, the presence of a γ^5 mass term would actually appear as a term proportional to $\gamma^0\gamma^5$, in the same way that our α_i correspond to $\gamma^0\gamma^i$. Thus, if we were able to show that our restricted Model 8 matrix

$$\begin{bmatrix} \boldsymbol{\sigma}\cdot\mathbf{p} & 0 & m_0\sigma_0 & -im_2\sigma_0 \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{p} & im_2\sigma_0 & m_0\sigma_0 \\ m_0\sigma_0 & im_2\sigma_0 & -\boldsymbol{\sigma}\cdot\mathbf{p} & 0 \\ -im_2\sigma_0 & m_0\sigma_0 & 0 & -\boldsymbol{\sigma}\cdot\mathbf{p} \end{bmatrix} \tag{4.143}$$

could be brought into the form

$$\begin{bmatrix} \begin{pmatrix} \boldsymbol{\sigma}\cdot\mathbf{p} & m_0\sigma_0 \\ m_0\sigma_0 & -\boldsymbol{\sigma}\cdot\mathbf{p} \end{pmatrix} + m_2\sigma_0\gamma^0\gamma^5 & 0 \\ 0 & \begin{pmatrix} \boldsymbol{\sigma}\cdot\mathbf{p} & m_0\sigma_0 \\ m_0\sigma_0 & -\boldsymbol{\sigma}\cdot\mathbf{p} \end{pmatrix} + m_2\sigma_0\gamma^0\gamma^5 \end{bmatrix}, \tag{4.144}$$

where γ^0 and γ^5 take the appropriate form of their 2×2 representation, then we would have a direct sum of two Hamiltonians of the form (4.139), which we know to be equivalent to an Hermitian representation of two particles with mass $\mu^2 = m_1^2 - m_2^2$. In the special case $m_1 = m_2$ these would be massless.

Given that we already know the transformation that would bring the Hermitian matrix

$$\begin{bmatrix} \boldsymbol{\sigma}\cdot\mathbf{p} & 0 & m_0\sigma_0 & 0 \\ 0 & \boldsymbol{\sigma}\cdot\mathbf{p} & 0 & m_0\sigma_0 \\ m_0\sigma_0 & 0 & -\boldsymbol{\sigma}\cdot\mathbf{p} & 0 \\ 0 & m_0\sigma_0 & 0 & -\boldsymbol{\sigma}\cdot\mathbf{p} \end{bmatrix} \tag{4.145}$$

into the form

$$\begin{bmatrix} \begin{pmatrix} \boldsymbol{\sigma}\cdot\mathbf{p} & m_0\sigma_0 \\ m_0\sigma_0 & -\boldsymbol{\sigma}\cdot\mathbf{p} \end{pmatrix} & 0 \\ 0 & \begin{pmatrix} \boldsymbol{\sigma}\cdot\mathbf{p} & m_0\sigma_0 \\ m_0\sigma_0 & -\boldsymbol{\sigma}\cdot\mathbf{p} \end{pmatrix} \end{bmatrix}, \tag{4.146}$$

and it so happens that this transformation leaves the non-Hermitian element

$$\begin{bmatrix} 0 & 0 & 0 & -im_2\sigma_0 \\ 0 & 0 & im_2\sigma_0 & 0 \\ 0 & im_2\sigma_0 & 0 & 0 \\ -im_2\sigma_0 & 0 & 0 & 0 \end{bmatrix} \tag{4.147}$$

invariant, the task remaining is to find a transformation that would bring the matrix (4.147) into the form

$$\begin{bmatrix} m_2\sigma_0\gamma^0\gamma^5 & 0 \\ 0 & m_2\sigma_0\gamma^0\gamma^5 \end{bmatrix} \tag{4.148}$$

without changing the form of (4.146). We note that because we require the resulting Hamiltonian to be $\mathcal{PT}$-symmetric, the transformation matrix is also required to be $\mathcal{PT}$ symmetric. We can see this as follows: if $H' = MHM^{-1}$, where H is $\mathcal{PT}$-symmetric, and we also require H' to be $\mathcal{PT}$-symmetric, then we have

