Massive neutrinos and cosmology

-a study of the effects of the mass hierarchy

Maren K. Grindstad

Thesis submitted for the degree of
Master of Astronomy

Institute of Theoretical Astrophysics
University of Oslo

June 2010

Acknowledgements

I started the work with neutrino cosmology over the Christmas holiday in 2008 by reading articles of which I understood only a very little. As the amount understood in the articles read increased and the number of weeks and days to the deadline of submission decreased, there have been a number of people who have helped me raise my understanding and generally kept me sane.

First of all I want to tank my supervisor, Øystein Elgarøy, who not only introduced me to a interesting field of study, but who always welcomed my questions, frustrations, ideas and problems. The motivation and help given, and his calming attitude, has been crucial for both my academic performance and my mental health. I am also very grateful for being introduced to neutrino cosmology through an interesting and instructive summer project!

I am also very greatful for the help I received from Jostein Riiser Kristiansen to learn how to use (and install!) the CosmoMC code.

On the "keeping me sane" part, I would like to thank the people in TENK for making me realise that there is meaning to life outside the study hall. Thanks to Siv, Therese, Kari, Marie, Katinka and other friends for the fun and welcomed diversions. All the people in Astrokjelleren who made the time spent there more enjoyable should get a great thanks, and particularly Eirik and Thomas for help, discussions and for making this thesis look prettier. Also, thanks to Mikjel, for being there in every way and for giving me a kick in the butt whenever needed.

Contents

Abstract

Neutrino cosmology is a very interesting field of research, where properties of some of the smallest constituents of the universe are probed by the very largest structures of the universe; large scale structures of the size of clusters of galaxies and the cosmic microwave background radiation. Cosmology has, over the last decade, provided strong limits on the total neutrino mass, assuming that the splitting of the neutrino mass contributes negligible to the effect of neutrinos on cosmology. There is, however, a splitting between the masses of the individual neutrino mass eigenstates, as shown by neutrino oscillation experiments, and although the mass square difference is measured, the ordering of the masses is still to be determined.

As cosmology has provided stronger limits on the total neutrino mass than other experiments, it is hoped that cosmology also can solve the mystery of the neutrino mass hierarchy. The goal of this thesis is to investigate the effect of neutrino mass hierarchy on cosmology, by translating the results of neutrino mass experiments to a hierarchy dependent prior on the total neutrino mass, which is then applied to cosmological parameter likelihood distributions. It is found that adding such a hierarchy dependent prior does not allow for a determination of the neutrino mass hierarchy from three chosen parameters; the spectral index n_s, the baryon acoustic oscillation parameter $\mathcal{A}(z)$ and the fluctuation amplitude on cluster scales, σ_8. This supports the general assumption that the neutrino mass hierarchy can be neglected in cosmology.

Introduction

This master's thesis is an investigation into the field of neutrino cosmology. Neutrino cosmology is a very interesting field of research, where properties of some of the smallest constituents of the universe are probed by the very largest structures of the universe; large scale structures of the size of clusters of galaxies and the cosmic microwave background radiation (the CMB for short) which fills the entire universe. The reason that this is possible, is the lightness of the neutrinos, having the smallest mass of all known massive particles they were relativistic when structure started forming in the universe, and possible also when the CMB was released. Thus the very light particles substantially affects the universe today, even though they make up just a very tiny part of the total energy content of the universe.

It might seem like quite an overkill to probe the mass of the tiny neutrino in the largest laboratory that we have; cosmos, but it turns out to be quite reasonable. Ever since the neutrino was "invented" in the 1930s by Wolfgang Pauli, particle-physicists have aspired to reveal its nature. As neutrinos only interact weakly, this has been a time consuming pursuit. The mechanism of neutrino oscillation was suggested by Bruno Pontecorvo in 1957. He suggested that if neutrinos have a mass this can lead to oscillations between flavour states. In the late 1990s, neutrino oscillations were indeed detected. This implies that neutrinos must have a mass. Other experiments has established that neutrinos are very light, much lighter than any other fermion.

Although it is established that neutrinos have a mass, establishing the size of this mass has proved to be troublesome. Oscillation experiments are sensitive to the difference between the masses of the different species of the neutrinos, but they are not sensitive to the mass scale. β-decay experiments and double β-decay experiments are sensitive to the mass scale, but such experiments are technically demanding, and the constraints found on the neutrino mass scale are poor.

Cosmological observables are sensitive to the neutrino mass, and it is the sum of neutrino masses that affects the observables to the largest extent. It is often assumed in such models that the neutrino masses are degenerate, ie. that it is proper to assume that the neutrinos all share the same mass. This approach to the neutrino mass scale has lately provided strong upper limits on the total neutrino mass, constraining the mass better than any laboratory decay experiments has achieved.

Although the differences in the neutrino masses are known from oscillation experiments, and the total neutrino mass is constrained from cosmology, the question

of which neutrino is the heaviest and lightest is left unanswered. The ordering of the neutrino masses is called the neutrino mass hierarchy, and this is a question of interest, as it can reveal information about the mechanism generating the neutrino masses, a mechanism that is affected by physics beyond the standard model of particle physics and thus can give further insight on the laws of nature.

When starting the work with this thesis, the aim was to place limits on parameters related to the generating processes of neutrino mass, with particular focus on the seesaw model. This did, however, turn out to be more complicated than anticipated, as the number of parameters involved were larger than the cosmological parameters to which they did relate, and many of the parameters were very weakly described by the model of which they were created. This lead to a degeneracy between the parameters which was impossible for me to resolve without guessing. Thus a change of course was performed.

The goal of the work was changed, and the new problem of discussion for the thesis was dual. The first part was to take a look at neutrino mass generating mechanisms, in particular the seesaw mechanism that makes use of the peculiar characteristic of neutrinos of being chargeless fermions, and thus might be described by the Majorana equation. The second part of the thesis were to be dedicated to the hierarchy problem of the neutrino masses. Cosmology has provided strong limits on the mass scale, can it also provide information about the hierarchy of the masses? Three cosmological observables was chosen for this study, the spectral index n_s, the baryon acoustic oscillation parameter $\mathcal{A}(z)$ and the fluctuation amplitude on cluster scales, σ_8. This was to be done by translating the results of neutrino mass experiments to a hierarchy dependent prior on the total neutrino mass, which was to be applied to the cosmological parameter likelihood distributions provided by the CosmoMC Markov Chain Monte Carlo parameter sampler.

In part I of this thesis, I have introduced the physics of the neutrino. In chapter 1 I looked at the neutrino as it was thought to be in the olden days, massless and easy to deal with. I also looked into neutrino oscillations, experimental constraints on neutrino parameters and the neutrino mass hierarchies. In chapter 2 I introduce mass generating mechanism, with emphasise on the seesaw models. Part II of the thesis looks at cosmology and how neutrinos effects the cosmological observables. The seesaw model is particularly interesting for my work, as it can provide information about the neutrino mass hierarchy. Chapter 3 deals with cosmology at the most basic level, with the Friedmann-Robertson-Walker universe and inflation. Chapter 4 introduces the perturbed metric, and the perturbed matter density giving rise to the matter power spectrum. Chapter 5 takes a look at the cosmological structure formation, looking first at a ΛCDM model, and then moving on to a universe with massive neutrinos. The effects of neutrinos on large scale structure, and hence the matter power spectrum is also discussed. The last chapter in part II, chapter 6, introduces the cosmic microwave background radiation, and the way this spectrum is affected by massive neutrinos. All the way through part II of the thesis, the approximation is used that the neutrino mass hierarchy does not affect cosmology, which is a common assumption in astrophysics.

Part III contains a brief summary of the methods used and the results and discussion. I introducing the publicly available CosmoMC code and the mass prior imposed on the neutrino mass from experiments and hierarchy assumptions in chapter 7. In chapter 8 I present the results of my work. It turns out that adding a hierarchy dependent neutrino mass prior to the cosmological parameter likelihood provided by CosmoMC provides no solution for resolving the hierarchy problem from the three cosmological parameters in question (n_s, $\mathcal{A}(z)$ and σ_8). This is likely to generalize to other cosmological parameters. In chapter 9 I conclude by giving a brief summary of my work, and by commenting on the future outlook of determining the neutrino mass hierarchy from cosmology.

Part I

Neutrinos from the particle physicist's point of view

Chapter 1

The Massive Neutrino

Cosmic Gall
Neutrinos, they are very small.
They have no charge and have no mass
And do not interact at all.
The earth is just a silly ball
To them, through which they simply pass,
Like dustmaids through a drafty hall
Or photons through a sheet of glass.
They snub the most exquisite gas,
Ignore the most substantial wall,
Cold-shoulder steel and sounding brass,
Insult the stallion in his stall,
And scorning barriers of class,
Infiltrate you and me! Like tall
And painless guillotines, they fall
Down through our heads into the grass.
At night, they enter at Nepal
And pierce the lover and his lass
From underneath the bed - you call
It wonderful; I call it crass.

John Updike, 1963

This chapter will explore the theoretical background for the neutrinos, in and out of the standard model of particle physics. The standard model predicts massless neutrinos, while experiments and observations have shown that the at least some of the neutrinos must have a quite small, though non-zero, mass. I will look at the place of the neutrino in the standard model, and discuss some of the evidence of a non-zero neutrino mass. I will also briefly discuss the current bounds on the neutrino mass. This chapter is mainly based on the references [1, 2, 3].

