Theoretical Neutrino Physics

Lecture Notes

Joachim Kopp

November 4, 2015

Contents

1 Notation and conventions

Throughout this lecture, we work in *natural units*, i.e. we set $\hbar = c = 1$. We express all energies and momenta in eV, and all lengths and times in eV^{-1}.

Since $\hbar c = 197\,\mathrm{MeV\,fm}$, this implies in particular

$$1\,\mathrm{cm} = 5.076 \times 10^4\,\mathrm{eV}^{-1} \tag{1.1}$$

$$1\,\mathrm{sec} = 1.523 \times 10^{15}\,\mathrm{eV}^{-1}\,. \tag{1.2}$$

When dealing with fermions and the Dirac equation, we work in the chiral basis where

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \qquad \gamma^5 = \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix}. \tag{1.3}$$

2
Neutrinos in the Standard Model

2.1 Field theory recap

Field $\phi(\mathbf{x}, t)$: function that maps every spacetime point to a field amplitude.

The dynamics of the field are described by a *Lagrangian density* $\mathcal{L}(\phi, \partial_\mu \phi)$ (in anaology to the Lagrange function $L(x, \dot{x})$ in classical mechanics). Example: for a real scalar field:

$$\mathcal{L}(\phi, \partial_\mu \phi) \equiv \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2} m^2 \phi^2 \,. \tag{2.1}$$

The equations of motion are obtained from the principle of stationary action, which states that

$$\delta \mathcal{S} = 0 \,, \tag{2.2}$$

where the action $\mathcal{S}$ is defined as

$$\mathcal{S} \equiv \int d^4x \, \mathcal{L}(\phi, \partial_\mu \phi) \tag{2.3}$$

and $\delta \mathcal{S}$ in eq. (2.2) means the variation of $\mathcal{S}$ with respect to ϕ and $\partial_\mu \phi$. Thus,

$$\delta \mathcal{S} = \int d^4x \left[\frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi)} \delta(\partial_\mu \phi) + \frac{\delta \mathcal{L}}{\delta \phi} \delta \phi \right] \tag{2.4}$$

$$= \int d^4x \left[- \partial_\mu \frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi)} + \frac{\delta \mathcal{L}}{\delta \phi} \right] \delta \phi \,, \tag{2.5}$$

where in the last step we have integrated by parts. Since eq. (2.5) is required to be satisfied for any variation $\delta\phi$, the term in square brackets must vanish. This leads to the *Euler-Lagrange equations*

$$\partial_\mu \frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi)} - \frac{\delta \mathcal{L}}{\delta \phi} = 0 \,. \tag{2.6}$$

For the scalar field Lagrangian from eq. (2.1), they lead to the *Klein-Gordon equation*

$$\partial^\mu \partial_\mu \phi + m^2 \phi = 0\,. \tag{2.7}$$

Similarly, the Lagrangian for a fermion field $\psi(\mathbf{x}, t)$,

$$\mathcal{L}_{\text{Dirac}}(\bar{\psi}, \psi, \partial_\mu \bar{\psi}, \partial_\mu \psi) = i\bar{\psi}\partial\!\!\!/\psi - m\bar{\psi}\psi \tag{2.8}$$

leads to the Dirac equation

$$i\partial\!\!\!/\psi - m\psi = 0\,. \tag{2.9}$$

It is often useful to separate the left-chiral and right-chiral components of the 4-component spinor ψ:

$$\psi_L \equiv P_L \psi \equiv \frac{1-\gamma^5}{2}\psi \tag{2.10}$$

$$\psi_R \equiv P_R \psi \equiv \frac{1+\gamma^5}{2}\psi\,, \tag{2.11}$$

which can be considered as independent fields. It follows that

$$\psi = \psi_L + \psi_R\,. \tag{2.12}$$

If you are used to thinking in terms of 4-component spinors in the chiral basis, $\psi = (\chi_1, \chi_2, \xi_1, \xi_2)^T$, P_L projects out the upper two components: $\psi_L = (\chi_1, \chi_2, 0, 0)^T$ and P_R projects out the lower two components: $\psi_R = (0, 0, \xi_1, \xi_2)^T$. Using the properties

$$P_L^2 = P_L\,, \qquad P_R^2 = P_R\,, \qquad P_L P_R = P_R P_L = 0\,, \tag{2.13}$$

te Lagrangian eq. (2.8) can be rewritten as

$$\mathcal{L}_{\text{Dirac}} = i\overline{(\psi_L + \psi_R)}\partial\!\!\!/(\psi_L + \psi_R) - m\overline{(\psi_L + \psi_R)}(\psi_L + \psi_R) \tag{2.14}$$

$$= i\bar{\psi}_L \partial\!\!\!/\psi_L + i\bar{\psi}_R \partial\!\!\!/\psi_R - m\bar{\psi}_L\psi_R - m\bar{\psi}_R\psi_L \tag{2.15}$$

Finally, the Lagrangian of quantum electrodynamics (a fermion field ψ coupled to a gauge boson field A_μ) is

$$\mathcal{L}_{\text{QED}} = i\bar{\psi}\partial\!\!\!/\psi - m\bar{\psi}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + e\bar{\psi}\gamma^\mu\psi A_\mu\,. \tag{2.16}$$

It contains the kinetic term for the fermion, the fermion mass term, the kinetic term for the gauge boson (with the electromagnetic field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$), and the gauge coupling term. The Euler-Lagrange equation for the gauge boson is just the inhomogeneous Maxwell equation in covariant formulation

$$\partial_\mu F^{\mu\nu} = -e\bar{\psi}\gamma^\nu\psi\,. \tag{2.17}$$

(The homogeneous Maxwell equation $\partial_{[\alpha}F_{\beta\gamma]} = \partial_\alpha(\epsilon_{\beta\gamma\rho\tau}F^{\rho\tau}) = 0$ is automatically satisfied.)

Quarks	u	c	t	$Q = +\frac{2}{3}$
	d	s	b	$Q = -\frac{1}{3}$
Leptons	ν_e	ν_μ	ν_τ	$Q = 0$
	e	μ	τ	$Q = -1$

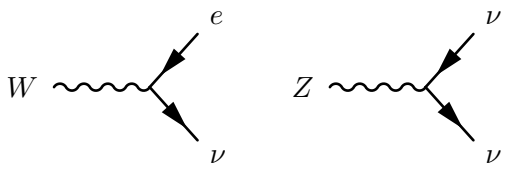
Figure 2.1: The elementary particle zoo

2.2 Neutrino masses and mixings

As shown in fig. 2.1, the three flavors of neutrinos, ν_e, ν_μ, ν_τ complete the zoo of elemetary particles of the Standard Model. Every neutrino is the partner of a charged lepton (electron, muon, tau), connected to it by the weak interaction:

$$\mathcal{L} = \sum_{\alpha=e,\mu,\tau} \left[\bar{\nu}_\alpha i\partial\!\!\!/ \nu_\alpha + \frac{g}{\sqrt{2}} \left(\overline{\nu_{\alpha,L}} \gamma^\mu e_{\alpha,L} W^+_\mu + h.c. \right) + \frac{g}{2\cos\theta_w} \overline{\nu_{\alpha,L}} \gamma^\mu \nu_{\alpha,L} Z_\mu \right] - \sum_{\alpha,\beta=e,\mu,\tau} \left(m_{\alpha\beta}\, \overline{\nu_{\alpha,L}} \nu_{\beta,R} + h.c. \right) \Bigg] . \tag{2.18}$$

Here, g is the weak coupling constant and θ_w is the Weinberg angle. Note that only left-handed neutrinos couple to the weak gauge bosons $W^\pm$ and Z. In terms of Feynman diagrams, the neutrino interaction vertices can be written as



Note that the mass term $\sum_{\alpha,\beta=e,\mu,\tau} m_{\alpha\beta}\, \overline{\nu_{\alpha,L}} \nu_{\beta,R}$ in eq. (2.18) is in general *off-diagonal* (i.e. $m_{\alpha\beta}$ can be non-zero even if $\alpha \neq \beta$). This means that the *flavor eigenstates* or *interaction eigenstates* ν_α ($\alpha = e, \mu, \tau$) do not have a definite mass.

The mass matrix m can be diagonalized according to

$$m = U m_D V^\dagger , \tag{2.19}$$

where $m_D = \mathrm{diag}(m_1, m_2, m_3)$ is a diagonal matrix and U, V are unitary matrices. We define the neutrino *mass eigenstates* according to

$$\nu_{j,L} \equiv \sum_\alpha U^*_{\alpha j} \nu_{\alpha,L} \tag{2.20}$$

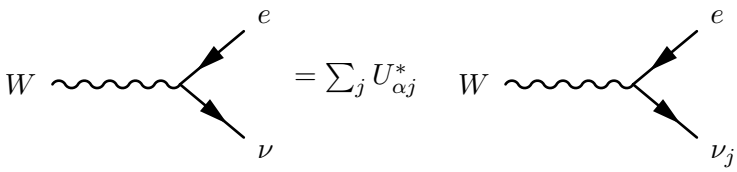
$$\nu_{j,R} \equiv \sum_\alpha V^*_{\alpha j} \nu_{\alpha,R} . \tag{2.21}$$

Figure 2.2: Illustration of why a Dirac mass term makes right handed neutrinos physical. Starting with a left-handed fermion, one can always perform a large boost, turning the particle into a right-handed statet.

In terms of the mass eigenstates, the Lagrangian (2.18) can be written as

$$\mathcal{L} = \sum_{j=1,2,3} \left[\bar{\nu}_j i\partial\!\!\!/ \nu_j + \frac{g}{\sqrt{2}} \left(\overline{\nu_{j,L}} U^*_{\alpha j} \gamma^\mu e_{\alpha,L} W^+_\mu + h.c. \right) + \frac{g}{2\cos\theta_w} \overline{\nu_{j,L}} \gamma^\mu \nu_{j,L} Z_\mu \right] - \sum_{j=1,2,3} \left(m_j \, \overline{\nu_{j,L}} \nu_{j,R} + h.c. \right) \Bigg] . \tag{2.22}$$

Thus, a charged current neutrino interaction produces a superposition of mass eigenstates, for instance



We will discuss neutrino mixing in much greater detail when we talk about neutrino oscillations. For now, let us focus on the mass term in the Lagrangian.

2.3 Dirac neutrino masses

Note that, without the mass term, RH neutrinos would be unphysical: they do not couple to any o the SM interactions and therefore cannot be produced in any particle reaction. The (Dirac) mass term

$$\mathcal{L}_{\text{Dirac}} \supset -m \, \overline{\nu_L} \nu_R + h.c. \, , \tag{2.23}$$

however, makes them physical because it couples left- and right-handed fields. An intuitive way of understanding this is by noting that, for a massive fermion which is left-handed in a given reference frame, one can always perform a boost along its direction of travel to a frame where it is right-handed. This is because the spin is invariant under such boosts, while the direction of the momentum vector can be reversed if the boost is large enough, see fig. 2.2.

2.4 Majorana neutrino masses

In the previous section, we have noted that fermion mass terms couple left handed and right handed fields. Since we know that the antiparticle of a left-handed neutrino is a right-handed field, we may ask the question whether the right handed neutrino in eq. (2.23) can be replaced by the antineutrino of the left handed neutrino.

To do this in a consistent way, we need to introduce the *charge conjugation* operation

$$\hat{\mathcal{C}} : \psi \to \psi^c \equiv -C\bar{\psi}^T \equiv -i\gamma^2\gamma^0\bar{\psi}^T = -i\gamma^2\psi^* \,. \tag{2.24}$$

Its effect on chirality is

$$\gamma^5\psi^c = -i\gamma^5\gamma^2\psi^* = +i\gamma^2\gamma^5\psi^* = +i\gamma^2(\gamma^5\psi)^* = -(\gamma^5\psi)^c \,, \tag{2.25}$$

i.e. the chirality of ψ^c is the opposite of the chirality of ψ. In other words, $\hat{\mathcal{C}}$ transforms left-handed states into right-handed states and vice-versa.

Some properties of the charge conjugation operation that will be useful below include

$$(\psi^c)^c = -i\gamma^2(-i\gamma^2\psi^*)^* = \psi \tag{2.26}$$

and

$$\begin{aligned}
\bar{\psi}\,\chi^c &= \psi^\dagger\gamma^0(-i\gamma^2\chi^*) \\
&= -i(\psi^*)^T\gamma^0\gamma^2\chi^* \\
&= +i(\psi^*)^T\gamma^2\gamma^0\chi^* \\
&= -i\chi^\dagger(\gamma^0)^T(\gamma^2)^T\psi^* \\
&= \bar{\chi}(-i\gamma^2\psi^*) \\
&= \bar{\chi}\psi^c \,.
\end{aligned} \tag{2.27}$$

Similarly,

$$\overline{\psi^c}\,\chi = \overline{\chi^c}\,\psi \,. \tag{2.28}$$

Moreover, we sometimes need the relation

$$\overline{\psi^c} = (-i\gamma^2\psi^*)^\dagger\gamma^0 = i[(\gamma^2)^*\psi]^T\gamma^0 = -i[\gamma^0\gamma^2\psi]^T \,. \tag{2.29}$$

Since $\hat{\mathcal{C}}$ transforms left-handed states into right-handed ones, one can hypothesize that

$$\nu_R \equiv (\nu_L)^c \,. \tag{2.30}$$

(We *cannot* simply define $\nu_R = \nu_L^*$ because ν_L^* would not be a right-handed field.) In 4-component notation, writing the spinor $\nu_L = (\chi, 0)^T$, where χ and 0 are two-component spinors):

$$\nu_R = -i\begin{pmatrix} 0 & \sigma^2 \\ -\sigma^2 & 0 \end{pmatrix}\begin{pmatrix} \chi^* \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ i\sigma^2\chi^* \end{pmatrix} \,. \tag{2.31}$$

This leads to a *new type of mass term*

$$\mathcal{L}_{\text{Majorana}} \supset -\frac{1}{2} m \, \overline{(\nu_L)^c} \, \nu_L + h.c. \tag{2.32}$$

$$= -\frac{1}{2} m \, \overline{(\nu_L)^c} \, \nu_L - \frac{1}{2} m^\dagger \overline{\nu_L} (\nu_L)^c \, . \tag{2.33}$$

In the second line, we have written out the Hermitian conjugate (*h.c.*) contribution explicitly.

Note that a Majorana mass term cannot be written down for any SM field except the neutrino because it would violate electric charge conservation. (For a charged fermion, f_L and $(f_L)^c$ carry opposite charge, so $\overline{(f_L)^c}$ carries the same charge as f_L.)

Even for neutrinos, one still has to think about ways of obtaining eq. (2.33) from an $SU(2)$-invariant theory. Remember that the left-handed neutrinos are in the same $SU(2)$ doublet as the charged leptons, so without the breaking of $SU(2)$, any term that exists for neutrinos must also exist for charged leptons. For the latter, however, a Majorana mass term is forbidden because they are charged.

Moreover, it is intriguing (and unexplained) that the neutrino masses are at least 6 orders of magnitude smaller than any other fermion masses in the SM. This may suggest that their mass terms are somehow special.

2.5 The seesaw mechanism

Consider again the Dirac mass term for a single species of neutrinos:

$$\mathcal{L}_{\text{Dirac}} \supset -m \, \overline{\nu_L} \nu_R + h.c. \, . \tag{2.34}$$

As noted above, ν_L and ν_R are independent degrees of freedom in this case.

The Dirac mass term in consistent with $SU(2)$-invariance, while a Majorana mass term for left-handed neutrinos would not be. A Majorana mass term for ν_R, on the other hand, is allowed. Including it is one of the main ideas behind the seesaw mechanism:

$$\mathcal{L}_{\text{seesaw}} \supset -m_D \, \overline{\nu_R} \nu_L - m_D \, \overline{\nu_L} \nu_R - \frac{1}{2} m_M \overline{(\nu_R)^c} \nu_R - \frac{1}{2} m_M \overline{\nu_R} (\nu_R)^c \tag{2.35}$$

$$= -\frac{1}{2} \overline{n^c} M n + h.c. \, , \tag{2.36}$$

where in the second line we have defined

$$n = \begin{pmatrix} \nu_L \\ (\nu_R)^c \end{pmatrix} , \qquad M = \begin{pmatrix} 0 & m_D \\ m_D & m_M \end{pmatrix} , \tag{2.37}$$

and used the properties that $(\nu^c)^c = \nu$ (see eq. (2.26)) and that $\overline{(\nu_R)^c} \, (\nu_L)^c = \bar{\nu}_L \nu_R$ (see eq. (2.27)).

Let us now diagonalize the mass matrix M by doing a transformation

$$\begin{pmatrix} \nu_L \\ (\nu_R)^c \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \chi_{1L} \\ \chi_{2L} \end{pmatrix} . \tag{2.38}$$

By computing

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 0 & m_D \\ m_D & m_M \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{2.39}$$

and requiring that the off-diagonal elements vanish, we find for the mixing angle

$$\tan 2\theta = \frac{2m_D}{m_M}\,. \tag{2.40}$$

The eigenvalues of M, on the other hand, are

$$m_{1,2} = \frac{m_M}{2} \mp \sqrt{\frac{m_M^2}{4} + m_D^2}\,. \tag{2.41}$$

The diagonalized mass term then reads

$$\mathcal{L}_{\text{seesaw}} \supset -\frac{1}{2} m_1 \overline{(\chi_{1L})^c}\, \chi_{1L} - \frac{1}{2} m_2 \overline{(\chi_{2L})^c}\, \chi_{2L} + h.c.\,, \tag{2.42}$$

i.e. it corresponds to two Majorana mass terms for the two Majorana fermion field χ_{1L} and χ_{2L}.

In the limiting case $m_M = 0$ (pure Dirac mass), we recover the eigenvalues $\pm m_D$. (The minus sign on the second can be removed by a field redefinition $\chi_{1L} \to i\chi_{1L}$.) The mixing angle in this case becomes $\theta = \pi/4$, and the mass eigenstates are

$$\chi_{1L} = \frac{1}{\sqrt{2}}(\nu_L + (\nu_R)^c)\,, \tag{2.43}$$

$$\chi_{2L} = \frac{1}{\sqrt{2}}(\nu_L - (\nu_R)^c)\,. \tag{2.44}$$

This demonstrates that a Dirac fermion can be viewed as being composed of two Majorana fermions with identical mass.

More interesting to us here is the case $\boxed{m_M \gg m_D}$. In this case, there is one Majorana neutrino with a very small mass

$$\boxed{m_1 \simeq \frac{m_D^2}{m_M}\,,} \tag{2.45}$$

and one very heavy Majorana neutrino with a mass of order m_M. The light Majorana neutrino is

$$\chi_{1L} \simeq \nu_L + \frac{m_D}{m_M}(\nu_R)^c\,, \tag{2.46}$$

i.e. it is almost identical to the SM neutrinos. (The small admixture of ν_R is not relevant experimentally.) In practice, it is natural (whatever that means) to assume that $m_D \sim 100$ GeV because this mass term comes from the Higgs mechanism, and the Higgs vev is 246 GeV. m_M, on the other hand, is completely arbitrary, so one can assume that it is very large. For $m_M \sim 10^{14}$ GeV, we then find $m_1 \sim 0.1$ eV.

3
Neutrino oscillations

3.1 Quantum mechanics of neutrino oscillation

Since neutrino flavor eigenstates—the states that are produced by the weak interaction—are superpositions of mass eigenstates—the states with well-defined kinematics and propagators—we expect quantum interference effects in neutrino experiments. These effects are the neutrino oscillations.

We start again from the weak interaction Lagrangian, written in the mass basis

$$\mathcal{L} = \sum_{j=1,2,3} \left[\bar{\nu}_j i\partial\!\!\!/ \nu_j + \frac{g}{\sqrt{2}} \left(\overline{\nu_{j,L}} U^*_{\alpha j} \gamma^\mu e_{\alpha,L} W^+_\mu + h.c. \right) + \frac{g}{2\cos\theta_w} \overline{\nu_{j,L}} \gamma^\mu \nu_{j,L} Z_\mu \right] - \sum_{j=1,2,3} \left(m_j \, \overline{\nu_{j,L}} \nu_{j,R} + h.c. \right) \Bigg] . \tag{3.1}$$

This implies that the neutrino state of flavor α ($\alpha = e,\ \mu,\ \tau$) produced in a weak interaction can be written as the following superposition of mass eigenstates

$$|\nu_\alpha\rangle = \sum_j U^*_{\alpha j} |\nu_j\rangle \, . \tag{3.2}$$

Note that even though the transformation of the field operators is $\nu_{\alpha,L} = \sum_j U_{\alpha j} \nu_{j,L}$, the transformation of the ket-states is determined by $U^\dagger$ rather than U. The reason is that these states are produced by the creation operator $\sum_j U^*_{\alpha j} \bar{\nu}_{j,L}$, not by the annihilation opearators $\sum_j U_{\alpha j} \nu_{j,L}$. Treating $|\nu_j\rangle$ as plane wave states, the wave function at a distance L from the production point, and at a time T after production, is given by

$$|\nu_\alpha(T, L)\rangle = \sum_j U^*_{\alpha j} e^{-iE_j T + ip_j L} |\nu_j\rangle \, . \tag{3.3}$$

Note that the energy E_j and the momentum p_j are in general different for the different mass eigenstates because the kinematics of the production process is different for different mass.

A neutrino detector measures the neutrino flavor, i.e. it detects the neutrino in a state

$$\langle \nu_\beta | = \sum_j U_{\beta j} \langle \nu_j | \,. \tag{3.4}$$

Therefore, the amplitude for a neutrino produced as $|\nu_\alpha\rangle$ to be detected as $\langle \nu_\beta|$ is

$$\langle \nu_\beta | \nu_\alpha(T, L)\rangle = \sum_{j,k} U^*_{\alpha j} U_{\beta k} e^{-iE_j T + ip_j L} \langle \nu_k | \nu_j \rangle \tag{3.5}$$

$$= \sum_j U^*_{\alpha j} U_{\beta j} e^{-iE_j T + ip_j L} \,. \tag{3.6}$$

The *oscillation probability* is thus

$$P_{\alpha\beta}(T, L) = |\, \langle \nu_\beta | \nu_\alpha(T, L)\rangle \,|^2 = \sum_{j,k} U^*_{\alpha j} U_{\beta j} U_{\alpha k} U^*_{\beta k} e^{-i(E_j - E_k)T + i(p_j - p_k)L} \,. \tag{3.7}$$

In a typical neutrino oscillation experiment, we do not know when precisely each neutrino is produced (the experimental uncertainty in the production time is much larger than the energy uncertainty of each individual neutrino). Therefore, we should integrate over T:

$$\begin{aligned} P_{\alpha\beta}(L) &= \frac{1}{N} \int dT \, |P_{\alpha\beta}(T, L)|^2 \\ &= \frac{1}{N} \sum_{j,k} U^*_{\alpha j} U_{\beta j} U_{\alpha k} U^*_{\beta k} \exp\left[i \Big(\sqrt{E^2 - m_j^2} - \sqrt{E^2 - m_k^2} \Big) L \right] \delta(E_j - E_k) \\ &\simeq \sum_{j,k} U^*_{\alpha j} U_{\beta j} U_{\alpha k} U^*_{\beta k} \exp\left[- i \frac{\Delta m^2_{jk} L}{2E} \right] . \end{aligned} \tag{3.8}$$

Here, N is a normalization constant, which is chosen such that $\sum_\beta P_{\alpha\beta}(L) = 1$. In the last line of eq. (3.8), we have made the approximation $|m_j - m_k| \ll E = E_j = E_k$ and carried out a Taylor expansion in the mass squared difference

$$\Delta m^2_{jk} \equiv m_j^2 - m_k^2 \,. \tag{3.9}$$

We could also have made the (somewhat unjustified) assumption that all neutrino mass eigenstates are emitted with the same momentum p, but different energies. This would have led to the same result, but with phase factor $\exp[-i\Delta m^2_{jk} T/(2E)]$ instead of $\exp[-i\Delta m^2_{jk} L/(2E)]$. Since neutrino travel at the speed of light (up to negligible corrections of order $\Delta m^2_{jk}/E^2$), we can set $L = T$, so that the two approaches become completely equivalent.

The expression for $P_{\alpha\beta}(L)$ becomes particularly simple in the 2-flavor approximation, where the mixing matrix U can be written as

$$U = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} . \tag{3.10}$$

For instance, if the two flavors are e and μ, we obtain

$$\begin{aligned} P_{e\mu}^{\text{2-flavor}}(L) &= |U_{e1}|^2|U_{\mu 1}|^2 + |U_{e2}|^2|U_{\mu 2}|^2 \\ &\quad + U_{e1}U_{\mu 1}U_{e2}U_{\mu 2}\left[\exp\left(-i\frac{\Delta m^2 L}{2E}\right) + \exp\left(+i\frac{\Delta m^2 L}{2E}\right)\right] && (3.11) \\ &= 2\cos^2\theta\sin^2\theta - 2\cos^2\theta\sin^2\theta\cos\left[\frac{\Delta m^2 L}{2E}\right] && (3.12) \\ &= \frac{1}{2}\sin^2 2\theta\left(1 - \cos\left[\frac{\Delta m^2 L}{2E}\right]\right) && (3.13) \\ &= \sin^2 2\theta\,\sin^2\left[\frac{\Delta m^2 L}{4E}\right] . && (3.14) \end{aligned}$$

Several comments are in order here:

- States with different energy and momentum (E_j, p_j), $j = 1, 2, 3$ can interfere only if the energy and momentum uncertainties associated with the production and detection processes are larger than $|E_j - E_k|$, $|p_j - p_k|$. This is always satisfied. To demonstrate this, let us compute E_j, p_j for a specific case: neutrinos from $\pi^\pm$ decay at rest,

$$\pi^\pm \to \mu^\pm + \overset{(-)}{\nu}_{1,2,3} . \tag{3.15}$$

  In the pion rest frame,

$$\begin{aligned} m_\pi^2 = (p_\mu + p_{\nu_j})^2 &= m_\mu^2 + m_j^2 + 2(E_\mu E_j + \mathbf{p}_\mu \mathbf{p}_{\nu_j}) && (3.16) \\ &= m_\mu^2 + m_j^2 + 2\Big(\sqrt{|\mathbf{p}_{\nu_j}|^2 + m_\mu^2}\sqrt{|\mathbf{p}_{\nu_j}|^2 + m_j^2} + |\mathbf{p}_{\nu_j}|^2\Big) . && (3.17) \end{aligned}$$

  Solving for $|\mathbf{p}_{\nu_j}|$ gives

$$\begin{aligned} 4\big(|\mathbf{p}_{\nu_j}|^2 + m_\mu^2\big)\big(|\mathbf{p}_{\nu_j}|^2 + m_j^2\big) &= (m_\pi^2 - m_\mu^2 - m_j^2 - 2|\mathbf{p}_{\nu_j}|^2)^2 && (3.18) \\ 4|\mathbf{p}_{\nu_j}|^2(m_\mu^2 + m_j^2 + m_\pi^2 - m_\mu^2 - m_j^2) &= (m_\pi^2 - m_\mu^2 - m_j^2)^2 - 4m_\mu^2 m_j^2 , && (3.19) \end{aligned}$$

  and thus

$$\begin{aligned} |\mathbf{p}_{\nu_j}|^2 &= \frac{m_\pi^2}{4}\left(1 - \frac{m_\mu^2}{m_\pi^2}\right)^2 - \frac{m_j^2}{2}\left(1 + \frac{m_\mu^2}{m_\pi^2}\right) + \frac{m_j^4}{4m_\pi^2} && (3.20) \\ E_{\nu_j}^2 &= \frac{m_\pi^2}{4}\left(1 - \frac{m_\mu^2}{m_\pi^2}\right)^2 + \frac{m_j^2}{2}\left(1 - \frac{m_\mu^2}{m_\pi^2}\right) + \frac{m_j^4}{4m_\pi^2} . && (3.21) \end{aligned}$$

To lowest order in m_j^2, we can write

$$E_{\nu_j} = E_0 + \xi \frac{m_j^2}{2E}\,, \tag{3.22}$$

$$|\mathbf{p}_{\nu_j}| = p_0 - (1-\xi)\frac{m_j^2}{2E}\,, \tag{3.23}$$

where E_0, p_0 are the energy and momentum that a massless neutrino ($m_j = 0$) would have, and $\xi = (1 - m_\mu^2/m_\pi^2)/2 \sim 0.25$. Therefore, even for unrealistically large neutrino masses of order 1 eV, we have typically $|E_{\nu_j} - E_{\nu_k}| \sim \big||\mathbf{p}_{\nu_j}| - |\mathbf{p}_{\nu_j}|\big| \sim 10^{-8}$ eV.

The typical momentum uncertainty associated with a neutrino production process is at least of the order of an inverse interatomic distance, i.e. of order keV. Therefore, the interference conditions for different neutrino mass eigenstates is easily satisfied.

- A related point: since the energy and momentum uncertainties are so important for interference to happen, it is not correct to treat neutrinos as plane waves. A wave packet formalism is more appropriate.
- The approximation $\sqrt{E^2 - m_j^2} - \sqrt{E^2 - m_k^2} \simeq \Delta m_{jk}^2/(2E)$ does not require $m_j, m_k \ll E$, but only $m_j^2 - m_k^2 \ll E$.
- For anitneutrinos, the above derivation goes through in exactly the same way, except that U should be replaced by U^* everywhere. This is because an antineutrino is created by the field operator ν rather than the operator $\bar{\nu}$, and the corresponding weak interaction term in the Lagrangian (3.1) is the hermitian conjugate of the term creating neutrinos. We denote oscillation probabilities for $\bar{\nu}_\alpha \to \bar{\nu}_\beta$ transitions by $P_{\bar{\alpha}\bar{\beta}}(L)$.

3.2 3-flavor neutrino oscillations

Let us consider the 3-flavor neutrino mixing matrix U:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}\,, \tag{3.24}$$

and count the number of parameters it has:

General 3×3 matrix	9 real parameters	9 complex phases
Unitarity:		
$\sum_j \lvert U_{\alpha j}\rvert^2 = 1$	-3	
$\sum_j U_{\alpha j} U^*_{\beta j} \overset{\alpha\neq\beta}{=} 0$	-3	-3
Field redfinitions:		
$\nu_j \to e^{i\phi_j}\nu_j$		-5
$\nu_\alpha \to e^{i\phi_\alpha}\nu_\alpha$		
$\sum$	3	1

Note that, even though there are 6 independent ways of rephasing the fields, only 5 complex phases from U can be absorbed this way. The reason is that applying a universal phase factor to all field (all ν_j and all ν_α) leaves U unchanged, as can be directly seen from eq. (3.24). So, in total there will be 4 independent physical parameters: 3 real ones (mixing angles θ_{12}, θ_{13}, θ_{23}) and one complex phase (δ) that leads to CP violation (see below).

