The Seesaw Mechanism

By Robert D. Klauber
www.quantumfieldtheory.info

1 Background

It may seem unusual to have such low values for masses of neutrinos, when all other particles like electrons, quarks, etc are much heavier, with their masses relatively closely grouped. Given that particles get mass via the Higgs mechanism, why, for example, should the electron neutrino be 10^5 times or more lighter than the electron, up and down quarks. That is, why would the coupling to the Higgs field be so many orders of magnitude less?

One might not be too surprised if the Higgs coupling were zero, giving rise to zero mass. One might likewise not be too surprised if the coupling resulted in masses on the order of the Higgs, or even the GUT, symmetry breaking scale.

Consider the quite reasonable possibility that after symmetry breaking, two types of neutrino exist, with one having zero mass (no Higgs coupling) and the other having (large) mass of the symmetry breaking scale. As we will see, it turns out that reasonable superpositions of these fields can result in light neutrinos (like those observed) and a very heavy neutrino (of symmetry breaking scale, and unobserved).

2 Fundamental Math Concept Underlying the Seesaw Mechanism

Consider a real, two dimensional space with a matrix (tensor) expressed in one set of orthonormal basis vectors (primed) for that space as

$$\tilde{\mathcal{M}} = \begin{bmatrix} 0 & 0 \\ 0 & 100 \end{bmatrix} . \qquad (1)$$

Now, if we consider a new set of basis vectors, rotated by an angle ϕ from the original basis, then the matrix components change, of course, and can be found by

$$\begin{aligned} \mathcal{M} &= \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 100 \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \\ &= \begin{bmatrix} 100\sin^2\phi & 100\cos\phi\sin\phi \\ -100\cos\phi\sin\phi & 100\cos^2\phi \end{bmatrix} . \end{aligned} \qquad (2)$$

Note how this matrix looks if ϕ is small, say $\phi = 2°$, with $\cos\phi = .99939$ and $\sin\phi = .03490$.

$$\mathcal{M} = \begin{bmatrix} .122 & 3.488 \\ -3.488 & 99.878 \end{bmatrix} , \qquad (3)$$

and we get the upper left diagonal term almost 3 orders of magnitude smaller than the lower right term, which is approximately the same as the original such term. The off diagonal terms equal the geometric mean of the diagonal terms, i.e., $\sqrt{\left(100\cos^2\phi\right)\left(100\sin^2\phi\right)}$, and are not as small as the upper left term, but significantly smaller that the lower right one.

The fundamental point is that by starting with a matrix of form like (1), and transforming to another basis, which is rotated by a small angle from the original, we get a matrix of form like (3).

3 Dirac vs Majorana Mass Terms in the Lagrangian

We don’t know a great deal, experimentally, about neutrino mass, but on general theoretical grounds, two distinct classes of neutrino mass terms are allowed in the Lagrangian of electroweak interactions. These are called Dirac and Majorana mass terms.

Note that Majorana mass terms have nothing to do with the Majorana representation in spinor space. One can use any representation for the fields of which Majorana and Dirac mass terms are composed. Neither do Majorana mass terms imply the associated particles/fields are Majorana fermions, of which you may have heard. Majorana fermions are their own anti-particles. More on this in Sect. 5. For now, we will assume that both Dirac and Majorana mass terms contain only Dirac type particles (in any representation we like.)

The **Dirac** mass terms, which are the usual terms dealt with in introductory quantum field theory (QFT), have form

$$-m_D\left(\bar{\nu}_L\nu_R+\bar{\nu}_R\nu_L\right), \tag{4}$$

and **Majorana** mass terms, which may look unfamiliar to the uninitiated, have form

$$-\tfrac{1}{2}m_M^L\left(\bar{\nu}_L\nu_L{}^c+\bar{\nu}_L{}^c\nu_L\right)-\tfrac{1}{2}m_M^R\left(\bar{\nu}_R\nu_R{}^c+\bar{\nu}_R{}^c\nu_R\right), \tag{5}$$

where sub/superscripts L and R designate left or right hand chirality, and the superscript c represents charge conjugation. That is,

ν_L destroys a LH chiral neutrino and creates a RH antineutrino,
$\bar{\nu}_L$ creates “ “ “ “ and destroys “ “ “ ,
$\nu_L{}^c$ creates “ “ “ “ and destroys “ “ “ (does same as $\bar{\nu}_L$),
$\bar{\nu}_L{}^c$ destroys “ “ “ “ and creates “ “ “ (does same as ν_L),

and for R subscript, interchange $L \leftrightarrow R$ everywhere above.