1.1 The Massless Neutrino of the Standard Model

In this section, I will look at the neutrino as described by the standard model of particle physics, without going into more detail than needed to understand the rest of the thesis. For simplicity of notation I will assume only one neutrino species whereever only one species is needed to understand the concepts.

The neutrino was introduced in the 1930's by Wolfgang Pauli as a "last resort" to solving the problem of the continuous spectrum of the electron in the beta decay. Seemingly undetectable it solved all the problems of the beta decay spectrum, but the resistance to accept a particle that could not be seen was large. Not untill the 1950's were the neutrinos observed in an experiment, when Cowans and Reines detected the electron anti neutrino $\overline{\nu_e}$.

The neutrinos are leptons of spin one half. The chargelessness and apparent masslessness of the particle deviates from all the other standard model fundamental fermions. Being chargeless, the neutrino can be described by two different equations of motions; while the Dirac equation conserves electric charge, the Majorana equation is only valid for chargeless particles like the neutrinos. Whether the neutrinos are Dirac of Majorana particles is yet to be determined.

In the standard model, the neutrino obeys the Dirac equation

$$\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi(x) = 0 \tag{1.1}$$

which follows from the free Langrangian density

$$\mathcal{L} = i\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\overline{\psi}\psi, \tag{1.2}$$

where the spinor field $\psi(x)$ has four components. The neutrino spinor ν might be represented as

$$\nu = \begin{pmatrix} \nu_L \\ \nu_R \end{pmatrix}, \tag{1.3}$$

where ν_L and ν_R are the two-component spinors which have left- and right-handed chirality states. The chirality states are projections of the full ν field, and the projection operators are

$$\left.\begin{matrix} P_R \\ P_L \end{matrix}\right\} = \frac{1}{2}\left(1 \pm \gamma^5\right), \tag{1.4}$$

such that

$$\begin{aligned} \nu_L &= P_L\nu \\ \nu_R &= P_R\nu. \end{aligned} \tag{1.5}$$

If the neutrino is massless, then the chirality eigenstate is also a helicity eigenstate.

Rewriting the free Langrangian density from equation (1.2) using the left-handed and righthanded states, we find

$$\mathcal{L} = \overline{\nu}_R\gamma^{\mu}i\partial_{\mu}\nu_R + \overline{\nu}_L\gamma^{\mu}i\partial_{\mu}\nu_L - m\left(\overline{\nu}_R\nu_L + \overline{\nu}_L\nu_R\right). \tag{1.6}$$

We can then write out the Dirac equation (equation 1.1) using the chiral states (see appendix A):

$$\begin{aligned} i\left(\partial_0 + \sigma \cdot \nabla\right)\nu_R &= m\nu_L \\ i\left(\partial_0 - \sigma \cdot \nabla\right)\nu_L &= m\nu_R. \end{aligned} \tag{1.7}$$

We can see that due to the mass term, there is a coupling between the righthanded and the left-handed state. If the mass was zero, then we would have separate equations for the left-handed and the righthanded neutrinos, and one could exist without the other. Only lefthanded neutrinos and righthanded anti-neutrinos interact in weak interactions, and the standard model leaves no room for righthanded neutrinos and lefthanded anti neutrinos: they are thought not to exist.

This has a profound consequence for the neutrino mass. Looking at the Lagrangian density in equation (1.6), we see that the mass term occurs in the coupling of a left-handed anti-neutrino and a right-handed neutrino. As these fields does not exist in the standard model, the neutrino must be massless. It is important to note that the mass is set to zero by choosing to omit the neutrino fields that does not take part in weak interactions, although nothing really forbids massive neutrinos. The photons must have zero mass by gauge symmetries, but the neutrino is massless by choice.

It is now known that the neutrinos are in fact massive, although the mass is very small. This was not known when the standard model of particle physics was created in the 1960's and 1970's, thus John Updike was not telling any lies in his poem on the previous page. The fact that neutrinos are massive, arises many questions in particle physics, and the process of generating the mass of the neutrinos is thought to be relate to physics beyond the standard model.

1.1.1 Dirac vs Majorana

In the case of massive neutrinos, we have the trouble of figuring out which mechanism that may provide with a mass, and why the mass is so small. There is, however, another puzzle of the neutrinos; the question of their nature, and the nature of their mass.

Most fermions have electric charge. Neutrinos do not. Thus the neutrinos can be described in a way that charged fermions can not, they can be described by the Majorana equation rather than the Dirac equation described above. If the neutrinos are Majorana particles, then they are their own anti-particle. The relation between the righthanded and lefthanded fields is then given by

$$\nu_R = \xi \mathcal{C}\overline{\nu_L}^T = \nu_L^C \tag{1.8}$$

where $\mathcal{C}$ is the charge conjugate operator, and ξ is a normalized phase factor that I will take to be $\xi = 1$ for simplicity. A particle can only be its own charge-conjugate if it is neutral.

The nature of the neutrino also has consequences for the mass term of the Lagrangian. We saw above that for the Dirac case, the mass term is given by

$$\mathcal{L}_{mass} = -m\overline{\psi}\psi = -m\overline{\nu_R}\nu_L - m\overline{\nu_L}\nu_R = -m\overline{\nu_R}\nu_L + H.c., \qquad (1.9)$$

where H.c. denotes hermittian conjugate. If, as predicted by the standard model, only lefthanded neutrinos and righthanded anti-neutrinos exists, they would have non-zero mass given the mass term above.

In the case of Majorana neutrinos, an additional mass term occurs:

$$\mathcal{L}^M_{mass} = -\frac{1}{2}m\overline{\nu^C_R}\nu_L + H.c. = \frac{1}{2}m\nu^T_R\nu_R + H.c. \qquad (1.10)$$

If the neutrinos were massless, the Dirac and Majorana description would be equivalent. I neutrinos are Majorana particles, they can have both the Majorana and the Dirac mass terms, or only the Majorana mass term.

Majorana description of neutrinos is tempting, as it cuts down on the number of particles needed, and Majorana description of the neutrinos is assumed in several theories explaining the nature of the neutrino mass. A problem with this description, however, is that the lepton can not be a conserved quantity in this scenario. This is because it does not make sense to talk of a lepton number of a neutrino that is its own anti-particle; sometimes it is perceived to have lepton number 1 and sometimes it has lepton number -1. Lepton number conservation is a conservation law that has been assumed in the standard model, and there is no strong evidence of this law being broken if the neutrinos possesses Dirac nature.

1.2 Neutrino Oscillations

The only evidence for a non-zero neutrino-mass is found in the detections of neutrino oscillations. Neutrino oscillations within three generations of neutrinos is only found if at least two of the three generations are massive. This section will look at the solar neutrino problem leading up to the discovery of the neutrino oscillations, and the theory of the oscillations. The references [1], [2], [4] and [5] were used in writing this section.

1.2.1 The solar neutrino problem

During the nineteenth century, Lord Rayleigh calculated the age of the sun, assuming that all the sun's energy were gravitational and that its radiation was constant. His result was disappointing, as the lifetime of the sun was found to be substantially shorter than both the age required by Darwin's theory of evolution and the estimates made by geologist. As radioactivity was discovered by the end of the same century, nuclear fission was launched as the energy source of the sun, but the sun didn't seem to contain any radioactive material. By 1920, Francec William Aston had found that four hydrogen atoms were heavier than one helium atom. Thus

energy would be released in a fusion of hydrogen to helium. Sir Arthur Eddington suggested that fusion of hydrogen would be the source of energy in the sun.

The proton-proton chain was later calculated, and it produces a vast amount of neutrinos. Measurements of the neutrinos from the sun turned out to be a good way of testing the fusion theory, and a good way to probe the interior of the sun. Photons from the sun have been scattered around for years before they reaches the surface of the sun and travels towards Earth. Neutrinos interact much weaker, and reach Earth basically unaffected by its travel from the core of the sun. Different stages in the proton-proton chain produce neutrinos of different energies, but they are all of electron flavour. The boron-8 neutrinos has the highest flux at high energies, and are therefore the most used reaction for experiments.

In 1968, Ray Davis et al. reported from the first solar neutrino experiment at the Homestake mine in South Dakota that the abundance of neutrinos measured in the process

$$^{37}Ch + \nu_e \rightarrow ^{37} Ar + e^- \tag{1.11}$$

was only a third of the expected amount, based on John Bahcall's calculations of the solar neutrino abundance. At the time, there were great doubts as to whether the experiment was correct. Several later experiments, some using Cherenkov light to detect neutrino interactions, has however also found a deficit. This is referred to as the solar neutrino problem.

Already in the early 1960's, experiments had been carried out in order to establish whether there could be more than one type of neutrinos. The results favoured more than one type of neutrinos. Bruno Pontecorvo made use of this when he came up with the conseptually simple solution of the solar neutrino problem (in fact he even predicted the problem the year before the Homestake experiment). If the neutrinos have a mass, contrary to the predictions of the Standard Model, the mass eigenstates might differ from the flavour eigenstates. In the quark sector, the Cabbibo angle was introduced in 1963 to preserve universality of the weak force, showing that the states taking part in weak interactions and the states forming the weak isospin doublets were not the same. In 1973 this was extended to three generations in the CKM matrix. Pontecorvo proposed neutrino oscillations in 1957, in a time where only one neutrino was known, the electron neutrino ν_e. He thus introduced a sterile (non-interacting fermion) neutrino in order to have a state for the electron neutrino to oscillate to and from. After the detection of the muon neutrino, the theory was adapted for two active neutrinos. If the electron neutrinos created in the proton proton chain in the fusion in the sun had a non-zero probability of oscillating into another flavour, then less neutrinos than expected would be detected in a detectionprocess like the one in equation (1.11), which is only sensitive to electron neutrinos.