By convention, the mixing matrix is parameterized as

$$U = \begin{pmatrix} 1 & & \\ & c_{23} & s_{23} \\ & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & & s_{13}e^{-i\delta} \\ & 1 & \\ -s_{13}e^{i\delta} & & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & \\ -s_{12} & c_{12} & \\ & & 1 \end{pmatrix} \tag{3.25}$$

$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \tag{3.26}$$

Here, $c_{ij} \equiv \cos\theta_{ij}$ and $s_{ij} \equiv \sin\theta_{ij}$. In the literature, the 3-flavor neutrino mixing matrix is often called MNS (Maki-Nakagawa-Sakata) matrix or PMNS (Pontecorvo-Maki-Nakagawa-Sakata) matrix.

3.2.1 2-flavor limits

It is clear that the expressions for the oscillation probabilities in the full 3-flavor case are quite complicated, and a numerical evaluation is usually necessary. Nevertheless, in a few special cases, a 2-flavor analysis can get us quite far, thanks the specific values of the oscillation parameters chosen by nature. We will now discuss three of these special cases. In all cases, our starting point is the general expression for $P_{\alpha\beta}(L)$:

$$P_{\alpha\beta}(L) \simeq \sum_{j,k} U^*_{\alpha j} U_{\beta j} U_{\alpha k} U^*_{\beta k} \exp\left[-i \frac{\Delta m^2_{jk} L}{2E} \right]. \tag{3.27}$$

In matrix form, it can be written also as

$$P_{\alpha\beta}(L) \simeq \left| U \begin{pmatrix} \exp[-i\frac{m_1^2 L}{2E}] & & \\ & \exp[-i\frac{m_2^2 L}{2E}] & \\ & & \exp[-i\frac{m_3^2 L}{2E}] \end{pmatrix} U^\dagger \right|^2 . \tag{3.28}$$

We can always pull out a matrix proportional to the identity matrix, therefore, this is equivalent to

$$P_{\alpha\beta}(L) \simeq \left| U \begin{pmatrix} 1 & & \\ & \exp[-i\frac{\Delta m_{21}^2 L}{2E}] & \\ & & \exp[-i\frac{\Delta m_{31}^2}{2E}] \end{pmatrix} U^\dagger \right|^2 . \tag{3.29}$$

1. **"Solar oscillations."** Consider ν_e disappearance (i.e. $P_{ee}(L)$), which are relevant if we have a pure ν_e or $\bar{\nu}_e$ source (such as the Sun or a nuclear reactor) and a detector sensitive only to ν_e or $\bar{\nu}_e$ (like a detector relying on some nuclear process, e.g. $\nu_e + {}^{37}\mathrm{Cl} \to {}^{37}\mathrm{Ar} + e^-$). We use the fact that θ_{13} is turns out to be numerically very small and therefore negligible in this context. Moreover, let us define

$$\begin{pmatrix} \nu_x \\ \nu_y \end{pmatrix} \equiv \begin{pmatrix} c_{23} & -s_{23} \\ s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \nu_\tau \end{pmatrix} . \tag{3.30}$$

   In this new basis, we have (neglecting θ_{13})

$$\begin{pmatrix} \nu_e \\ \nu_x \\ \nu_y \end{pmatrix} = \begin{pmatrix} c_{12} & s_{12} & \\ -s_{12} & c_{12} & \\ & & 1 \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} . \tag{3.31}$$

   In other words, the state ν_y does not participate in oscillations in the approximation $\theta_{13} = 0$, and oscillations can be described in a two-flavor framework involving only ν_e and ν_x, with mixing angle θ_{12}.

2. **"Atmospheric oscillations."** Neutrino oscillation experiments tell us that $\Delta m_{21}^2 \ll \Delta m_{31}^2$. Looking at eq. (3.27) or eq. (3.29), we see that for this reason, the oscillating exponentials involving Δm_{21}^2 can be set to 1 as long as $L \ll 2E/\Delta m_{21}^2$. This condition is usually satisfied in accelerator experiments. Therefore, the 12-rotation matrix which is part of U (see eq. (3.25) can be commuted past the diagonal matrix with the complex exponentials in eq. (3.29) and cancels against the corresponding piece from $U^\dagger$. Moreover, we will again set $\theta_{13} = 0$. Then,

$$P_{\alpha\beta}(L) \simeq \left| \begin{pmatrix} 1 & & \\ & c_{23} & s_{23} \\ & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} 1 & & \\ & \exp[-i\frac{\Delta m_{21}^2 L}{2E}] & \\ & & \exp[-i\frac{\Delta m_{31}^2}{2E}] \end{pmatrix} \begin{pmatrix} 1 & & \\ & c_{23} & -s_{23} \\ & s_{23} & c_{23} \end{pmatrix} \right|^2 , \tag{3.32}$$

   i.e. we're back to a two-flavor problem.

3. **Reactor neutrinos.** Reactor neutrino experiments measure $P_{\bar{e}\bar{e}}(L)$, i.e. $\bar{\nu}_e \to \bar{\nu}_e$ oscillations. Most of them are carried out at a relatively short baseline, which satisfies $\Delta m^2_{21} L/(2E) \ll 1$. Argue that these experiments are sensitive to the mixing angle θ_{13}, and that their data can be approximately analyzed in a two-flavor framework.

3.2.2 CP violation in neutrino oscillations

We have argued that the leptonic mixing matrix contains a complex phase δ, which violates the particle–antiparticle symmetry CP. This effect cannot be seen in a 2-flavor approximation. (The reason is that, by repeating the argument given below eq. (3.24) in the 2-flavor case, we obtain *no* complex phase.) In the 3-flavor case, however, we can show that in general

$$\Delta P_{\alpha\beta}(L) \equiv P_{\alpha\beta}(L) \neq P_{\bar{\alpha}\bar{\beta}}(L)\,. \tag{3.33}$$

Using the parameterization (3.26) for the leptonic mixing matrix, we can derive from the general expression for the oscillation probability, eq. (3.27),

$$\Delta P_{\alpha\beta}(L) \simeq \sum_{j,k} U^*_{\alpha j} U_{\beta j} U_{\alpha k} U^*_{\beta k} \exp\left[-i\frac{\Delta m^2_{jk} L}{2E}\right] - \sum_{j,k} U_{\alpha j} U^*_{\beta j} U^*_{\alpha k} U_{\beta k} \exp\left[-i\frac{\Delta m^2_{jk} L}{2E}\right] \tag{3.34}$$

$$= \sum_{j\neq k} \left(U^*_{\alpha j} U_{\beta j} U_{\alpha k} U^*_{\beta k} - U_{\alpha j} U^*_{\beta j} U^*_{\alpha k} U_{\beta k}\right) \exp\left[-i\frac{\Delta m^2_{jk} L}{2E}\right] \tag{3.35}$$

$$= \sum_{j\neq k} 2i\,\mathrm{Im}\left(U^*_{\alpha j} U_{\beta j} U_{\alpha k} U^*_{\beta k}\right) \exp\left[-i\frac{\Delta m^2_{jk} L}{2E}\right] \tag{3.36}$$

$$\tag{3.37}$$

We immediately see that for $\alpha = \beta$, this expression vanishes. This makes perfect sense: for $\alpha = \beta$, the oscillation process is identical to the T-reversed (time-reversed) process, i.e. the process where the initial and final states are interchanged. Thus, T symmetry cannot be violated, and therefore, due to CPT invariance, also CP cannot violated.

Let us consider in more detail the quantities

$$J^{jk}_{\alpha\beta} \equiv \mathrm{Im}\left(U^*_{\alpha j} U_{\beta j} U_{\alpha k} U^*_{\beta k}\right), \qquad \alpha \neq \beta,\ j \neq k \tag{3.38}$$

which measure the amount of CP violation in the leptonic mixing matrix. There are 9 such quantities, and they have the following properties:

- $J^{jk}_{\alpha\beta}$ is invariant under any rephasing of the fields, $\nu_\alpha \to e^{i\phi}\nu_\alpha$ or $\nu_j \to e^{i\phi}\nu_j$.

- Due to unitarity of U, all $J^{jk}_{\alpha\beta}$ are equal up to signs. Unitarity implies

$$U^*_{\alpha1}U_{\beta1} + U^*_{\alpha2}U_{\beta2} + U^*_{\alpha3}U_{\beta3} = 0 \tag{3.39}$$
$$\Leftrightarrow U^*_{\alpha1}U_{\beta1}U_{\alpha1}U^*_{\beta1} + U^*_{\alpha2}U_{\beta2}U_{\alpha1}U^*_{\beta1} + U^*_{\alpha3}U_{\beta3}U_{\alpha1}U^*_{\beta1} = 0\,. \tag{3.40}$$

  The first term is real. Therefore, it follows

$$\mathrm{Im}\big(U^*_{\alpha2}U_{\beta2}U_{\alpha1}U^*_{\beta1}\big) = -\mathrm{Im}\big(U^*_{\alpha3}U_{\beta3}U_{\alpha1}U^*_{\beta1}\big)\,. \tag{3.41}$$

  Similar relation can be shown for all the $J^{jk}_{\alpha\beta}$, therefore we can write

$$J^{jk}_{\alpha\beta} \equiv \mathrm{Im}\big(U^*_{\alpha j}U_{\beta j}U_{\alpha k}U^*_{\beta k}\big) = \pm J\,. \tag{3.42}$$

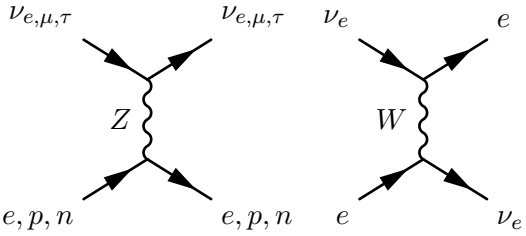
  J is called the *Jarlskog invariant*. In the standard parameterization of the leptonic mixing matrix eq. (3.26), it is given by

$$J = c_{12}s_{12}c_{23}s_{23}c^2_{13}s_{13}\sin\delta\,. \tag{3.43}$$

- The Jarlskog invarinat J is non-zero only if all three mixing angles are $\neq 0$. This shows again that CP violation is a true 3-flavor effect.

3.3 Neutrino oscillations in matter

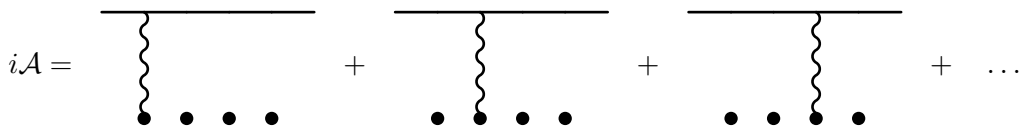
As neutrino travel through matter, they can interact through the following Feynman diagrams:



Note that the first diagram exists for all neutrino flavors, and can couple neutrinos to electrons, protons and neutrinos, while the second exits only for ν_e and coupled them only to electrons.

Processes mediated by the above diagrams in which momentum is exchanged between the neutrino and the backgrounds matter are usually negligible. However, there is the possiblity that a W or Z boson is exchanged without any momentum transfer. This is analogous to what happens to photons travelling through water: the photons interact with water molecules by being continuously absorbed and reemitted with excatly the same momentum. Therefore, the speed of light in water is different from the one in vacuum.

The crucial point is that there is no way of telling with which of the many particles in the background matter the neutrino has interacted. Therefore, the amplitudes for interactions with different particles must be summed up coherently:



Therefore, the interaction probability $|\mathcal{A}|^2$ is proportional to the density of n target particles *squared*, whereas for "real" scattering processes with momentum transfer are proportional just to n. The extra factor of n leads to a huge enhancement and makes coherent forward scattering phenomenologically relevant in spite of the general weakness of W and Z mediated processes.

For the mathematical treatment of coherent forward scattering, we start weak interaction Hamiltonian.

$$\mathcal{H} \supset -\sum_{\alpha=e,\mu,\tau}\left[\frac{g}{\sqrt{2}}\left(\overline{\nu_{\alpha,L}}\gamma^\mu e_{\alpha,L}W^+_\mu + h.c.\right) + \frac{g}{2\cos\theta_w}\overline{\nu_{\alpha,L}}\gamma^\mu\nu_{\alpha,L}Z_\mu\right]. \tag{3.44}$$

(The interaction terms in the Hamiltonian $\mathcal{H}$ are just the negative of the interaction terms in the Lagrangian $\mathcal{L}$, as follows from the Legendre transform.)

Let us consider first the interactions of ν_e with electrons. At energies $\ll M_W \sim 80$ GeV, the W boson can be integrated out (i.e. its propagator $[q^2 - M_W^2]^{-1}$ can be replaced simply by M_W^{-2}). This leads to an effective Hamiltonian (i.e. a Hamiltonian that yields Feynman rules in which the W propagator no longer appears, only M_W^{-2}):

$$\begin{aligned}\mathcal{H}_{\text{eff}} &\supset \frac{g^2}{2M_W^2}\left[\overline{\nu_{e,L}}\gamma^\mu e_L\right]\left[\overline{e_L}\gamma^\mu\nu_{e,L}\right]\\ &= \frac{G_F}{\sqrt{2}}\left[\overline{\nu_e}\gamma^\mu(1-\gamma^5)e\right]\left[\overline{e}\gamma^\mu(1-\gamma^5)\nu_e\right]\\ &= \frac{G_F}{\sqrt{2}}\left[\overline{e}\gamma^\mu(1-\gamma^5)e\right]\left[\overline{\nu_e}\gamma^\mu(1-\gamma^5)\nu_e\right].\end{aligned} \tag{3.45}$$

Here,

$$G_F \equiv \frac{\sqrt{2}g^2}{8M_W^2} \tag{3.46}$$

is the Fermi constant, and the theory described by eq. (3.45) is precisely Fermi's theory of weak interactions. In the last step of eq. (3.45), we have used the so-called *Fierz identity*, (see e.g. [1]). Its proof involves te properties of the Dirac gamma matrices (which in turn follow from the properties of the Pauli matrixces) and some gymnastics with spinor indices. Care must be taken not to forget that fermion fields *anticommute*.

For the purposes of describing neutrino propagation through matter, the electrons in the background matter can be considered static. We can thus take the expectation value of $\mathcal{H}$ over the electron fields:

$$\mathcal{H}_{\text{eff}} \supset \frac{G_F}{\sqrt{2}} \left\langle \overline{e}\gamma^\mu(1-\gamma^5)e \right\rangle \left[\overline{\nu_e}\gamma^\mu(1-\gamma^5)\nu_e\right] . \tag{3.47}$$

To evaluate $\left\langle \overline{e}\gamma^\mu(1-\gamma^5)e \right\rangle$, we use the Fourier expansion of the electron field:

$$e = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2E_{\mathbf{p}}}} \sum_s \left(u^s(p) e^{-ipx} a^s_{\mathbf{p}} + v^s(p) e^{ipx} b^{s\dagger}_{\mathbf{p}} \right) . \tag{3.48}$$

Here, $u^s(p)$ and $v^s(p)$ are Dirac spinors (solutions of the Dirac equation for a free particle with 4-momentum p), $a^s_{\mathbf{p}}$ is the particle annihilation operator, $b^{s\dagger}_{\mathbf{p}}$ is the antiparticle creation operator, and the sum runs over spins. A 1-particle state can be written as

$$|e(\mathbf{p},s)\rangle = a^{s\dagger}_{\mathbf{p}}|0\rangle , \tag{3.49}$$

where we use the canonical quantum mechanical normalization convention rather than the Lorentz invariant normalization convention often employed in QFT textbooks [1]. Assuming the background electrons to follow a momentum distribution $f(\mathbf{p})$, normalized to the electron number density:

$$\int d^3p\, f(\mathbf{p}) = n_e , \tag{3.50}$$

we can write

$$\begin{aligned}
\left\langle \overline{e}\gamma^\mu(1-\gamma^5)e \right\rangle &= \frac{1}{2}\sum_s \int d^3p\, f(\mathbf{p}) \langle e(\mathbf{p},s)|\overline{e}\gamma^\mu(1-\gamma^5)e|e(\mathbf{p},s)\rangle && (3.51)\\
&= \frac{1}{2}\sum_s \int \frac{d^3p}{2E_{\mathbf{p}}} f(\mathbf{p})\, \bar{u}^s(p)\gamma^\mu(1-\gamma^5)u^s(p) && (3.52)\\
&= \frac{1}{2}\int \frac{d^3p}{2E_{\mathbf{p}}} f(\mathbf{p}) \operatorname{tr}\left[(\not{p}+m_e)\gamma^\mu(1-\gamma^5)\right] && (3.53)\\
&= \frac{1}{2}\int \frac{d^3p}{2E_{\mathbf{p}}} f(\mathbf{p}) \operatorname{tr}\not{p}\gamma^\mu && (3.54)\\
&= \begin{cases} n_e & \text{for } \mu = 0 \\ 0 & \text{for } \mu = 1,2,3 \end{cases} && (3.55)
\end{aligned}$$

The factor $1/2$ in front of the sum comes from the average over electron spins. We thus obtain for the effective Hamiltonian

$$\begin{aligned}
\mathcal{H}_{\text{eff}} &\supset \frac{G_F}{\sqrt{2}} n_e \left[\overline{\nu_e}\gamma^0(1-\gamma^5)\nu_e\right] && (3.56)\\
&= \underbrace{\sqrt{2}G_F n_e}_{\equiv V_{CC}} \overline{\nu_{eL}}\gamma^0\nu_{eL} . && (3.57)
\end{aligned}$$

This is just a potential energy term for neutrino. (Cf., as an analogy, the interaction of an electron with an electrostatic potential, $\bar{e}\gamma^0 e\phi$.) In other words, the energy of a ν_e propagating through matter, is enhanced by V_{CC}.

In a similar way, one can show that all neutrino flavors also receive a potential from Z exchange with electrons, protons and neutrons. This neutral current potential has the form

$$V_{NC} = -\frac{1}{2}\sqrt{2}G_F n_n \,, \tag{3.58}$$

where n_n is the number density of neutrons. (The contributions of electrons and protons to NC interactions exactly cancel each other.)

What is the effect of these potentials on neutrino oscillations? Remember that the neutrino oscillation probability for a stationary source,,

$$P_{\alpha\beta}(L) = \frac{1}{N}\sum_{j,k} U^*_{\alpha j}U_{\beta j}U_{\alpha k}U^*_{\beta k}\exp\left[i\big(\phi_j - \phi_k\big)\right]\delta(E_j - E_k)\,, \tag{3.59}$$

depends on the neutrino energy through the oscillation phase

$$\phi_j \equiv p_j L\,, \tag{3.60}$$

where j labels the different states of definite momentum. In vacuum, these are just the neutrino mass eigenstates. In matter, however, the flavor structure of the matter potentials V_{CC} and V_{NC} (the former acting only on electron neutrinos, the latter acting on all neutrino flavors) changes this. To determine the new momentum eigenstates (and the new mixing matrix) in matter, we have to diagonalize the matrix

$$\hat{p} \equiv \sqrt{(\hat{H} - \hat{V})^2 - \hat{M}^2} \tag{3.61}$$

$$\simeq \hat{H} - \frac{\hat{M}^2}{2\hat{H}} - \hat{V}\,. \tag{3.62}$$

In the flavor basis, and using the 2-flavor approximation, $\hat{V}$ and $\hat{M}$ are given by

$$\hat{V} \equiv \begin{pmatrix} V_{CC} + V_{NC} & \\ & V_{NC} \end{pmatrix} \tag{3.63}$$

and

$$\hat{M} \equiv \begin{pmatrix} c & s \\ -s & c \end{pmatrix}\begin{pmatrix} m_1 & \\ & m_2 \end{pmatrix}\begin{pmatrix} c & -s \\ s & c \end{pmatrix}\,, \tag{3.64}$$

respectively. We have used the abbreviations $c = \cos\theta$ and $s = \sin\theta$. In the second line of eq. (3.62), we have used the fact that neutrino masses and matter potentials are much smaller than the neutrino energy. Since we have assumed a stationary source, i.e. E is the same for all neutrino momentum eigenstates, we can replace $\hat{H} \to E$. Using moreover

the fact that contributions proportional to the identity matrix, (and thus in particular V_{NC}) do not affect neutrino oscillations (since only phase differences are measured), we find that we have to diagonalize the matrix

$$-\underbrace{\frac{\Delta m^2}{4E}}_{\equiv\Delta/2}\begin{pmatrix} c & s \\ -s & c \end{pmatrix}\begin{pmatrix} -c & s \\ s & c \end{pmatrix} - \begin{pmatrix} \frac{V_{CC}}{2} & \\ & -\frac{V_{CC}}{2} \end{pmatrix} = -\frac{\Delta}{2}\begin{pmatrix} -c^2+s^2 & 2sc \\ 2sc & -s^2+c^2 \end{pmatrix} - \begin{pmatrix} \frac{V_{CC}}{2} & \\ & -\frac{V_{CC}}{2} \end{pmatrix} \tag{3.65}$$

$$= -\begin{pmatrix} -\frac{\Delta}{2}\cos 2\theta + \frac{V_{CC}}{2} & \frac{\Delta}{2}\sin 2\theta \\ \frac{\Delta}{2}\sin 2\theta & \frac{\Delta}{2}\cos 2\theta - \frac{V_{CC}}{2} \end{pmatrix}. \tag{3.66}$$

The characteristic polynomial is

$$\left(\frac{V_{CC}}{2} - \frac{\Delta}{2}\cos 2\theta - \lambda\right)\left(-\frac{V_{CC}}{2} + \frac{\Delta}{2}\cos 2\theta - \lambda\right) - \frac{\Delta^2}{4}\sin^2 2\theta = 0\,, \tag{3.67}$$

or

$$\lambda^2 - \left(\frac{V_{CC}}{2} - \frac{\Delta}{2}\cos 2\theta\right)^2 - \tfrac{\Delta^2}{4}\sin^2 2\theta = 0\,. \tag{3.68}$$

This yields for the eigenvalues

$$\lambda_{1,2} = \pm\frac{1}{2}\sqrt{(V_{CC} - \Delta\cos 2\theta)^2 + \Delta^2\sin^2 2\theta}\,. \tag{3.69}$$

It thus makes sense to define the effective mass squared difference in matter as

$$\Delta m^2_{\text{eff}} = 2E\sqrt{(V_{CC} - \Delta\cos 2\theta)^2 + \Delta^2\sin^2 2\theta}\,, \tag{3.70}$$

so that the Hamiltonian after diagonalization (and subtraction of a contribution propotional to the identity matrix, $(E + V_{CC}/2)\mathbb{1}_{2\times 2}$), has the same structure as in vacuum. The behavior of the energy eigenvalues (after re-inserting the contribution proportional to $\mathbb{1}_{2\times 2}$)) is plotted in fig. 3.1. We see that, for increasing matter density, the two energy eigenvalues approach each other and then separate again. The point where they are closest to each other is the point where Δm^2_{eff} is minimal. This happens for

$$V_{CC} = \Delta\cos 2\theta\,. \tag{3.71}$$

We also define an effective mixing angle in matter, θ_{eff}, by the requirement that it diagonalized the Hamiltonian in matter:

$$-\begin{pmatrix} \cos\theta_{\text{eff}} & -\sin\theta_{\text{eff}} \\ \sin\theta_{\text{eff}} & \cos\theta_{\text{eff}} \end{pmatrix}\begin{pmatrix} -\frac{\Delta}{2}\cos 2\theta + V_{CC} & \frac{\Delta}{2}\sin 2\theta \\ \frac{\Delta}{2}\sin 2\theta & \frac{\Delta}{2}\cos 2\theta \end{pmatrix}\begin{pmatrix} \cos\theta_{\text{eff}} & \sin\theta_{\text{eff}} \\ -\sin\theta_{\text{eff}} & \cos\theta_{\text{eff}} \end{pmatrix} = \begin{pmatrix} -\frac{\Delta m^2_{\text{eff}}}{4E} & \\ & \frac{\Delta m^2_{\text{eff}}}{4E} \end{pmatrix}. \tag{3.72}$$

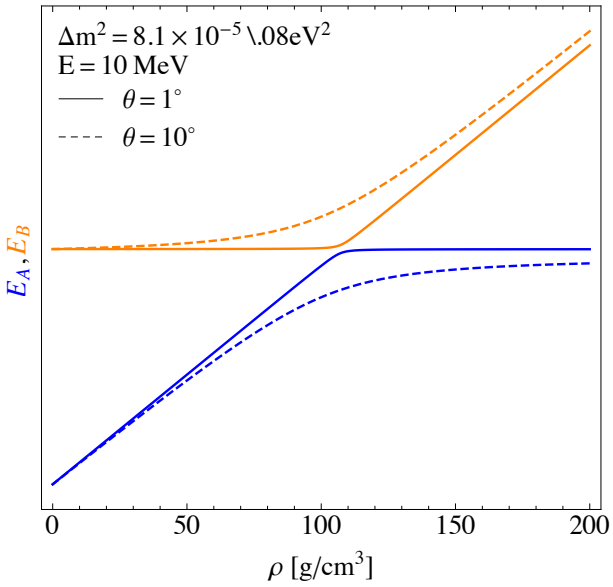


Figure 3.1: The eigenvalues of the Hamiltonian in matter, $E + V_{CC}/2 \pm \Delta m^2_{\text{eff}}/(2E)$.

It is sufficient to consider the off-diagonal entries of the matrix on the left and requiring them to vanish. This yields

$$\tan 2\theta_{\text{eff}} = \frac{\Delta \sin 2\theta}{V_{CC} - \Delta \cos 2\theta}\,, \tag{3.73}$$

or, equivalently,

$$\sin 2\theta_{\text{eff}} = \frac{\tan 2\theta_{\text{eff}}}{\sqrt{1 + \tan^2 \theta_{\text{eff}}}} \tag{3.74}$$

$$= \frac{\Delta \sin 2\theta}{\sqrt{V_{CC} - \Delta \cos 2\theta)^2 + \Delta^2 \sin^2 2\theta}}\,. \tag{3.75}$$

In terms of the effective mass squared difference Δm^2_{eff} and the effective mixing angle θ_{eff}, the neutrino oscillation probability can be written in analogy to eq. (3.29) as

$$P_{\alpha\beta}(L) \simeq \left| \begin{pmatrix} \cos\theta_{\text{eff}} & \sin\theta_{\text{eff}} \\ -\sin\theta_{\text{eff}} & \cos\theta_{\text{eff}} \end{pmatrix} \begin{pmatrix} \exp\left[\frac{i\Delta m^2_{\text{eff}} L}{4E}\right] & \\ & \exp\left[-\frac{i\Delta m^2_{\text{eff}} L}{4E}\right] \end{pmatrix} \begin{pmatrix} \cos\theta_{\text{eff}} & -\sin\theta_{\text{eff}} \\ \sin\theta_{\text{eff}} & \cos\theta_{\text{eff}} \end{pmatrix}_{\beta\alpha} \right|^2 . \tag{3.76}$$

The expression for θ_{eff}, eq. (3.75), has a remarkable phenomenological consequence: if

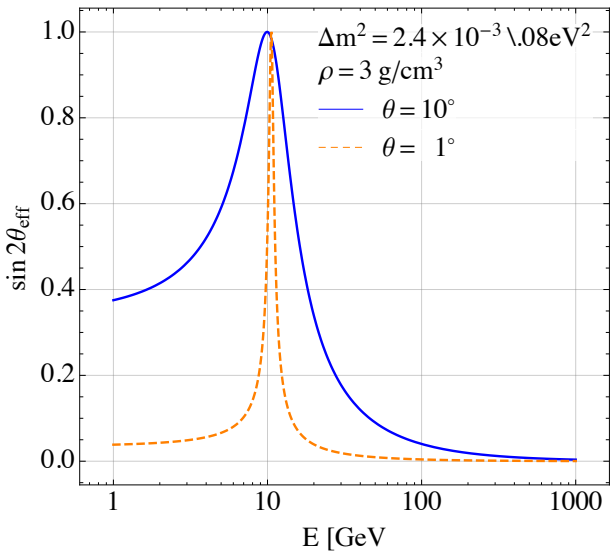


Figure 3.2: Illustration of the MSW resonance: sine of twice the effective mixing angle in matter, $\sin 2\theta_{\text{eff}}$, as a function of energy for matter density $\rho = 3$ g/cm^3, corresponding roughly to the density of the Earth mantle.

the so-called *Mikheyev-Smirnov-Wolfenstein (MSW) resonance* condition,

$$V_{CC} = \Delta \cos 2\theta \tag{3.77}$$

$$\Leftrightarrow \quad \sqrt{2} G_F n_e = \frac{\Delta m^2}{2E} \cos 2\theta \,, \tag{3.78}$$

is fulfilled, $\sin 2\theta_{\text{eff}} = 1$, i.e. the mixing angle in matter is *maximal*, irrespective of how small the vacuum mixing angle θ is! This is illustrated in fig. 3.2. Note that the MSW resonance condition is identical to the condition from eq. (3.71), which described the point where the two energy eigenvalues of the Hamiltonian are closest to each other.

We observe the following

- Below the resonance, $\theta_{\text{eff}} \simeq \theta$.
- At the resonance, $\theta_{\text{eff}} = \pi/4$.
- Well above the resonance, $\theta_{\text{eff}} \to 0$ since $\sin 2\theta_{\text{eff}} \propto \Delta m^2/(2E)$ in this limit.
- The width of the resonance is smaller for smaller vacuum mixing angle θ.