Note that the subscript always refers to particles. For a non conjugated field, no overbar means destroys particles, overbar means creates particles, and antiparticle actions for the same field are just reversed from particle actions (particle ↔ antiparticle, LH ↔ RH, destroy ↔ create).

Charge conjugating a field has the same effect on particle/antiparticle and creation/destruction as an overbar (overbar is effectively a complex conjugate transpose [plus a γ_0 multiplication]). That is, the overbar and the superscript “*c*” have the same effect. The charge conjugation merely lets us have the overbar (row) operator effect in a non overbar (column) vector. In fact, the symbol $\nu_L\nu_L$ is used by some for the $\bar{\nu}_L{}^c\nu_L$ term of (5), with similar changes for other terms, where one must keep in mind for such notation that inner product in spinor space is implied, even though there is no obvious transpose term (row vector on left) in $\nu_L\nu_L$.

Note that the first term in (4) destroys a RH particle and creates a LH one. The Feynman diagram for this term shows a RH particle disappearing at a point and a LH particle appearing. Thus weak (chiral) charge is not conserved, as a LH neutrino has +1/2 weak charge and a RH neutrino has zero weak charge. Lepton number, however, is conserved, as we started with a neutrino (not an anti-neutrino) and ended up with a neutrino.

Somewhat similarly, the first term in (5) creates two LH neutrinos out of the vacuum and thus also does not conserve weak charge. But, importantly, it does not conserve lepton number (which the Dirac terms do.) We started with zero neutrinos and ended up with two neutrinos.

Wholeness Chart 1. Weak Charge and Lepton Number Conservation

	Dirac mass terms	Majorana mass terms
Conserves weak charge?	No	No
Conserves lepton number?	Yes	No

Mathematically, charge conjugation of the field, where C is the charge conjugation operator, can be expressed as

$$\nu^c = C\nu = i\gamma^2\nu^* \qquad \bar{\nu}^c = \nu^T C = \nu^T i\gamma^2 \ , \tag{6}$$

which needs some study in spinor space to fully understand, but doing so would lead us astray from the task at hand.

With all this in mind, we can then express (4) and (5) in terms of a mass matrix $\mathcal{M}$ as (where "*h.c.*" means hermitian conjugate of the prior term)

$$\mathcal{L}_{\substack{mass\\terms}} = -\frac{1}{2}\left(\bar{\nu}_L \bar{\nu}_R{}^c\right)\mathcal{M}\begin{pmatrix}\nu_L{}^c\\ \nu_R\end{pmatrix} + h.c. \tag{7}$$

with

$$\mathcal{M} = \begin{bmatrix} m_M^L & m_D \\ m_D & m_M^R \end{bmatrix} . \tag{8}$$

(As we will see, the matrix in (8) is the neutrino space analog of the matrix in (3) of Section 2.)

Hermitian conjugates of fields are as follows

$$\nu_L \xleftrightarrow{h.c.} \bar{\nu}_L \qquad \nu_R \xleftrightarrow{h.c.} \bar{\nu}_R \qquad \nu_L^c \xleftrightarrow{h.c.} \bar{\nu}_L^c \qquad \nu_R^c \xleftrightarrow{h.c.} \bar{\nu}_R^c ,$$

so (7) becomes (taking the complex conjugate transpose of the first term in (7) for the second)

$$\mathcal{L}_{\substack{mass\\terms}} = -\frac{1}{2}\left(\bar{\nu}_L \ \bar{\nu}_R{}^c\right)\mathcal{M}\begin{pmatrix}\nu_L{}^c\\ \nu_R\end{pmatrix} - \frac{1}{2}\left(\bar{\nu}_L{}^c \ \bar{\nu}_R\right)\mathcal{M}\begin{pmatrix}\nu_L\\ \nu_R{}^c\end{pmatrix} , \tag{9}$$

which, for our identification of the effects of ν_L^c and ν_R (both destroy RH particles and create LH antiparticles), $\bar{\nu}_R^c$ and $\bar{\nu}_L$, ν_R^c and ν_L, $\bar{\nu}_L^c$ and $\bar{\nu}_R$, yields (4) plus (5).