If the mass eigenstates and the flavour eigenstates are not the same, this allows for neutrino oscillation, meaning that what starts out as one flavour (say electron) will end up as another flavour (say muon) with a given probability at another time and place. The description of how this work is rather simple, using only quantum

mechanics.

1.2.2 The theory of oscillations

In this subsection I will look at the theory behind neutrino oscillations. The following assumptions has been made

1. The neutrinos are assumed to be ultra-relativistic.
2. The flavour neutrinos has a definite momentum $\mathbf{p}$, meaning that all the massive neutrino components share the same momentum.

Neutrinos couple to charged fermions via the left-handed leptonic charged current

$$j^{\rho}_{W,L} = 2 \sum_{\alpha=e,\mu,\tau} \overline{\nu_{\alpha L}} \gamma^{\rho} l_{\alpha L} \tag{1.12}$$

where $l_{\alpha L}$ is a left-handed charged fermion of type α. $\nu_{\alpha L}$ is the left-handed neutrino which couples to $l_{\alpha L}$; the *flavour neutrino field*. The neutrino mass term looks like

$$\mathcal{L}_{\mu_\nu} = -\overline{\nu_L} M \nu_R + h.c. \tag{1.13}$$

in the Dirac scenario, where

$$\nu = \begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix}. \tag{1.14}$$

If the mass matrix M is not diagonal, then the left-handed neutrino $\nu_{\alpha L}$ must be a superposition of mass-eigenstates:

$$\begin{aligned} |\nu_\alpha\rangle &= \sum_k U^*_{\alpha k} |\nu_k\rangle \, (\alpha = e, \mu, \tau) \qquad (a) \\ |\nu_k\rangle &= \sum_\alpha U_{\alpha k} |\nu_\alpha\rangle \, (\alpha = e, \mu, \tau) \qquad (b). \end{aligned} \tag{1.15}$$

The U-matrix is the mixing matrix that gives the weights of $|\nu_k\rangle$ in $|\nu_\alpha\rangle$, and $|\nu_k\rangle$ is the mass-eigenstate neutrino field. The U-matrix is a product of $V_L^{l\dagger}$, which is the unitary matrix that participates in the diagonalization of the charged lepton mass matrix and V_L^{ν} is the unitary matrix participating in the diaginalization of the neutrino mass matrix [1]:

$$U_{\alpha k} = \sum_{\beta=e,\mu,\tau} (V_L^{l\dagger})_{\alpha\beta} (V_L^{\nu})_{\beta k}. \tag{1.16}$$

Thus U is not unitary (unless it is a square matrix), as it might be a rectangular matrix if the number of flavours states do not equal the number of mass states. But it still has the property

$$\begin{aligned} UU^\dagger &= 1 \\ U^\dagger U &\neq 1 \end{aligned} \tag{1.17}$$

Applying the Schrödinger equation in the restframe of the neutrino, one obtains

$$|\nu_k(\tau)\rangle = e^{-im_k c^2\tau/\hbar}\,|\nu_k(0)\rangle\,. \tag{1.18}$$

In the laboratory frame, one gets

$$|\nu_k(\tau)\rangle = e^{-i(E_k t/\hbar - p_k L/c\hbar)}\,|\nu_k(0)\rangle\,, \tag{1.19}$$

and hence it follows that

$$\begin{aligned} |\nu_\alpha(\tau)\rangle &= \sum_k U^*_{\alpha k} e^{-i(E_k t/\hbar - p_k L/c\hbar)}\,|\nu_k(0)\rangle \\ &= \sum_\beta \sum_k U^*_{\alpha k} U_{\beta k} e^{-i(E_k t/\hbar - p_k L/c\hbar)}\,|\nu_\beta\rangle\,. \end{aligned} \tag{1.20}$$

The transition amplitude is found from

$$A_{\nu_\alpha\to\nu_\beta} \equiv \langle\nu_\beta|\nu_\alpha(\tau)\rangle\,, \tag{1.21}$$

where only $|\nu_\alpha(\tau)\rangle$ is time dependent. Orthogonality of the flavor states then gives

$$A_{\nu_\alpha\to\nu_\beta} = \sum_k U^*_{\alpha k} U_{\beta k} e^{-i(E_k t/\hbar - p_k L/c\hbar)}. \tag{1.22}$$

In the ultra-relativistic approximation, the mass is very small compared to the momentum, and the length L travelled is to good accuracy given by the time travelled times the speed of light, ie $L \approx ct$. By using

$$E_k = \sqrt{p_k^2 + m_k^2 c^4} \approx p + \frac{1}{2}\frac{m_k^2 c^4}{p} \tag{1.23}$$

we find

$$A_{\nu_\alpha\to\nu_\beta} \approx \sum_k U^*_{\alpha k} U_{\beta k} e^{-i\frac{1}{2}\frac{m_k^2 c^4 t}{p\hbar}}. \tag{1.24}$$

Using $L \approx ct$ and $E \approx p$, as L and E are easier to measure than t and p, one finds

$$A_{\nu_\alpha\to\nu_\beta} \approx \sum_k U^*_{\alpha k} U_{\beta k} e^{-i\frac{1}{2}\frac{m_k^2 c^3 L}{E\hbar}}. \tag{1.25}$$

One could also assume that all the ν_ks had the same energy rather than the same momentum, but in our approximation of a ultra-relativistic neutrino the two assumptions gives the same result.

The transition probability is given by

$$\begin{aligned} P_{\nu_\alpha\to\nu_\beta} &= |A_{\nu_\alpha\to\nu_\beta}|^2 \\ &= \sum_{k,j} U^*_{\alpha k} U_{\beta k} e^{-i\frac{1}{2}\frac{m_k^2 c^3 L}{E\hbar}} U_{\alpha j} U^*_{\beta k} e^{-i\frac{1}{2}\frac{m_j^2 c^3 L}{E\hbar}} \\ &= \sum_{k,j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} e^{\frac{-i(m_k^2 - m_j^2)c^3 L}{2E\hbar}} \end{aligned} \tag{1.26}$$

as the neutrino flavour state ν_β must have the same energy as the neutrino flavour state ν_α given the assumption that all ν_k must have the same energies; the energy of ν_α. This result is quite interesting in itself, as the probability depends on the difference of the square of the masses $\Delta m^2_{kj} = m^2_k - m^2_j$ and the length travelled L. The sum of matrices is, however, worth getting rid of.

$$\begin{aligned} P_{\nu_\alpha \to \nu_\beta} &= \sum_{k,j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta k} e^{\frac{-i\Delta m^2_{kj} c^3 L}{2E\hbar}} \\ &= \sum_{j=k} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} e^{\frac{-i\Delta m^2_{kj} c^3 L}{2E\hbar}} + \sum_{k\neq j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta k} e^{\frac{-i\Delta m^2_{kj} c^3 L}{2E\hbar}} \\ &= \sum_{k} U^*_{\alpha k} U_{\beta k} U_{\alpha k} U^*_{\beta k} + \sum_{k\neq j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta k} e^{\frac{-i\Delta m^2_{kj} c^3 L}{2E\hbar}} \end{aligned} \tag{1.27}$$

The relation given in equation (1.17) implies

$$UU^\dagger = 1 \Leftrightarrow \sum_k U_{\alpha k} U^*_{\beta k} = \delta_{\alpha\beta}, \tag{1.28}$$

which will become useful as it simplifies the first term in equation (1.27):

$$\begin{aligned} \sum_k U^*_{\alpha k} U_{\beta k} U_{\alpha k} U^*_{\beta k} &= \sum_k |U_{\alpha k}|^2 |U_{\beta k}|^2 \\ &= \sum_{k=j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} \\ &= \sum_{k,j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} - \sum_{k\neq j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j} \\ &= \sum_k U_{\beta k} U^*_{\beta j} \delta_{\alpha\beta} - 2 \sum_{k>j} \mathcal{R}\left[U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j}\right] \\ &= \delta_{\alpha\beta} - 2 \sum_{k>j} \mathcal{R}\left[U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta j}\right]. \end{aligned} \tag{1.29}$$

The last term can be simplified in a similar manner. For $k \to j$ the exponential function gives the conjugate value, and so does the matrix product. Thus the complex parts of $k > j$ cancel the complex parts of $k < j$, and we can write

$$\begin{aligned} \sum_{k\neq j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta k} e^{\frac{-i\Delta m^2_{kj} c^3 L}{2E\hbar}} &= 2\mathcal{R} \sum_{k>j} U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta k} e^{\frac{-i\Delta m^2_{kj} c^3 L}{2E\hbar}} \\ &= 2 \sum_{k>j} \mathcal{R}\left[U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta k}\right] \cos\left(\frac{-i\Delta m^2_{kj} c^3 L}{2E\hbar}\right) \\ &\quad -2 \sum_{k>j} \mathcal{I}\left[U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta k}\right] \sin\left(\frac{-i\Delta m^2_{kj} c^3 L}{2E\hbar}\right). \end{aligned} \tag{1.30}$$

Taking it all together we then find

$$P_{\nu_\alpha \to \nu_\beta} = \delta_{\alpha\beta} - 2\sum_{k>j} \mathcal{R}\left[U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta k}\right] \left(1 - \cos(\frac{-i\Delta m^2_{kj} c^3 L}{2E\hbar})\right) + 2\sum_{k>j} \mathcal{I}\left[U^*_{\alpha k} U_{\beta k} U_{\alpha j} U^*_{\beta k}\right] \sin(\frac{-i\Delta m^2_{kj} c^3 L}{2E\hbar}). \quad (1.31)$$

(1.32)

This is the probability that a neutrino of flavour α and energy E propagates into a neutrino of flavour β when it has traveled a length L. It can be calculated exact only if we know the difference of the squared masses of all the neutrinos, but is not sensitive to the absolute mass of the mass eigenstates.