3.4 Adiabatic flavor transitions in matter of varying density

The derivations in the previous section allowed us to understand how neutrino oscillations are modified in matter of constant density. If the density is slowly varying along the

neutrino trajectory (as is the case for instance for neutrinos propagating out of the Sun), further complications arise. To understand those, let us go back to eq. (3.3). There, we had used the fact that in vacuum (or in matter of constant density), the time and space evolution of a neutrino energy eigenstate with energy E_j and momentum p_j is given by

$$|\nu_j(T,L)\rangle = e^{-iE_jT+ip_jL}|\nu_j\rangle\,. \tag{3.79}$$

In the following, we will for simplicity work in the 2-flavor framework, and we will consider only the evolution in time, i.e. we assume all p_j to be equal. (Assuming more realistically evolution only in space would lead to exactly the same results.) Since an overall phase factor $\exp[ipL]$ is irrelevant, we will thus simply write

$$|\nu_j(T)\rangle = \underbrace{e^{-iE_jT}}_{\equiv\nu_j(T)}|\nu_j\rangle\,. \tag{3.80}$$

We deonte the coefficients e^{-iE_jT} as $\nu_j(T)$ (without $|\cdot\rangle$). They are just the wave functions solving the Schrödinger equation

$$i\frac{d}{dT}\begin{pmatrix}\nu_1(T)\\ \nu_2(T)\end{pmatrix} = \underbrace{\begin{pmatrix}E_1 & \\ & E_2\end{pmatrix}}_{\equiv\hat{H}}\begin{pmatrix}\nu_1(T)\\ \nu_2(T)\end{pmatrix} = \left[E + \begin{pmatrix}-\Delta m^2/(4E) & \\ & \Delta m^2/(4E)\end{pmatrix}\right]\begin{pmatrix}\nu_1(T)\\ \nu_2(T)\end{pmatrix} \tag{3.81}$$

with $E \equiv E_1 + \Delta m^2/(4E)$ and with the Hamilton operator $\hat{H}$. The term containing just E leads to a common phase factor for ν_1 and ν_2, which does not affect oscillations. Therefore, this term will be neglected in the following. In flavor space, the Schrödinger equation reads

$$i\frac{d}{dT}\begin{pmatrix}\nu_e(T)\\ \nu_\mu(T)\end{pmatrix} = U\begin{pmatrix}-\Delta m^2/(4E) & \\ & \Delta m^2/(4E)\end{pmatrix}U^\dagger\begin{pmatrix}\nu_e(T)\\ \nu_\mu(T)\end{pmatrix} \tag{3.82}$$

Let us now consider oscillations in matter, and in particular in matter with varying density and thus varying MSW potential $V_{CC}(T)$. Then, the Schrödinger equation becomes

$$i\frac{d}{dT}\begin{pmatrix}\nu_e(T)\\ \nu_\mu(T)\end{pmatrix} = \left[U\begin{pmatrix}-\frac{\Delta m^2}{4E} & \\ & \frac{\Delta m^2}{4E}\end{pmatrix}U^\dagger + \begin{pmatrix}\frac{V_{CC}(T)}{2} & \\ & -\frac{V_{CC}(T)}{2}\end{pmatrix}\right]\begin{pmatrix}\nu_e(T)\\ \nu_\mu(T)\end{pmatrix}\,. \tag{3.83}$$

We have again removed a flavor-universal contribution $V_{CC}/2 + V_{NC}$ from the matter potential because it only leads to an overall phase factor, common to ν_e and ν_μ. It is convenient to work in a basis $\nu_A(T)$, $\nu_B(T)$ of *instantaneous matter eigenstates*, i.e. a basis in which the Hamiltonian is diagonal at any given time T. We write

$$\begin{pmatrix}\nu_e(T)\\ \nu_\mu(T)\end{pmatrix} = \tilde{U}(T)\begin{pmatrix}\nu_A(T)\\ \nu_B(T)\end{pmatrix}\,, \tag{3.84}$$

where $\tilde{U}(T)$ satisfies

$$[\tilde{U}(T)]^\dagger \left[U \begin{pmatrix} -\frac{\Delta m^2}{4E} & \\ & \frac{\Delta m^2}{4E} \end{pmatrix} U^\dagger + \begin{pmatrix} \frac{V_{CC}(T)}{2} & \\ & -\frac{V_{CC}(T)}{2} \end{pmatrix} \right] \tilde{U}(T) = \begin{pmatrix} -\frac{\Delta m^2_{\text{eff}}}{4E} & \\ & \frac{\Delta m^2_{\text{eff}}}{4E} \end{pmatrix} . \tag{3.85}$$

The effective mass squared difference in matter, Δm^2_{eff}, is given by eq. (3.70), which we repeat here:

$$\Delta m^2_{\text{eff}} = 2E\sqrt{(V_{CC} - \Delta\cos 2\theta)^2 + \Delta^2 \sin^2 2\theta}\,. \tag{3.86}$$

In the basis $(\nu_A(T), \nu_B(T))$, the Schrödinger equation becomes

$$i\frac{d}{dT}\left[\tilde{U}(T)\begin{pmatrix}\nu_A(T)\\ \nu_B(T)\end{pmatrix}\right] = \left[U\begin{pmatrix} -\frac{\Delta m^2}{4E} & \\ & \frac{\Delta m^2}{4E}\end{pmatrix}U^\dagger + \begin{pmatrix}\frac{V_{CC}(T)}{2} & \\ & -\frac{V_{CC}(T)}{2}\end{pmatrix}\right]\tilde{U}(T)\begin{pmatrix}\nu_A(T)\\ \nu_B(T)\end{pmatrix} \tag{3.87}$$

$$\Leftrightarrow i\frac{d}{dT}\begin{pmatrix}\nu_A(T)\\ \nu_B(T)\end{pmatrix} = \left[\begin{pmatrix} -\frac{\Delta m^2_{\text{eff}}}{4E} & \\ & \frac{\Delta m^2_{\text{eff}}}{4E}\end{pmatrix} - i[\tilde{U}(T)]^\dagger\frac{d}{dT}\tilde{U}(T)\right]\begin{pmatrix}\nu_A(T)\\ \nu_B(T)\end{pmatrix}. \tag{3.88}$$

The second term on the right hand side can be explicitly evaluated:

$$i[\tilde{U}(T)]^\dagger\frac{d}{dT}\tilde{U}(T) = \begin{pmatrix}\cos\theta_{\text{eff}} & -\sin\theta_{\text{eff}}\\ \sin\theta_{\text{eff}} & \cos\theta_{\text{eff}}\end{pmatrix}\begin{pmatrix}-\sin\theta_{\text{eff}} & \cos\theta_{\text{eff}}\\ -\cos\theta_{\text{eff}} & -\sin\theta_{\text{eff}}\end{pmatrix} i\dot{\theta}_{\text{eff}} \tag{3.89}$$

$$= \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix} i\dot{\theta}_{\text{eff}}\,, \tag{3.90}$$

so that we obtain

$$i\frac{d}{dT}\begin{pmatrix}\nu_A(T)\\ \nu_B(T)\end{pmatrix} = \begin{pmatrix} -\frac{\Delta m^2_{\text{eff}}}{4E} & i\dot{\theta}_{\text{eff}}\\ -i\dot{\theta}_{\text{eff}} & \frac{\Delta m^2_{\text{eff}}}{4E}\end{pmatrix}\begin{pmatrix}\nu_A(T)\\ \nu_B(T)\end{pmatrix}. \tag{3.91}$$

The interpretation of this equation is the following: if the change in the matter potential and thus the change in θ_{eff} is very slow, the off-diagonal terms on the right hand side can be neglected, and ν_A and ν_B do not mix.

Consider for instance electron neutrinos produced at the center of the Sun, where V_{CC} is so large that $\sin 2\theta_{\text{eff}} \simeq 0$ ($\theta_{\text{eff}} \simeq 90°$). Then, $\nu_e \simeq \nu_B$. On the way out of the Sun, the neutrino always stays a ν_B until it reaches the vacuum outside the Sun. In vacuum, however, $\nu_B = \nu_2 = \sin\theta\,\nu_e + \cos\theta\,\nu_\mu$. In other words, the ν_e has been *adiabatically* converted into a mixture of ν_e and ν_μ. If the vacuum mixing angle θ is small, the conversion is nearly complete.

Looking back at fig. 3.1, this means that the neutrino starts out entirely on the upper (ν_B) curve and stays on it all the way out of the Sun. If transitions were not adiabatic

(i.e. if $\dot{\theta}_{\rm eff}$ was not negligible), the probability for the neutrino to "jump" from the upper to the lower curve would be non-negligible.

In the more realistic case, that, at the center of the Sun, neutrinos are produced at $T = 0$ with $\sin 2\theta_{\rm eff}(0) \neq 0$, we have

$$\begin{pmatrix} \nu_A(0) \\ \nu_B(0) \end{pmatrix} = \begin{pmatrix} \cos\theta_{\rm eff} \\ \sin\theta_{\rm eff} \end{pmatrix} \tag{3.92}$$

After propagation out of the Sun, this state has evolved at time T into

$$\begin{pmatrix} \nu_A(T) \\ \nu_B(T) \end{pmatrix} = \begin{pmatrix} \cos\theta_{\rm eff}\, e^{i\phi} \\ \sin\theta_{\rm eff}\, e^{-i\phi} \end{pmatrix} , \tag{3.93}$$

with

$$\phi \equiv \int_0^T dt\, \frac{\Delta m^2_{\rm eff}(t)}{4E} . \tag{3.94}$$

In the flavor basis, this is

$$\begin{pmatrix} \nu_e(T) \\ \nu_\mu(T) \end{pmatrix} = \begin{pmatrix} \cos\theta\, \cos\theta_{\rm eff}\, e^{i\phi} + \sin\theta\, \sin\theta_{\rm eff}\, e^{-i\phi} \\ -\sin\theta\, \cos\theta_{\rm eff}\, e^{-i\phi} + \cos\theta\, \sin\theta_{\rm eff}\, e^{i\phi} \end{pmatrix} . \tag{3.95}$$

Thus, the oscillation probability becomes

$$P_{e\mu} = \left| -\sin\theta\, \cos\theta_{\rm eff}\, e^{i\phi} + \cos\theta\, \sin\theta_{\rm eff}\, e^{-i\phi} \right|^2 \tag{3.96}$$

$$= \sin^2\theta\, \cos^2\theta_{\rm eff} + \cos^2\theta\, \sin^2\theta_{\rm eff} - \frac{1}{2} \sin 2\theta\, \sin 2\theta_{\rm eff} \cos\phi \tag{3.97}$$

$$= \frac{1}{2}\big(1 - \cos 2\theta\, \cos 2\theta_{\rm eff} - \sin 2\theta\, \sin 2\theta_{\rm eff} \cos\phi\big) . \tag{3.98}$$

The third (oscillating) term is negligible if either the vacuum mixing angle θ or the mixing angle in matter at the neutrino production point, $\theta_{\rm eff}$ is small. Moreover, if the region in which neutrinos are produced is much larger than the oscillation length (as is the case at the center of the Sun), this term averages to zero when the oscillation probability is integrated over all possible T.

Let us briefly come back to the condition for the validity of the above derivation: the off-diagonal elements of the Hamiltonian matrix on the right hand side of eq. (3.91) had to be small. More precisely, in order for mixing between $|\nu_A\rangle$ and $|\nu_B\rangle$ to be negligible, the condition

$$\gamma^{-1} \equiv \left| \frac{2\dot{\theta}}{\Delta m^2/(2E)} \right| \ll 1 \tag{3.99}$$

must be satisfied. (The quantity γ, called the *adiabaticity parameter*, is just the inverse of $\tan 2\tilde{\theta}$, where $\tilde{\theta}$ is the mixing angle that diagonalizes the Hamiltonian matrix on the right hand side of eq. (3.91).)

4
Sterile neutrinos

4.1 Evidence for a 4-th neutrino state?

Several neutrino oscillation experiments have reported results that are inconsistent with the standard 3-flavor oscillation framework:

- **LSND** reports appearance of ν_e in a beam of ν_μ at energies $\sim$ few $\times$ 10 GeV and a baseline of 30 m. If confirmed, this would imply oscillations could develop over very short distances:

$$\frac{\Delta m^2 L}{4E} \simeq \frac{\pi}{2} \quad \Rightarrow \Delta m^2 \simeq 1 \text{ eV}^2 \,. \tag{4.1}$$

  This is inconsistent with the two known mass squared differences $\Delta m^2_{21} \sim 8 \times 10^{-5}$ eV^2 and $\Delta m^2_{31} \sim 2 \times 10^{-3}$ eV^2. (With three neutrino mass eigenstates, there are only two independent mass squared difference.)

- **MiniBooNE** sees an excess of ν_e ($\bar{\nu}_e$) in a beam of ν_μ ($\bar{\nu}_\mu$) at energies $\sim$ 1 GeV and a baseline $\sim$ 1 km. Thus, L/E, the quantity describing the oscillation phase, is similar to the one observed in LSND, pointing again to oscillations with $\Delta m^2 \simeq 1$ eV^2.

- **Short-baseline reactor experiments** have measured the neutrino flux from a nuclear reactor, but found their results to be inconsistent at the 3σ level with theoretical predictions. (Actually, until 2011, they appeared consistent with theory, but then the theory was improved.) While this unexpected results could be due to a problem with the theoretical prediction of the reactor neutrino flux or an underestimation of systematic uncertainties, it could also be interpreted in terms of disappearance of $\bar{\nu}_e$ into a yet-unknown (anti)neutrino species $\bar{\nu}_s$ with an oscillation probability $\sim 10\%$ and $\Delta m^2 \gtrsim 1$ eV^2.

- **Experiments with radioactive sources** have also found a neutrino deficit. In these experiments, a very intense radioactive source was placed close to a neutrino detector and the neutrino flux from the source was measured. The observed deficit has an overall significance $\sim 3\sigma$ and appears consistent with the deficit observed in reactor experiments.

Thus, there are several completely independent experiments that observe anomalies which could be explained by the existence of a fourth neutrino flavor ν_s, with a mass ~ 1 eV and $\mathcal{O}(10\%)$ mixing with the active neutrinos.

If a fourth light neutrino state ν_s exists, it cannot couple to SM weak interactions because this would imply that the decay $Z \to \bar{\nu}_s \nu_s$ is allowed and would contribute to the "invisible width of the Z", Γ_{inv}, i.e. the total width of the Z minus the partial widths of all "visible" decay channels observed at the LEP collider. The constraint on the invisible Z width leads to the conclusion that there should be 2.92 ± 0.05 light neutrino species coupled to the Z. This clearly precludes the existence of a fourth weakly charged neutrino. Therefore, a fourth light neutrino must be a total singlet under the SM gauge group and is therefore called *sterile neutrino*.

4.2 Predicting the reactor neutrino spectrum

One if the most important motivations for considering sterile neutrinos is the reactor neutrino anomaly. It relies crucially on theoretical predictions for reactor antineutrino spectra. Therefore, we will in the following discuss how such predictions are obtained [2–4].

Neutrinos are produced in nuclear reactors when neutron-rich fission products decay via β^- decay:

$$(A, Z) \to (A, Z+1) + e^- + \bar{\nu}_e \,. \tag{4.2}$$

Neglecting the nuclear recoil energy, the decay energy Q is distributed among the electron and the neutrino. For a single isotope (A, Z), one could thus determine the neutrino spectrum by measuring the electron spectrum $dN_e(E_e)/dE_e$ and inverting it:

$$\frac{dN_\nu(E_\nu)}{dE_\nu} \equiv \frac{dN_e(Q - E_e)}{dE_e} \,, \tag{4.3}$$

where E_e and E_ν are the electron and neutrino energies, respectively. The maximum decay energy Q can also be determined by the same measurement from the cutoff energy of the e^- spectrum.

If we know Q, we can also compute $dN_\nu(E_\nu)/dE_\nu$ analytically. Assuming that the nuclear matrix element for the beta decay is energy-independent, the neutrino spectrum

is given simply by the phase space factor.

$$\begin{aligned} dN_\nu &\propto d^3p_e\, d^3p_\nu\, \delta(E_e + E_\nu - Q) \\ &\propto p_e^2 dp_e\, p_\nu^2 dp_\nu\, \delta(E_e + E_\nu - Q) \\ &= p_e E_e\, p_\nu E_\nu dE_\nu \\ &= \sqrt{(Q - E_\nu)^2 - m_e^2}\,(Q - E_\nu)\, E_\nu^2\, dE_\nu\,. \end{aligned} \tag{4.4}$$

In practice, there are several important corrections to this simple formula:

- The final state electron interacts with the Coulomb field of the nucleus. This leads to a correction factor called the *Fermi function* $F(Z, A, E)$.
- The electron doesn't see the full nuclear charge, but reduced charge due to screening by the other electrons.
- The fact that the nucleus is not a point-particle but has a non-zero radius leads to further corrections of the Fermi function.
- There is the possibility of final state radiation.
- Also for the weak interaction, the nucleus cannot be treated as an exactly point-like object. Remember that the weak interaction vertex is

  $$\mathcal{L}_{\text{weak}} \supset \frac{g}{\sqrt{2}} J_W^\mu W_\mu + h.c. \tag{4.5}$$

  with the weak current J_W^μ. For interactions with point-like particles (e.g. up and down quarks), J_W^μ has the form

  $$J_{W,\text{point-like}}^\mu = \bar{u}\gamma^\mu \frac{1 - \gamma^5}{2} d\,. \tag{4.6}$$

  For interactions with extended objects, this has to be extended to

  $$J_{W,\text{extended}}^\mu = \bar{u}\left[c_V(q^2)\gamma^\mu + c_A(q^2)\gamma^\mu\gamma^5 + F_2(q^2)\frac{i\sigma^{\mu\nu}q_\nu}{2M}\right] d\,, \tag{4.7}$$

  where M is the mass of the nucleus, q^ν is the 4-momentum of the virtual W boson and c_V, c_A, F_2 are called the vector (or Fermi), axial vector (or Gamow-Teller) and magnetic form factors. This correction is called *weak magnetism*.
- The approximation of energy-independent nuclear matrix elements is valid for allowed β decays (i.e. decays in which the electron and the neutrino are emitted without orbital angular moment L), but for forbidden β decays ($L > 0$), the matrix element is energy dependent.

The discussion so far was for a single β decay reaction. However, the reactor neutrino spectrum receives contributions from $\mathcal{O}(10\,000)$ different β decay branches (taking into account that β decay of an individual isotope can go to various excited states of the daughter nucleus). Some of these involve very short-lived parent nuclei - so short-lived that they cannot be studied individually in the laboratory. Therefore, the reactor spectrum is traditionally obtained in a different way [4–7]:

1. Starting from the measured e^- spectrum from nuclear fission, pick the s highest data points and fit the spectrum for an allowed β decay to them. The resulting fit function defines a "virtual" β decay.

2. Subtract this fit function from the spectrum. This removes the s highest bins to good accuracy.

3. Repeat with the s highest data points among the residual spectrum. Continue until the full spectrum has been fitted.

4. For each virtual β decay, the neutrino spectrum is easily obtained (see above). Add up the neutrino spectra from all virtual β decays.

Alternatively, one can use data from nuclear data tables to account for those decays which are well studied, and apply the above fitting procedure only to the residual spectrum after subtracting these known β decays. However, care must be taken since also nuclear data tables may contain incorrect entries [2].

A careful treatment of systematic uncertainties shows that the above derivations are accurate at the few per cent level.

Interestingly, older calculations of the reactor neutrino spectrum [5–7] were in very good agreement with the data, while newer evaluations [2, 4] are larger than the measured fluxes by about 3% on average. Since the deviation is larger towards the higher end of the neutrino spectrum (few MeV), and the neutrino interaction cross section scales $\propto E_\nu^2$, the event rate differs by $\sim 6\%$ from the prediction (see fig. 4.1). This discrepancy could be explained if there is an extra, sterile, neutrino species exists, into which reactor $\bar{\nu}_e$ can oscillate with an oscillation length $\lesssim 10$ m.

4.3 Global fits to sterile neutrino data

As we have seen, the possible hints for sterile neutrinos are coming from experiments at very different energies and baselines, using very different detector technologies, very different neutrino sources and different oscillation channels. Especially the last two points imply that the dependence of their data on the 4-flavour oscillation parameters (3+3 mixing angle, 1+2 complex phases, 2+1 mass squared differences) is highly non-trivial. In particular, no individual experiment can constrain all parameters, and often they depend only on combinations of parameters.

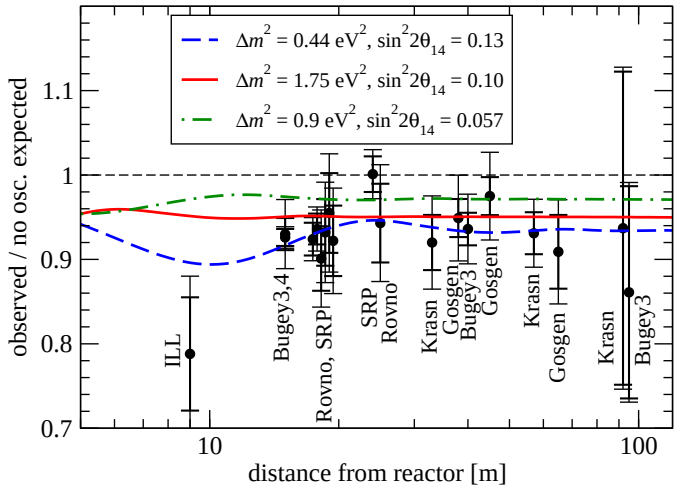


Figure 4.1: Comparison of measured reactor antineutrino fluxes to theoretical predictions of the flux [8].

To make this statement more clear, consider a 3+1 model (3 active neutrinos, 1 sterile neutrino) with a sterile neutrino mass of order 1 eV, i.e. much larger than the active neutrino masses. The oscillation probability is as usual

$$P_{\alpha\beta}(L) = \sum_{j,k} U^*_{\alpha j} U_{\beta j} U_{\alpha k} U^*_{\beta k} \exp\left[-i \frac{\Delta m^2_{jk} L}{2E} \right] . \tag{4.8}$$

If we assume that the baseline is much shorter than $2E/\Delta m^2_{21}$, $2E/\Delta m^2_{31}$, but much larger than $2E/\Delta m^2_{41}$ (an approximation that is well satisfied for LSND, MiniBooNE, reactor and gallium experiments in much of the relevant parameter space), we can set the oscillating exponentials involving Δm^2_{21}, Δm^2_{31}, Δm^2_{32} to 1, while the ones containing Δm^2_{4j} $(j = 1, 2, 3)$ average to zero when the finite experimental resolution in L and E is taken into account.

For the probability for ν_e disappearance, we then have

$$P_{ee}(L) = \sum_{j,k=1,2,3} |U_{ej}|^2 |U_{ek}|^2 + |U_{e4}|^4 \tag{4.9}$$

$$= 1 - 2|U_{e4}|^2 (1 - |U_{e4}|^2) . \tag{4.10}$$

In the second line, we have used the unitarity of the mixing matrix. Similar, for ν_μ disappearance:

$$P_{\mu\mu}(L) = 1 - 2|U_{\mu 4}|^2 (1 - |U_{\mu 4}|^2) . \tag{4.11}$$

For $\nu_\mu \to \nu_e$ transitions, on the other hand, we have

$$P_{\mu e}(L) = \sum_{j,k=1,2,3} U^*_{ej} U_{\mu j} U_{ek} U^*_{\mu k} + |U_{e4}|^2 |U_{\mu 4}|^2 \tag{4.12}$$

$$= 2|U_{e4}|^2 |U_{\mu 4}|^2 \,. \tag{4.13}$$

Here, we have again used the unitarity of U. We see that ν_e disappearance experiments, ν_μ disappearance experiments and $\nu_\mu \to \nu_e$ appearance experiments are sensitive to different combinations of the mixing parameters.

However, eqs. (4.10), (4.11) and (4.13) also show that having measurements in all three channels, we can overconstrain the system. In particular, the relation

$$2P_{\mu e} \simeq (1 - P_{ee})(1 - P_{\mu\mu}) \tag{4.14}$$

holds approximately, for small mixing between the active and sterile neutrinos.. This is exploited in *global fits*. Some results from such a fit are shown in fig. 4.2. We see that consistent fits are obtained for certain subsets of the data (panels (a) and (b)), while taking into account all oscillation channels leads to severe tension. We thus conclude that a 3+1 model cannot account for the global data. In fact, models with several sterile neutrinos are not free of tension eitehr. This implies that either the positive hints for sterile neutrinos are wrong, or some of the null results are wrong. Only future experiments will be able to definitely discover or rule out eV-scale sterile neutrinos.

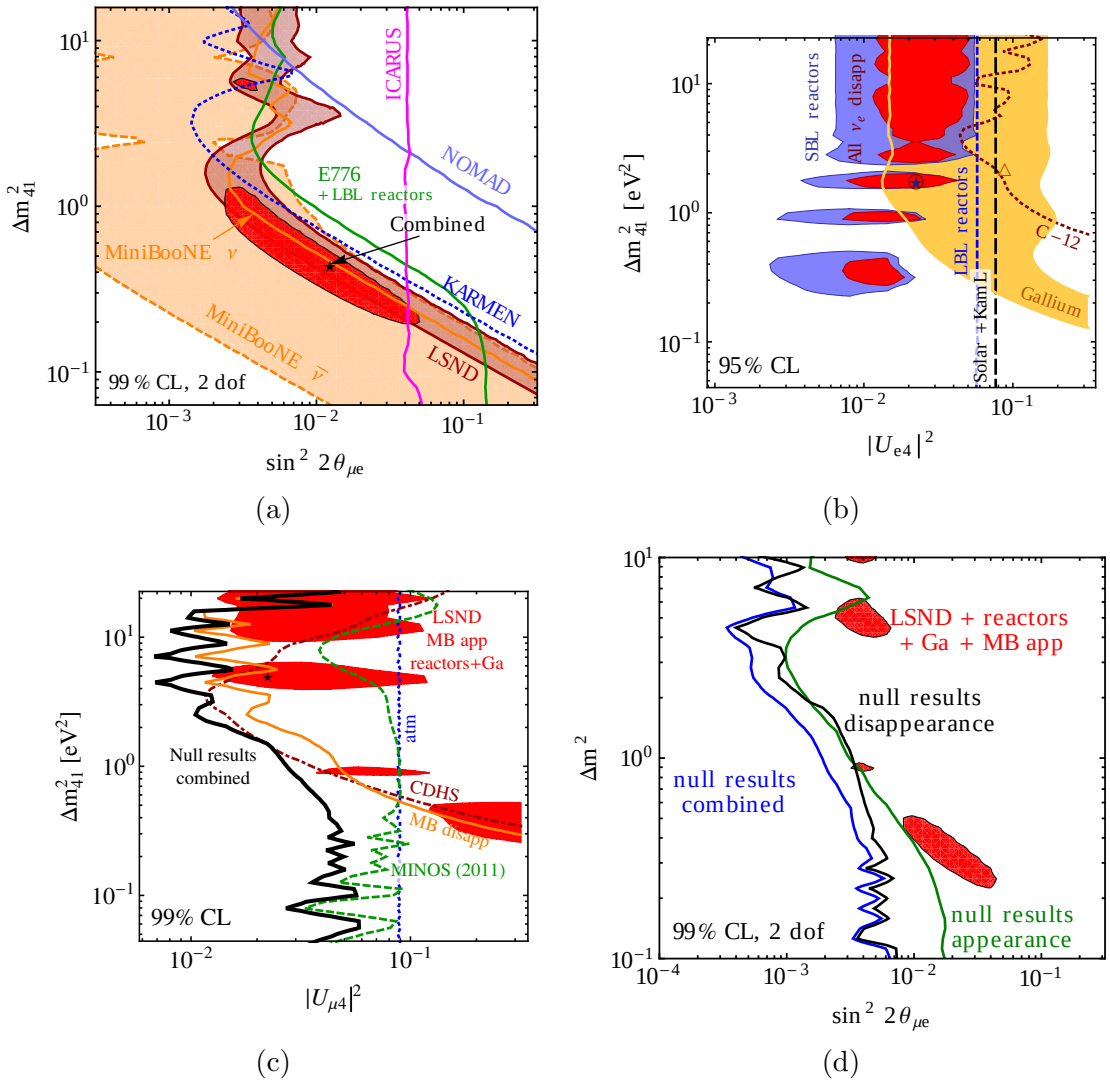


Figure 4.2: Results from a global fit to neutrino oscillation data probing the possible existence of eV-scale sterile neutrinos [8]. (a) shows a fit to $\overset{(-)}{\nu}_\mu \to \overset{(-)}{\nu}_e$ appearance data, including the LSND and MiniBooNE signals. The fit to this subset of the global data alone is consistent. (b) shows the data on $\overset{(-)}{\nu}_e$ disappearance, which also is without tension. In (c), we compare the data on $\overset{(-)}{\nu}_\mu$ disappearance to the favored values of the mixing angle θ_{24} (which determines the mixing matrix element $U_{\mu 4}$) obtained from the data shown in (a) and (b) and using the relation (4.14). This is where tension appears, demonstrating that a 3+1 model cannot account for all the data. The same can be seen in panel (d), which shows the results of the global fit as a function of Δm^2_{41} and $\sin^2 2\theta_{e\mu}$, where $\sin^2 2\theta_{e\mu} \equiv \frac{1}{4}|U_{e4}|^2|U_{\mu 4}|^2$ is the effective mixing angle that determines the ammplitude of ν_s-induced $\nu_\mu \leftrightarrow \nu_e$ oscillations.

5

Direct neutrino mass measurements

The most direct way of measuring neutrino masses is by studying the kinematics of β decay processes like

$$^3\mathrm{H} \to {}^3\mathrm{He} + e^- + \bar{\nu}_e \,. \tag{5.1}$$

The maximum energy of the electron is

$$E_e^{\max} = Q - \min_j(m_j) \,, \tag{5.2}$$

where $Q = m_\mathrm{H} - m_\mathrm{He}$ is the decay energy and m_j are the neutrino mass eigenstates.

The spectrum of electron energies from β decay is given by

$$\frac{dN_e(E_e)}{dE_e} \propto F(Z, E_e)\sqrt{E_e^2 - m_e^2}\, E_e (Q - E_e)\sqrt{(Q - E_e)^2 - m_{\bar{\nu}_e}^2}\, \theta(Q - E_e - m_{\bar{\nu}_e}) \,. \tag{5.3}$$

Here, the *Fermi function* $F(Z, E_e)$ accounts for the interaction of the produced electron with the Coulomb field of the nucleus. By $\theta(x)$ we denote the Heaviside function.

To make the dependence on the neutrino mass more visible, one defines the *Kurie functions*

$$K(E_e) \equiv \sqrt{\frac{dN_e(E_e)/dE_e}{F(Z, E_e) p_e E_e}} \,. \tag{5.4}$$

For $m_{\bar{\nu}_e} = 0$, it should be a straight line $\propto Q - E_e$, while for $m_{\bar{\nu}_e} \neq 0$, it has a cutoff, see fig. 5.1.

It is important that $\bar{\nu}_e$ does not have a definite mass, i.e. when we write $m_{\bar{\nu}_e}$ above, this is actually ill-defined. We should therefore regard the process (5.1) as the combination

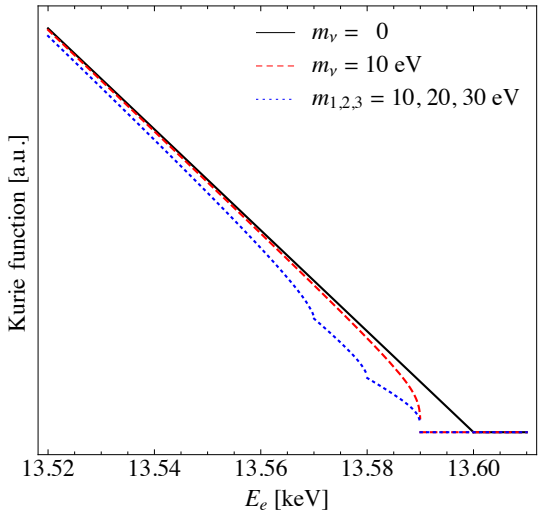


Figure 5.1: The Kurie plot.

of the processes

$$^3\mathrm{H} \to {}^3\mathrm{He} + e^- + \bar{\nu}_1\,, \tag{5.5}$$

$$^3\mathrm{H} \to {}^3\mathrm{He} + e^- + \bar{\nu}_2\,, \tag{5.6}$$

$$^3\mathrm{H} \to {}^3\mathrm{He} + e^- + \bar{\nu}_3\,.. \tag{5.7}$$

Since we can only detect the electron, we do not know on an event-by-event basis which of these processes has occured, hence we always have to take all three of them into account. Thus, in particular, the maximum energy of the electron $E_e^{\max}$ is given by the maximum of the kinematic endpoints of the three processes. For neutrino mass $m_j \gg \max_k |m_j - m_k|$, this subtlety is negligible.