$$\mathcal{PT}H'\mathcal{PT} = H' = \mathcal{PT}MHM^{-1}\mathcal{PT} \tag{4.149}$$

$$= \mathcal{PT}M\mathcal{PT}\mathcal{PT}H\mathcal{PT}\mathcal{PT}M^{-1}\mathcal{PT} \tag{4.150}$$

$$= \mathcal{PT}M\mathcal{PT}H\mathcal{PT}M^{-1}\mathcal{PT} \tag{4.151}$$

$$= \mathcal{PT}M\mathcal{PT}H(\mathcal{PT}M\mathcal{PT})^{-1} \tag{4.152}$$

$$\rightarrow \mathcal{PT}M\mathcal{PT} = M. \tag{4.153}$$

Determining the existence or non-existence of such a transformation is key in establishing the validity of this seemingly new type of particle.

5 Closing remarks

We have seen that the Standard Model of particle physics, as it stands, is unable to account for the observed oscillation of neutrinos between flavour eigenstates. Allowing neutrinos to have a small but non-zero mass is widely accepted as the mechanism by which this phenomenon can be explained, and with its assumed particle content and requirements for gauge invariance and renormalisablity, no such mass generating terms are permitted within the Standard Model Lagrangian. In order to accommodate these non-zero masses, we are therefore required to make extensions to the Standard Model, and we have considered two possible examples.

In the first example, the particle content of the Standard Model was supplemented with a light, sterile, right-handed neutrino, which allowed for neutrino masses to be generated in exactly the same way as for the other fermions - via the Higgs mechanism. Problems with this model were the lack of evidence for the proposed right-handed neutrinos and the fact that it required some of its parameters to be tuned very close to zero.

The second example was one in which the requirement for renormalisability was relinquished, which amounts to conceding that the Standard Model is not a complete theory for energy scales greater than that of the electroweak sector. We saw that the only Lorentz and gauge invariant term to 'first-order' of non-renormalisability was exactly that required to generate a neutrino mass without introducing any new, light particles. Furthermore, the neutrino mass generated was inversely proportional to the energy scale of the new physics, so naturally came out small. As a specific example of what the new, high-energy physics might be, we looked at the possibility that there exist very massive right-handed neutrino states, and we saw how these could give rise to the dimension-5, non-renormalisable term that in turn generates the neutrino mass. Given their large mass, we would not expect to have observed these proposed states directly, but non-zero neutrino masses could be an indication of their existence. Although this second type of mechanism is perhaps more appealing, we noted that due to the high-energy nature of the new physics, it is not possible for us to determine its exact nature through experiment.

Having established the uncertainty that remains in trying to incorporate neutrino oscillations and their associated non-zero mass into the Standard Model, we then turned to a recent result in the field of $\mathcal{PT}$-symmetric quantum mechanics that could potentially offer an interesting new possibility. Following the work of Jones-Smith *et al.*, we saw that the 8-dimensional solution to the $\mathcal{PT}$-symmetric Dirac equation, dubbed the restricted Model 8 solution, might be able to allow for oscillations between between two flavours of *massless* particle. This was by virtue of the fact that the Dirac equation contained a non-zero mass matrix, despite the massless nature of the particles it described, which in turn leads to the possibility that flavour and mass eigenstates do not coincide. We stressed, however, that the exact mechanism behind the oscillations could not be the standard one used to describe the phenomenon of massive neutrino oscillations, as this requires a non-zero Δm_{ij}^2, which would not be the case in the restricted Model 8 scenario. We also stressed that before we can conclude for certain that a new type of particle has indeed been found, an explicit calculation of the equivalent Hermitian Hamiltonian and its solutions should be calculated.