The *survival probability*; the probability that a neutrino measured after travelling a length L is still in the same state, is given by

$$P_{\nu_\alpha \to \nu_\alpha} = 1 - 4\sum_{k>j} \mathcal{R}\left[|U_{\alpha k}|^2 |U_{\alpha j}|^2\right] \sin^2(\frac{-i\Delta m^2_{kj} c^3 L}{4E\hbar}) \quad (1.33)$$

because the $\mathcal{I}\left[|U_{\alpha k}|^2 |U_{\alpha j}|^2\right]$ vanishes as the expression is real.

1.2.3 Two neutrino mixing

Looking at equation (1.33) in the case of just two mass eigenstates, we get

$$\begin{aligned} P_{\nu_\alpha \to \nu_\alpha} &= 1 - 4\mathcal{R}\left[|U_{\alpha 2}|^2 |U_{\alpha 1}|^2\right] \sin^2(\frac{-i\Delta m^2_{21} c^3 L}{4E\hbar}) \\ &= 1 - C\sin^2(\frac{-i\Delta m^2_{21} c^3 L}{4E\hbar}), \end{aligned} \quad (1.34)$$

where C is a constant determined by the mixing matrix. The survival probability equals one minus an oscillating function (which depends on the square of the mass difference, the length of travel and the energy).

For two generations, a unitary mixing matrix will be on the form

$$U = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \quad (1.35)$$

and, for mixing between electron and muon flavour, we have

$$|\nu_\alpha> = \sum_k U^*_{\alpha k}(\theta)|\nu_k> = \sum_k U_{\alpha k}(\theta)|\nu_k> \qquad \alpha = e, \mu \qquad k = 1, 2 \quad (1.36)$$

and

$$|\nu_k(E, L)> = e^{\frac{-im^2_k c^3 L}{2E\hbar}} |\nu_k(E, L = 0)> . \quad (1.37)$$

From this we find

$$
\begin{aligned}
&|\nu_e> = \cos\theta|\nu_1> + \sin\theta|\nu_2> \\
&|\nu_e(L)> = \cos\theta e^{\frac{-im_k^2c^3L}{2E\hbar}}|\nu_1(E, L=0)> + \sin\theta e^{\frac{-im_k^2c^3L}{2E\hbar}}|\nu_2(E, L=0)> \\
&|\nu_\mu> = -\sin\theta|\nu_1> + \cos\theta|\nu_2>
\end{aligned}
\tag{1.38}
$$

and the transition probability is defined by

$$
\begin{aligned}
P_{\nu_e\to\nu_\mu} = |<\nu_\mu|\nu_e(L)>|^2 &= \left|(-\sin\theta, \cos\theta)\begin{pmatrix} \cos\theta e^{\frac{-im_1^2c^3L}{2E\hbar}} \\ \sin\theta e^{\frac{-im_2^2c^3L}{2E\hbar}} \end{pmatrix}\right|^2 \\
&= \sin^2\theta\cos^2\theta(e^{\frac{-im_2^2c^3L}{2E\hbar}} - e^{\frac{-im_1^2c^3L}{2E\hbar}})(e^{\frac{im_2^2c^3L}{2E\hbar}} - e^{\frac{im_1^2c^3L}{2E\hbar}}) \\
&= 2\sin^2\theta\cos^2\theta\left[1 - \cos(\frac{\Delta m_{12}c^3L}{2E\hbar})\right] \\
&= \left[\sin(2\theta)\sin\left(\frac{\Delta m_{12}^2c^3L}{4E\hbar}\right)\right]^2.
\end{aligned}
\tag{1.39}
$$

It is easy to verify that we would get the same result from equation (1.32) with $\alpha \neq \beta$, as the imaginary part is zero.

The probability for mixing depends on the mixing angle θ as well as the difference of the squared masses, the length travelled and the energy of the neutrino. If the mixing angle is $\theta = \pi/4$, we have so called *perfect mixing*. If the mixing angle is zero, we have $P_{\nu_e\to\nu_\mu} = 0$, and the mixing matrix U is diagonal.

1.2.4 Three generations

With three generations of massive and flavour neutrinos, the mixing matrix is a bit more complex. In the Dirac case, it can be parametrized as

$$
U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}s_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{1.40}
$$

where $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$. θ_{ij} is in the range $0 \leq \theta_{ij} \leq \pi/2$, and the CP-violating phase δ is in the interval $0 \leq \delta \leq 2\pi$. It is worth noting that the CP-violating phase δ always occurs in a product with $sin\theta_{13}$, thus measuring δ is hard if the angle θ_{13} is small.

In the case of Majorana neutrinos, the mixing matrix looks different, as it contains three CP-violating phases and not one, as in the Dirac case. The mixing matrix can be written as a matrix product

$$
U^M = UD^M, \tag{1.41}
$$

where U is the mixing matrix in the case of Dirac neutrinos, and D^M is a diagonal matrix containing two extra phase factors:

$$D^M = diag[e^{i\lambda_1}, e^{i\lambda_2}, e^{i\lambda_3}] \qquad \lambda_1 = 0. \tag{1.42}$$

As the mixing matrix in the case of Dirac neutrinos is unitary if it is square, it follows U^M is unitary too. The Majorana phases does not affect the neutrino oscillations, so whenever they are not needed they will be omitted. The Majorana phases does however have a role in the violation of lepton number allowed by Majorana behavior of neutrinos.

There are only three light ($m_\nu < m_Z/2$) weakly interacting neutrinos. This has been shown by studying the decay of Z-bosons at LEP. If the number of massive eigenstates exceeds the three expected states, there must be more than three possible superpositions of these states. The "extra" superpositions will, however, not be active flavor-states (weakly interacting states), they will be *sterile* states (singlets of the electro-weak contribution to the Standard Model, and hence not interacting with other particles).

1.3 Neutrino Mass Schemes

Neutrino oscillation experiments shows that there is a difference in mass between the different states. It is commonly believed that there are three massive states, $\nu_i,\ \ i = 1, 2, 3$. Neutrino oscillation experiments are however not sensitive to the absolute mass scale of the neutrinos. They tell the difference in mass, but not which neutrino is the lighter, and they say nothing about if the mass is of the order 10^{-7} eV or 10^4 eV.

There are a number of different "mass schemes", or orderings of the masses. The *normal hierarchy* is when $m_1 < m_2 < m_3$. *Inverted hierarchy* has $m_3 < m_1 < m_2$. These two hierarchies is shown in figure 1.1. This drawing is made assuming that the neutrinos are not very heavy, and that $\Delta m^2_{21} = \Delta m^2_{sun}$, $\Delta m^2_{32} = \Delta m^2_{atm}$ and $|\Delta m^2_{32}| >> |\Delta m^2_{21}|$. Thus $\Delta m^2_{31} \approx \Delta m^2_{32} = \Delta m^2_{atm}$.

If the neutrinos all are very heavy, such that the mass square difference of the neutrino masses becomes insignificant, and $m_1 \approx m_2 \approx m_3$, then the masses is said to be *degenerate*.

1.4 Experimental Neutrino Mass Limits

As the observations of neutrino oscillations confirmed that at least some of the mass eigenstates does have a non-zero mass, effort has been made to find the mass with as high accuracy as possible. As the neutrinos only interact weakly, and since the mass is indeed very little, the determination of the neutrino mass has been a complicated process that is still short of accurate results. In this section I will look into three methods of finding limits on the neutrino mass; direct limits from end

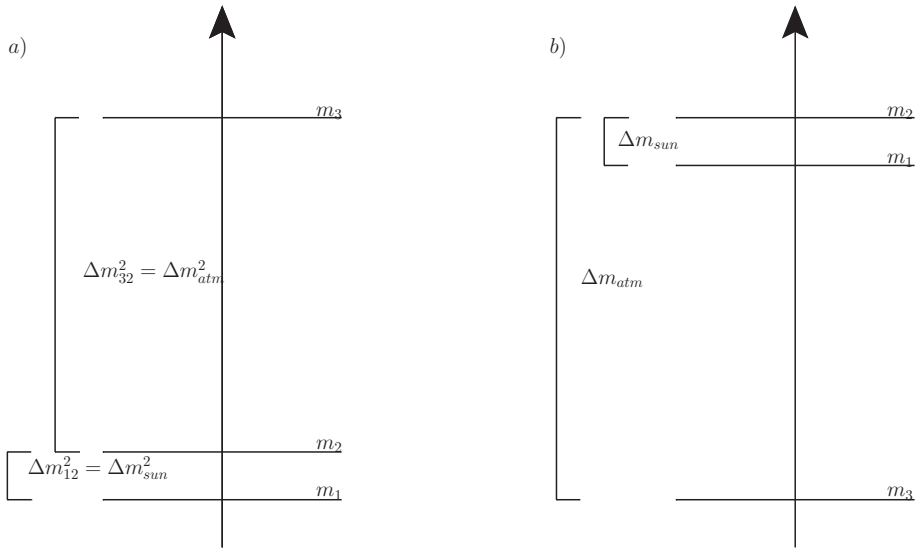


Figure 1.1: Normal (a) and inverted (b) hierarchies

point measurements, indirect limits from double β-decay and finally cosmological observations. I will start with the results from the neutrino oscillation experiments determining the mass square difference of the neutrino masses. This section is based on the reference [6, 5, 1], other references are listed where needed.