For smaller neutrino masses, the Kurie function is

$$K(E_e) = \sqrt{|U_{e1}^2|[K_1(E_e)]^2 + |U_{e2}^2|[K_2(E_e)]^2 + |U_{e3}^2|[K_3(E_e)]^2}\,, \tag{5.8}$$

with

$$K_j(E_e) \equiv \mathcal{C}\sqrt{(Q-E_e)\sqrt{(Q-E_e)^2 - m_j^2}\,\theta(Q-E_e-m_j)}\,, \tag{5.9}$$

and $\mathcal{C}$ being a j-independent normalization constant. For E_e not too close to the endpoint, we can expand

$$\sqrt{(Q-E_e)^2 - m_j^2} \simeq Q - E_e - \frac{m_j^2}{2(Q-E_e)}\,. \tag{5.10}$$

Plugging this into eq. (5.9) and (5.8), we get

$$K(E_e) \simeq \mathcal{C}\sqrt{(Q-E_e)\left[Q-E_e+\sum_j |U_{ej}^2|\frac{m_j^2}{2(Q-E_e)}\theta(Q-E_e-m_j)\right]}\,. \tag{5.11}$$

At energies below $Q-\max_j(m_j)$ (where all the θ functions are 1), this can be written as

$$K(E_e) \simeq \mathcal{C}\sqrt{(Q-E_e)\sqrt{(Q-E_e)^2-m_{\bar{\nu}_e}^2}}\,, \tag{5.12}$$

with the *definition* of the *effective mass*

$$m_{\bar{\nu}_e} \equiv \sqrt{\sum_j |U_{ej}^2| m_j^2}\,. \tag{5.13}$$

Very close to the endpoint, this description breaks down. Instead, several kinks are expected in the spectrum.

6 Neutrinoless double beta decay

6.1 The rate of neutrinoless double beta decay

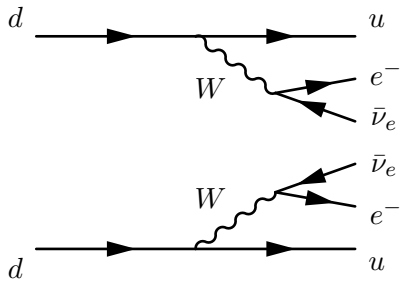
Consider the level scheme in fig. 6.1. Note that the isotopes in general follow two parabolas: a higher energy one for the odd–odd nuclei (i.e. nuclei with an odd number of protons and an odd number of neutrons), and a lower energy one for the even–even nuclei. This leads to the situation that the energy of ^{136}Xe is lower than that of the isotope directly to its right, ^{136}Cs, so that the direct β^- decay of ^{136}Xe to ^{136}Cs is energetically forbidden. On the other hand, two simultaneous β^- decays are allowed:

$$^{136}\mathrm{Xe} \rightarrow {}^{136}\mathrm{Ba} + 2e^- + 2\bar{\nu}_e\,. \tag{6.1}$$

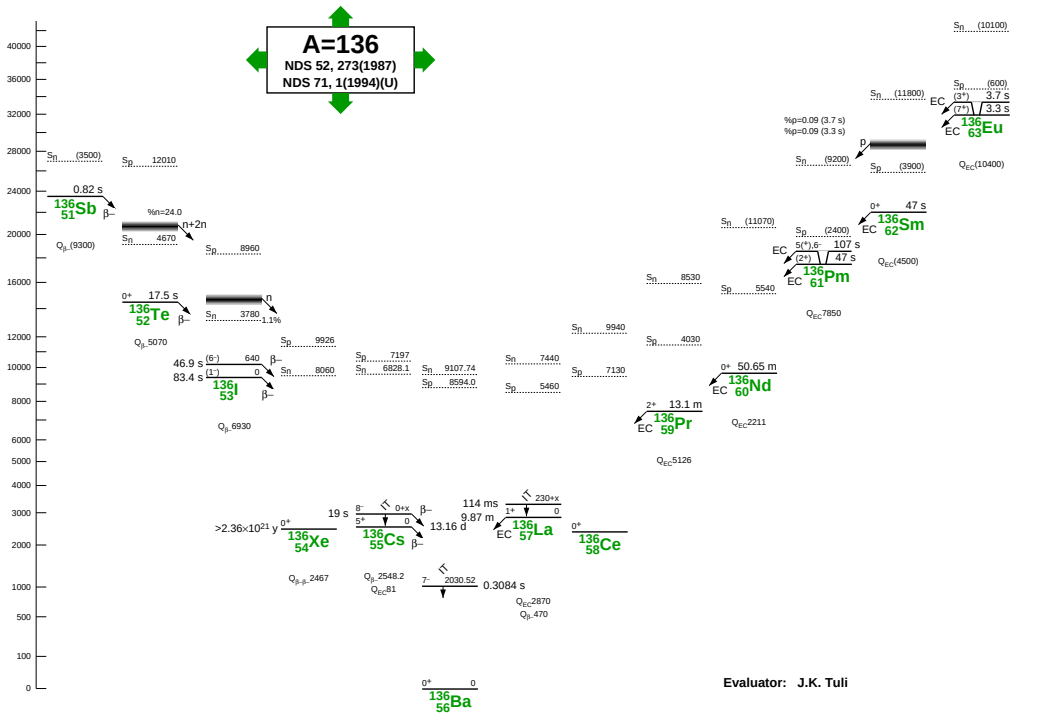
The Feynman diagram for such a process of the form

$$(A, Z) \rightarrow (A, Z-2) + 2e^- + 2\bar{\nu}_e \tag{6.2}$$

is



Remember that replacing an outgoing antiparticle in a Feynman diagram by an ingoing particle yields a valid Feynman diagram as well ("crossing symmetry"). Therefore, if the



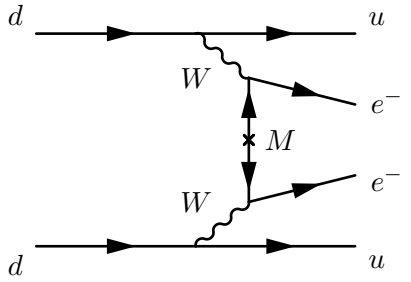
4632

Figure 6.1: Level scheme for $A = 136$ nuclei. Note that ^{136}Xe cannot undergo direct β^- decay to ^{136}Cs, but $0\nu2\beta$ decay to ^{136}Ba is allowed. Similarly, ^{136}Ce cannot decay via β^+ decay to ^{136}La, but $0\nu2\beta^+$ decay to ^{136}Ba is allowed.

neutrino is its own antiparticle (Majorana neutrinos), the two neutrino lines in the above diagram can be connected. In other words, the neutrino emitted in the decay of the first down quark can be absorbed by the second down quark. This leads to *neutrinoless double beta decay*

$$(A, Z) \to (A, Z-2) + 2e^- \,, \tag{6.3}$$

the Feynman diagram for which is



This diagram is sometimes called "lobster diagram". Since my artistic skills are not sufficient to explain why, I have to appeal to your imagination ...

While in two-neutrino double beta decay, part of the decay energy is carried away by the neutrinos, in neutrinoless double beta decay it is all carried by the electrons and thus visible to a detector. The telltale signature of neutrinoless double beta decay is thus a monoergetic peak at the Q value of the decay.

To compute the rate for neutrinoless double beta decay, we start from the weak interaction Lagrangian with a Majorana mass term for the neutrinos

$$\begin{aligned}\mathcal{L} = \sum_{j=1,2,3} \Bigg[& \bar{\nu}_{jL} i \partial\!\!\!/ \nu_{jL} + \frac{g}{\sqrt{2}} \Big(\overline{\nu_{j,L}} U^*_{\alpha j} \gamma^\mu e_{\alpha,L} W^+_\mu + h.c. \Big) + \frac{g}{2\cos\theta_w} \overline{\nu_{j,L}} \gamma^\mu \nu_{j,L} Z_\mu \Bigg] \\ & - \sum_{j=1,2,3} \frac{1}{2} m_j \Big(\overline{(\nu_{j,L})^c}\, \nu_{j,L} + \overline{\nu_{j,L}}\, (\nu_{j,L})^c \Big) \Bigg] \,. \end{aligned} \tag{6.4}$$

Note that the mass term can also be written as

$$- \sum_{j=1,2,3} \frac{1}{2} m_j \Big(- i[\gamma^0\gamma^2 \nu_{j,L}]^T\, \nu_{j,L} + \overline{\nu_{j,L}}\, (-i\gamma^2\gamma^0) \overline{\nu_{j,L}}^T \Big) \tag{6.5}$$

Already at this stage, we can see that the rate for neutrinoless double beta decay will have the form

$$\Gamma_{0\nu2\beta} \propto G_F^4 |\tilde{M}_{0\nu2\beta}|^2 \, \Big| \sum_j U_{ej}^2 m_j \Big|^2 p_e^2 \,. \tag{6.6}$$

The factor $G_F^4 = [g^2/(4\sqrt{2}M_W^2)]^4$ comes from the two W boson propagators $\sim 1/M_W^2$ and the vertices to which they connect, each of which is proportional to the weak coupling

constant g. $\tilde{M}_{0\nu2\beta}$ is a nuclear matrix element that describes the probability for the parent nucleus (A, Z) to emit two virtual W bosons, thus transforming to the daughter nucleus $(A, Z-2)$. The leptonic mixing matrix elements U_{ej} come from the vertices connecting the W boson to the leptons. The factor m_j comes from the vertex labeled with $\times$ in the lobster diagram above—the conversion of a left-handed Majorana neutrino into its (right-handed) antiparticle can happen only through the mass term, so the amplitude must be proportional to the neutrino mass. Since we cannot tell which neutrino mass eigenstate is propagating in the diagram, we have to sum over all them. The expression

$$m^{\text{eff}}_{0\nu2\beta} \equiv \sum_j U^2_{ej} m_j \tag{6.7}$$

is called the effective neutrino mass in neutrinoless double beta decay, and it is the main (only) neutrinophysics observable to which the process is sensitive. Finally, p_e in eq. (6.6) is the momentum of the final state electrons, which is used here as a proxy for the typical energy scales in the problem. We have omitted $\mathcal{O}(1)$ factors.

The parameters entering the effective mass $m^{\text{eff}}_{0\nu2\beta}$ are the mixing angles, CP violating phases and neutrino masses. Of these, the three mixing angles are known. Regarding the masses, we only know the mass squared differences $|\Delta m^2_{31}|$ and Δm^2_{21} from oscillation experiments, but not the absolute mass scale (or, equivalently, the value of the lightest neutrino mass m_1). The CP phases are completely unconstrained.

Note that neutrinoless double beta decay depends not only on the phase δ to which also oscillation experiments are sensitive, but also on two *Majorana phases*. To understand these, we have to once again count the parameters of the 3×3 mixing matrix, but this time for Majorana particles. Let us make a table similar to the one in chapter 3.2:

General 3×3 matrix	9 real parameters	9 complex phases
Unitarity:		
$\sum_j \lvert U_{\alpha j}\rvert^2 = 1$	-3	
$\sum_j U_{\alpha j} U^*_{\beta j} \overset{\alpha\neq\beta}{=} 0$	-3	-3
Field redfinitions:		
$\nu_\alpha \to e^{i\phi_\alpha}\nu_\alpha$		-3
$\sum$	3	1

Note that, for Majorana neutrinos, only 3 complex phases can be removed by field rephasings. The freedom to rephase $\nu_j \to e^{i\phi_j}\nu_j$ is lost because this would make the masses complex. In the Dirac case, where the diagonal mass terms have the form $m_j \bar{\nu}_{jL}\nu_{jR}$, any rephasing of ν_{jL} can be compensated by a corresponding rephasing of ν_{jR}, leaving m_j unchanged. For a Majorana mass term $m_j \overline{\nu^c_{jL}}\nu_{jL}$ this is not possible—any rephasing $\nu_{jL} \to e^{i\phi_j}\nu_j$ would have to be compensated by making m_j complex, but masses should always be real and positive. The leptonic mixing matrix including the Majorana CP

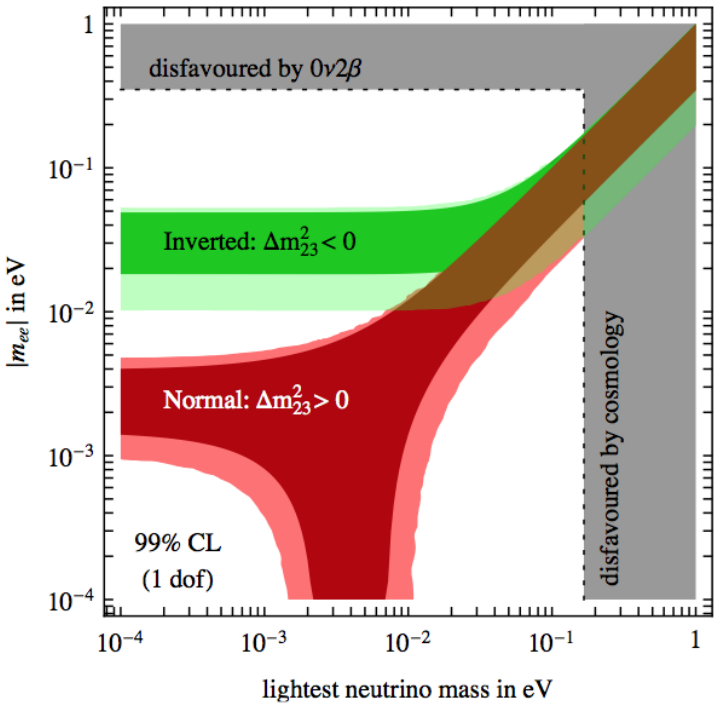


Figure 6.2: Allowed values for the effective mass $m^{\text{eff}}_{0\nu2\beta}$ measured in neutrinoless double beta decay as a function of the lightest neutrino mass m_1. Figure taken from [9].

phases can be written as

$$U = \begin{pmatrix} 1 & & \\ & c_{23} & s_{23} \\ & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & & s_{13}e^{-i\delta} \\ & 1 & \\ -s_{13}e^{i\delta} & & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & \\ -s_{12} & c_{12} & \\ & & 1 \end{pmatrix} \begin{pmatrix} e^{i\alpha} & & \\ & e^{i\beta} & \\ & & 1 \end{pmatrix} . \quad (6.8)$$

Taking into account what we know about the mixing angles and the mass squared differences, and what we do not know about the complex phases and about the absolute mass scale, $m^{\text{eff}}_{0\nu2\beta}$ is restricted to lie within the bands shown in fig. 6.2.

Applying the Feynman rules to the lobster diagram, we obtain for the amplitude of

neutrinoless double beta decay the expression

$$i\mathcal{A} = \frac{g^2}{2}\int d^4x_1\, d^4x_2 \int \frac{d^4q}{(2\pi)^4}\, e^{-iq(x_1-x_\nu)}\, \langle A, Z-2|J^\mu(x_1)J^\nu(x_2)|A,Z\rangle \tag{6.9}$$

$$\times \sum_j \left[\bar{u}(p_{e1})\frac{ig}{\sqrt{2}}U^*_{ej}\gamma^\mu P_L \frac{i}{\not{q}}\right]_\alpha \tag{6.10}$$

$$\times \frac{im_j}{2}\left[\bar{v}(p_{e2})\frac{ig}{\sqrt{2}}U^*_{ej}\gamma^\nu P_l \frac{i}{-\not{q}}(-i\gamma^2\gamma^0)\right]_\alpha \left(\frac{i}{M_W^2}\right)^2 - (p_{e1}\leftrightarrow p_{e2})\,. \tag{6.11}$$

This expression probably requires some explanations.

- We treat the neutrino as massless everywhere except in the vertex marked with $\times$. Normally in perturbation theory, we treat the kinetic term and the mass term as the zeroth order terms, and all other terms as small perturbations. Here, we also treat the mass term as a perturbation, corresponding to a Feynman vertex with just two lines connected to it. This is justified because the neutrino mass is much smaller than the typical $\mathcal{O}$(MeV) energy and momentum scale of $0\nu2\beta$ decay.

- The term $[\bar{u}(p_{e1})\frac{ig}{\sqrt{2}}U^*_{ej}\gamma^\mu P_L\frac{i}{\not{q}}]_\alpha$ corresponds to the upper lepton line, from the neutrino mass vertex to the upper external electron. The index α is a spinor index.

- The term $[\bar{v}(p_{e2})\frac{ig}{\sqrt{2}}U^*_{ej}\gamma^\nu P_l\frac{i}{\not{q}}(-i\gamma^2\gamma^0)]_\alpha$ corresponds to the lower lepton line, from the neutrino mass vertex to the lower external electron. The index α is again a spinor index, which is contracted with the corresponding index α on the other fermion term. Note the extra factor $-i\gamma^2\gamma^0$ from the charge conjugation operation in the mass term: $[(\nu_{j,L})^c]_\alpha = [-i\gamma^2\gamma^0\overline{\nu_{j,L}}]_\alpha$. The minus sign in the fermion propagator, $i/(-\not{q})$ comes from the fact that the momentum q flows against the direction of the arrow on this line.

- $g/\sqrt{2}$ is the coupling of the fermions to the W boson.

- The factor $im_j/2$ is the mass vertex.

- There is a sum over the three neutrino mass eigenstates. The reason is that there are effectively three Feynman diagrams: one where the internal fermion line corresponds to a ν_1, one where it is a ν_2, and one where it is a ν_3. Since there is now way tof distinguishing these three cases, all three diagrams must be summed coherently.

- The term $(i/M_W)^2$ comes from the two W boson propagators, using the fact that the W mass is much larger than the momentum transfer, so that we can approximate $-i/(q_W^2 - M_W^2) \to i/M_W^2$.

- The contribution from the quark lines in the Feynman diagrams is encoded in the *nuclear matrix element* $\langle A, Z-2|J^\mu(x_a)J^\nu(x_2)|A,Z\rangle$, where $J^\mu(x_1)$ and $J^\nu(x_2)$ are

the hadronic currents corresponding to the two interacting quarks. The nuclear matrix element describes the amplitude for the initial state nucleus (A, Z) to emit two (virtual) W bosons, thus converting to the final state nucleus $(A, Z-2)$. The coordinates x_1, x_2 denote the coordinates of the two nucleons participating in the decay. Obtaining the nuclear matrix elements is highly nontrivial because of the complicated hadron physics and nuclear physics involved. It requires sophisticated nuclear physics calculations, and even the best calculations available today still have a very large uncertainty.

- Since the nuclear matrix elements are given in terms of x_1 and x_2, we Fourier transform the rest of the amplitude from momentum space to coordinate space (hence the d^4q integral). We then integrate over x_1 and x_2 to account for the fact that the two nucleons can be located anywhere in the nucleus.

- Since the two electrons in the final state are indistinguishable, we have to allow for a set of equivalent diagrams with p_{e1} and p_{e2} exchanged. There is a relative minus sign between these two sets of diagrams because an odd number of fermion anticommutations is needed in the second case.

6.2 Nuclear matrix elements

As mentioned above, determining the nuclear matrix elements of neutrinoless double beta decay is highly non-trivial. They are *not* identical to the matrix elements for 2-neutrino double beta decay (which can be experimentally measured) because for $0\nu2\beta$ decay, it is favorable if the two decaying nucleons are close together—the suppression due to the virtuality of the neutrino propagator is less severe in this case. For $2\nu2\beta$ decay, no such restriction exists. Therefore, the nuclear matrix elements need to be obtained from nuclear theory.

The basic goal of nuclear matrix element calculations is to evaluate

$$\sum_m \frac{\langle A, Z-2|J^\mu(x_1)|m\rangle\langle m|J^\nu(x_2)|A, Z\rangle}{E_m - (M_i + M_f)/2}\,. \tag{6.12}$$

The two factors in the numerator correspond to the amplitude for the first β^- decay, which transforms the initial state $|A, Z\rangle$ into an intermediate state $|m\rangle$, and the second β^- decay, which transforms the intermediate state $|m\rangle$ into the final state $|A, Z-2\rangle$. The denominator describes the off-shellness of the intermediate state, and the sum runs over all intermediate states. The art is now to have expressions for the initial and final state wave functions, and to have a suitable basis of all intermediate states.

The two most common methods to achieve this are:

- **Nuclear shell model calculations.** The nuclear shell model is similar to the QM model of the hydrogen atom. One starts from hypothesizing a certain shape for the potential in which the nucleons reside and then computes the energy eigenvalues.

The goal is that the states have the correct quantum numbers, as measured from nuclear excitation spectra. However, the model works well only for small nuclei. For heavier nuclei, the outer shells are not described well, and also multiparticle effects (e.g. giant resonances) are not properly included. Moreover, it is impossible to precisely determine the exact wave functions of the initial and final states in $0\nu2\beta$ decay.

- **Quasiparticle Random Phase Approximation (QRPA).** [10] Here, one starts with a system of nucleon states, where nucleons are created and annihilated by operators c_j, $c_j^\dagger$. Since different nucleons interact quite strongly, these operators are not appropriate to describe the actual states of the nucleus, though. Therefore, one defines quasiparticle creation and annihilation operators α_j, $\alpha_j^\dagger$, which are linear combinations of c_j, $c_j^\dagger$:

$$\begin{aligned} \alpha_j^\dagger &= \sum_k (v_{kj} c_k + u_{kj} c_k^\dagger) \,, \\ \alpha_j &= \sum_k (u_{kj}^* c_k + v_{kj}^* c_k^\dagger) \,. \end{aligned} \tag{6.13}$$

If a particular $\alpha_j^\dagger$ is composed mostly of creation operators $c_k^\dagger$, it can be interpreted as an operator creating some superposition of one-nucleon states. If an $\alpha_j^\dagger$ is composed mostly of annihilation operators c_k, it can be understood as creating a hole state (as in solid state physics), i.e. creating an excited state by removing a particle from the position it is occupying in the ground state. The linear combinations $\alpha_j^\dagger$ and α_j are defined such that the ground state wave function $|0\rangle$ (i.e. the wave function that satisfies $\alpha_i|0\rangle = 0$) yields the minimum energy

$$E_0 = \frac{\langle 0|H|0\rangle}{\langle 0|0\rangle} \,, \tag{6.14}$$

where H is the Hamiltonian, which is written as

$$H = E_0 + \sum_{j,j'} t_{jj'} c_j^\dagger c_{j'} + \frac{1}{4} \sum_{jj'kk'} v_{jj'kk'} : c_j^\dagger c_j^\dagger c_{k'} c_k : \,. \tag{6.15}$$

Here, the parameters $t_{jj'}$ describe the energy of individual nucleons, and $v_{jj'kk'}$ describes interactions between nucleons. Determining these parameters is a topic we will not address here. The symbol $: \cdots :$ indicates normal ordering of the operators, with all creation operators to the left of all annihilation operators, and with the smaller indices (lower lying states) to the left. In other words, one needs to vary $\langle 0|H|0\rangle / \langle 0|0\rangle$ with respect to u_{kj} and v_{kj}. (Note that this can be done without having an explicit expression for $|0\rangle$, using the property $\alpha_i|0\rangle = 0$.) Once this is done, the Hamiltonian can be rewritten as

$$H = E_0 + \sum_{j,j'} \bar{t}_{jj'} \alpha_j^\dagger \alpha_{j'} + \frac{1}{4} \sum_{jj'kk'} \bar{v}_{jj'kk'} : \alpha_j^\dagger \alpha_{j'}^\dagger \alpha_k \alpha_{k'} : \,. \tag{6.16}$$

The next step is to find a way of computing transition matrix elements of the form $\langle m|J^\mu(x)|0\rangle$ without having an explicit expression for the complicated ground state $|0\rangle$. To achieve this, define formally the creation operator for the state $|m\rangle$,

$$Q_m^\dagger \equiv |m\rangle\langle 0| \,. \tag{6.17}$$

This implies for instance $Q_m|0\rangle = 0$. Then

$$[H, Q_m^\dagger]|0\rangle = (E_m - E_0)Q_m^\dagger|0\rangle \equiv \Omega_m Q_m^\dagger|0\rangle \,, \tag{6.18}$$

and therefore, after commuting both sides with an arbitrary operator G and multiplying by $\langle 0|$ from the left,

$$\langle 0|[G,[H, Q_m^\dagger]]|0\rangle = \Omega_m \langle 0|[G, Q_m^\dagger]|0\rangle \,. \tag{6.19}$$

Choosing specifically $G = \alpha_k^\dagger \alpha_{k'}^\dagger$ or $G = \alpha_j \alpha_{j'}$, this leads to

$$\begin{aligned} \langle 0|[\alpha_k^\dagger \alpha_{k'}^\dagger,[H, Q_m^\dagger]]|0\rangle &= \Omega_m \langle 0|[\alpha_k^\dagger \alpha_{k'}^\dagger, Q_m^\dagger]|0\rangle \,, \\ \langle 0|[\alpha_j \alpha_{j'},[H, Q_m^\dagger]]|0\rangle &= \Omega_m \langle 0|[\alpha_j \alpha_{j'}, Q_m^\dagger]|0\rangle \,. \end{aligned} \tag{6.20}$$

The next step is to make an *ansatz* for $Q_m^\dagger$:

$$Q_m^\dagger = \sum_{kk'} X_{kk'}^m \alpha_k^\dagger \alpha_{k'}^\dagger - Y_{kk'}^m \alpha_k \alpha_{k'} \,. \tag{6.21}$$

In other words, we assume that $Q_m^\dagger$ always creates or destroys two states simultaneously. This makes sense because whenever we create a hole state by removing a nucleon from its place in the ground state, that nucleon has to go somewhere else, namely into some higher lying one-particle state. The interpretation of this ansatz is as follows: the coefficients $X_{kk'}^m$ ($Y_{kk'}^m$) describe the amplitude of a particular combination of states $\alpha_k^\dagger \alpha_{k'}^\dagger$ ($\alpha_k \alpha_{k'}$) within the complicated state $|m\rangle$. This can be seen by using

$$\langle 0|\alpha_k^\dagger \alpha_{k'}^\dagger|m\rangle = \langle 0|[\alpha_k^\dagger \alpha_{k'}^\dagger, Q_m^\dagger]|0\rangle \simeq X_{kk'}^m \,, \tag{6.22}$$

$$\langle 0|\alpha_k \alpha_{k'}|m\rangle = \langle 0|[\alpha_k \alpha_{k'}, Q_m^\dagger]|0\rangle \simeq Y_{kk'}^m \,. \tag{6.23}$$

Here, we have made a crucial assumption: we assume that the pairs of operators $\alpha_k^\dagger \alpha_{k'}^\dagger$ and $\alpha_k \alpha_{k'}$ (each of which creates or annihilates a two-fermion state) behave like *bosonic* creation and annihilation operators, i.e. that they obey commutation relations of the form

$$[\alpha_k^\dagger \alpha_{k'}^\dagger, \alpha_j \alpha_{j'}] = \delta_{kj}\delta_{k'j'} \,, \tag{6.24}$$

$$[\alpha_k^\dagger \alpha_{k'}^\dagger, \alpha_j^\dagger \alpha_{j'}^\dagger] = 0 \,, \tag{6.25}$$

$$[\alpha_k \alpha_{k'}, \alpha_j^\dagger \alpha_{j'}^\dagger] = 0 \,. \tag{6.26}$$

This approximation is called the *quasi-boson approximation*. With this approximation we can proceed and rewrite eqs. (6.20) as

$$\sum_{kk'} \left(X^m_{kk'} \langle 0|[\alpha^\dagger_j \alpha^\dagger_{j'}, [H, \alpha^\dagger_k \alpha^\dagger_{k'}]]|0\rangle - Y^m_{kk'} \langle 0|[\alpha^\dagger_j \alpha^\dagger_{j'}, [H, \alpha_k \alpha_{k'}]]|0\rangle \right) = -\Omega_m Y^m_{jj'} \,, \tag{6.27}$$

$$\sum_{kk'} \left(X^m_{kk'} \langle 0|[\alpha_j \alpha_{j'}, [H, \alpha^\dagger_k \alpha^\dagger_{k'}]]|0\rangle - Y^m_{kk'} \langle 0|[\alpha_j \alpha_{j'}, [H, \alpha_k \alpha_{k'}]]|0\rangle \right) = -\Omega_m Y^m_{jj'} \,. \tag{6.28}$$

These equations can be solved for $X^m_{kk'}$ and $Y^m_{kk'}$. Given an arbitrary operator—e.g. $J^\mu(x_1)$ given in terms of nucleon creation and annihilation operators c_j, $c^\dagger_j$—we can write it in terms of α_j, $\alpha^\dagger_j$ using eq. (6.13). We can then use eqs. (6.27) and (6.28) to compute the matrix elements $\langle 0|J^\mu(x)|m\rangle$.

6.3 The Schechter-Valle theorem

We have seen above that Majorana neutrino masses lead to neutrinoless double beta decay. However, one could imagine other mechanisms that lead to this process, in particular the exchange of hypothetical new particles other than heavy right handed neutrinos (see fig. 6.3 for a few examples).

The *Schechter-Valle Theorem* states that *any* mechanism generating neutrinoless double beta implies that neutrinos are Majorana particles. This statement is best illustrated by the generic Feynman diagram in fig. 6.4

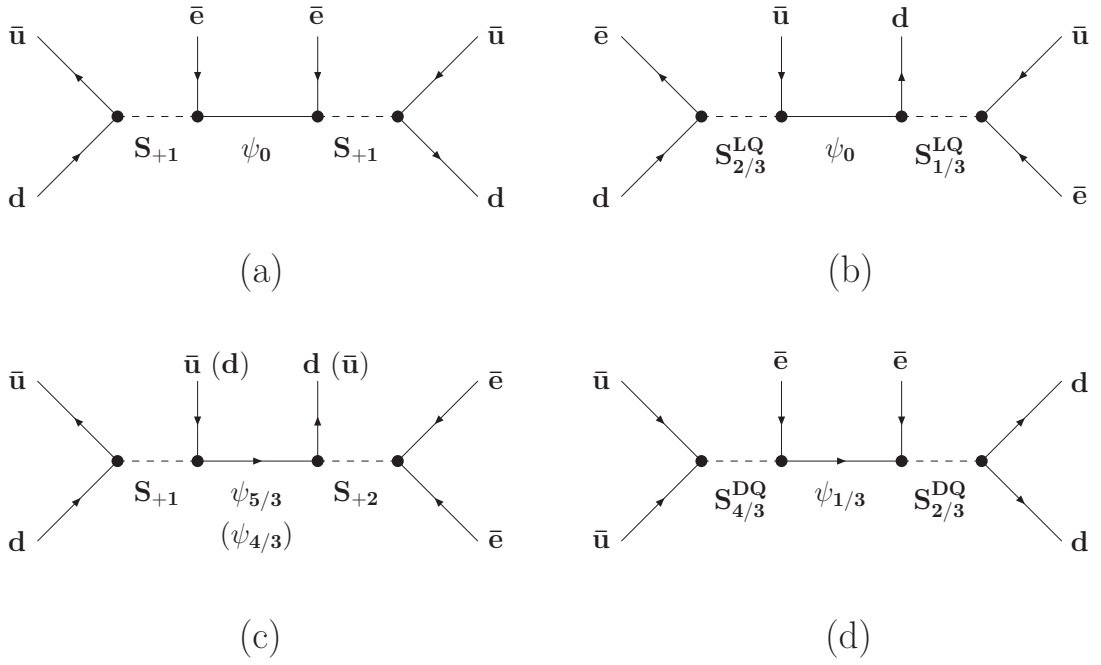


Figure 6.3: Examples for various mechanisms leading to neutrinoless double beta decay. Solid lines correspond to fermions, dashed lines can be scalar or vector bosons. (a) This is the standard mechanism if S_{+1} is the W boson and ψ_0 a heavy Majorana neutrino. (b) A mechanism involving *leptoquarks* (particles that couple a quark to a lepton). (c) involves weird particles with unusual charge assignments. (d) involves *diquarks*, i.e. fields that carry the quantum numbers of two quark (color octect, unusual charge assignments). Figure taken from [11]. (Note that vertices with two incoming or two outgoing arrows imply that one of the fermion fields comes with a charge conjugation operator.)