If the validity of this new type of particle is confirmed, and an alternative mechanism for describing the oscillations can be determined, then its potential relevance in the context of neutrino oscillations is certainly of great interest. The first natural question is whether or not a higher dimensional solution can be found that would allow for 3-flavour oscillations. Our initial analysis would suggest that such a solution does not exist, but a more complete analysis that places fewer restrictions on the form of the mass matrix is certainly worth pursuing. Jones-Smith *et al.* have already set out a reformulation of the restricted Model 8 as a quantum field theory in [1] and [2], but the question of how to incorporate this new type of particle into the Standard Model is one that remains to be addressed. Lastly, seeing as the new type of particle comes out of the *restricted* Model 8, an exploration of other, potentially new solutions where these restrictions are lifted should also be considered. Equally, given that a new type of solution was found when considering an 8-dimensional representation of the $\mathcal{PT}$-symmetric Dirac equation, this might suggest that there will be additional new types of solution if we consider higher-dimensional representations. Even if these solutions don't describe 3-flavour oscillations, they are still of great interest.

Bibliography

[1] Jones-Smith K. and Mathur H. “Relativistic Non-Hermitian Quantum Mechanics”, *arXiv: 0908.4257v1 [hep-th]* (2009).

[2] Jones-Smith K. “Non-Hermitian Quantum Mechanics”, PhD Thesis 2010, Case Western Reserve University.

[3] Bender C. M. “Making sense of non-Hermitian Hamiltonians”, Rep. Prog. Phys. **70**, 947 - 1018 (2007).

[4] Burgess C. P. and Moore G. D. “The Standard Model: A Primer”, Cambridge University Press (2007).

[5] Cottingham W. N. and Greenwood D. A. “An Introduction to the Standard Model of Particle Physics”, Second Edition, Cambridge University Press 2008.

[6] Rajantie A. “Advanced Quantum Field Theory”, Imperial College MSc course lecture notes (2009/10).

[7] Stelle K. S “Unification”, Imperial College course lecture notes (2009/10).

[8] Gribov V. “Neutrino astronomy and lepton charge”, Phys. Lett. B **28**, 493 (1969).

[9] Gonzalez-Garcia M. C. “Phenomenology with Massive Neutrinos”, *arXiv:0704.1800v2 [hep-ph]*, (2007).

[10] Nakamura K. “Neutrino Mass, Mixing and Oscillations”, PDG Review, JPG 37, 075021, (2010), (http://pdg.lbl.gov).

[11] Beuthe M. “Oscillations of neutrinos and mesons in quantum field theory”, Physics Reports 375, 105218, (2003).

[12] Blasone M. "Observables in the Quantum Field Theory of neutrino mixing and oscillations", *arXiv:hep-ph/0308009v1*, (2003).

[13] Mostafazadeh A. "Pseudo-Hermiticity versus PT Symmetry: The necessary condition for the reality of the spectrum of a non-Hermitian Hamiltonian", *arXiv 0107001v3 [math-ph]*, (2001).

[14] Weinberg S. "The Quantum Theory of Fields", Volume I, Cambridge University Press 1995.

[15] Japaridze G. S. "Space of state vectors in PT symmetrized quantum mechanics", LANL Archives: *quant-ph/0104077*, (2001).

[16] Mostafazadeh A. "Exact PT-symmetry is equivalent to Hermiticity", J. Phys. A: Math. Gen. **36** 7081, (2003).

[17] Mostafazadeh A. "Pseudo-Hermiticity versus PT Symmetry: Equivalence of pseudo-Hermiticity and the presence of anti-linear symmetries", *arXiv: math-ph/0203005v2*, (2002).

[18] Mostafazadeh A. "Pseudo-Hermiticity and Generalized PT- and CPT -Symmetries", J. Math. Phys. **44** 974 (Preprint math-ph/0209018), (2003).

[19] Jones-Smith K. and Mathur H. "A New Class of non-Hermitian Quantum Hamiltonians with PT symmetry", *arXiv: 0908.4255v3 [hep-th]*, (2009).

[20] Aitchison I. J. R. "Relativistic Quantum Mechanics", The Macmillan Press Ltd, 1972.

[21] Dirac P. A. M. "The Quantum Theory of the Electron", Proc. Roy. Soc **A117** 610, (1928).

[22] Bender C. M. "Dual $\mathcal{PT}$-Symmetric Quantum Field Theories", *arXiv: hep-th/0508105v1*, (2005).