1.4.1 Mass square difference

The mass square differences are found from neutrino oscillation experiments as described in section 1.2. A number of different experiments has been performed, and the results varies quite a bit.

Atmospheric neutrinos

Atmospheric neutrinos are created by the interaction of cosmic rays with the particles (nuclei) in the atmosphere. The cosmic rays primarily consist of photons, and the interactions in the atmosphere typically produces pions, which typically decays to μ and ν_μ. Electron flavours of neutrinos are also created, but to a lesser extent. From the atmosphere and to observatories below sea level, the neutrinos have time to oscillate.

The first strong evidence of neutrino oscillations was obtained by the Kamiokande experiment in the 1990's, when they discovered that the upward-going and downward-going muon-neutrinos had a different flux. This indicates that some muon neutrinos disappears during their travel through Earth. The most reliable explanation is that the muon neutrinos oscillates into other flavours. Both Kamiokande and its successor Super-Kamiokande has provided good data on atmospheric neutrinos.

The most recent results form the Super-Kamiokande [7](April 2010) experiment suggests that

$$\begin{aligned} \Delta m_{32} &= 2.1 \times 10^{-3} eV^2 && \text{Best fit} \\ \sin^2\theta_{23} &= 0.5 && \text{Best fit} \\ \Delta m_{32} &= (1.9 - 2.6) \times 10^{-3} eV^2 && \text{Normal Hierarchy, 90\% CL} \\ \Delta m_{32} &= (1.7 - 2.7) \times 10^{-3} eV^2 && \text{Inverted Hierarchy, 90\% CL} \\ \sin^2\theta_{23} &\epsilon [0.407 - 0.583] && \text{90\% CL.} \end{aligned} \tag{1.43}$$

The Super-Kamiokande is a huge tank of water surrounded by photon multiplier. It detects neutrinos by Cerenkov radiation, and this works for all types of neutrinos, but is more efficient for the electron neutrino.

Solar neutrinos

Huge numbers of electron neutrinos are produced in the sun, by fusion of H to He, fusion of He and H to ^7Be and the decay of ^8Be. The SNO (Sudbury Neutrino Observatory) observes neutrinos, using a tank of heavy water (D_20) rather than ordinary water. This allowes for the detection of all types of neutrinos, through the reactions

$$\begin{aligned} \nu_{e,\mu,\tau} + e &\rightarrow \nu_{e,\mu,\tau} + e \\ \nu_{e,\mu,\tau} + d &\rightarrow \nu_{e,\mu,\tau} + p + n \\ \nu_e + d &\rightarrow e^- + p + p. \end{aligned} \tag{1.44}$$

In 2001 SNO published results showing that the total neutrino flux corresponds well with the predicted electron neutrino flux from the sun, confirming both neutrino oscillations and the standard solar model.

Being able to measure all types of neutrinos, measurements at SNO of the solar neutrinos detected what is sometimes referred to as "smoking-gun" evidences for the oscillation of electron neutrinos into other types. From the two deuteron breakup reactions, SNO finds that the flux of electron neutrinos $\phi(\nu_e)$ relates to the flux of other flavours $\phi(\nu_{\mu,\tau})$ from ^8B-production like

$$\frac{\phi(\nu_e)}{\phi(\nu_e) + \phi(\nu_{\mu,\tau})} = 0.340 \pm 0.023(stat)^{+0.029}_{-0.031}(syst). \tag{1.45}$$

This has given clearer evidence of neutrino-oscillations than ever before.

The total neutrino flux from the sun originating from ^8B-production has been calculated to amount to $5.49^{+0.95}_{-0.81} \cdot 10^6 cm^{-1}s^{-1}$. Detection of neutral current deuteron breakup has found a total active flux of $4.94 \pm 0.21(stat)^{+0.38}_{-0.34}(syst)$, which is in reasonable agreement. Thus it seems like the production of neutrino flux from the sun is correctly understood, although the deviation from the expected result obtained in measurements need further investigation.

Combined efforts using both solar neutrinos and accelerator based experiments has concluded that [6]:

$$\begin{aligned} \sin^2 2\theta_{12} &= 0.86^{+0.03}_{-0.04} \\ \Delta m^2_{21} &= (8.0 \pm 0.3) \times 10^{-5} eV^2. \end{aligned} \tag{1.46}$$

Ground based experiments, neutrinos from reactors

Although neutrinos come in from space all the time, there are good sources of neutrinos at Earth as well. Nuclear reactors produce lots of neutrinos. In this section, I have made use of the reference [2] as well as the references given at the beginning of this section.

Early reactor-based on $\overline{\nu}_e$ in the 1980's and early 1990 were not able to detect neutrino oscillations with their short base-line detections. Later experiments like CHOOZ in 1999 and experiments by the Palo Verde Nuclear Generating Station of oscillations in $\overline{\nu}_e$ in 2000 [8, 9] later confirmed the earlier results. These experiments used long base-lines (the Palo Verde experiment measured the neutrino flux at a distance of 800 meters from the reactor).

Later experiments, with longer base-lines has later been performed, and these have been able to detect oscillations. The KamLAND experiment in Japan was designed to detect $\overline{\nu}_e$ from reactors at a distance fro 80 to 800 km, with an average distance of about 180 km. The KamLAND experiment has measured a deficit in the $\overline{\nu}_e$ flux relative to the non-oscillatory expectations. Combined with solar neutrino experiments, the results yield [10]:

$$\begin{aligned} \Delta m^2_{sun} &= 7.9^{+0.6}_{-0.5} \times 10^{-5} eV^2 \\ \tan^2 \theta_{sun} &= 0.4^{+0.10}_{-0.07} \end{aligned} \tag{1.47}$$

Also accelerators are used to perform ground-based experiments, using muon-neutrino sources. The K2K is a long base-line experiment in Japan, giving further information on atmospheric oscillations. It has reported on a best fit value of the mass square difference [1]:

$$\Delta m^2_{atm} = 2.8 \times 10^{-3} eV^2. \tag{1.48}$$

1.4.2 The determination of absolute neutrino masses

Although neutrino oscillation experiments are a very successful tools for determining the mass square difference of the neutrinos, they are not able to find the absolute neutrino mass, nor the scale of the masses. Knowing the mass of one of the neutrino mass states, or one of the neutrino flavour states and the mixing angles, plus the hierarchy of the neutrino masses, one can get quite far in computing the individual neutrino masses. In this section I will look into some of the methods used to determine the neutrino mass scale. β-decays are used, and have been used for years when it comes to determining neutrino properties, but more recently

cosmological methods has been able to provide good limits on the neutrino mass scale. As well as the references listed above, reference [11] has been useful.

End point β spectrum of tritium decay

The end point of the energy spectrum of the electron in a β decay is in theory a very good way of measuring the effective electron neutrino mass. It can be shown that the decay rate of a β-decay

$$X_M^N \rightarrow X_{M+1}^{\prime N} + e^- + \overline{\nu_e} \tag{1.49}$$

has a decay-rate that goes like

$$\frac{1}{|\mathbf{p}_e|}\sqrt{\frac{d\Gamma}{d|\mathbf{p}_e|}} \propto \sqrt{(m_n - m_p - E_e\sqrt{(m_n - m_p - E_e)^2 - m_{\nu_e}^2})}. \tag{1.50}$$

A somewhat simplified derivation of this relation is shown in appendix B.

In a β-decay, the amount of energy carried away by the electron and the anti-neutrino depends on the difference of mass of the mother and daughter nuclei. The electron and the anti-neutrino shares the released energy, but the amount carried by the electron is restricted by the mass of the anti-neutrino. If the anti-neutrino has a mass, then the electron cannot take all the energy released. In fact, finding an energy deficit in the process of a β-decay did inspire the realisation that another particle must be created in the decay process, which lead to the "invention" of the neutrino. The maximal energy that the electron carries gives information about the effective electron-neutrino mass.

From equation (1.50) we see that the squarer root differential decay-rate divided by the electron momentum exhibits a linear behaviour if the neutrino is massless. If, however, the neutrino has a mass, there is a deviation from linearity. In order to be able to do such an experiment, one have to be able to cope with the fact that the decay-rate for electrons with maximal energy is very small, and one need a high energy resolution on the electron energy.

In particular, the tritium β-decay has been used to find upper limits on the electron neutrino mass.

$$^3_1T \rightarrow ^3_2 He + e^- + \overline{\nu_e}, \tag{1.51}$$

which produces 18.6 keV of energy. This process has quite a short halflife (12.3 years), making it appropriate for decay studies.

The Mainz ans Troitsk experiments give upper limits on the effective electron neutrino mass of $m_{\nu_e} < 2.3$ eV (95% C.L.) and $m_{\nu_e} < 2.05$ eV (95% C.L.) [11]. The KATRIN experiment (the Karlsruhe Tritium Neutrino experiment) is designed to measure the neutrino mass with a sensitivity of 0.2 eV and detection limit at 0.35 eV. Data collection is scheduled to start in 2012 [12].

Neutrinoless double β Decay

Neutrinoless double β-decay is a field of great interest as a successful experiment would bot enable a good way to calculate the absolute mass of the neutrino and a way to decide upon the nature of neutrinos. It does, however, have some quite bothersome complications, as the process of double β-decay is a very rare second order process. In this subsection, I have used the references [2, 3] in addition to the general references for this section.