4 See-sawing

Suppose, as suggested earlier, that Higg's or GUT symmetry breaking only gave Majorana mass to neutrinos. That is, coupling to the Higgs (or Higges) was not done in a way that led to Dirac mass terms. So, the mass matrix would be diagonal, unlike (8), of form

$$\tilde{\mathcal{M}} = \begin{bmatrix} m_\nu & 0 \\ 0 & M \end{bmatrix}, \tag{10}$$

and our Lagrangian mass terms would look like

$$\mathcal{L}_{\substack{mass\\terms}} = -\frac{1}{2}\begin{pmatrix} \bar{\nu} & \bar{N} \end{pmatrix} \tilde{\mathcal{M}} \begin{pmatrix} \nu \\ N \end{pmatrix} + h.c., \tag{11}$$

where we have represented the fields directly coupled to the Higgs by $(\nu \;\; N)^T$. In other words, ν and N are the mass eigenstates for our neutrinos.

On the other hand, the weak eigenstates ν_L and ν_R (and their conjugates) of (7), which are linear superpositions of ν and N, interact directly via the weak force, and represent what we detect in weak interaction experiments (ignoring in this context the fact that ν_R has zero weak charge and does not so interact.)

Finding (10) from (8) is just an eigenvalue problem, with m_ν and M the eigenvalues. That is, we could think of our fields in two different, but essentially equivalent, ways: 1) a mix of Majorana and Dirac mass terms with the column vector of fields in (7), or 2) pure Majorana mass terms associated with the mass matrix of (10), whose associated fields are represented by the different column vector $(\nu \;\; N)^T$.

Heuristically, finding $(\nu \;\; N)^T$ from $(\nu_L^c \;\; \nu_R)^T$ can be thought of as "rotating" our basis vectors in an abstract space until we find an alignment giving the fields vector the components $(\nu \;\; N)^T$.

Assuming that is the case in the real world (we have no way of knowing via experiments to date), what would the mass matrix (10) look like in order to give us the kind of masses (either m_D or perhaps m_M^L) that we see? Remember we are looking for a reason why neutrino mass is so much lower than that of other particles.

That reason posits that the field components of the vector in (11) are the ones directly coupled to the Higgs field. It works best if the mass $m_\nu = 0$, as that means there is no Higgs coupling for the ν field, but there is such coupling for the N. (And (10) then becomes the analog of (1) in Section 2.) Note that if we took $m_\nu \neq 0$, but $m_\nu << M$, we would still be left with our original problem, which is "why is one mass so much smaller than the others?". Having zero mass is easier to explain (no coupling) than extremely low mass (extremely small coupling.)

4.1 The "Shortcut" Analysis

Given the treatment of Section 2, we can immediately draw conclusions about the magnitudes of the four components of (8), given (10) with the upper left component equal to zero and the $(\nu \;\; N)^T$ basis being close to the $(\nu_L^c \;\; \nu_R)^T$ basis. That is, we have the <u>mass hierarchy</u> we need,

$$M \approx m_M^R >> m_D > m_M^L \approx 0 \;, \tag{12}$$

where the Dirac mass m_D is the <u>geometric mean</u> of the left and right Majorana masses, the diagonal components of (8). That is,

$$m_M^R m_M^L = m_D{}^2 \;. \tag{13}$$

Note that for given value of m_D, a higher value for m_M^R means a lower the value for m_M^L and vice versa. This is the reason for the name "<u>see-saw mechanism</u>".

4.2 The Formal Eigenvalue Analysis

The simple "deduce by analogy" method of the prior section allows us to see, relatively easily, the essence of the see-saw mechanism. But to fully quantify it, we need the following more rigorous analysis.