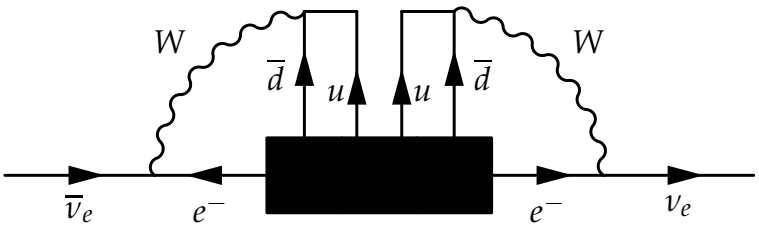


Figure 6.4: Illustration of the Schechter-Valle theorem: any mechanism that generates neutrinoless double beta decay (shown here as a black box) can be turned into a diagram that generates a Majorana mass term for neutrinos.

7 Neutrino mass models

The neutrino sector of the Standard Model is quite unusual: neutrino masses are at least 6 orders of magnitude smaller than the masses of the other fermions, and the mixing angles are much larger than those of the quarks. In some sense, the situation is reminiscent of the situation that chemistry was in the late 19th century: a lot of structure was observed in the system of chemical elements (the periodic table), but the origin of this structure was completely unknown. Naturally, theorists attempt to understand the structure observed in neutrino masses and mixings (and also in the masses and mixing angles of other elementary particles) and relating them to a more fundamental principle.

7.1 The seesaw mechanism

We have already encountered one attempt in this direction: the seesaw mechanism (see sec. 2.5), which we quickly recall here. The Lagrangian was

$$\mathcal{L}_{\text{seesaw}} = -\frac{1}{2}\overline{n^c} M n + h.c.\,, \tag{7.1}$$

with the neutrino fields

$$n = \begin{pmatrix} \nu_L \\ (\nu_R)^c \end{pmatrix} \tag{7.2}$$

and the mass matrix

$$M = \begin{pmatrix} 0 & m_D \\ m_D & m_M \end{pmatrix}. \tag{7.3}$$

Note that the right handed neutrinos ν_R are singlets under the SM gauge group and are therefore, by definition, sterile neutrinos. (However, in conventional seesaw models, they

are considered very heavy and thus not relevant to oscillation experiments.) Assuming $m_M \gg m_D$, we can integrate out the right handed singlet neutrinos ν_R and obtain for the effective Majorana mass of the light neutrinos the seesaw formula

$$m_\nu \simeq \frac{m_D^2}{m_M} \,. \tag{7.4}$$

The heavy neutrino mass is of order m_M. The mixing angles between the light and heavy neutrino states is of order m_D/m_M. In chapter 2.5, we had considered just one generation of neutrinos, but the formalism easily generalizes to 3 generation of left handed neutrinos ($\nu_L \to (\nu_{L,e}, \nu_{L,\mu}, \nu_{L,\tau})$) and k generations of right handed neutrinos ($\nu_r \to (\nu_{R,1}, \cdots \nu_{R,k})$). The mass matrix is then

$$M = \begin{pmatrix} 0 & m_D \\ m_D^T & m_M \end{pmatrix} \,, \tag{7.5}$$

where m_D is a general $3 \times k$ complex matrix and m_M is a $k \times k$ complex symmetric matrix. M is diagonalized according to

$$M_{\text{diag}} = \begin{pmatrix} m_\nu & 0 \\ & m_N \end{pmatrix} = U M U^T \tag{7.6}$$

by the (approximate) matrix

$$U \simeq \begin{pmatrix} 1 & -m_D m_M^{-1} \\ m_D m_M^{-1} & 1 \end{pmatrix} \,. \tag{7.7}$$

The seesaw formula for the effective light neutrino mass matrix becomes

$$m_\nu \simeq -m_D m_M^{-1} m_D^T \,. \tag{7.8}$$

(Minus signs can be removed by field rephasings if necessary.)

For sufficiently large m_M (which one can consider "natural" because it is not protected by any symmetry), the seesaw mechanism explains why neutrino masses are so small, but does not say anything about the hierarchy among the neutrino masses, and it also does not provide information on the mixing angles.

The above mechanism is often called type I seesaw mechanism to distinguish from other, related, mechanism which we will discuss next.

7.2 Variants of the seesaw mechanism

Besides the type I seesaw mechanism discussed in the previous section, several alternative mechanism have been proposed. These are summarized in fig. 7.1.

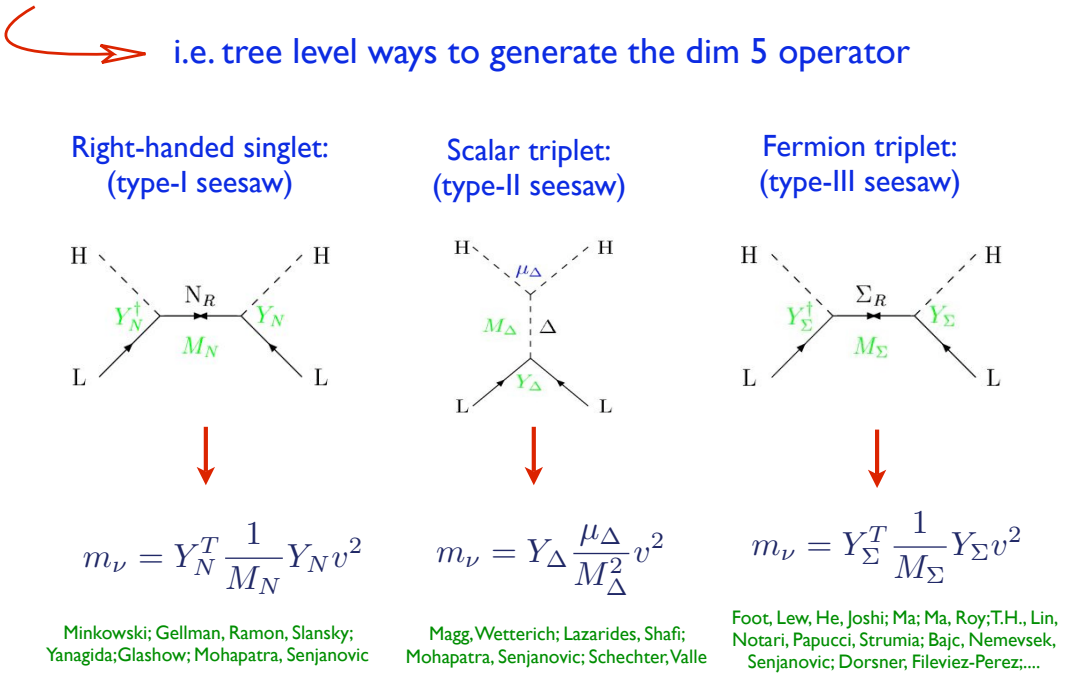


Figure 7.1: The conventional (type I) seesaw mechanism and two of its variants (type II, mediated by a scalar triplet, and type III, mediated by a fermion triplet). All three mechanisms generate the dimension-5 Weinberg operator $\mathcal{L}_{\text{Weinberg}} = \frac{1}{2} c_{\alpha\beta} (\overline{L_\alpha^c} \tilde{H}^*)(\tilde{H}^\dagger L_\beta$, where $\tilde{H} = i\sigma^2 H^*$. Picture taken from [12].

7.2.1 Type II seesaw

In the type II seesaw scenario, there are no heavy (RH) sterile neutrinos. Instead, the Standard Model is augmented by a second Higgs field Δ that is a triplet under $SU(2)_L$ with a hypercharge of $Y = 2$. Since the LH leptons and the Higgs boson are $SU(2)_L$ doublets with $Y = -1$, the following couplings can be constructed:

$$\mathcal{L}_{\text{type II}} \supset y_{\alpha\beta}\overline{L^c_\alpha}\Delta L_\beta + \mu_\Delta \tilde{H}^\dagger \Delta H \,. \tag{7.9}$$

(There are additional couplings in the Higgs sector, which will not be relevant to us here.) Here, we use the definition $\tilde{H} = i\sigma^2 H^*$, which swaps the upper and lower components of the Higgs doublet (putting the vev in the upper one) while making sure the resulting object still transforms in the doublet representation of $SU(2)_L$. The reason these couplings are allowed by the symmetries is that two $SU(2)$ doublets and one triplet can combine into a singlet. To see this, recall that $SU(2)$ representations describe "spins" of elementary particles. Here, of course, the $SU(2)_L$ has nothing to do with the real spin, but rather with the weak isospin. A doublet corresponds to an isospin 1/2 state, a triplet to an isospin 1 state. Just like real spins, isospins can be added according to the usual rules, known from introductory QM courses. And those rules tell us that two isospin 1/2 states and one isospin 1 state can be combined to form an isospin 0 state — an $SU(2)$ singlet. In matrix notation, we can write

$$H = \begin{pmatrix} H^+ \\ H^0 \end{pmatrix} , \qquad L_\alpha = \begin{pmatrix} \nu_{L,\alpha} \\ e_{L,\alpha} \end{pmatrix} , \quad \text{and} \quad \Delta = \begin{pmatrix} \delta^+/\sqrt{2} & \delta^{++} \\ \delta^0 & -\delta^+/\sqrt{2} \end{pmatrix} . \tag{7.10}$$

Under an $SU(2)$ transformation $U = e^{i\alpha^a\sigma^a/2}$ (written here in the doublet representation), these states transform as $H \to UH$, $L_\alpha \to UL_\alpha$ and $\Delta \to U\Delta U^\dagger$.[1]

When H^0 acquires a vacuum expectation value, a Majorana neutrino mass term of the form

$$\mathcal{L}_{\text{mass,type II}} \sim y_{\alpha\beta}\mu_\Delta \frac{v^2}{M_\Delta^2}\overline{\nu^c_{L\alpha}}\nu_{L\beta} \tag{7.11}$$

is generated. If M_Δ is very large, the mass $y_{\alpha\beta}\mu_\Delta$ is naturally small.

7.2.2 Type III seesaw

The type III seesaw scenario replaces the RH singlet neutrinos in the type I seesaw by new fermionic $SU(2)$ triplets $\Sigma_{R,\beta}$. The relevant couplings are

$$\mathcal{L}_{\text{type III}} \supset Y_{\Sigma,\alpha\beta}\, H\, \overline{L^c_\alpha}\, \Sigma_{R,\beta} + \frac{1}{2} M_{\Sigma,\alpha\beta}\, \overline{\Sigma^c_{R,\alpha}}\, \Sigma_{R,\beta} \,, \tag{7.12}$$

and the resulting neutrino mass term is

$$\mathcal{L}_{\text{mass,type III}} \sim v^2 (y_\Sigma^T M_\Sigma^{-1} y_\Sigma)_{\alpha\beta}\, \overline{\nu^c_{L\alpha}}\nu_{L\beta} \,. \tag{7.13}$$

[1] It is an interesting mathematical exercise to show that writing the triplet $\Delta = (\delta^{++}, \delta^+, \delta^0)^T$ as a 2×2 matrix as in eq. (7.10) leads to the transformation rule $\Delta \to U\Delta U^\dagger$.

7.3 Light sterile neutrinos in seesaw scenarios

Within the type-I seesaw model, the assumption that all eigenvalues of m_M are very heavy is a pure theoretical prejudice. Let us assume that, instead, the sterile neutrinos are light, $\mathcal{O}$(eV) and the Yukawa couplings that determine m_D are $\mathcal{O}(10^{-12})$, so that the entries of m_D are of order 0.1 eV. Then, we can still use the approximation $\|m_M\| \gg \|m_D\|$, and we obtain from eq. (7.8) the correct light neutrino masses, while the heavy neutrino mass is of order eV and the mixing is of order $\|m_D\|/\|m_M\| \sim 0.1$. One can argue whether this simple model is theoretical beautiful, but considering only the hard facts, we have to admit that there is nothing that forbids it.

7.4 Flavor symmetries

To understand the origin of the neutrino mixing angles, we have to consider possible relations between the three generations of fermions. This leads to the topic of *flavor symmetries*.

If the fermion mass terms were absent, the Standard Model (including RH neutrinos) would possess a $U(3)^6 = U(3)_Q \times U(3)_U \times U(3)_D \times U(3)_L \times U(3)_E \times U(3)_N$ flavor symmetry. The individual factors in this product group transform the left-handed quarks, the RH up-type quarks, the RH down-type quarks, the LH charged leptons, the RH charged leptons, and the RH neutrinos into each other. To see that this flavor symmetry exists (in the absence of mass terms), just consider that, in the flavor basis, there is no term that transform a particle from one generation to another.

Obviously, the mass matrices break this beautiful symmetry. However, one can hypothesize that, whatever mechanism breaks the flavor group (e.g. some highly complicated new scalar sector) leaves some subgroups intact, just like the breaking of the $SU(2)_L \times U(1)_Y$ symmetr in the SM by the Higgs field leaves the $U(1)_{\rm em}$ subgroup intact. Since $U(3)^6$ has many subgroups, this is a large playground for theorists.

For the lepton sector, written in the *symmetry basis* denoted by indices $a, b, c, \cdots$, the (Dirac) mass terms read

$$\mathcal{L}_\ell \supset (m_\ell)_{ab} \bar{e}_{aL} e_{bR} + (m_D)_{ab} \bar{\nu}_{aL} \nu_{bR} + h.c.\,. \tag{7.14}$$

The mass matrices are diagonalized by writing

$$e_{L,a} = V_{L,ai} e_{Li}\,, \qquad e_{R,a} = V_{R,ai} e_{Ri}\,, \qquad \nu_{L,a} = U_{L,ai} \nu_{Li}\,, \qquad \nu_{R,a} = U_{R,ai} \nu_{Ri}\,, \tag{7.15}$$

with suitably chosen unitary matrices V_L, V_R, U_L, and U_R. Here, we have used the theorem that any matrix M can be diagonalized according to $M_d = V^\dagger M U$, where M_d is diagonal and U, V are unitary. To go to the flavor basis, we would diagonalize only the charged lepton Yukawa matrix. The above transformations imply that the gauge coupling term transforms as

$$\mathcal{L}_W = \frac{g}{\sqrt{2}} \bar{e}_{La} \gamma^\mu \nu_{La} W^-_\mu \quad \to \quad \frac{g}{\sqrt{2}} \bar{e}_{Li} (V_L^\dagger)_{ia} U_{L,aj} \gamma^\mu \nu_{Lj} W^-_\mu\,. \tag{7.16}$$

In other words, the leptonic mixing matrix is given by

$$U_{\text{PMNS}} = V_L^\dagger U_L \,. \tag{7.17}$$

The same result holds for Majorana neutrino masses, except that there is no matrix U_R in this case, and the Majorana mass matrix is diagonalized according to $M_d = U^T M U$.

7.4.1 ν_μ–ν_τ reflection symmetry

As an example, consider a Z_2 symmetry corresponding to the exchange $\nu_\mu \leftrightarrow \nu_\tau$. Let us assume that, for the charged leptons, the symmetry basis, the flavor basis and the mass basis coincide. (Since, by $SU(2)_L$ invariance, the left handed charged leptons should transform in the same way as the left handed neutrinos, obtaining different muon and tau masses requires the right handed charged leptons to transform differently.) The general Majorana neutrino mass term

$$\mathcal{L}_{\nu,\text{general}} \supset (\overline{\nu_e^c}, \overline{\nu_\mu^c}, \overline{\nu_\tau^c})^T \begin{pmatrix} m_{ee} & m_{e\mu} & m_{e\tau} \\ m_{e\mu} & m_{\mu\mu} & m_{\mu\tau} \\ m_{e\tau} & m_{\mu\tau} & m_{\tau\tau} \end{pmatrix} \begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} \tag{7.18}$$

is restricted by the symmetry to have the form

$$\mathcal{L}_{\mu-\tau} \supset (\overline{\nu_e^c}, \overline{\nu_\mu^c}, \overline{\nu_\tau^c})^T \begin{pmatrix} m_{ee} & m_{e\mu} & m_{e\mu} \\ m_{e\mu} & m_{\mu\mu} & m_{\mu\tau} \\ m_{e\mu} & m_{\mu\tau} & m_{\mu\mu} \end{pmatrix} \begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} \,. \tag{7.19}$$

The mass matrix m_ν in this expression is diagonalized via $m_{\nu,\text{diag}} = U^T m_\nu U$ by the unitary mixing matrix

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \cos\theta_{12} & \sin\theta_{12} & \\ -\sin\theta_{12} & \cos\theta_{12} & \\ & & 1 \end{pmatrix} \tag{7.20}$$

with $\arctan 2\theta_{12} = 2\sqrt{2} m_{e\mu}/(m_{ee} - m_{\mu\mu} - m_{\mu\tau})$. Thus, μ–τ symmetry predicts maximal (23) mixing, in agreement with the data.

7.4.2 Bimaximal and tribimaximal mixing

With more elaborate flavor symmetries, more structure is predicted for the leptonic mixing matrix. Commonly encountered patterns include *bimaximal mixing*,

$$U_{\text{bimaximal}} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{7.21}$$

with $\theta_{12} = \pi/4$, $\theta_{13} = 0$ and $\theta_{23} = \pi/4$, and *tribimaximal mixing*

$$U_{\text{tribimaximal}} = \begin{pmatrix} \sqrt{\frac{2}{3}} & \sqrt{\frac{1}{3}} & 0 \\ -\sqrt{\frac{1}{6}} & \sqrt{\frac{1}{3}} & -\frac{1}{\sqrt{2}} \\ -\sqrt{\frac{1}{6}} & \sqrt{\frac{1}{3}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{7.22}$$

with $\tan^2\theta_{12} = 1/2$, $\theta_{13} = 0$, $\theta_{23} = \pi/4$.

8

High energy astrophysical neutrinos

The physics of ultra-high energy cosmic rays (UHE CRs) is of great interest in particle physics and astrophysics. First, cosmic rays can have energies up to $\sim 10^{19}$ eV (10 EeV) significantly larger than what can be achieved in particle accelerators. Thus, their interactions with nucleons in the atmosphere probe QCD in an unexplored energy range. For instance, the center of mass energy for the collision of a $E_{\rm CR} = 10^{19}$ eV cosmic ray proton with an atmospheric proton at rest is $\sqrt{2m_p E_{\rm CR}} \sim 100$ TeV.

Second, the objects that can accelerate particles to 10^{19} eV are of great astrophysical interest. In fact, it is not known which objects these are, with candidates being active galactic nuclei (AGNs) and gamma ray bursts (GRBs). The reason the sources of cosmic rays are still unknown is that the highest energy particles are mostly protons and/or atomic nuclei, which are deflected in galactic and intergalactic magnetic fields. Hence, their arrival direction does not point back to their sources. Photons and neutrinos, on the other hand, would point back to the source. Photons are more likely to interact and undergo electromagnetic cascades while they propagate. Morever, photons can be produced by electrons and hadrons, while neutrinos come only from hadronic processes. Thus, photons and neutrinos can provide complementary information about cosmic ray sources, and the search for UHE cosmic neutrinos is among the most promising ways of reveiling the sources of UHE CRs.

In the following, we first discuss how charged particles are accelerate in astrophysical sources (for obvious reasons, neutral particles cannot be directly accelerated). Then, we show how interactions of charged cosmic rays within the CR sources produce UHE neutrinos.

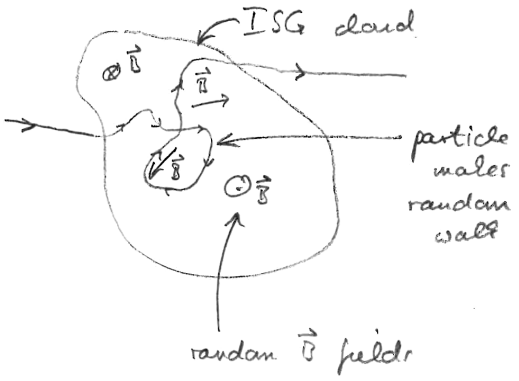


Figure 8.1: An ultra high energy cosmic ray particle undergoing a random walk due to the magnetic fields in an interstellar gas cloud.

8.1 Acceleration of cosmic rays: the Fermi mechanism

Consider a dilute ($\sim 1\,\text{particle/cm}^3$) astrophysical gas cloud permeated by stochastic magnetic fields. A particle travelling through the gas cloud will be affected by the magnetic fields, thus its direction will change while its energy will remain constant (see fig. 8.1. Now, assume the gas cloud is moving with a velocity $\mathbf{u}$ relative to the observer frame.

8.1.1 Non-relativistic toy model

Assume a particle enters the gas cloud with observer frame velocity $\mathbf{v}_i$, performs a random walk inside the cloud, and eventually leaves it with velocity $\mathbf{v}_f$ (see fig. 8.1). The corresponding quantities in the rest frame of the cloud are denoted by

$$\mathbf{v}'_i = \mathbf{v}_i - \mathbf{u} \tag{8.1}$$

$$\mathbf{v}'_f = \mathbf{v}_f - \mathbf{u}\,. \tag{8.2}$$

Since magnetic fields change only the direction of momentum, but not its modulus, we have

$$|\mathbf{v}'_i| = |\mathbf{v}'_f|\,. \tag{8.3}$$

Therefore,

$$\Delta E = \frac{1}{2}m(\mathbf{v}_f^2 - \mathbf{v}_i^2) \tag{8.4}$$

$$= \frac{1}{2}m(\mathbf{v}_f'^2 + 2\mathbf{v}'_f\mathbf{u} + \mathbf{u}^2 - \mathbf{v}_i'^2 - 2\mathbf{v}'_i\mathbf{u} - \mathbf{u}^2) \tag{8.5}$$

$$= m(\mathbf{v}'_f - \mathbf{v}'_i)\mathbf{u}\,. \tag{8.6}$$

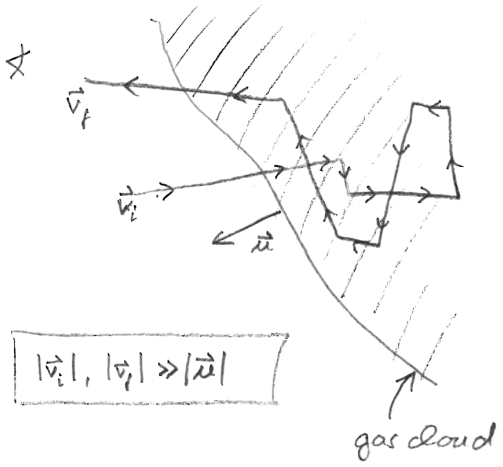


Figure 8.2: An ultra high energy cosmic ray entering a gas cloud, scattering multiple times on random magnetic fields, and leaving again. The figure shows the notation used for nonrelativistic toy model of Fermi accelartion, see sec. 8.1.1.

Let us consider the following limiting cases

1. If the particle does not change direction, i.e. $\mathbf{v}'_i \uparrow\uparrow \mathbf{v}'_f$, we have $\Delta E = 0$.
2. For head-on collisions with reflections ($\mathbf{v}'_i \uparrow\downarrow \mathbf{v}'_f$ and $\mathbf{v}_i \uparrow\downarrow \mathbf{u}$), we have $\mathbf{v}'_i = -\mathbf{v}'_f$ and therefore
$$\Delta E = 2m(\mathbf{u}^2 - \mathbf{v}_i\mathbf{u}) > 0\,. \tag{8.7}$$
3. For rear-end collisions with reflections ($\mathbf{v}'_i \uparrow\downarrow \mathbf{v}'_f$ and $\mathbf{v}_i \uparrow\uparrow \mathbf{u}$), we have also $\mathbf{v}'_i = -\mathbf{v}'_f$, but
$$\Delta E = 2m(\mathbf{u}^2 - \mathbf{v}_i\mathbf{u}) < 0\,. \tag{8.8}$$

The crucial observation is that the energy gain in head-on collisions is larger than the energy loss in rear-end collisions. Thus, on average, the particle gains energy, and after many collisions with gas clous, it can reach very high energies.

8.1.2 Relativistic model

In the relativistic regime, we need to be a bit more careful. Using again primed symbols for quantities expressed in the rest frame of the gas cloud and unprimed symbols for quantities expressed in the lab frame, we have for the energies of the particle when

Figure 8.3: An ultra high energy cosmic ray entering a gas cloud, scattering multiple times on random magnetic fields, and leaving again. The figure shows the notation used in the relativistic description of Fermi acceleration, see sec. 8.1.2.

entering the cloud (E_i, E_i') and when leaving the cloud (E_f, E_f'):

$$E_i' = E_i\gamma - u\gamma p_i \cos\theta_i \tag{8.9}$$
$$\simeq E_i\gamma(1 - u\cos\theta_i) \tag{8.10}$$

and

$$E_f \simeq E_f'\gamma(1 + u\cos\theta_f')\,. \tag{8.11}$$

(See fig. 8.3.) Here, $\gamma = (1-u^2)^{-1/2}$ is the boost factor of the cloud, $p_i \simeq E_i$ is the particle's lab frame momentum when entering the cloud, and θ_i, θ_f' are angles relative to the velocity vector $\mathbf{u}$. Since scattering processes with the magnetic fields inside the cloud are elastic, the relation

$$E_i' = E_f' \tag{8.12}$$

has to hold. Thus,

$$E_f = E_i\gamma^2(1 - u\cos\theta_i)(1 + u\cos\theta_f')\,, \tag{8.13}$$

which leads to

$$\frac{\Delta E}{E} = \gamma^2\left[1 - \frac{1}{\gamma^2} + u(\cos\theta_f' - \cos\theta_i) - u^2\cos\theta_i\cos\theta_f'\right]. \tag{8.14}$$

We are interested in the average energy gain $\langle \Delta E/E\rangle$, i.e. we need the average angles. By assumption, $\left\langle \cos\theta_f' \right\rangle = 0$ since particles are assumed to undergo a random walk inside the cloud, so that the distribution of their momentum directions is isotropic afterwards. Assuming that the momentum directions of particles outside the cloud are also isotropically distributed, the number of particles reaching the cloud per time interval δt under an angle θ_i is

$$dN \propto (1 - u\cos\theta_i)\delta t\,. \tag{8.15}$$

This equation means that the number of particles chasing the cloud and entering it from behind (rear-end collision) is slightly smaller than the number of particles entering the cloud from the front. Therefore,

$$\langle\cos\theta_i\rangle = \frac{\int_{-1}^{1} d(\cos\theta_i)\ \cos\theta_i(1 - u\cos\theta_i)}{\int_{-1}^{1} d(\cos\theta_i)(1 - u\cos\theta_i)} \tag{8.16}$$
$$= -\frac{\frac{2}{3}u}{2} = -\frac{u}{3}\,, \tag{8.17}$$

and thus

$$\left\langle \frac{\Delta E}{E} \right\rangle = \frac{1}{1-u^2}\left[u^2 + \frac{u^2}{3}\right] \simeq \frac{4}{3}u^2 > 0\,. \tag{8.18}$$

Again, we find on averag an energy *gain*. Since $\langle \Delta E/E \rangle$ is second order in u, this mechanism is called *second order Fermi acceleration*.

Typically, $u \sim 10\,\mathrm{km/sec}$, which implies $\langle \Delta E/E \rangle \sim 10^{-8}$. The typical distance between gas clouds is of order light years, so that the total duration of a typical acceleration process is of order $\tau_{\mathrm{acc}} \sim 10^8\,\mathrm{yrs}$.

8.1.3 Final energy spectrum

To obtain the energy spectrum $N' \equiv \partial N(E,t)/\partial E$ of cosmic rays accelerated by second order Fermi acceleration, we consider the diffusion equation, assuming an equilibrium of particle gain and loss processes:

$$0 = \frac{d}{dt}\frac{\partial N(E,t)}{\partial E} = \frac{1}{\partial E}\frac{dN(E,t)}{dt} \tag{8.19}$$

$$= \frac{\partial^2 N(E,t)}{\partial t\,\partial E} + \frac{\partial}{\partial E}\left(\frac{\partial N(E,t)}{\partial E}\frac{dE}{dt}\right) \tag{8.20}$$

$$\simeq -\frac{1}{\tau_{\mathrm{esc}}}\frac{\partial N(E,t)}{\partial E} - \frac{\partial}{\partial E}\left(\frac{E}{\tau_{\mathrm{acc}}}\frac{\partial N(E,t)}{\partial E}\right), \tag{8.21}$$

where τ_{esc} is the escape time, i.e. the average time it takes a particles to leave the acceleration region. With the abbreviation $N' \equiv= \partial N(E,t)/\partial E$, we obtain

$$-\frac{N'}{\tau_{\mathrm{esc}}} - \frac{N'}{\tau_{\mathrm{acc}}} = \frac{E}{\tau_{\mathrm{acc}}}\frac{dN'}{dE} \tag{8.22}$$

or

$$\frac{dN'}{dE} = -\frac{N'}{E}\left(1 + \frac{\tau_{\mathrm{acc}}}{\tau_{\mathrm{esc}}}\right). \tag{8.23}$$

The solution of the differential equation is

$$N' \propto E^{-\alpha} \quad \text{with} \quad \alpha \simeq 1 + \frac{\tau_{\mathrm{acc}}}{\tau_{\mathrm{esc}}}. \tag{8.24}$$

Therefore, the generic expectation for the cosmic ray spectrum is a power law. This is indeed the case experimentally. The power law index α depends on the velocity u of the accelerating gas clouds ($\tau_{\mathrm{acc}} \sim 1/u^2$) and on the density and size of the acceleration region (through τ_{esc}).