An ordinary double β-decay is a process that can occur in a nuclei where a single beta decay is not energetically favoured, but a double decay where two neutrons decays simultaneously is allowed:

$$X_M^N \rightarrow X_{M+2}'^N + e_1^- + e_2^- + \overline{\nu_{e_1}} + \overline{\nu_{e_2}}. \tag{1.52}$$

Measuring the energy spectrum of the emitted electrons, the decay rate as a function of the emitted energy would yield a continuous spectrum, as the neutrinos share the emitted energy with the electrons. This process is, however, not very good for determining any neutrino mass.

A similar process that is useful for determining the neutrino mass is the neutrinoless double β-decay:

$$X_M^N \rightarrow X_{M+2}'^N + e_1^- + e_2^-. \tag{1.53}$$

A measurement of this process, would give a decayrate that peaks at the electron energy that equals the released energy of the process, as there are no neutrinos to share the energy with. In this process, the conservation of lepton number is violated, which points to it being a Majorana process, only occurring if the neutrinos are massive Majorana particles. The reason why this process is allowed given Majorana nature of the neutrinos is dual; first of all, in the Majorana case, the lepton number is not a conserved quantity. The second reason is a bit more complicated. For Majorana particles, the particle and antiparticle are equal. Thus the anti-neutrino emitted in the one β-decay can be absorbed as a neutrino causing the other β-decay. It is, however, so that the anti-neutrino emitted is righthanded, and the neutrino absorbed by the second reaction must be lefthanded. This is solved by the mass of the neutrino. Only massless neutrinos are helicity eigenstates. Thus the anti-neutrino emitted has a small lefthanded component. The lefthanded component is proportional to the neutrino mass. This means that there is a probability proportional to the neutrino mass that the second neutron perceives the emitted righthanded anti-neutrino as a lefthanded neutrino, and thus " absorbs" the neutrino, starting the second β-decay. Thus, the neutrino plays a "virtual" role in the process.

The halflife of the neutrinoless double β-decay is given proportional to $\left(\frac{|m_{\nu_e,\beta\beta}|}{m_e}\right)^2$, where $|m_{\nu_e,\beta\beta}|$ is the effective neutrino mass given by

$$|m_{\nu_e,\beta\beta}| = |\sum_i U_{ei}^2 m_i|. \tag{1.54}$$

U is the mixing matrix in the case of Majorana nature (it includes the Majorana phases). Thus a positive detection of a neutrinoless double beta decay would give us a good indication of the neutrino masses scale. Knowing the mass hierarchy, the mixing angles and Majorana phases, it would give us enough information to find all the neutrino masses. ans most importantly, it would reveal the neutrino nature.

The nuclei studied and found to have been undergoing double β-decays are ^{48}Ca, ^{76}Ge, ^{83}Se, ^{96}Zr, ^{100}Mo, ^{116}Cd and ^{150}Nd. No unquestionable signal has been observed, but the Heidelberg-Moscow ^{78}Ge experiment claims to have set a 95% C.L. lower bound to the halflife of ^{76}Ge to neutrinoless double β-decay to [13]

$$T_{1/2}^{0+nu}(^{76}Ge) > 1.9 \times 10^{25} y, \tag{1.55}$$

and the IGEX ^{76}Ge experiment claim to have found a comparable limit. This gives

$$|m_{\nu_e,\beta\beta}| = (0.2 - 0.6) eV (99.73\% C.L.) \tag{1.56}$$

One of the problems is that the nuclear matrix elements used in determining the half life of the process gives the halflife calculation a systematic uncertainty of about a factor of three. Allowing for this uncertainty [13] finds the limits to be

$$|m_{\nu_e,\beta\beta}| = (0.1 - 0.9) eV (99.73\% C.L.). \tag{1.57}$$

Also other later experiments has found lower bounds to halflifes of double β-decay of nuclei, but none has an undisputed observation of a neutrinoless double β-decay. The results of the experiments are heavily debated because of the uncertainty in the nuclear matrix elements, the background noise and the small statistics provided by the experiments.

Cosmology

The neutrinos inhabit the universe in numerous amounts. The understanding of the physics of the early universe is good, and this allows for a good understanding of the number density of the neutrinos in the neutrino background. There are more than 300 neutrinos for every cm^3 in the universe!

The impact of neutrinos on cosmology is mainly to reduce structure formation on small scales, as neutrino free-streaming suppresses structure formation on a scales smaller than a mass dependent scale called the free-streaming scale. Neutrinos also affect the cosmic expansion and the cosmic microwave background radiation (the CMB) in a way that also depends on the neutrino mass. This effects will be discussed to greater detail in part II of this thesis, but the interesting effect at this point is that the neutrino mass affects cosmology, allowing for cosmology to probe the mass. Cosmology is in a good approximation only sensitive to the sum of the neutrino masses, $M_\nu = \sum_i m_i$. As oscillation experiments only measure mass differences, and the β-decay experiments have not yet provided very strong results, cosmology has became an important tool for estimating neutrino masses.

For the last decade, various cosmological data has been used to estimate neutrino masses. The results varies according to the data sets used and the priors on other cosmological parameters assumed. The CMB, mostly provided by the Wilkinsom Microwave Anisotropy Probe (WMAP) has been used, often in combination with large scale structure data by the Sloan Digital Sky Survey (SDSS) and 2dFGRS. Also data from Supernova type 1a data, and priors from big bang nucleosyntesis have been applied. The limits on the total neutrino mass generally gives lower masses when adding more data and more priors, but in doing so the result obtained is limited by the validity of the priors applied. Upper limits on neutrino masses have given upper limits on the total neutrino mass in the interval $M_\nu < 0.17 - 2.0$ eV. The limit of $M_{nu} < 2.0$ eV is obtaied using CMB alone (WMAP) [14], while the lower upper limits has been obtained combining WMAP with large scale structure surveys and sometimes also other observations. A reasonable upper limit seems to be about $M_\nu < 0.6 - 0.7$ eV [15, 16]. These limits are much lower than the ones obtained from single β-decay experiments. An easily accessible review on the connection between cosmology and neutrino physics is found in [15].

Although cosmology gives limits on the neutrino mass that are substantially lower than the limits provided by β-decay experiments, the limits obtained are not made by direct observations. What is observed, to first order, is the effect of the total neutrino mass on cosmology, and not the mass itself. Thus the limits obtained are limited by our understanding of the evolution of the universe. For example, the big band model and a ΛCDM model of a universe dominated by cold dark matter and a cosmological constant is assumed. If this turns out to be wrong, the neutrino mass limits are worthless. Small corrections to the model assumed today can also have profound consequences for the mass limits. An other issue is the degeneracy between the cosmological parameters. As an example, [17] showed that using CMB data alone there is a considerable degeneracy between the Hubble constant and the neutrino mass. Degeneracy of parameters makes the individual parameters harder to estimate. Also, most methods used to deduct a neutrino mass from cosmological observations assumes that the neutrinos can be assumed to all have the same weight, ie. that they are degenerate. This is, however, not entirely true, as we know from oscillation experiments that there is a mass splitting. To what extent this approximations influences the mass-limits obtained from cosmology is not fully understood.

Although there are uncertainties attached to the limits on the neutrino mass from cosmology, it is in great interest to cosmologist both to determine the effect of neutrinos on cosmology and to nail a value of the neutrino mass. As the total mass of the cosmic neutrinos are affecting cosmology and thus the observations we have of the universe, other cosmological parameters and the overall cosmological model cannot be determined without knowing how and to what extent neutrinos influence the cosmic evolution, and this depends on the size of the mass.

Chapter 2

The Seesaw Mechanism

A question that has puzzled particle physicists since the first detections of the neutrino, is its mass. The neutrino is very light, compared to other fermions. While the lightest charged fermion, the electron, has a rest energy of about 0.5 MeV, the rest energy of a neutrino seems to be at most of the order of 1 eV. Why is the neutrino so light?

The standard model assumes massless neutrinos, and although an inclusion of massive neutrinos to the standard model is possible[1], the standard model is not able to provide a plausible explanation of the smallness of the neutrino mass compared to charged fermions. Thus the generating process of neutrino masses takes us beyond the standard model, and can shed light on the nature of fundamental quantum field theory and fundamental physics not yet discovered. A good neutrino mass model must both explain the generating process of the neutrino mass, and why the mass is so small. In this section, I will look at neutrino mass generating theories, focusing on the seesaw mechanism, which is one of the most popular models for explaining the neutrino mass. The seesaw theory does not only account for the smallness of the mass, it is also able to describe different neutrino mass hierarchies. This section is mostly based on the references [1] and [3].

2.1 Grand Unification Theory

The standard model has provided a powerful tool for describing three of the four fundamental forces of nature through a gauge quantum field theory. It has, however, not given a full description of neither particle physics nor nature as a whole, and one of the shortcomings of the theory is its inability to describe the neutrino mass.

The electromagnetic force, the weak force and the strong force each have individual coupling "constants" in the standard model, and these couplings are energy-dependent. The electromagnetic and weak forces are united in a electroweak theory

[1] Such models allows for righthanded neutrinos, so that the neutrinos can couple to the Higgs field. These models generally provides no explanation of the smallness of the neutrino mass.

at high energies. The energy scale at which the two forces are united, is called the electroweak scale, and this scale is at the order of 100 GeV. At everyday-temperatures and energies, the electroweak symmetry is broken, leaving us with the weak and the electromagnetic force. Thus the apparent symmetry of a theory might not be the actual symmetry, and we might only see a small part of a bigger picture.