The characteristic equation for the eigenvalue problem solution of (8) is

$$\left(m_M^L - \lambda\right)\left(m_M^R - \lambda\right) - \left(m_D\right)^2 = 0 \,, \tag{14}$$

with eigenvalues,

$$\lambda_{1,2} = \tfrac{1}{2}\left(m_M^R + m_M^L\right) \pm \tfrac{1}{2}\sqrt{\left(m_M^R + m_M^L\right)^2 - 4\left(m_M^R m_M^L - m_D{}^2\right)} \;. \tag{15}$$

For $\lambda_1 = m_\nu = 0$, we must have the minus sign in (15) and

$$m_M^R m_M^L = m_D{}^2 \;, \tag{16}$$

which, not surprisingly, is the same as (13).

Then, we would have, with the plus sign in (15) for $\lambda_2 = M$,

$$\begin{aligned} \lambda_1 &= m_\nu = 0 \\ \lambda_2 &= M = m_M^R + m_M^L \,. \end{aligned} \tag{17}$$

We'll work out the eigenvector N (i.e., for λ_2) expressed in the $(\nu_L{}^c\ \nu_R)^T$ basis and leave the simpler case ν eigenvector (i.e., for λ_1) for the reader.

From the eigenvalue problem for (7) and (8), with the eigenvalue λ_2 of (17) we get the two equations

$$\begin{aligned} \left(m_M^L - \left(m_M^R + m_M^L\right)\right)\nu_L^c \quad &+ \quad m_D \nu_R = 0 \\ m_D \nu_L^c \quad &+ \quad \left(m_M^R - \left(m_M^R + m_M^L\right)\right)\nu_R = 0 \,. \end{aligned} \tag{18}$$

This yields

$$\nu_L^c = \frac{m_D}{m_M^R}\nu_R \;, \tag{19}$$

and an eigenvector

$$N = \begin{bmatrix} \dfrac{m_D}{m_M^R}\nu_R \\ \nu_R \end{bmatrix} \;. \tag{20}$$

Some care is needed to note that the top component here is really the $\nu_L{}^c$ field with the fractional factor indicating the size of the $\nu_L{}^c$ field compared to the ν_R field. That is, N is really a superposition of the two fields, such that if ν_R has a coefficient of one in that superposition, then the $\nu_L{}^c$ field has a coefficient of m_D / m_M^R. In other words, in (18), the symbol ν_R really stands for the coefficient (effectively, the magnitude) of the ν_R field, not the field itself (which the location in the column vector denotes.)

Note also that, up to here, we have ignored the Hermitian conjugate half of (7), which we will have to include. So our true N will also include that, and is, in terms of the fields themselves, rather than as a two component vector, expressed as

$$N = \left(\nu_R + \nu_R^c\right) + \frac{m_D}{m_M^R}\left(\nu_L + \nu_L^c\right) \ . \tag{21}$$

Similarly, the other eigenvector is found to be

$$\nu = \left(\nu_L + \nu_L^c\right) - \frac{m_D}{m_M^R}\left(\nu_R + \nu_R^c\right) \ . \tag{22}$$

If we now assume (to be justified below)

$$m_M^R >> m_D \ , \tag{23}$$

then N is composed almost entirely of ν_R (and its similar sibling ν_R^c), from (17) and (23) is very heavy, and is thus effectively sterile. Conversely, ν_R can be thought of as composed almost entirely of N. Similarly, ν is composed almost entirely of ν_L (and ν_L^c), and conversely, ν_L is almost entirely composed of the weightless ν.

From (16), one sees that for a given value of m_D, a higher value for m_M^R means a lower the value for m_M^L, and vice versa, and thus, the name "see-saw mechanism". Further, from (21) and (22), the higher the value for m_M^R, the more $\nu_R \rightarrow N$ and $\nu_L \rightarrow \nu$.

Approached in a different way, given m_M^R and m_M^L, m_D will be the geometric mean of those two masses, and will generally be closer to the lower of the two. (If $m_M^R = 100$ and $m_M^L = 1$, then m_D = 10.) Further, if (23) holds, from (16), we have

$$m_M^L \approx 0 \quad \text{(but not 0)} \ , \tag{24}$$

and from (17),

$$m_M^R \approx M \ . \tag{25}$$

Thus, the mass hierarchy appears naturally as

$$M \approx m_M^R >> m_D > m_M^L \approx 0 \ . \tag{26}$$

These results match those of the simpler approach of Section 4.1.