8.1.4 Diffusive shock acceleration (first order Fermi acceleration)

Consider now a spherical shock wave rather than random clouds, e.g. from a supernova remnant (see fig. 8.4). Particles reaching the shock wave from the outside experience only head-on collisions, no rear-end collision. From eq. (8.14), we have

$$\frac{\Delta E}{E} = \gamma^2 u(\cos\theta'_f - \cos\theta_i) + \mathcal{O}(u^2), \tag{8.25}$$

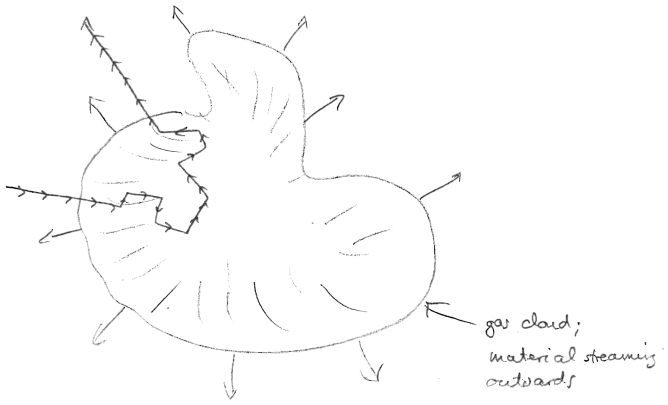


Figure 8.4: Schematic illustration of first order Fermi acceleration. Cosmic ray particles approaching an expanding gas cloud (e.g. a supernova remnant) from far will always experience head-on collisions. Therefore, they gain energy in each collision, making this mechanism more efficient than second order Fermi acceleration, where rear-end collisions are possible, leading to energy loss.

where $u \ll 1$ should be intepreted as the difference of the matter velocities upstream and downstream of the shock wave. The number of particles reaching the shock wave per time interval dt per solid angle interval $d\Omega$ is now

$$dN \propto \cos\theta_i \, d\Omega \, dt \,, \tag{8.26}$$

which leads to

$$\langle \cos\theta_i \rangle \simeq \frac{\int_{-1}^{0} d(\cos\theta_i) \, \cos^2\theta_i}{\int_{-1}^{0} d(\cos\theta_i) \, \cos\theta_i} = -\frac{2}{3} \,. \tag{8.27}$$

By a similar argument,

$$\left\langle \cos\theta'_f \right\rangle = +\frac{2}{3} \,, \tag{8.28}$$

and thus

$$\left\langle \frac{\Delta E}{E} \right\rangle = \frac{4}{3} u > 0 \,. \tag{8.29}$$

Since $\langle \Delta E/E \rangle$ is first order in u, this mechanism is called *first order Fermi acceleration.* The average enegry gain is much larger than for second order Fermi acceleration.

After n acceleration cycles, the particle's energy is

$$E_n = \left(1 + \frac{4}{3} u\right)^n E_0 \equiv (1+k)^n E_0 \,, \tag{8.30}$$

where E_0 is its initial energy. To achieve a final energy E, we therefore need $\log(E/E_0)/\log(1+k)$ acceleration cycles. The number of particles with energy $\geq E$ is then

$$N(\geq E) = N_0(1-P_{\rm esc})^{\frac{\log E/E_0}{\log(1+k)}} \tag{8.31}$$

$$= N_0\left(\frac{E}{E_0}\right)^{\frac{\log(1-P_{\rm esc})}{\log(1+k)}} . \tag{8.32}$$

Expanding $\log(1-P_{\rm esc}) \simeq -P_{\rm esc}$ and $\log(1+k) \simeq k$ in the exponent, we obtain

$$N(\geq E) \simeq N_0\left(\frac{E}{E_0}\right)^{-P_{\rm esc}/k} \tag{8.33}$$

and thus

$$N(E) = \frac{dN(\geq E)}{dE} \propto \left(\frac{E}{E_0}\right)^{-\alpha} \quad \text{width} \quad \alpha = \frac{P_{\rm esc}}{k} + 1\,. \tag{8.34}$$

Using hydrodynamic conservation laws, one can show that, for strong shocks ($v_{\rm upstream} \gg v_{\rm downstream}$), $\alpha \simeq 2$.

8.2 Neutrino production and the Waxman-Bahcall bound

From observations of UHE charged cosmic rays, one can derive that the production rate of these charged cosmic rays (mainly protons) in the Universe is

$$E_p^2 \frac{d\dot{N}(E_p)}{dE_p} \simeq 10^{44}\ {\rm erg\,Mpc^{-3}\,yr^{-1}}\,. \tag{8.35}$$

We consider here in particular cosmic ray protons with energy $> 10^{19}$ eV, whose main interaction channel is with CMB photons and photons produced in the same source

$$p + \gamma \to p/n + \pi^0/\pi^+ \tag{8.36}$$

(mainly through the Δ resonance). From this, one can derive that the energy density in (muon) neutrinos today is of order

$$E_{\nu_\mu}^2 \frac{dN(E_{\nu_\mu})}{dE_{\nu_\mu}} \simeq 0.25\,\epsilon\,H_0^{-1} E_p^2 \frac{d\dot{N}(E_p)}{dE_p}\,, \tag{8.37}$$

where $H_0^{-1} \sim 10^{10}$ yrs (the inverse of the Hubble constant) is a measure for the age of the Universe and ϵ is the average fraction of the proton energy that is lost in interactions of the form (8.36). (We assume this fraction to be energy-independent.) The factor 0.25 accounts for the fact that the probability for neutral meson production (which does not yield any neutrinos) and for charged meson production is roughly the same, and that, in

the decay chain $\pi^+ \to \mu^+ + \nu_\mu \to e^+ + \bar{\nu}_\mu + \nu_e + \nu_\mu$, roughly half of the pion's energy goes to muon neutrinos. (We consider only muon neutrinos here because they are easiest to distinguish from background experimentally.)

By setting $\epsilon = 1$ in eq. (8.37), we obtain an upper bound on $E_\nu^2\, dN(E_\nu)/dE_\nu$ and thus on the muon neutrino flux Φ_{ν_μ}:

$$E_{\nu_\mu}^2 \Phi_{\nu_\mu} \lesssim 0.25 \frac{c}{4\pi} H_0^{-1} E_p^2 \frac{d\dot{N}(E_p)}{dE_p} \tag{8.38}$$

$$\simeq 1.5 \times 10^8 \,\mathrm{GeV\ cm^{-2}\ sec^{-1}\ sr^{-1}}\,. \tag{8.39}$$

This constraint is called the *Wavman-Bahcall bound* [13].

9
Neutrinos in cosmology

Neutrinos play an important role in the evolution of the early Universe. An understanding of neutrino physics is therefore necessary to understand the origins of the Universe. Vice-versa, cosmological observations can be used to constrain neutrino properties. In this chapter, we will first give a brief overview of cosmology and then discuss two types of observations crucial to neutrino physics: the measurement of the energy density in relativistic particles N_{eff} using cosmic microwave background (CMB) data and the measurement of the neutrino mass using large scale structure observations. Much of this chapter will be based on [14].

9.1 A brief overview of Big Bang cosmology

The early history of the Universe is summarized in fig. 9.1. According to the best current theories, the earliest epoch relevant to the world today is *inflation*, a phase of very rapid expansion, during which the Universe was expanding according to an exponential law (size of Universe $\propto \exp(Ht)$. After the end of inflation, the expansion slowed down to a power law.

The Universe was initially extremely hot and then cooled down adiabatically as it expanded (same principle as in a refrigerator). Due to the high temperature, all known particle species, plus possible new particle species such as dark matter, were in thermal equilibrium very early on. Thermal equilibrium means that the energy density was evenly distributed among all degrees of freedom (i.e. all particle species). As the Universe cooled down to a temperature below $\sim 0.1\,\text{GeV}$, the binding energy of nucleons, the plasma of quarks and gluons combined into protons and neutrons.

At temperatures below 0.1 MeV (nuclear binding energy), these nucleons combine to form nuclei. This process, which happened a few minutes after the Big Bang, is called *Big Bang Nucleosynthesis* (BBN). It resulted mostly in deuterium and helium nuclei, but

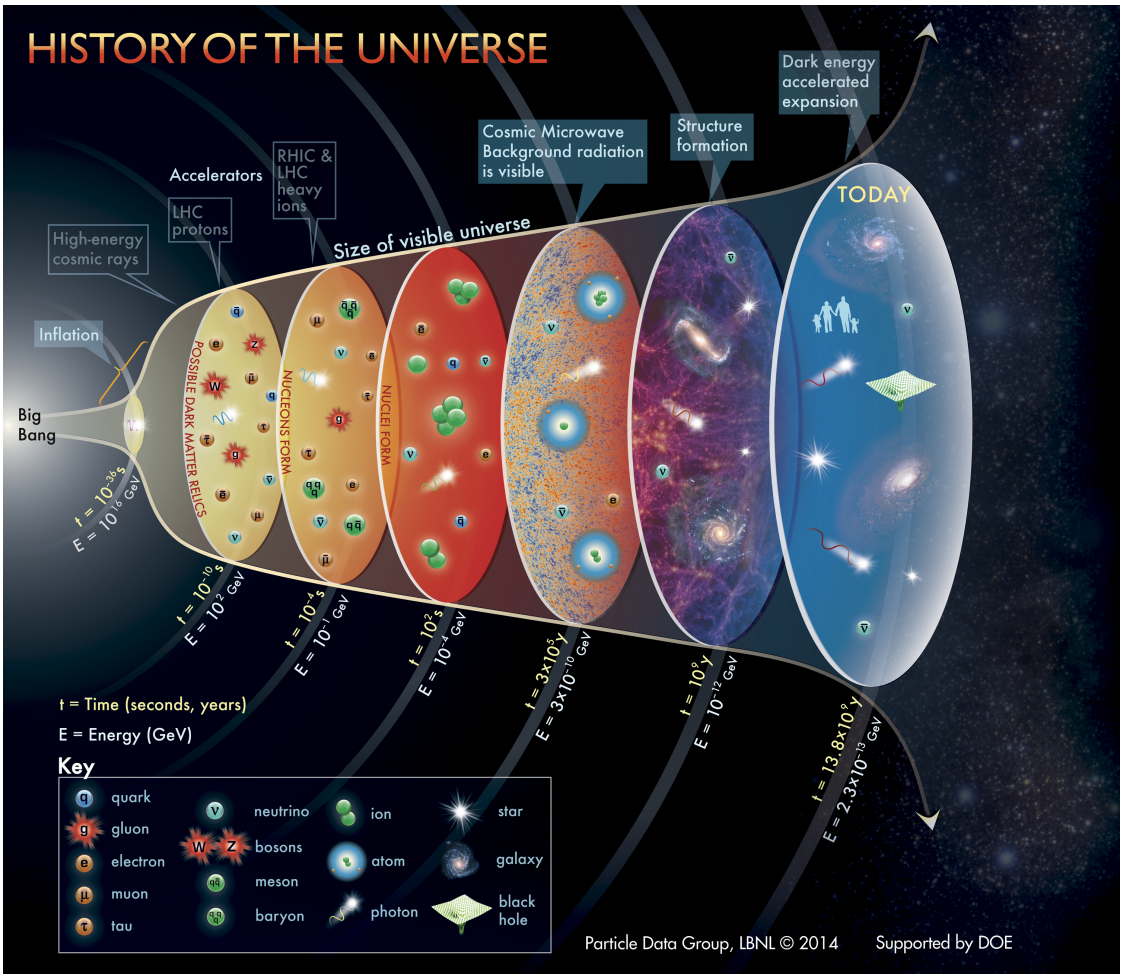


Figure 9.1: History of the early Universe. Image created by the Particle Data Group.

also some heavier elements were produced.

The next important step occured at energies $\lesssim 1\,\text{eV}$, about 300 000 years after the Big Bang. At this time, free electrons and protons *recombined* to form hydrogen nuclei. This is particularly important because this meant that the Universe turned from an opaque plasma to a transparent medium. In other words, after recombination, light could travel unimpeded. Therefore, this is the earliest epoch from which we get a photon signal: the cosmic microwave background (CMB). The temperature of this radiation was originally 0.26 eV, corresponding to the temperature of the Universe at recombination. Due to the expansion of the Universe, which also streteches electromagnetic (or any other) waves, its wavelength has increased by a factor $\sim 1\,100$ since then and its energy or temperature has correspondingly dropped to 2.725 K.

Already before recombination, small fluctuations in the energy density of the early Universe start to grow since overdense regions tend to gravitationally attract more and more matter. But it is not until long after recombination, at about 10^9 yrs after the Big Bang, that density contrasts have grown to the extent that the first galaxies and stars form.

9.2 Big Bang Theory

At the foundation of cosmology is the *Friedmann equation*, which described the expansion of the Universe. It is based on the (observationally well-motivated) assumption that the Universe on large scales is *isotropic* and *homogeneous*. Homogeneity implies that the metric $g_{\mu\nu}$ should be independent of the space coordinate $\mathbf{x}$, while isotropy means that no spatial direction should be special. This implies that $g_{\mu\nu}$ should have the form

$$(g_{\mu\nu}) = \begin{pmatrix} 1 & & & \\ & -a(t) & & \\ & & -a(t) & \\ & & & -a(t) \end{pmatrix} , \tag{9.1}$$

where $a(t)$ is the *scale factor* of the Universe. This metric is called the *Friedmann-Robertson-Walker metric*. We take $a(t_0) = 1$ today ($t = t_0$), and $a(t) < 1$ at earlier times. This implies that two objects that are, say, 1 Mpc apart today where $a(t)$ Mpc apart at an earlier time t. Note that it is not the objects that are moving, but the space between them is expanding. This is analogous to drawing two dots on a rubber sheet and then stretching the rubber sheet. The dots are fixed to the sheet, but the sheet between them is expanding, so their distance increases.

The next step is to plug the ansatz (9.1) into the Einstein equations

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} \, . \tag{9.2}$$

One then uses the fact that also the energy-momentum tensor $T_{\mu\nu}$ should obey homogeneity and isotropy, i.e. it should have the form

$$(T^{\mu}_{\nu}) = \begin{pmatrix} \rho(t) & & & \\ & p(t) & & \\ & & p(t) & \\ & & & p(t) \end{pmatrix} , \tag{9.3}$$

where $\rho(t)$ is the energy density and $p(t)$ is the pressure. As before homogeneity implies that ρ and p should be independent of $\mathbf{x}$, and isotropy implies that all the diagonal spatial components of $T_{\mu\nu}$ should be the same and the off-diagonal ones should be zero. One can then solve the Einstein equations. Their (00) component leads to the *Friedmann equation*, probably the most important equation in cosmology:

$$\left(\frac{\dot{a}}{a} \right)^2 = \frac{8}{3} \pi G \rho \, . \tag{9.4}$$

It tells us that the expansion rate $\dot{a}/a$ is larger the larger the energy density ρ is. Since the energy density was highest as very early times, the expansion was fastest then and has since slowed down. The expansion rate is also known as the *Hubble parameter* and its value today is also called the *Hubble constant*.

The spatial (*ii*) components of the Einstein equations lead to

$$2\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} = -8\pi G p\,. \tag{9.5}$$

Subtracting (9.4) from (9.5), one finds the *Friedmann-Lemaître equation*

$$\frac{\ddot{a}}{a} = -\frac{4}{3}\pi G(\rho + 3p)\,. \tag{9.6}$$

Let us now solve the Friedmann equation in two special cases: the first one is a *radiation-dominated* Universe, i.e. a Universe in which most of the energy is in the form of kinetic energy of relativistic particles. This is the state that our Universe was in at early times (until temperatures $\sim 1\,\mathrm{eV}$). In this case, the energy density scales as $\rho \propto a^{-4}$. Three powers of a are coming from the fact that any given volume element expans as a^3, and the fourth power of a comes from the redshifting of the de Broglie waves of the particles. We thus have from eq. (9.4) $\dot{a} \propto a^{-1}$, which can be solved by separation of variables and leads to

$$a(t) \propto t^{1/2}\,. \qquad \text{(radiation-dominated Universe)} \tag{9.7}$$

The second special case is a Universe filled with non-relativistic matter, where most of the energy is in the form of particle masses. In such a Universe, the energy density scales simply as $\rho \propto a^{-3}$, a simple geometric scaling. We then have $\dot{a} \propto a^{-1/2}$ and thus

$$a(t) \propto t^{2/3}\,. \qquad \text{(matter-dominated Universe)} \tag{9.8}$$

The fact that the energy density scales as a^{-4} for radiation, but only as a^{-3} for non-relativistic matter also has the important consequence that, early on, the Universe was radiation-dominated. At $\sim 100\,000$ years after the Big Bang, the radiation density was redshifted away to the extent that the non-relativistic matter took over as the dominant contributor to the energy density.

9.3 The Cosmic Neutrino Background

As the Universe cools down and the primordial gas becomes more dilute, interactions among particles become weaker and weaker. This happens on the one hand because the probably for two particles to get close enough to interact with each other. On the other hand, many interaction cross sections get smaller at lower energy. This is in particular true for weak interactions between neutrinos and other particles, whose cross section scale as $\sigma \propto G_F^2 E^2$.

Early on ($T \gtrsim$ MeV), neutrino–quark and neutrino–lepton interactions kept the neutrinos in full thermal equilibrium with the primordial soup. This in particular means that their phase space distribution was given by the Fermi-Dirac distribution

$$f_{\mathrm{FD}}(\mathbf{p}) = \frac{1}{\exp(p/T)+1}\,, \tag{9.9}$$

This implies that the neutrino number density was given by

$$n_\nu = \frac{g}{(2\pi)^3}\int_0^\infty d^3p\,\frac{1}{\exp(p/T)+1} \tag{9.10}$$

$$= \frac{3}{4}\frac{\zeta(3)}{\pi^2}gT^3\,. \tag{9.11}$$

and their energy density was

$$\rho_\nu = \frac{g}{(2\pi)^3}\int_0^\infty d^3p\,\frac{p}{\exp(p/T)+1} \tag{9.12}$$

$$= \frac{7}{8}\frac{\pi^2}{30}gT^4\,. \tag{9.13}$$

Here, g denotes the number of internal degrees of freedom (2 for a left-handed neutrino plus its right-handed antineutrino) and ζ is the Riemann zeta function, which happens to be the result of the integral.

Very roughly, the interaction rate for processes like

$$\nu + N \leftrightarrow \nu + N \tag{9.14}$$

$$\nu + \nu \leftrightarrow e^+e^- \tag{9.15}$$

$$\nu + N \leftrightarrow e + N' \tag{9.16}$$

is given by

$$\Gamma_\nu \sim n_\nu G_F^2 T^2 \sim G_F^2 T^5\,. \tag{9.17}$$

Interactions cease when their rate becomes smaller than the Hubble expansion rate H, i.e. the rate at which particles move away from each other due to cosmic expansion. H, can be written according to eq. (9.4) as

$$H \sim \sqrt{g_* G}\,T^2\,, \tag{9.18}$$

where g_* is the number of relativistic degrees of freedom. (Each relativistic degree of freedom has an energy density similar to (9.13), i.e. proportional to T^4. We use the fact that, at the epoch of interest to us, the Universe is radiation dominated, so non-relativistic degrees of freedom are negligible.) Setting $g_* \sim 10$ (electrons, positrons, photons, neutrinos) and equating eqs. (9.17) and (9.18), we can solve for T and find[1]

$$T_{\mathrm{fo}} \sim \mathrm{MeV}\,. \tag{9.19}$$

[1] Here and in the following, an index "fo" on a quantity denotes the value of that quantity at the time of neutrino freeze-out.

After neutrino interactions freeze out, they simply stream freely through the Universe, and their momenta get redshifted due to Hubble expansion. If the Universe has expanded by a factor $a/a_{\rm fo}$, the momentum distribution of neutrinos is thus

$$f'_{\rm FD}(\mathbf{p}) = \frac{1}{\exp[(a/a_{\rm fo}p/T_{\rm fo}] + 1} = \frac{1}{\exp[p/T'] + 1} \tag{9.20}$$

with

$$T' = T_{\rm fo}\frac{a_{\rm fo}}{a}\,. \tag{9.21}$$

In other words, while neutrinos are still relativistic, they still follow a Fermi-Dirac distribution and their number density and energy density are given by expressions of the form (9.11) and (9.13), respectively, but with T replaced by T'.

Note that, as we have argued above, $\rho \sim T^4$ and $\rho \sim a^{-4}$ for relativistic particles, also the temperature of particles still in thermal equilibrium changes as $T \propto a^{-1}$. Therefore, the neutrino temperature T' actually remains identical to the temperature of the thermal bath even after neutrinos have decoupled!

Modifications to this rule arise only when particles disappear from the thermal bath and their energy density gets distributed among the remaining particles. This happens in particular when the temperature drops significantly below the mass of the positron. Then, most electron–positron pairs will annihilate away and their energy gets converted into photons. This slows down the cooling of the photons compared to the cooling of the neutrinos. A more detailed calculation shows that, after e^+e^- annihilation, the temperature ratio between neutrinos and photons is

$$\frac{T_\nu}{T_\gamma} = \left(\frac{4}{11}\right)^{1/3}. \tag{9.22}$$

We know that the temperature of the primordial photons (i.e. the photons that are left over from the time of the Big Bang, known as the *Cosmic Microwave Background (CMB)*) today is 2.73 Kelvin. We thus conclude that the temperature of the *Cosmic Neutrino Background (CνB)* is $(4/11)^{1/3} \times 2.73$ Kelvin = 1.95 Kelvin. The CNB density is thus, acccording to eq. (9.11) of order 340 cm^{-3}. (Here, we use the fact that there are three neutrino flavors, each of which has two degrees of freedom.)

9.4 The Cosmic Microwave Background and the effective number of neutrino species

The richest treasure trove in modern cosmology is the cosmic microwave background (CMB). This is the thermal radiation that was emitted when electrons and atomic nuclei had recombined into neutral atoms, so that the Universe became transparent. This happend when the Universe was between 300 000 and 400 000 years old and its temperature was $\sim$ 0.3 eV. Since the photons, the electrons and the atomic nuclei were, to a very

good approximation, in thermal equilibrium at the time, the emitted radiation had a black body form, i.e. it was given by the Bose-Einstein distribution. As we have argued above, even after a relativistic particle species like the photons have decoupled, their distribution remains thermal, and only the effective temperature changes due to redshift. Therefore, today, the CMB temperature has dropped to 2.73 Kelvin, i.e. the redshift of the CMB is $\sim 1\,100$.

In a very rough approximation, the dynamics of CMB decoupling can be described in the following way: the number density of a non-relativistic particle species in thermal equilibrium is

$$n = g\left(\frac{mT}{2\pi}\right)^{3/2} \exp[-m/T]\,, \tag{9.23}$$

where g is the number of internal degrees of freedom (2 for non-relativistic electrons and protons (spin up, spin down), 4 for a hydrogen atom), T is the temperature, and m is the particle's mass. Eq. (9.23) follows easily from the Maxwell-Boltzmann, Fermi-Dirac or Bose-Einstein distribution in the non-relativistic limit. The number densities of free electrons n_e, free protons n_p, and bound hydrogen atoms $n_{\rm H}$ are thus related by

$$\frac{n_e n_p}{n_{\rm H}} \simeq \left(\frac{m_e T}{2\pi}\right)^{3/2} \exp[-E_0/T]\,, \tag{9.24}$$

where $E_0 = 13.6$ eV is the hydrogen binding energy. This equation is called the *Saha equation*. It allows us to compute, as a function of temperature, the ionization fraction $x \equiv n_e/(n_e + n_{\rm H})$. (Note that $n_p = n_e$ due to the conservation of electric charge and the overall charge neutrality of the Universe.)

The crucial feature about the CMB is that it is not exactly a perfect and boring black body spectrum. There are direction-dependent (anisotropic) fluctuations at the 10^{-5} level which reflect. These reflect tiny fluctuations in the temperature, induced by quantum fluctuations in the primordial plasma. Overdense regions are slightly hotter, underdense regions are slightly colder. A map of the CMB anisotropis is shown in fig. 9.2.

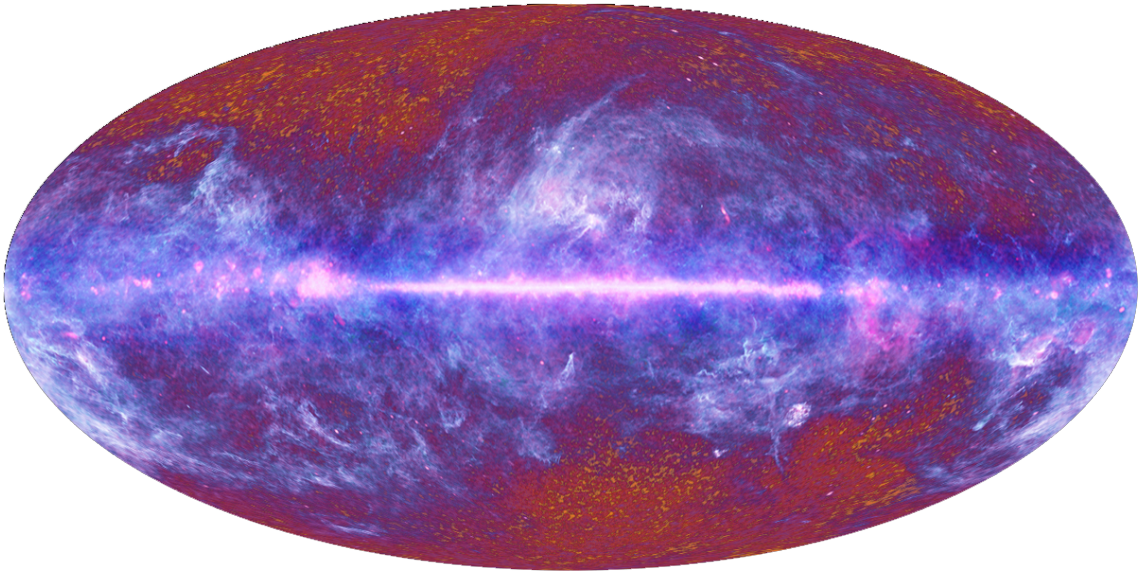
Studying the CMB anisotropies allows to learn a great deal about the early Universe. In doing this analysis, it is convenient to expand the angle dependent temperature map $T(\theta, \phi)$ in spherical harmonics (this is the spherical analogue to a Fourier transform):

$$T(\theta, \phi) = \sum_{l,m} a_{lm} Y_{lm}(\theta, \phi)\,. \tag{9.25}$$

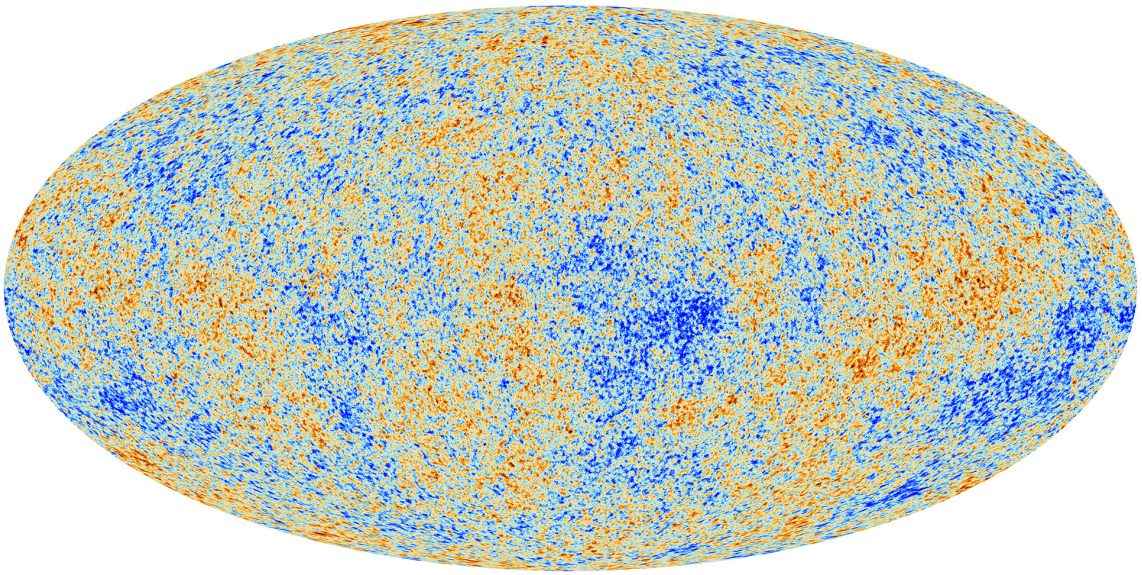
The coefficients a_{lm} are given by

$$a_{lm} = \int d(\cos\theta)\, d\phi\, Y^*_{lm}(\theta, \phi) T(\theta, \phi)\,, \tag{9.26}$$

as follows from the orthogonality of different spherical harmonics. Each coefficient a_{lm} describes the magnitude of the temperature fluctuations over an angular scale given



(a)



(b)

Figure 9.2: The Planck microwave skymap (a) before removal of foreground contamination and (b) after removal of foregrounds. Panel (b) shows the currently best map of the anisotropies in the cosmic microwave background.

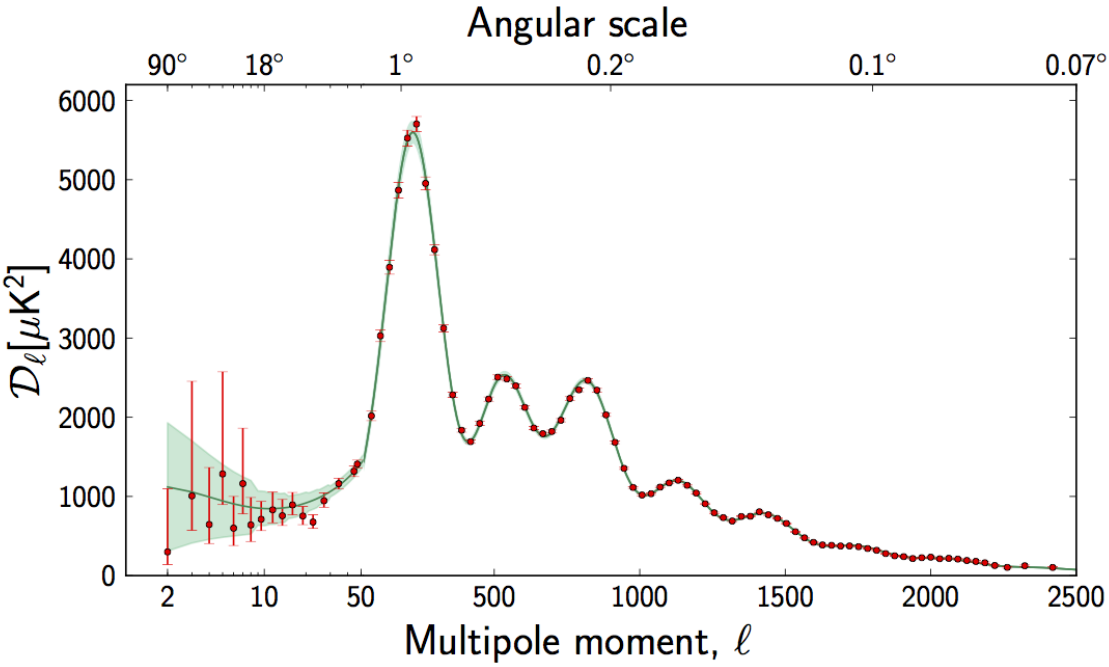


Figure 9.3: The CMB power spectrum from Planck.

by l, where $l = 2$ (dipole) would correspond for instance to a north–south or east–west asymmetry, and $l \sim 180$ corresponds to a variation of scales of order 1 degree (the distance between two minima would then be 2 degress, i.e. 1/180-th of the circumference of the total range). The quantum number m indicates the orientation of a given multipole. It distinguishes for instance between a north–sputh and east–west asymmetry. We are more interested in the scale of the fluctuations, therefore one defines the *CMB power spectrum*

$$c_l \equiv \frac{1}{2l+1} \sum_m |a_{lm}|^2 \,. \tag{9.27}$$

The current measurement of this power spectrum, as obtained by the Planck satellite mission, is shown in fig. 9.3.