There exists theories where it is assumed that the coupling constants of the forces mentioned above will unite into one scale at high enough energies, and that our perception of three forces is only due to a spontaneous symmetry breaking. These theories are called grand unified theories (GUT). At energies higher than the GUT scale, the forces will be described by a common theory. The GUT scale is assumed to be of the order $10^{14} - 10^{16}$ GeV.

2.2 The General Idea

The general idea of the seesaw mechanism is not too complicated. As the neutrinos carry no electric charge, they are allowed to possess both Dirac and Majorana behaviour. Thus they can have both Majorana and Dirac mass terms. The Majorana mass term consists of a mass associated with the lefthanded neutrinos and a mass associated with the righthanded neutrinos. It can be shown that (and I will do this later in this chapter) the full mass matrix can be written as

$$M = \begin{pmatrix} m_L & m_D \\ m_D & m_R \end{pmatrix}. \tag{2.1}$$

Diagonalizing this matrix to find the physical masses (demanding $m_L = 0$), we find two eigenstates

$$\begin{aligned} |m_1| &\approx \left|\frac{m_D^2}{m_R}\right| \\ m_2 &\approx m_R. \end{aligned} \tag{2.2}$$

The standard model places constraints on the fermion masses by the electroweak symmetry-breaking scale. This applies to the charged fermions, and to the Dirac neutrino mass. The righthanded neutrino is not a part of the standard model, and thus the standard model places no constraints on the mass associated with it, m_R. It is often assumed that this mass is generated in the breaking of the GUT scale, which is a lot larger than the electroweak scale, and thus m_R is limited from above by the GUT energy. Hence one of the masses, m_D, could be very small, and the other mass, m_R, very large. If one assume that the large mass, m_2, is mostly associated with the righthanded neutrinos, which are sterile, as they do not interact with the other particles of the standard model, this mechanism gives out a small observed neutrino mass quite naturally, by introducing a very heavy righthanded neutrino as a counterweight the electroweak scale.

2.3 The Seesaw Mechanism for one Generation

From the earlier discussions in section 1.1.1 it is clear that if the neutrinos have Dirac behaviour, the mass term will be the Dirac mass term

$$\mathcal{L}^D_{mass} = -m_D \overline{\nu}\nu. \tag{2.3}$$

If the neutrinos are Majorana particles, an additional massterm will be generated:

$$\mathcal{L}^M_{mass} = -\frac{1}{2} m_L \overline{\nu^C_L} \nu_L - \frac{1}{2} m_R \overline{\nu^C_R} \nu_R + h.c. \tag{2.4}$$

where the right handed and left handed neutrinos generate separate mass terms with different masses. The factor $1/2$ not present in the Dirac mass terms originates in the Majorana condition which states that ν^C_L and $\overline{\nu_L}$ are not independent. The hermitian conjugate of $\overline{\nu^C_L}\nu_L$ is

$$\begin{aligned}
\left(\overline{\nu^C_L}\nu_L\right)^\dagger &= \nu^\dagger_L \overline{\nu^C_L}^\dagger \\
&= \nu^\dagger_L \left(\left(\nu^C_L\right)^\dagger \gamma_0\right)^\dagger \\
&= \nu^\dagger_L \gamma^\dagger_0 \nu^C_L \\
&= \nu^\dagger_L \gamma_0 \nu^C_L \\
&= \overline{\nu_L}\nu^C_L.
\end{aligned} \tag{2.5}$$

The same kind of calculations gives

$$\left(\overline{\nu^C_R}\nu_R\right)^\dagger = \overline{\nu_R}\nu^C_R, \tag{2.6}$$

and the hermitian conjugate of $\overline{\nu_R}\nu_L$ is

$$\begin{aligned}
(\overline{\nu_R}\nu_L)^\dagger &= \nu^\dagger_L \overline{\nu_R}^\dagger \\
&= \nu^\dagger_L \left(\nu^\dagger_R \gamma_0\right)^\dagger \\
&= \nu^\dagger_L \gamma^\dagger_0 \nu_R \\
&= \nu^\dagger_L \gamma_0 \nu_R \\
&= \overline{\nu_L}\nu_R.
\end{aligned} \tag{2.7}$$

The full mass term, including both Majorana man Dirac mass terms, will then be

$$\mathcal{L}_{mass} = -\frac{1}{2} m_L \overline{\nu^C_L} \nu_L - \frac{1}{2} m_R \overline{\nu^C_R} \nu_R - \frac{1}{2} m_D \left(\overline{\nu_R}\nu_L + \overline{\nu_L}\nu_R\right) + h.c. \tag{2.8}$$

or

$$\mathcal{L}_{mass} = -\frac{1}{2} m_L \overline{\nu^C_L} \nu_L - \frac{1}{2} m_R \overline{\nu^C_R} \nu_R - m_D \overline{\nu_R}\nu_L + h.c. \tag{2.9}$$

Rephasing of the fields ν_R and ν_L can ensure two positive real mass terms, so I'll choose m_R and m_D positive and real, leaving m_L complex.

To write this out more efficiently, it is useful to introduce a lefthanded field column matrix

$$N_L = \begin{pmatrix} \nu_L \\ \nu_R^C \end{pmatrix} = \begin{pmatrix} \nu_L \\ \mathcal{C}\overline{\nu_R}^T \end{pmatrix}, \tag{2.10}$$

giving

$$N_L^C = \begin{pmatrix} \nu_L \\ \nu_R^C \end{pmatrix}^C = \begin{pmatrix} \nu_L^C \\ \nu_R \end{pmatrix} \tag{2.11}$$

and a mass matrix

$$M = \begin{pmatrix} m_L & m_D^\dagger \\ m_D & m_R \end{pmatrix} = \begin{pmatrix} m_L & m_D \\ m_D & m_R \end{pmatrix}. \tag{2.12}$$

The mass term can then be written

$$\begin{aligned}
\mathcal{L}_{mass} &= -\frac{1}{2}\overline{N_L^C} M N_L + h.c. \\
&= -\frac{1}{2}\begin{pmatrix} \overline{\nu_L^C} & \overline{\nu_R} \end{pmatrix}\begin{pmatrix} m_L & m_D \\ m_D & m_R \end{pmatrix}\begin{pmatrix} \nu_L \\ \nu_R^C \end{pmatrix} + h.c. \\
&= -\frac{1}{2}\begin{pmatrix} \overline{\nu_L^C} & \overline{\nu_R} \end{pmatrix}\begin{pmatrix} m_L\nu_L + m_D\nu_R^C \\ m_D\nu_L + m_R\nu_R^C \end{pmatrix} + h.c. \\
&= -\frac{1}{2}\left(\overline{\nu_L^C} m_L \nu_L + \overline{\nu_L^C} m_D \nu_R^C \overline{\nu_R} m_D \nu_L + \overline{\nu_R} m_R \nu_R^C\right) + h.c. \\
&= -\frac{1}{2} m_L \overline{\nu_L^C}\nu_L - \frac{1}{2} m_D \overline{\nu_L^C}\nu_R^C - \frac{1}{2} m_D \overline{\nu_R}\nu_L - \frac{1}{2} m_R \overline{\nu_R}\nu_R^C + h.c. \\
&= -\frac{1}{2} m_L \overline{\nu_L^C}\nu_L - m_D \overline{\nu_R}\nu_L - \frac{1}{2} m_R \overline{\nu_R}\nu_R^C + h.c.
\end{aligned} \tag{2.13}$$

which is exactly the same as in equation (2.9). The matrix form of this equation becomes useful in the next section.

2.3.1 Diagonalizing the mass matrix

As the mass matrix in equation (2.13) is not diagonal, the chiral fields ν_L and ν_R does not have a well-defined mass. In order to find states of defined mass and the field masses, the mass matrix must be diagonalized to find the eigenvalues. But we don't want just any eigenvalues, we went real eigenvalues, as the masses must be real, and finding real eigenvalues demands a bit more work than just finding any eigenvalues. As we deal with a complex matrix, we have to do it the hard way, but we warm up with the simpler procedure of a real matrix.

The diagonalizing of a symmetric matrix is done with unitary matrices[2], such

[2] If the mass matrix is square, meaning that the number of flavour states equal the number of mass states, the matrix U that relates the flavour and mass states is unitary. This is assumed whenever otherwise is not explicitly stated.

that

$$N_L = Un_L \qquad\qquad n_L = \begin{pmatrix} \nu_{1L} \\ \nu_{2L} \end{pmatrix} \tag{2.14}$$

where the column matrix n_L contains the mass eigenstates and where the matrix U fulfills

$$U^T M U = M' \qquad\qquad M'_{ij} = \delta_{ij} m_i, \tag{2.15}$$

where m_i are the mass eigenstates. This gives, from equation (2.13)

$$\mathcal{L}_{mass} = -\frac{1}{2} \sum_{k=1,2} m_k \overline{\nu^C_{kL}} \nu_{kL} + h.c., \tag{2.16}$$

which clearly shows that the massive neutrinos are in fact Majorana neutrinos. The eigenvalues of M are

$$m_{2,1} = \frac{m_L + m_R}{2} \pm \frac{\sqrt{(m_L - m_R)^2 + 4m_D^2}}{2}, \tag{2.17}$$

but these eigenvalues are not in general real, as m_L may be complex and the term inside the square root can be negative.