<u>The $m_M^R >> m_D$ Assumption</u>

An astute reader, who hadn't read Sections 2 and 4.1, might question if we have gained anything. We originally sought a reason why the known Dirac mass m_D is so small compared to other masses. We got that via the eigenvalues analysis above, but in the process, we had to make another, seemingly arbitrary, assumption (23). With this assumption, we appear merely to be substituting one mass hierarchy problem for another. That is, we now have to ask why m_D turns out to be so much smaller than m_M^R.

The answer is this. If we start with the mass matrix (10) with one field having zero mass (uncoupled to Higgs particle(s)),

$$\tilde{\mathcal{M}} = \begin{bmatrix} 0 & 0 \\ 0 & M \end{bmatrix}, \tag{27}$$

and do a slight "rotation" in the 2D space of $(\nu \;\; N)^{\mathrm{T}}$, we end up with a matrix like (8) with the characteristic (23), which served as our initial assumption, but which is justified if we started with (27). Our assumption boils down simply to assuming a small transformation.

5 Distinction between Majorana Mass Terms, Particles, and Representation

The adjective "Majorana" is applied to three distinctly different things, which we need to distinguish between.

The first use most people see of this term is for one of three representations of Dirac matrices and spinors. The three representations are Dirac-Pauli (the Standard Rep), Weyl, and Majorana. As noted at the beginning, this use of "Majorana" has nothing to do with the Majorana mass terms of this article. Everything in this article can be done in any one of the three representations.

Herein, we so far have been dealing with the second use of the term regarding Majorana vs. Dirac type mass terms in the Lagrangian, i.e., (4) and (5). The neutrinos and Dirac matrices in these terms can be represented by any one of the three representations above.

The third use of the term refers to type of neutrino. A Majorana *particle* is defined as a particle that is its own antiparticle. A Dirac particle, on the other hand, has an antiparticle that is distinctly different from it. Typically, in almost all of one's study of QFT, one deals with Dirac type particles.

Neutrinos are the only particles that can be either Dirac or Majorana types. All other fermions are known, from experiment, to be Dirac fermions. No experiments to date (Jan 2012) have been able to determine if neutrinos are Majorana or Dirac particles. Double beta decay experiments may one day be able to do this.

As an aside, Majorana particles are easiest to handle mathematically in the Majorana representation.

Note that the neutrinos we deal with in our mass terms can be either Dirac or Majorana neutrinos, but both type mass terms would need to involve the same particle type. From (4) and (5), we see that the particles in each type term are represented by the same symbols, i.e., they represent the same particle type (Dirac or Majorana) in *both* type mass terms (Dirac and Majorana).

<u>In summary</u>, we can have

- Majorana representation in spinor space (it or one of other 2 reps can be used for any of below)
- Majorana vs Dirac mass terms in Lagrangian (both together can be used with either particle type below)
- Majorana vs Dirac type particles (Majorana is its own antiparticle)

6 Comments on Lepton Number Conservation

With regard to Majorana vs. Dirac type mass terms in the Lagrangian, we saw (see Wholeness Chart 1, pg. 3) how both types of terms do not conserve weak charge. We also saw that the Majorana mass terms lead to non-conservation of lepton number, whereas the Dirac mass terms lead to conservation of lepton number. These results were specifically for Dirac neutrinos in both types of mass term, where Dirac neutrinos have a lepton number +1, and Dirac antineutrinos have a lepton number of –1.

However, what if the neutrinos we are dealing with in experiment are actually Majorana neutrinos? Then neutrinos and anti-neutrinos would have the same lepton number, since they are the same particle. But this number would have to be its own negative, since quantum numbers for anti-particles have opposite sign of those for particles. Zero is the only number that works, so we could conclude that Majorana particles have lepton number zero.