Let us discuss several of the features of this spectrum. First, we observe an oscillatory pattern. This is related to plasma oscillations in the primordial soup. An overdense region accretes more matter due to its larger gravitational pull. As more matter falls in, pressure counteracts this infall, and material flows out again. Each Fourier mode of these oscillations has its characteristic period. It so happens that the modes at $l \sim 200$ have just reached their first oscillation maximum at the time of recombination, hence the peak in the CMB power spectrum. The second peak corresponds to a mode that has compressed once and then rarified again, and so on.

An important lesson we can learn from the position of the first peak in the CMB power spectrum is the geometry of the Universe. The length of the oscillation mode that has just had time to compress once is known from plasma dynamics. The angular scale under which we observe this length scale, however, depends on the geometry of the Universe. This is illustrate in fig. 9.4. In a closed Universe, a given distance scale is observed under a much larger angle today than in an open Universe. Our Universe happens to be very flat, but if the matter and energy content were different from what it is, it would be closed or open.

The ratio of the amplitudes of the peaks in the CMB power spectrum contains information on the baryon (ordinary matter as opposed to dark matter and dark energy) content of the Universe. If there are more baryons, there is more mass that oscillates, hence the plasma gets more compressed when matter falls into gravitational wells. This enhances the odd-numbered peaks (which correspond to a compression of the plasma). When the plasma rebounds out of the gravitational well, more baryons means that the photons have to work harder to generate enough pressure, hence the rebounce is less strong. Therefore, when the amount of baryons is increased, the odd-numbered peaks become larger compared to the even-numbered peaks.

Let us now discuss how neutrinos affect the cosmic microwave background. In fact, one of the cosmological parameters that can be inferred from the CMB power spectrum is N_{eff}, the "effective number of neutrino species". Even though it is called that way, it does not directly measure anything about neutrinos. Rather, it measures the total energy density in relativistic degrees of freedom—and one assumes that neutrinos are the only such degrees of freedom in the early Universe besides the photons. In the Standard Model, $N_{\text{eff}} \simeq 3$, of course. The precise value is 3.046, where the small deviation from 3 comes from the fact that the neutrino energy distribution is not perfectly Fermi-Dirac after neutrino decoupling. The reason for this is that neutrino decoupling is not completely over when electron–positron annihilation begins. Moreover, the neutrino interaction cross sections scale as E^2. therefore high-energy neutrinos freeze out a little later than the low energy ones [15]. Both effects together lead to spectral disortions in the neutrino population.

We have seen in sec. 9.2 that relativisic particles affect the Hubble expansion rate H of the Universe in a different way than non-relativistic matter. Therefore, the expansion rate before and during recombination depends on N_{eff}, and this has many phenomenological consequences. A major difficulty in isolating the effect of N_{eff} arises because the effect of changing N_{eff} is degenerate with many other cosmological parameters. In the case of N_{eff}, the breaking of these degeneracies works as follows [16]. First, note that the parameters ρ_b (baryon energy density), z_{eq} (the redshift[2] of matter–radiation equality) and θ_s (angular scale of the sound horizon at recombination) are very precisely determined by CMB measurements. We have already seen above that ρ_b is measured by considering the ratio of the heights of the odd and even peaks in the CMB power spectrum. The redshift

[2]Redshift is a measure of cosmological time. It is defined as $1 + z \equiv T/T_0$, where T is the photon temperature at the given epoch and T_0 is the photon temperature today, 2.73 K.

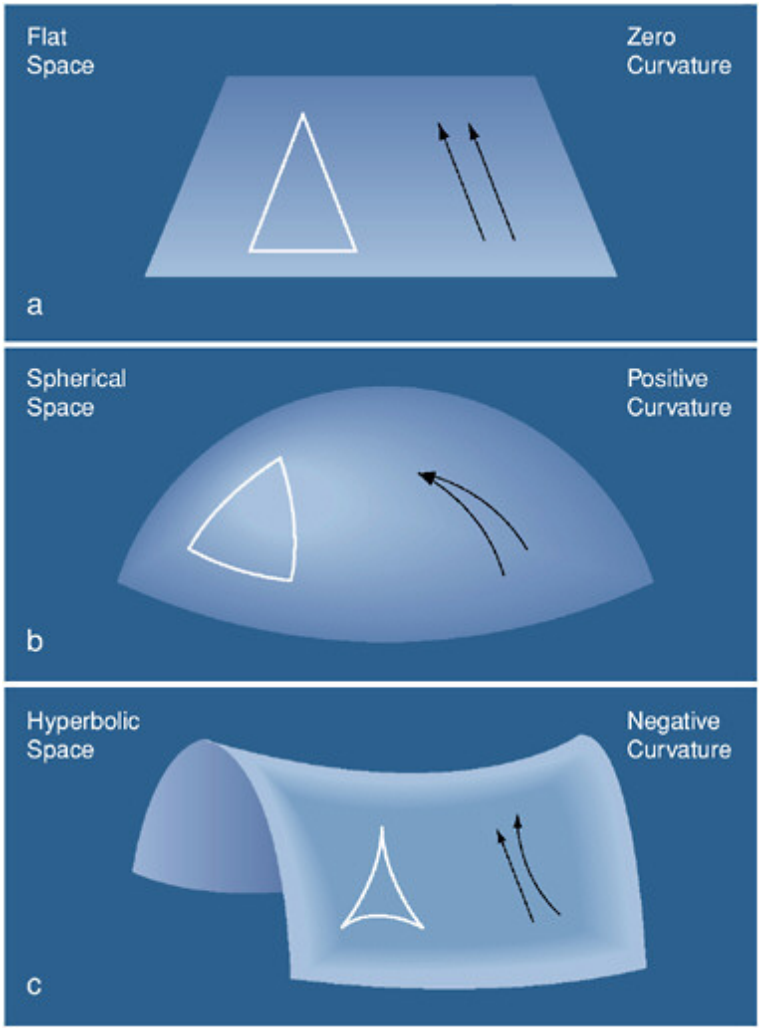


Figure 9.4: Illustration of the three possible topologies of the Universe: flat, closed (positive curvature) and open (negative curvature). Figure from `http://hendrix2.uoregon.edu/~imamura/123/lecture-5/topology.html`

of matter–radiation equality $z_{\rm eq}$ is determined by considering the overall amplitude of the peaks, which depends on the ratio of baryon to photon energy density and therefore on how long before recombination matter–radiation equality had occured. Finally, the sound horizon (a measure for how far perturbations in the plasma can have propagated by the time of recombination) is given by the position of the peaks, which measures the wavelength of the plasma oscillations that in turn depends on the sound speed. Fig. 9.5 shows how the CMB power spectrum changes when $N_{\rm eff}$ is varied while ρ_b, $z_{\rm eq}$ and θ_s are kept fixed.

The main effect here is due to modified *Silk damping*. Silk damping means that, just before decoupling completely, photons already have a non-negligible mean free path. This allows them to undergo a random walk (and thus carry energy) over non-negligible distances, thus washing out density and temperature inhomogeneities. Of course, this affects mostly relatively small scale (large l) perturbations. When $N_{\rm eff}$ is larger, H increases and this leads to a *decrease* in the photon diffusion (or damping) scale r_d because the epoch of Silk damping is shorter. On the other hand, also the sound horizon r_s becomes smaller (in comoving coordinates) when H increases. [3] It turns out that $r_d \propto H^{-1/2}$, while $r_s \propto H^{-1}$. To keep the well-measured θ_s fixed, the distance D_A the CMB photons have travelled since recombination (a measure for the age of the Universe) has to decrease as well. This means that the angular damping scale $\theta_d \equiv r_d/D_A$ *increases* because the change in D_A overcompensates the chaneg in r_d. Thus, Silk damping becomes more efficient when the expansion rate is larger, as is the case for larger $N_{\rm eff}$.

9.5 Structure formation and the neutrino mass

Cosmological observations—in particular observations of large scale structures in the Universe—can constrain the mass of neutrinos. Structures like galaxy clusters, galaxies, etc. form in the early Universe when region of space that have a slight overdensity of matter (due to quantum fluctuations in the primordial plasma) accrete more and more matter from underdense regions due to their higher gravitational pull. Given a model for the initial fluctuations (from theory or from the CMB), the growth of structures can be simulated, and global properties of the final Universe (power spectrum of the matter density map) can be compared to data from galaxy surveys.

During structure formation (starting just before recombination and going on until today), neutrinos can transport energy efficiently from overdense regions to regions with a lower-than-average density. This washes out density differences on the scales over which neutrinos can travel, smoothing out structures at these scales. If neutrinos have a nonzero mass, they carry more energy and the smoothing is more efficient. This effect can be constrained observationally.

[3] To understand the concept of comoving coordinates, imagine spacetime as a rubber surface with grid lines drawn onto it. Comoving distance is measured in coordinates defined by these grid lines. The comoving distance between two grid lines remains the same even when the surface is stretched due to the expansion of the Universe. Comoving coordinates are defined such that, at the present epoch $t = t_0$, comoving and physical coordinates coincide.

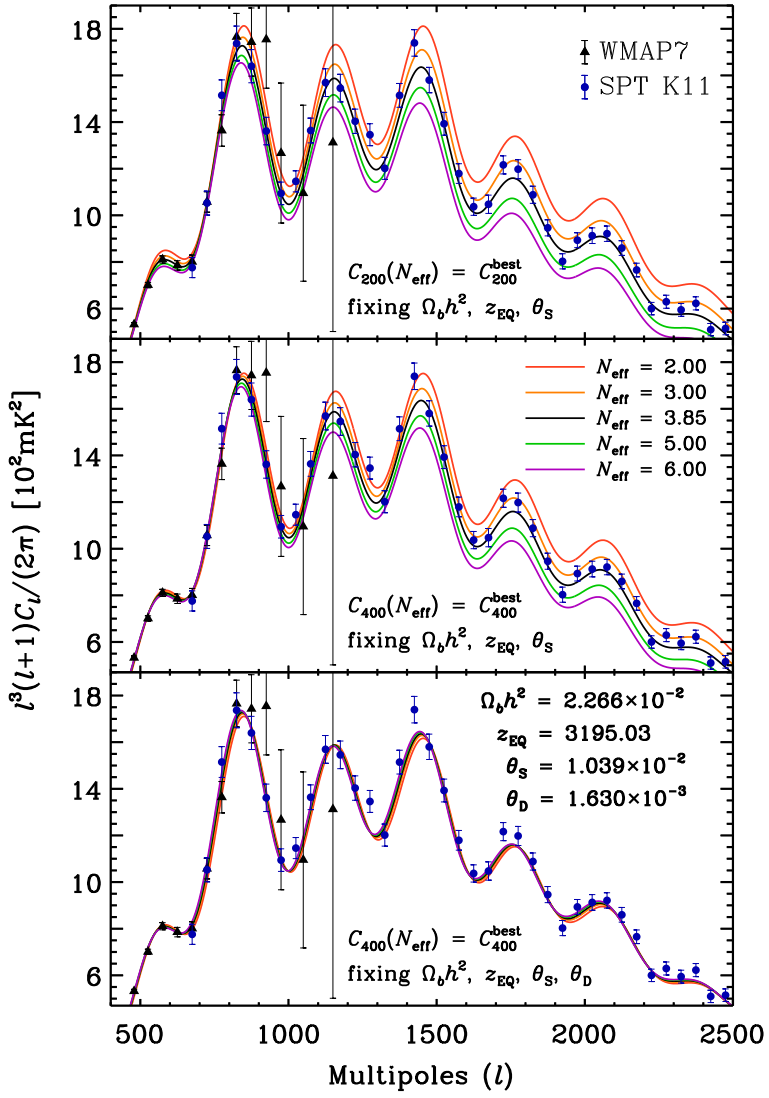


Figure 9.5: Effect of varying $N_{\rm eff}$ on the CMB power spectrum. We assume the well-measured parameters ρ_b, $z_{\rm eq}$ and θ_s are kept fixed. Figure taken from [16].

9.5.1 Formalism for structure formation in the linear regime

To understand how structure formation works, we will employ a toy model that does not include the expansion of the Universe. We follow [14] here. The basic equations are

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \mathbf{v}) = 0\,, \qquad \text{Continuity equation} \tag{9.28}$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \boldsymbol{\nabla})\mathbf{v} + \frac{1}{\rho}\boldsymbol{\nabla} p + \boldsymbol{\nabla}\phi = 0\,, \qquad \text{Euler equation} \tag{9.29}$$

$$\nabla^2 \phi = 4\pi G \rho\,. \qquad \text{Poisson equation} \tag{9.30}$$

Here, $\rho(\mathbf{x}, t)$ is the matter density at a given point $\mathbf{x}$ and time t, $\mathbf{v}(\mathbf{x}, t)$ is the velocity , $p(\mathbf{x}, t)$ is the pressure, $\phi(\mathbf{x}, t)$ is the gravitational potential, and G is Newton's constant.

The continuity equation says that any change of the density in a small volume element d^3x around point $\mathbf{x}$ must come from matter flowing in our out of this volume element. The flux is $\rho\mathbf{v}$ and its gradient gives the difference between inflow and outflow. Euler's equation says that the material is accelerated ($\partial\mathbf{x}/\partial t$) by three effects: 1) convective acceleration ($(\mathbf{v} \cdot \boldsymbol{\nabla})\mathbf{v}$, similar to the effect that when a water pipe narrows, the fluid velocity inside increases), 2) pressure gradients ($1/\rho \cdot \boldsymbol{\nabla} p$), 3) gradients of the gravitational field ($\boldsymbol{\nabla}\phi$). Finally, the Poisson equation relates the gravitational potential to the matter density.

We would like to write the dynamical quantities ρ, p, $\mathbf{v}$ and ϕ as some simple 0-th order term, plus a small perturbation:

$$\rho = \rho_0 + \rho_1\,, \tag{9.31}$$

$$p = p_0 + p_1\,, \tag{9.32}$$

$$\mathbf{v} = \mathbf{v}_0 + \mathbf{v}_1\,, \tag{9.33}$$

$$\phi = \phi_0 + \phi_1\,. \tag{9.34}$$

We take the 0-th order solution to eqs. (9.28)–(9.30) to be the static one: $\rho_0 = \text{const}$, $p_0 = \text{const}$, $\mathbf{v}_0 = 0$, $\phi_0 = 0$. The last condition, $\phi_0 = 0$ clearly violates the Poisson equation. We will ignore this fact—called the Jeans swindle—here. In an analysis that takes into account the expansion of the Universe, it would not be nessary to invoke this swindle.

To first order in ρ_1, p_1, $\mathbf{v}_1$, ϕ_1, eqs. (9.28)–(9.30) become

$$\frac{\partial \rho_1}{\partial t} + \rho_0 \boldsymbol{\nabla} \mathbf{v} = 0\,, \tag{9.35}$$

$$\frac{\partial \mathbf{v}_1}{\partial t} + \frac{v_s^2}{\rho_0}\boldsymbol{\nabla}\rho_1 + \boldsymbol{\nabla}\phi_1 = 0\,, \tag{9.36}$$

$$\nabla^2 \phi_1 = 4\pi G \rho_1\,. \tag{9.37}$$

In the second equation we have used the definition of the speed of sound

$$v_s^2 \equiv \left(\frac{\partial p}{\partial \rho}\right)_{\text{adiabatic}} = \frac{p_1}{\rho_1} \tag{9.38}$$

We can now take the derivative of eq. (9.35) with respect to time and plug in eqs. (9.36) and (9.37) to obtain

$$\frac{\partial^2 \rho_1}{\partial t^2} - v_s^2 \nabla^2 \rho_1 = 4\pi G \rho_0 \rho_1 \,, \tag{9.39}$$

which has solutions of the form

$$\rho_1(\mathbf{x}, t) = A \exp\left[i\omega t - i\mathbf{k}\mathbf{x}\right] , \tag{9.40}$$

where ω and $\mathbf{k}$ obey

$$\omega^2 = v_s^2 \mathbf{k}^2 - 4\pi G \rho_0 \,. \tag{9.41}$$

The actual matter density field in the Universe can be written as a superposition of solutions of the form (9.40). The large k modes (small scale structure) have real ω and are oscillating. The small k modes (very larger structures) have imaginary ω. These are the collapsing modes.

Remember that this was only a toy scenario without inclusion of the expansion of the Universe. The realistic formalism, including the expansion, and avoiding the Jeans swindle, is given in [14].

9.5.2 Impact of neutrinos on structure formation

As discussed in the introduction to this section, it is the *free streaming* of neutrinos that leads to wash-out of structures on small scales. Let us estimate the distance scales affected by free streaming. The coordinate distance (in comoving coordinates) travelled by a neutrino after decoupling is

$$\lambda_{\mathrm{fs}} = \int_{t_{\mathrm{dec}}}^{t_0} dt \, \frac{v(t)}{a(t)} \,, \tag{9.42}$$

where t_{dec} is the decoupling time and $v(t)$ is the neutrino velocity. The equation accounts for the fact that neutrinos eventually become non-relativistic. However, after they become non-relativistic, they stream over much smaller distances than before. Solving eq. (9.42), we find that the typical free streaming scale is of order few $\times$ 100 Mpc. (For comparison: the size of the currently observable Universe is about 30 Gpc, the diameter of the Milky Way disc is about 10 kpc, galaxy clusters—the largest gravitationally bound objects in the Universe—have sizes of order several Mpc.)

The neutrino-induced wash-out of structures is most easily observable in the *matter power spectrum*. This function is obtained by taking the matter density distribution $\rho(\mathbf{x})$, measured in galaxy surveys like the Sloan Digital Sky Survey (SDSS) and applying a 3-dimensional Fourier transform to it. Let us call the Fourier transform $\tilde{\rho}(\mathbf{k})$, where the wave number k is related to a distance scale λ according to $k = 2\pi/\lambda$. The 3D matter power spectrum is then defined as

$$P(\mathbf{k}) \equiv |\tilde{\rho}(\mathbf{k})|^2 \,. \tag{9.43}$$

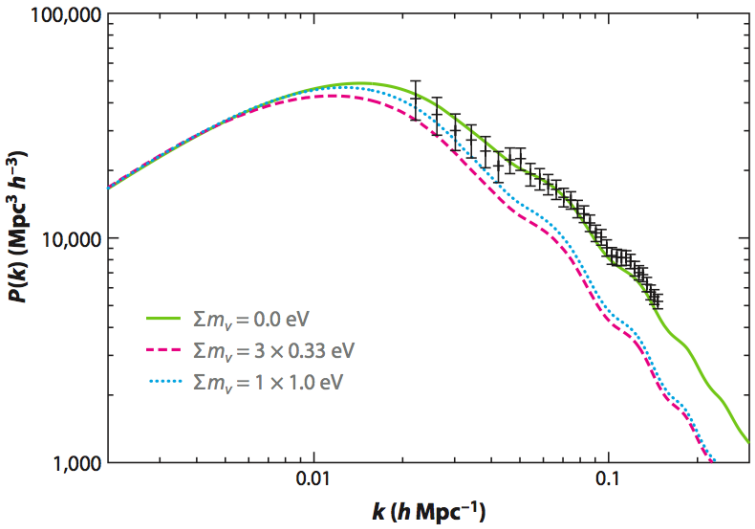


Figure 9.6: A plot of the matter power spectrum for scenarios with massive and massless neutrinos. We see the neutrino-induced suppression becomes active at $k \gtrsim 0.01h/\text{Mpc}$, corresponding to distance scales of order Gpc. Figure taken from [17].

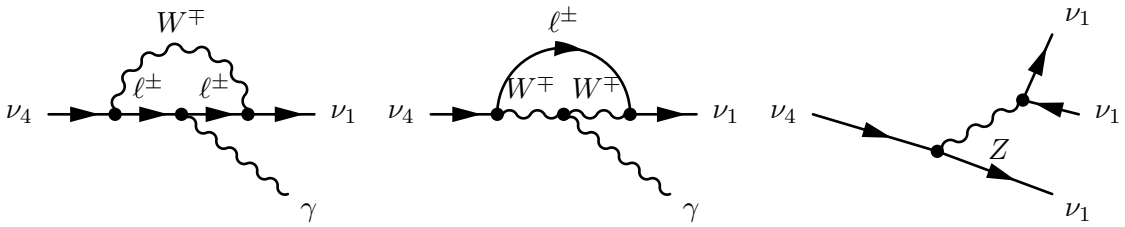
Due to isotropy, it is sufficient to plot $P(k)$ as a function of only the modulus $k \equiv |\mathbf{k}|$. This is done in fig. 9.6. We see clearly the suppression at $k \gtrsim 0.01$ h/Mpc. (Note the units, h/Mpc. Here, h is defined via $H_0 \equiv h \times 100$ km sec Mpc^{-1}. Since in cosmology, one usually measures some combination of an observable and the Hubble constant H_0, but H_0 is poorly known, it is convenient to rescale units by a factor of h.

9.6 Sterile neutrinos as dark matter candidates

Sterile neutrinos with masses $>$ keV have all the properties required to account for the DM in the Universe: they are electrically neutral, become non-relativistic early on (thus forming cold dark matter), can have very weak couplings with other particles (if the relevant mixing angles are small), and are stable over cosmological time scales.

9.6.1 Sterile neutrino decay

Regarding the last point, note that they are actually not absolutely stable. A massive, mostly sterile neutrino ν_4 with a small admixture of a light, mostly active neutrino state ν_1 can decay through the following diagrams:



The third of these is phenomenologically irrelevant because the decay products are invisible. It can only be used to impose the constraint that the lifetime of ν_4 should be much larger than the age of the Universe to provide a successful DM candidate. The first two diagrams, on the other hand, lead to radiative neutrino decay $\nu_4 \to \nu_1\gamma$. The rate for radiative sterile neutrino decay is [18]

$$\Gamma(\nu_4 \to \nu_{1,2,3}\gamma)^{\mathrm{D}} = \frac{9\alpha_{\mathrm{em}}G_F^2 m_4^5}{2^{11}\pi^4} \sum_{j=1,2,3} \left(1 - \frac{m_j^2}{m_4^2}\right)^3 \left(1 + \frac{m_j^2}{m_4^2}\right) \left| \sum_{\alpha=e,\mu,\tau} \frac{m_\alpha^2}{M_W^2} V_{\alpha 4} V_{\alpha j}^* \right|^2 \tag{9.44}$$

for Dirac neutrinos, and

$$\begin{aligned}\Gamma(\nu_4 \to \nu_{1,2,3}\gamma)^{\mathrm{M}} = \frac{9\alpha_{\mathrm{em}}G_F^2 m_4^5}{2^{10}\pi^4} \sum_{j=1,2,3} \left(1 - \frac{m_j^2}{m_4^2}\right)^3 \Bigg\{ \left(1 + \frac{m_j^2}{m_4^2}\right)^2 \left[\sum_{\alpha=e,\mu,\tau} \frac{m_\alpha^2}{M_W^2} \mathrm{Im}(V_{\alpha 4} V_{\alpha j}^*) \right]^2 \\ + \left(1 - \frac{m_j^2}{m_4^2}\right)^2 \left[\sum_{\alpha=e,\mu,\tau} \frac{m_\alpha^2}{M_W^2} \mathrm{Re}(V_{\alpha 4} V_{\alpha j}^*) \right]^2 \Bigg\}\end{aligned} \tag{9.45}$$

for Majorana neutrinos. In the above expressions, m_α denotes the mass of the charged lepton ℓ_α. The fact that the expression is different for the two cases comes from the fact that, for Dirac neutrinos, only an ℓ^- and a W^+ can propagate in the loop (opposite for Dirac antineutrinos), while for Majorana neutrinos, also the combination ℓ^+ and W^- is possible.

The radiative decay mode implies that sterile neutrino DM leads to potentially observable monoenergetic $\mathcal{O}$(keV) x-ray emission in regions of high DM density (Galactic Center, galaxy clusters, etc.). Searches for such signatures have been carried out, and results are shown in fig. 9.7.

9.6.2 Sterile neutrino production: the Dodelson–Widrow mechanism

An important question for any DM candidate is how the DM abundance observed in the Universe is determined. For the case of sterile neutrinos, the leading mechanism is the *Dodelson–Widrow* mechanism [21], which we will outline now.

The assumption is that, very early on, no sterile neutrinos exist. Later, they are produced via active-to-sterile ($\nu_a \to \nu_s$) neutrino oscillations. For $\mathcal{O}$(keV) masses, the

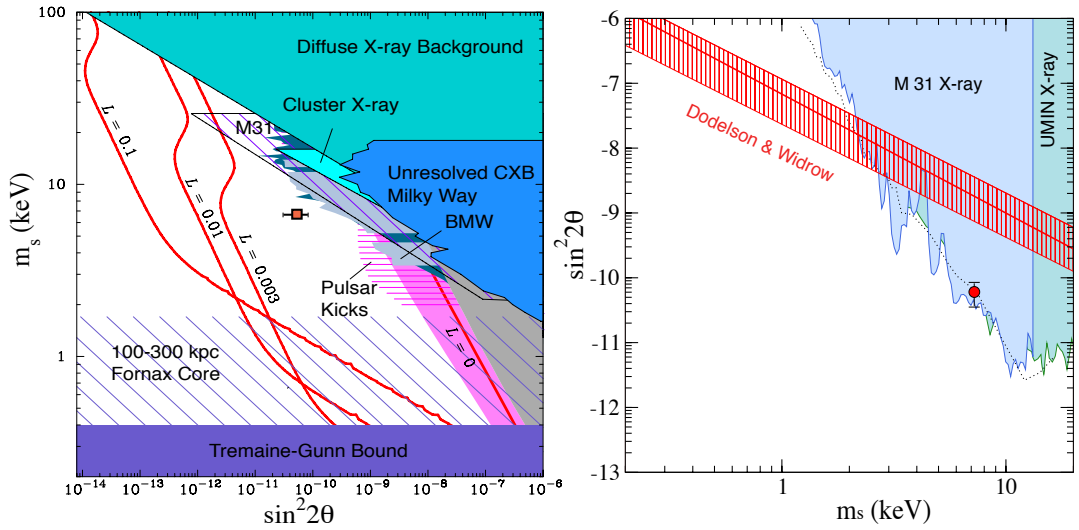


Figure 9.7: Constraints on sterile neutrino dark matter. Figure taken from [19]. The constraints in the left hand plot are from [20]. The horizontal band at the bottom is the Tremaine-Gunn (phase space/Pauli blocking) bound. The other shaded regions show constraints from searches for monoenergetic x-ray lines in various astrophysical objects. The horizontally hatched region is preferred by so-called pulsar kicks (the observation that neutron stars emerge from a supernova explosion with a very high momentum—this could be due to asymmetric emission of neutrinos, with the recoil momentum kicking the neutron star). The red curve labelled "$L = 0$" is favored by the Dodelson-Widrow production mechanism (no lepton asymmetry). The other red curves show where the correct DM density is achieved if the lepton asymmetry is nonzero (Shi-Fuller mechanism). The RH plot shows newer x-ray constraints from Chandra and XMM-Newton observations of the andromeda galaxt (M31) and from Suzaku observations of the Ursa Minor dwarf galaxy. See [19] and [20] for a complete list of references.

oscillation length/oscillation time scale is very small, so a ν_a–ν_s superposition is produced very quickly. For small mixing angle, it consists mostly of ν_a, with a small ($\sim \frac{1}{2}\sin^2 2\theta$) admixture of ν_s. Neutrino collisions with other particles act as quantum mechanical "measurements", collapsing the wave function either into ν_s (with a probability of $\frac{1}{2}\sin^2 2\theta$), or into ν_a (with a probability of $1 - \frac{1}{2}\sin^2 2\theta$). Afterwards, oscillations start again. Active neutrinos again acquire a ν_s component $\sim \frac{1}{2}\sin^2 2\theta$, and sterile neutrinos acquire a ν_a component of the same magnitude. However, since ν_s are much less abundant than ν_a, the back-conversion is negligible. After many collisions, the sterile neutrino abundance has increased to the level observed today. Eventually, collisions cease because the primordial gas becomes too dilute, and the ν_s abundance present at this time "freezes in". Note that, before freeze-in, active neutrinos are continuously replenished via pair production or CC interactions.

Dodelson–Widrow production of sterile neutrinos is described by the Boltzmann equation

$$\left(\frac{\partial}{\partial t} - H\,E\,\frac{\partial}{\partial E}\right) f_s(E,t) = \left[\frac{1}{2}\sin^2(2\theta_M(E,t))\Gamma(E,t)\right] f_a(E,t)\,, \tag{9.46}$$

where $f_s(E,t)$ and $f_a(E,t)$ are the time-dependent momentum distribution functions of sterile and active neutrinos, respectively. Before the interactions between active neutrinos and other SM particle freeze out (the epoch relevant here because the mechanism relies on these collisions), $f_a(E,t)$ is just a Fermi-Dirac distribution

$$f_a(E,t) = \frac{1}{e^{p/T}+1}\,. \tag{9.47}$$

The quantity

$$\Gamma(E,t) \simeq \frac{7\pi}{24} G_F^2 T^4 E \tag{9.48}$$

in eq. (9.46) is the active neutrino interaction rate. The expression in square brackets is thus the probability for the neutrino state to collapse to ν_s in a collision.[4] $\theta_M(E,t)$ denotes the mixing angle in matter. The second term on the left hand side of eq. (9.46) describes the change in the energy spectrum due to redshift. Indeed, we have

$$\frac{d}{dt} f_s(E,t) = \left(\frac{\partial}{\partial t} + \frac{dE}{dt}\frac{\partial}{\partial t}\right) f_s(E,t)\,, \tag{9.49}$$

and $dE/dt = d(E_0 a^{-1})/dt = E_0 a^{-2}\dot{a} = H\,E$.

From eq. (9.46), we can compute an evolution equation also for the ratio of number densities of sterile and active neutrinos, $r(t) \equiv n_s(t)/n_a(t)$, with $n_i(t) = 2\int d^3p\, f_i(E,t)/(2\pi)^3$.

[4] It may seem odd that the neutrino can collapse into ν_s even though only ν_a interact. This paradox can only be resolved in a more careful treatment of the Dodelson–Widrow mechanism using the density matrix formalism.

In doing so, it is convenient to go from derivatives with respect to t to derivatives with respect to $a(t)$. We use

$$\frac{d}{da}n_s(t) = \frac{d}{da}2\int\frac{d^3p}{(2\pi)^3}f_s(E,t) \tag{9.50}$$

$$= 2\frac{d}{da}\int\frac{4\pi E^2 dE}{(2\pi)^3}f_s(E,t) \tag{9.51}$$

$$= \frac{2}{\dot{a}}\int\frac{4\pi E^2 dE}{(2\pi)^3}\frac{\partial}{\partial t}f_s(E,t) + 2\int\frac{4\pi E^2 dE}{(2\pi)^3}\frac{dE}{da}\frac{\partial}{\partial E}f_s(E,t) + 6\int\frac{4\pi E\, dE}{(2\pi)^3}\frac{dE}{da}f_s(E,t) \tag{9.52}$$

$$= \frac{2}{\dot{a}}\int\frac{4\pi E^2 dE}{(2\pi)^3}\frac{\partial}{\partial t}f_s(E,t) - 2\int\frac{4\pi E^2 dE}{(2\pi)^3}\frac{E}{a}\frac{\partial}{\partial E}f_s(E,t) - 6\int\frac{4\pi E\, dE}{(2\pi)^3}\frac{E}{a}f_s(E,t)\,, \tag{9.53}$$

or, equivalently,

$$\dot{a}\frac{d}{da}n_s = 2\int\frac{4\pi E^2}{(2\pi)^3}\frac{\partial}{\partial t}f_s(E,t) - 2\int\frac{4\pi E^2 dE}{(2\pi)^3}H\,E\frac{\partial}{\partial E}f_s(E,t) - 3Hn_s\,. \tag{9.54}$$

We can thus rewrite eq. (9.46) into

$$\dot{a}\frac{d}{da}n_s + 3Hn_s = \gamma n_a\,, \tag{9.55}$$

where we have defined

$$\gamma \equiv \frac{1}{n_a}\int\frac{d^3p}{(2\pi)^3}\,\sin^2(2\theta_M)\Gamma(E,t)\frac{1}{e^{p/T}+1}\,. \tag{9.56}$$

Since, moreover,

$$\frac{d}{da}n_a = -\frac{3}{a}n_a\,, \tag{9.57}$$

we obtain

$$\dot{a}\frac{d}{da}r + \dot{a}\frac{r}{n_a}\frac{d}{da}n_a + 3Hr(t) = \gamma\,, \tag{9.58}$$

$$\Leftrightarrow \quad \dot{a}\frac{d}{da}r = \gamma\,, \tag{9.59}$$

$$\Leftrightarrow \quad aH\frac{d}{da}r = \gamma\,, \tag{9.60}$$

$$\Leftrightarrow \quad \frac{dr}{d\ln a} = \frac{\gamma}{H}\,. \tag{9.61}$$

Note that, in the above derivation, we have neglected the time-dependence of the effective number of relativistic degrees of freedom, g_*. At epochs where g_* changes, the dependence

of energy on the scale factor is no longer simply $E \propto a^{-1}$ because the energy of degrees of freedom that disappear is distributed among those remaining in thermal equilibrium. When this is taken into account, eq. (9.46) turns into [21]

$$\frac{d}{d\ln a} r = \frac{\gamma}{H} + r \frac{d}{d\ln a} g_* \,. \tag{9.62}$$

10
Supernova neutrinos

10.1 General timeline of a supernova explosion

Whem stars reach the end of their lives, they eventually run out of hydrogen to fuse into helium. The equilibrium between gravity and thermal pressure that has kept the star stable so far is then destroyed and the star contracts. In the process, its core heats up, until eventually the temperature is so high that the reaction $3\alpha \to {}^{12}\mathrm{C}$ becomes possible. When the helium runs out, the star contracts further, heats up even more at the core, and eventually starts burning carbon to produce even heavier elemnts. This continues until the core of the star consists mainly of ${}^{56}\mathrm{Fe}$, the most strongly bound nucleus in the table of isotopes. No further exothermic fusion reactions are possible at this stage, and eventually the burning stops. During the last stages of burning, the density at the core of the star is $\sim 2 \times 10^9$ grams/cm^3 and the temperature at the core is of order 0.5 MeV. The star is mainly stabilized by degeneracy pressure, i.e. by the fact that fermions cannot be arbitrarily densely packed due to the Pauli principle.

When the burning stops, even the degeneracy pressure is no longer sufficient to stabilized the star, and it collapses into a neutron star or a black hole. To overcome the degeneracy pressure, electrons are captured by protons to form neutrons, $e^- + p \to n + \nu_e$, releasing a large number of neutrinos. As the core of the star compresses further, it eventually becomes so dense that neutrinos can no longer escape freely. This happens at a density of order 10^{12} grams/cm^3, corresponding to 1% of the nuclear density. The energy released after this happens is trapped within the nascent neutron star, heating it up.

As the matter at the core of the star is compressed further, it eventually reaches nuclear density, 10^{14} grams/cm^3. At this point, the matter becomes rather incompressible, i.e. it suddenly stiffens. As more material is falling in from outside, this material rebounces, sending a shock wave outwards through the star. It is this shock wave thate eventually expels the outer shells of the star.

Neutrinos trapped in the core undergo a random walk and only diffuse outwards over timescales of several seconds, much longer than the core collapse. The star effectively emits "neutrino black body radiation", with roughly equal energies in all flavors.

The radius at which the density drops so low that neutrinos start to free stream again is called the neutrinosphere. An important feature is that different flavors start to free stream at different radii, therefore at different temperatures, i.e. their spectra are different. ν_μ, $\bar{\nu}_\mu$, ν_τ and $\bar{\nu}_\tau$ (collectively called ν_x in the literature) start free streaming first because they are trapped only by neutral current interactions. (Their energies are too low for CC production of muons or taus.) They are followed by $\bar{\nu}_e$, which, interact also via CC reactions like $\bar{\nu}_e + N \leftrightarrow e^+ + N'$. The ν_e start to free stream at the largest radii because they can undergo CC scattering with electrons, $\nu_e + e^- \to \nu_e + e^-$. Thus, ν_x have the hardest spectrum, followed by $\bar{\nu}_e$ and then by ν_e.

As neutrinos propagate out of the supernova, the density drops continuously, so they will go adiabatically through MSW resonances, similar to solar neutrinos.

The following link shows an example for a numerical simulation of a supernova explosion: `https://www.youtube.com/watch?v=2RxIwtxdEnQ`. As the text in the movie mentions, one problem of these simulations is that the simulated "supernovae" actually fail to explode. Instead, the shock wave propagating out from the core eventually gets stalled. It is presumed that the intense flux of neutrinos streaming out, in spite of the small neutrino interaction cross sections, deposits sufficient energy in the stalled shock wave to drive it all the way out. However, this behavior, if it is correct, seems to be not fully captured by current simulations yet.

10.2 Supernova 1987A

Astrophysicists estimate that supernovae in the Milky Way happend every 30–100 yrs on average, with large error bars. Actually, no galactic supernova has been seen in neutrinos yet. However, in 1987, a supernova exploded in the Large Magellanic Cloud, a dwarf galaxy accompanying the Milky Way at a distance of about 50 kpc. (For comparison, the Sun is about 8 kpc away from the galactic center, at the outer fringes of one of the spiral arms.)

At the time, three detectors observed neutrinos from the supernovae: Kamiokande-II (the predecessor of SuperK) in Japan, IMB in a mine in the US, and the Baksan experiment in the Soviet Union. The events are shown in fig. 10.1.

Even though there are only 25 events in total, more than 1 000 papers have been written about them. For instance, the observation allowed us to place a limit on the neutrino mass differences (not mass squared difference!) long before observations were observed: if the mass differences were too large, different neutrino mass eigenstates would propagate at different velocities and reach the detector at different times. From the fact that there was only one burst of supernova neutrinos, and not two or three, one can derive constraints.

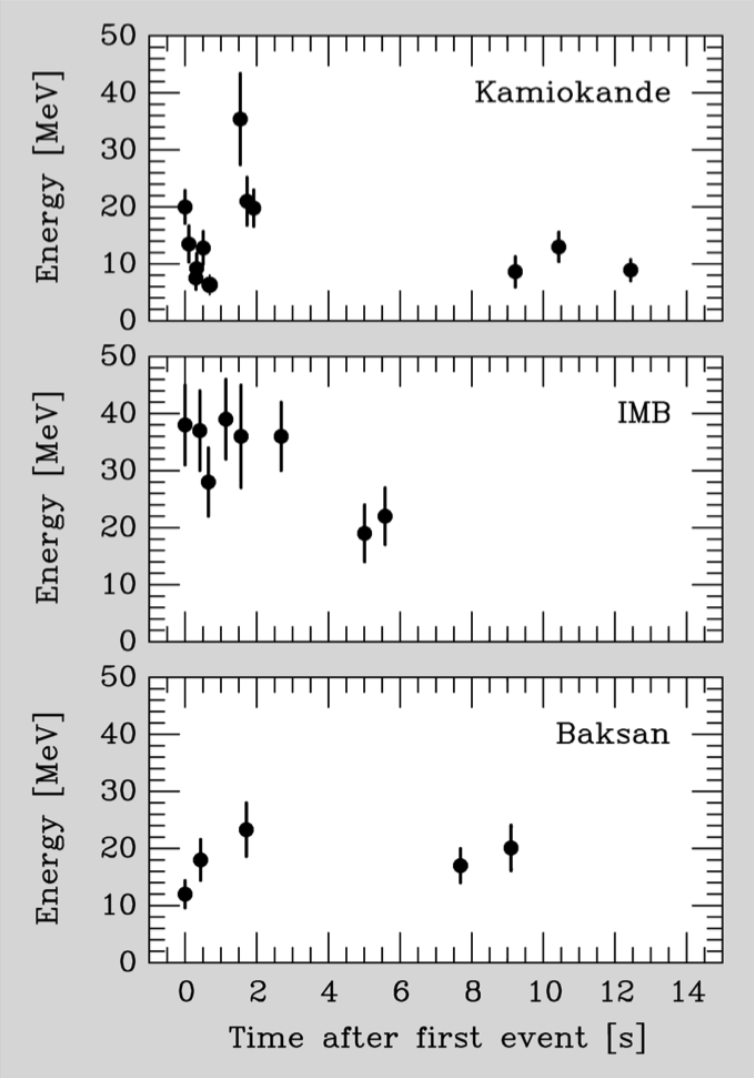


Figure 10.1: The neutrino events from supernova 1987A.

10.3 Determining the neutrino mass hierarchy using supernova neutrinos

One example for what we can learn from a future galactic supernova explosion is the neutrino mass hierarchy. Remember that MSW matter effects affect only neutrinos or only antineutrinos, depending on the sign of Δm^2. As neutrinos stream out from the dense supernova core into the vacuum of space, they pass through two MSW resonance, one depending on Δm^2_{21} and one depending on Δm^2_{31}. This leaves imprints on the neutrino and antineutrino spectra. However, the adiabaticity of the resonances depends strongly on the exact dynamics of the supernova, therefore determining the mass hierarchy that way depends on at least some rough knowledge of these dynamics.

There is, however, another way of determing the mass hierarchy using supernova neutrinos, namely by exploiting matter effects inside the Earth. We focus here on the imprint of such matter effects on the $\bar{\nu}_e$ spectrum specifically because at the typical energies of supernova neutrinos, few $\times$ 10 MeV, those are the only ones that can be efficiently observed in CC interactions, namely in inverse β decay.

We have to take into account that supernova neutrinos arrive at the Earth as mass eigenstates. This is because they have travelled over such long distances that the individual wave packets corresponding to the ν_1, ν_2 and ν_3 admixtures of an initial flavor eigenstates have become spatially separated by the time they arrive here. We thus have to compute the probability for a vacuum mass eigenstate ν_j to oscillate in the Earth and be detected as a flavor eigenstate β afterwards. This probability is given by (cf. also eq. (3.76))

$$P_{j\beta} \simeq \left| \begin{pmatrix} \cos\theta_{\text{eff}} & \sin\theta_{\text{eff}} \\ -\sin\theta_{\text{eff}} & \cos\theta_{\text{eff}} \end{pmatrix} \begin{pmatrix} e^{\frac{i\Delta m^2_{\text{eff}} L}{4E}} & \\ & e^{-\frac{i\Delta m^2_{\text{eff}} L}{4E}} \end{pmatrix} \begin{pmatrix} \cos(\theta_{\text{eff}} - \theta_{12}) & -\sin(\theta_{\text{eff}} - \theta_{12}) \\ \sin(\theta_{\text{eff}} - \theta_{12}) & \cos(\theta_{\text{eff}} - \theta_{12}) \end{pmatrix}_{\beta j} \right|^2 . \tag{10.1}$$

As usual, θ_{12} is the vacuum mixing angle, θ_{eff} and Δm^2_{eff} are the effective mixing angle and mass squared difference in matter. Oscillations described by (10.1) imprint characteristic peaks and dips onto the neutrino spectrum that are independent of the details of the supernova model and are very unlikely to be mimicked by it.

Things get even easier when supernova neutrinos are observed by two detectors, with the neutrinos observed in each of them travelling a different distance inside the Earth.

10.4 Collective neutrino oscillations and flavor polarization vectors

The feature that makes supernova neutrino oscillation a formidable theoretical challenge is the fact that, in addition to the regular MSW matter potentials neutrinos also feel a matter potential induced *by other neutrinos*. This is because the neutrino density is so exceptionally large.

We now discuss the formalism for dealing with these self-induced matter potentials, following ref. [**?**]. Let $\rho_{\mathbf{p}}$ be the density matrix in flavor space for neutrinos of momentum $\mathbf{p}$, i.e., in the 2-flavor approximation, a 2×2 matrix. If all neutrinos had the same wave function $\psi_{\mathbf{p}}$, we would simply have $\rho_{\mathbf{p}} = |\psi_{\mathbf{p}}\rangle\langle\psi_{\mathbf{p}}|$. For an ensemble of neutrinos with many different wave functions $\psi_{i,\mathbf{p}}$, we have instead $\rho_{\mathbf{p}} = \sum_i |\psi_{i,\mathbf{p}}\rangle\langle\psi_{i,\mathbf{p}}|$. The individual elements of the density matrix are

$$(\rho_{\mathbf{p}})_{\alpha\beta} = \langle\alpha|\rho_{\mathbf{p}}|\beta\rangle\,. \tag{10.2}$$

The diagonal elements give the number density of neutrinos of a given momentum $\mathbf{p}$ and flavor α, while the off-diagonal elements describe flavor correlations arising due to oscillations. For antineutrinos, we use the convention

$$(\bar{\rho}_{\mathbf{p}})_{\alpha\beta} = \langle\beta|\bar{\rho}_{\mathbf{p}}|\alpha\rangle\,, \tag{10.3}$$

i.e. the flavor indices are swapped. This may seem strange at first sight, but it will simplify the equations later on.

The density matrix satisfies the von Neumann equation

$$i\dot{\rho}_{\mathbf{p}} = [\hat{H}_{\mathbf{p}}, \rho_{\mathbf{p}}]\,. \tag{10.4}$$

Here, the Hamilton operator is

$$\hat{H}_{\mathbf{p}} = \pm\frac{\Delta m^2}{2E}\hat{B} + \sqrt{2}G_F n_e \hat{L} + \sqrt{2}G_F \int d^3p' \left(1 - \frac{\mathbf{p}\cdot\mathbf{p}'}{|\mathbf{p}||\mathbf{p}'|}\right)(\rho_{\mathbf{p}'} - \bar{\rho}_{\mathbf{p}'})\,. \tag{10.5}$$

The plus sign is for neutrinos, the minus sign for antineutrinos. If we had *not* defined $\bar{\rho}_{\mathbf{p}}$ with swapped flavor indices in eq. (10.3), the sign change would be not in the first term, but in the second and third terms of the Hamiltonian.

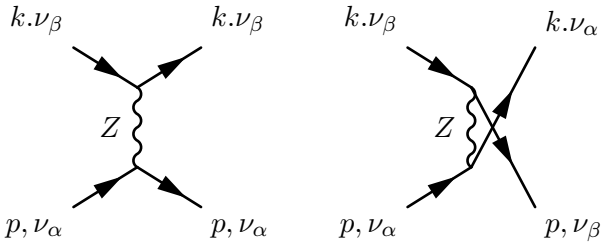
The first two terms in eq. (10.5) describe standard oscillation in matter. The first term is the standard vacuum oscillations term. In the flavor eigenstate basis, the matrix $\hat{B}$ is given by

$$\hat{B} = \frac{1}{2}U\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}U^\dagger = \frac{1}{2}\begin{pmatrix}-\cos 2\theta & \sin 2\theta\\ \sin 2\theta & \cos 2\theta\end{pmatrix}\,, \tag{10.6}$$

and U is the leptonic mixing matrix, with the vacuum mixing angle θ The second term in eq. (10.5) describes standard MSW matter effects. It involves the matrix

$$\hat{L} = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\,. \tag{10.7}$$

The thrid term in eq. (10.5) describes neutrino–neutrino coherent forward scattering. The flavor diagonal and flavor off-diagonal elements of this term correspond to the Feynman diagrams



respectively. In the second of these, the two neutrinos exchange their flavor. The self-interaction term depends on the density and momentum distribution of the background neutrinos, described by the density matrices $\rho_{\mathbf{p}'}$ of neutrinos and $\bar{\rho}_{\mathbf{p}'}$ of antineutrinos.

The factor $1 - \mathbf{p} \cdot \mathbf{p}'/(|\mathbf{p}||\mathbf{p}'|)$ in the last term of eq. (10.5) comes from the momentum dependent terms ins the MSW potential. In chapter 3, we have neglected those because the background matter was assumed to be at rest. In a supernova environment, however, this is not the case because particles are streaming outwards. The following discussion will focus on toy scenarios, where we are still going to make the unrealistic assumption that $\mathbf{p} \cdot \mathbf{p}'$ averages to zero. The density matrix then depends just on the neutrino energy, not on the momentum direction, and we will parameterize it by

$$\omega \equiv \frac{\Delta m^2}{2E}\,, \tag{10.8}$$

following conventions in the literature.

The difficulty with eq. (10.4) is that it is nonlinear in $\rho_{\mathbf{p}}$ (now ρ_ω) because the Hamiltonian itself depends on the density matrix. An elegant way of turning the von Neumann equation into a linear equation is provided by the formalism of *flavor polarization vectors*. The basic idea is to expand a 2×2 matrix (like the density matrices, the Hamiltonian, etc.) in $U(2)$ generators, e.g. for the Hamiltonian:

$$\hat{H}_\omega \equiv \frac{1}{2}(H_{\omega,0}\,\mathbb{1} + \mathbf{H}_\omega \cdot \boldsymbol{\sigma})\,, \tag{10.9}$$

where $\mathbb{1}$ is the 2×2 identity matrix, $\boldsymbol{\sigma}$ is the vector of Pauli matrices, and $H_{\omega,i}$ ($i = 0, 1, 2, 3$) are the coefficients that describe the actual physics. Similarly, we write for the density matrix ρ:

$$\rho_\omega \equiv \frac{1}{2}(P_{\omega,0}\,\mathbb{1} + \mathbf{P}_\omega \cdot \boldsymbol{\sigma}) \qquad \text{for neutrinos,} \tag{10.10}$$

$$\bar{\rho}_\omega \equiv \frac{1}{2}(P_{-\omega,0}\,\mathbb{1} - \mathbf{P}_{-\omega} \cdot \boldsymbol{\sigma}) \qquad \text{for antineutrinos.} \tag{10.11}$$

The 0-components $P_{\omega,0}$ and $P_{-\omega,0}$ contain information on the total neutrino number density, while the spatial components describe the flavor structure of the neutrino ensemble. A sample of pure ν_e corresponds to $\mathbf{P}_\omega \propto (0, 0, 1)$, while a sample of pure ν_μ corresponds to $\mathbf{P}_\omega \propto (0, 0, -1)$. The elements $P_{\omega,1}$ and $P_{\omega,2}$ contain information on the oscillation phase. A neutrino ensemble with an incoherent mixture of equal amounts of ν_e and ν_μ is described by the vector $(0, 0, 0)$.

Note that we distinguish the flavor polarization vector $\mathbf{P}_\omega$ for neutrinos from the one for antineutrinos by the sign of the index ω. Note also the sign convention chosing for $\mathbf{P}_{-\omega}$ in the second line. We will see below that these convention, together with the convention chosen for the definition of $\bar{\rho}_\omega$ (see eq. (10.3)) will allow us to unify the von Neumann equations for neutrinos and antineutrinos into just one equation.

The von Neumann equation (10.4) can then be written (for neutrinos) as

$$\begin{aligned} i\dot{P}_{\omega,0}\,\mathbb{1} + i\dot{\mathbf{P}}_\omega\cdot\boldsymbol{\sigma} &= \frac{1}{2}[H_{\omega,0}\,\mathbb{1} + \mathbf{H}_\omega\cdot\boldsymbol{\sigma},\quad P_{\omega,0}\,\mathbb{1} + \mathbf{P}_\omega\cdot\boldsymbol{\sigma}] \\ &= \frac{1}{2}[\mathbf{H}_\omega\cdot\boldsymbol{\sigma}, \mathbf{P}_\omega\cdot\boldsymbol{\sigma}]\,. \end{aligned} \tag{10.12}$$

Comparing the coefficients on the left hand side and on the right hand side of eq. (10.12), we obtain for $P_{\omega,0} = \mathrm{tr}\rho$ the trivial equation $\dot{P}_{\omega,0} = 0$. This just means that the total density of neutrinos with a given momentum (summed over flavors) does not change with time. This makes sense because coherent forward scattering does not lead to momentum exchange. More interesting is the equation we obtain for $P_{\omega,i}$, $i = 1, 2, 3$:

$$i\dot{\mathbf{P}}_\omega\cdot\boldsymbol{\sigma} = \frac{1}{2}[\mathbf{H}_\omega\cdot\boldsymbol{\sigma}, \mathbf{P}_\omega\cdot\boldsymbol{\sigma}]\,. \tag{10.13}$$

Using the identity $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$, it becomes

$$i\dot{P}_{\omega,k}\cdot\sigma_k = \epsilon_{ijk}H_{\omega,i}P_{\omega,j}\,, \tag{10.14}$$

or, in vector notation,

$$\dot{\mathbf{P}}_\omega = \mathbf{H}_\omega\times\mathbf{P}_\omega \qquad \text{(for neutrinos)}. \tag{10.15}$$

The corresponding equation for antineutrinos reads

$$\dot{\mathbf{P}}_{-\omega} = \mathbf{H}_{-\omega}\times\mathbf{P}_{-\omega} \qquad \text{(for antineutrinos)}. \tag{10.16}$$

Writing out the Hamiltonian, this gives

$$\dot{\mathbf{P}}_\omega = (\omega\mathbf{B} + \lambda\mathbf{L} + \mu\mathbf{D})\times\mathbf{P}_\omega \qquad \text{(for neutrinos)}, \tag{10.17}$$

$$\dot{\mathbf{P}}_{-\omega} = (-\omega\mathbf{B} + \lambda\mathbf{L} + \mu\mathbf{D})\times\mathbf{P}_{-\omega} \qquad \text{(for antineutrinos)}. \tag{10.18}$$

Here, $\omega \equiv \Delta m^2/(2E)$, $\lambda = \sqrt{2}G_F n_e$, $\mu = \sqrt{2}G_F n_\nu$ (where n_ν is the total neutrino number density), $\mathbf{B} = (\sin 2\theta, 0, -\cos 2\theta)$, $\mathbf{L} = (0, 0, 1)$, and

$$\mathbf{D} \equiv \int_{-\infty}^{\infty} d\omega\,\mathbf{P}_\omega\,. \tag{10.19}$$

This is the point where our at first sight unusual convention for the ordering of the flavor indices on $\bar{\rho}_{\mathbf{p}}$ in eq. (10.3) and the sign convention in eq. (10.11) pays off: eqs. (10.17) and (10.18) are actually identical, and in defining $\mathbf{D}$, we can deal with both neutrinos and antineutrinos in one go by simply integrating from $\omega = -\infty$ to $\omega = +\infty$.

The motion described by eq. (10.17) is the precession of a 3-dimensional vector $\mathbf{P}_\omega$ around a vector $\mathbf{H}_\omega \equiv \omega\mathbf{B} + \lambda\mathbf{L} + \mu\mathbf{D}$. If self-interactions are negligible ($\mu = 0$), $\mathbf{H}_\omega$ is constant, for $\mu \neq 0$, it is time-dependent. This is illustrated in fig. 10.2.

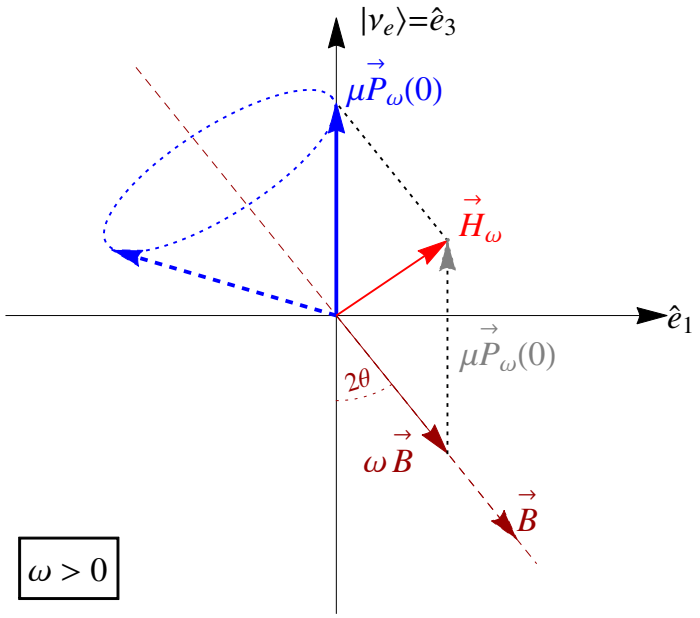


Figure 10.2: Illustration of the motion of a flavor polarization vector.

10.5 Synchronized oscillations

Let us now discuss the phenomenology of the solutions of eq. (10.17). First, let us consider a toy system with $\lambda = 0$ and $\mu = \text{const}$. We define the vector

$$\mathbf{S} \equiv \int_{-\infty}^{\infty} d\omega\, \omega \mathbf{P}_\omega \,. \tag{10.20}$$

We can then show that the quantity

$$\mathcal{E} \equiv \mathbf{B} \cdot \mathbf{S} + \frac{\mu}{2} \mathbf{D}^2 \tag{10.21}$$

is conserved. To see this, simply take the time derivative and apply (10.17):

$$\dot{\mathcal{E}} = \int_{-\infty}^{\infty} d\omega\, \omega\mu \big[\mathbf{B} \cdot (\mathbf{D} \times \mathbf{P}_\omega) + \mathbf{D} \cdot (\mathbf{B} \times \mathbf{P}_\omega) \big] = 0 \,. \tag{10.22}$$

At $\mu \gg \omega$ (strong self-interactions), this means that the length of $\mathbf{D}$ is conserved. Imagine that all neutrinos start out in the same flavor, e.g. $\mathbf{P}_\omega(0) = (0, 0, 1)$ for all ω. In the definition of $\mathbf{D}$, there is then maximal positive interference between the different $\mathbf{P}_\omega$. Due to the conservation of the length, this must remain true at later times. This can only be the case, if all $\mathbf{P}_\omega$ precess around $\mathbf{B}$ with the *same* frequency—in spite of the different neutrino energies which they describe! These so-called *synchronized oscillations* are a first example we encounter for collective behavior in dense neutrino gases.

Bibliography

[1] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Perseus Books, Cambridge, Massachusetts, 1995.

[2] T. Mueller, D. Lhuillier, M. Fallot, A. Letourneau, S. Cormon, *et al.*, *Improved Predictions of Reactor Antineutrino Spectra*, *Phys.Rev.* **C83** (2011) 054615, [1101.2663].

[3] G. Mention, M. Fechner, T. Lasserre, T. Mueller, D. Lhuillier, *et al.*, *The Reactor Antineutrino Anomaly*, *Phys.Rev.* **D83** (2011) 073006, [1101.2755].

[4] P. Huber, *On the determination of anti-neutrino spectra from nuclear reactors*, *Phys.Rev.* **C84** (2011) 024617, [1106.0687].

[5] F. Von Feilitzsch, A. Hahn, and K. Schreckenbach, *EXPERIMENTAL BETA SPECTRA FROM PU-239 AND U-235 THERMAL NEUTRON FISSION PRODUCTS AND THEIR CORRELATED ANTI-NEUTRINOS SPECTRA*, *Phys.Lett.* **B118** (1982) 162–166.

[6] K. Schreckenbach, G. Colvin, W. Gelletly, and F. Von Feilitzsch, *DETERMINATION OF THE ANTI-NEUTRINO SPECTRUM FROM U-235 THERMAL NEUTRON FISSION PRODUCTS UP TO 9.5-MEV*, *Phys.Lett.* **B160** (1985) 325–330.

[7] A. Hahn, K. Schreckenbach, G. Colvin, B. Krusche, W. Gelletly, *et al.*, *Anti-neutrino Spectra From ^{241}Pu and ^{239}Pu Thermal Neutron Fission Products*, *Phys.Lett.* **B218** (1989) 365–368.

[8] J. Kopp, P. A. N. Machado, M. Maltoni, and T. Schwetz, *Sterile Neutrino Oscillations: The Global Picture*, *JHEP* **1305** (2013) 050, [1303.3011].

[9] F. Zwirner, *Theory Summary*, 1310.3292.

[10] P. Ring and P. Schuck, *The Nuclear Many-Body Problem*. Physics and astronomy online library. Springer, 2004.

[11] J. Helo, M. Hirsch, H. Ps, and S. Kovalenko, *Short-range mechanisms of neutrinoless double beta decay at the LHC*, *Phys.Rev.* **D88** (2013) 073011, [1307.4849].

[12] T. Hambye, *Phenomenology of the type iii seesaw model*, 2009.

[13] E. Waxman and J. N. Bahcall, *High-energy neutrinos from astrophysical sources: An Upper bound*, *Phys.Rev.* **D59** (1999) 023002, [hep-ph/9807282].

[14] E. W. Kolb and M. S. Turner, *The Early Universe*, .

[15] **Topical Conveners: K.N. Abazajian, J.E. Carlstrom, A.T. Lee Collaboration**, K. Abazajian *et al.*, *Neutrino Physics from the Cosmic Microwave Background and Large Scale Structure*, *Astropart.Phys.* **63** (2015) 66–80, [1309.5383].

[16] Z. Hou, R. Keisler, L. Knox, M. Millea, and C. Reichardt, *How Additional Massless Neutrinos Affect the Cosmic Microwave Background Damping Tail*, 1104.2333.

[17] Y. Y. Wong, *Neutrino mass in cosmology: status and prospects*, *Ann.Rev.Nucl.Part.Sci.* **61** (2011) 69–98, [1111.1436].

[18] Z. Xing and S. Zhou, *Neutrinos in Particle Physics, Astronomy and Cosmology*. Advanced Topics in Science and Technology in China. Springer Berlin Heidelberg, 2011.

[19] E. Bulbul, M. Markevitch, A. Foster, R. K. Smith, M. Loewenstein, *et al.*, *Detection of An Unidentified Emission Line in the Stacked X-ray spectrum of Galaxy Clusters*, 1402.2301.

[20] K. N. Abazajian, *Detection of Dark Matter Decay in the X-ray*, 0903.2040.

[21] S. Dodelson and L. M. Widrow, *Sterile-neutrinos as dark matter*, *Phys.Rev.Lett.* **72** (1994) 17–20, [hep-ph/9303287].