Therefore, for Majorana neutrinos in both types of mass terms, all interactions solely from mass terms of either form will result in no change of lepton number. So, if we are dealing with Majorana neutrinos, the "No" we have in the last row, last column of Wholeness Chart 1 will change to a "Yes". Prior to this, we had been assuming we were working with Dirac neutrinos.

However, consider a typical interaction such as

$$n \rightarrow p + e^{-} + \bar{\nu} \qquad \bar{\nu} = \nu \text{ for Majorana neutrino} \tag{28}$$

where what we usually consider a Dirac anti-neutrino with lepton number –1, is now a Majorana neutrino with lepton number 0. Thus, we started with a neutron having zero lepton number, but end up with products having a net +1 lepton number (from the electron in (28)). We conclude that even though Majorana neutrinos in the Lagrangian mass terms (both Dirac and Majorana mass terms) will not lead to lepton number violation, interactions of Majorana neutrinos will.

Thus, we will have lepton number non-conservation for i) Dirac neutrinos if, and only if, Majorana mass terms exist in the Lagrangian or ii) Majorana fermions regardless of what mass terms are in the Lagrangian.

7 Possible Physical Scenarios

There are three possible scenarios, assuming both neutrino types exist.

Possibilities for both Dirac and Majorana neutrinos existing in nature

1) Dirac and Majorana fermions both interact weakly, and what we see in experiments is a blend of both. (Not considered likely by most.)

2) Only Dirac neutrinos interact weakly, and we don't ever see Majorana neutrinos in any experiments.

3) Only Majorana neutrinos interact weakly, and we don't ever see Dirac neutrinos in any experiments.

Possibilities if only one type exists in nature

4) Dirac neutrinos exist, but no Majorana ones.

5) Majorana neutrinos exist, but no Dirac ones.

If the See-Saw Mechanism is True

If the see-saw mechanism exists, then we have both type mass terms of (4) and (5), and with

$$M \approx m_M^R >> m_D > m_M^L \approx 0\,, \tag{29}$$

and for which we could have, in one scenario, Dirac neutrinos represented by ν_L and ν_R if only Dirac neutrinos exist. Alternatively, we could instead have Majorana neutrinos represented by those symbols. In either case, our interaction terms would include the symbols ν_L and ν_R, along with intermediate vector boson fields.

For m_D much larger than m_M^L, the m_D mass term would not play a role in the theory at energy levels of the present day. So we would effectively see neutrinos, be they Dirac or Majorana neutrinos, as having mass m_M^L, i.e., as having mass of the Majorana mass terms in $\mathcal{L}$.

8 Summary of See-Saw Mechanism

See-saw MechanismTheory

The (common textbook) treatment covered in Section 4.2 began with a general, non-diagonal mass matrix, looked at finding the mass eigenvalues of that matrix, and examined the relationships engendered between the masses. However, looking at it somewhat in reverse, as in Section 4.1, can be helpful pedagogically.

That is, start with the mass eigenstates fields ν and N, the ones coupled directly to the Higgs field (with ν having zero coupling), and the diagonal mass matrix (27). The weak eigenstates fields ν_L and ν_R (and their charge conjugation fields) are superpositions of the ν and N fields.

We then ask "If we transform $(\nu \;\; N)^{\mathrm{T}}$ into the $(\nu_L{}^{c} \;\; \nu_{\mathrm{R}})^{\mathrm{T}}$, what would the transformed mass matrix look like?" Well, if nature has chosen to make this a slight transformation (a small "rotation" in the 2D space of the fields), which is reasonable, then we would get a mass matrix with a very small upper left diagonal term m_M^L, a very large lower right diagonal term m_M^R, and off diagonal terms m_D which are each the geometric mean of the diagonal ones, as in (16). We would have a "see-saw" relation between the masses, and could readily have neutrino masses m_M^L of the order observed. For a greater "rotation" in 2D fields space, the greater would be the "see-saw" effect (bigger m_M^L and lower m_M^R), and also the greater the value of m_{D}.

The neutrinos we see in experiments could be either Dirac or Majorana types, though either type would have both forms for Majorana and Dirac mass terms in the Lagrangian.

For further pedagogic explanations of topics in quantum field theory by the same author, see www.quantumfieldtheory.info .