Since the mass-eigenstates has to be real, and one of the components of the mass matrix in equation (2.13) is complex, the diagonalizing procedure has to be a bit more thorough than for a real matrix. From equation (2.15) one obtains (since U is unitary)

$$\begin{aligned} &U^T M U = M' \\ &(U^T)^\dagger U^T M U = (U^T)^\dagger M' \\ &M U = U^* M'. \end{aligned} \tag{2.18}$$

This can be expressed as

$$M v^{(j)} = m_j v^{(j)*}, \tag{2.19}$$

where I have introduced the vector $v^{(j)}$ which is the column j of the matrix U:

$$v^{(j)}_k = U_{kj}. \tag{2.20}$$

From equation (2.19) I am searching for real values of m_j. I separate the equation into real and imaginary parts

$$(\mathcal{R}(M) + i\mathcal{I}(M)) \left(\mathcal{R}(v^{(j)}) + i\mathcal{I}(v^{(j)})\right) = m_j \left(\mathcal{R}(v^{(j)}) - i\mathcal{I}(v^{(j)})\right) \tag{2.21}$$

which can be expressed as

$$\begin{pmatrix} \mathcal{R}(M) & -\mathcal{I}(M) \\ -\mathcal{I}(M) & -\mathcal{R}(M) \end{pmatrix} \begin{pmatrix} \mathcal{R}(v^{(j)}) \\ \mathcal{I}(v^{(j)}) \end{pmatrix} = m_j \begin{pmatrix} \mathcal{R}(v^{(j)}) \\ \mathcal{I}(v^{(j)}) \end{pmatrix}. \tag{2.22}$$

We can then write the four dimensional eigenvalue equation

$$\mathcal{M}V^{(j)} = m_j V^{(j)} \tag{2.23}$$

where

$$V^{(j)} = \begin{pmatrix} \mathcal{R}(v^{(j)}) \\ \mathcal{I}(v^{(j)}) \end{pmatrix} \tag{2.24}$$

is now a four dimensional vector. The eigenvalue equation (2.23) is now an equation that will provide real eigenvalues m_j, since both the matrix $\mathcal{M}$ and the eigenvectors $V^{(j)}$ are real.

Our job, in order to find the masses, is to solve for the eigenvalues of the matrix

$$\mathcal{M} = \begin{pmatrix} \mathcal{R}(M) & -\mathcal{I}(M) \\ -\mathcal{I}(M) & -\mathcal{R}(M) \end{pmatrix} = \begin{pmatrix} \mathcal{R}(m_L) & m_D & -\mathcal{I}(m_L) & 0 \\ m_D & m_R & 0 & 0 \\ -\mathcal{I}(m_L) & 0 & -\mathcal{R}(m_L) & -m_D \\ 0 & 0 & -m_D & -m_R \end{pmatrix}. \tag{2.25}$$

As this is a four dimensional matrix, one expect to find four (possibly degenerate) eigenvalues. It is, however, straightforward to show that if m_j is an eigenvalue, then so is $-m_j$ thus, we are looking for the positive eigenvalues. I will also show that the eigenvalues are in fact real.

The positive eigenvalues are [1]

$$\begin{aligned} m_{2,1} = \frac{\sqrt{2}}{2} \Big[|m_L|^2 + m_R^2 + 2m_D^2 \pm \Big[(\mathcal{R}(m_L) + m_R)^2 \left((\mathcal{R}(m_L) - m_R)^2 + 4m_D^2 \right) \\ + \mathcal{I}(m_L)^4 + 2\mathcal{I}(m_L)^2 \left(\mathcal{R}(m_L)^2 - m_R^2 + 2m_D^2 \right) \Big]^{1/2} \Big]^{1/2}. \end{aligned} \tag{2.26}$$

Rewriting the the term in the inner square-root gives:

$$\begin{aligned} (\mathcal{R}(m_L) + m_R)^2 \left((\mathcal{R}(m_L) - m_R)^2 + 4m_D^2 \right) \\ + \mathcal{I}(m_L)^4 + 2\mathcal{R}(m_L)^2 \left(\mathcal{I}(m_L)^2 - m_R^2 + 2m_D^2 \right) \end{aligned} = \begin{aligned} \left(\left(\mathcal{R}(m_L)^2 + \mathcal{I}(m_L)^2 \right) - m_R^2 \right)^2 \\ + 4m_D^2 \left(\mathcal{R}(m_L) + m_R \right)^2 \\ + 4\mathcal{I}(m_L)^2 m_D^2 \end{aligned}, \tag{2.27}$$

which is always positive. To ensure that the mass eigenvalues are real, I must check that

$$\begin{aligned} |m_L|^2 + m_R^2 + 2m_D^2 \geq \Big[\left(\left(\mathcal{R}(m_L)^2 + \mathcal{I}(m_L)^2 \right) - m_R^2 \right)^2 \\ + 4m_D^2 \left(\mathcal{R}(m_L) + m_R \right)^2 + 4\mathcal{I}(m_L)^2 m_D^2 \Big]^{1/2}. \end{aligned} \tag{2.28}$$

Since both terms obviously are positive, I square both sides:

$$\begin{aligned} (|m_L|^2 + m_R^2 + 2m_D^2)^2 \geq \left(\left(\mathcal{R}(m_L)^2 + \mathcal{I}(m_L)^2 \right) - m_R^2 \right)^2 \\ + 4m_D^2 \left(\mathcal{R}(m_L) + m_R \right)^2 + 4\mathcal{I}(m_L)^2 m_D^2, \end{aligned} \tag{2.29}$$

giving

$$\begin{array}{cc} |m_L|^4 + m_R^4 + 4m_D^4 & \geq \\ +2|m_L|^2 m_R^2 + 4|m_L|^2 m_D^2 + 4m_R^2 m_D^2 & \end{array} \begin{array}{c} \left(|m_L|^2 - m_R^2\right)^2 + 4|m_L|^2 m_D^2 \\ +4m_D^2 \left(2m_R \mathcal{R}(m_L) + m_R^2\right) \end{array} \tag{2.30}$$

thus

$$4m_D^4 + 4|m_L|^2 m_R^2 - 8m_D^2 m_R \mathcal{R}(m_L) \geq 0. \tag{2.31}$$

Since $\mathcal{R}(m_L)^2 \leq |m_L|^2$ we have

$$\begin{aligned} 4m_D^4 + 4|m_L|^2 m_R^2 - 8m_D^2 m_R \mathcal{R}(m_L) &\geq 4m_D^4 + 4\mathcal{R}(m_L)^2 m_R^2 - 8m_D^2 m_R \mathcal{R}(m_L) \\ &= 4\left(m_D^2 - m_R \mathcal{R}(m_L)\right)^2 \\ &\geq 0 \end{aligned} \tag{2.32}$$

which always holds as all the terms $\mathcal{R}(m_L), m_R, m_D$ are real.

This shows that this way of diagonalizing the mass-matrix ensures real eigenvalues, and we choose the positive terms:

$$\begin{aligned} m_{2,1} = \frac{\sqrt{2}}{2} \Big[&|m_L|^2 + m_R^2 + 2m_D^2 \pm \\ &\left[\left(|m_L|^2 - m_R^2\right)^2 + 4m_D^2 \left(\mathcal{R}(m_L) + m_R\right)^2 + 4\mathcal{I}(m_L)^2 m_D^2\right]^{1/2} \Big]^{1/2}. \end{aligned} \tag{2.33}$$

The Dirac limit is the case when the Majorana mass terms disappear, when $m_L = m_R = 0$. Looking at equation (2.33), we then get

$$\begin{aligned} m_{2,1} &= \frac{\sqrt{2}}{2} \left[2m_D^2\right]^{1/2} \\ &= \pm m_D, \end{aligned} \tag{2.34}$$

as we would have expected.

In the case of a real m_L, we obtain

$$m_{2,1} = \frac{1}{2} \left| m_L + m_R \pm \sqrt{(m_L - m_R)^2 + 4m_D^2} \right| \tag{2.35}$$

which is the exact same expression as I found for the 2×2 matrix diagonalized in equation (2.17).

2.3.2 The mixing angle

The mixing matrix U, which relates the flavour eigenstates with the mass eigenstates, $\Psi_L = U\Psi'_L$, is unitary. Hence it's columns are orthonormal. Such a matrix

can generally be written as

$$U = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} e^{i\lambda} & 0 \\ 0 & 1 \end{pmatrix} = OP = \begin{pmatrix} \cos\theta e^{i\lambda} & \sin\theta \\ -\sin\theta e^{i\lambda} & \cos\theta \end{pmatrix} \tag{2.36}$$

where only the physical phase factor is included, and $0 \leq \theta \leq \frac{\pi}{2}$ and $0 \leq \lambda < 2\pi$. O is the matrix containing the real rotational matrix, and P is the matrix with the phase factor. Note that O is unitary, and that P is not contributing to the diagonalization, only to the phases of the matrix diagonalized by O.

The mixing angle θ is determined from our eigenvalue equation (2.23) saying $\mathcal{M}V^{(j)} = m_j V^{(j)}$. For $j = 2$ I get

$$V^{(2)} = \begin{pmatrix} \sin\theta \\ \cos\theta \\ 0 \\ 0 \end{pmatrix} \tag{2.37}$$

and

$$\begin{pmatrix} \mathcal{R}(m_L) & m_D & -\mathcal{I}(m_L) & 0 \\ m_D & m_R & 0 & 0 \\ -\mathcal{I}(m_L) & 0 & -\mathcal{R}(m_L) & -m_D \\ 0 & 0 & -m_D & -m_R \end{pmatrix} \begin{pmatrix} \sin\theta \\ \cos\theta \\ 0 \\ 0 \end{pmatrix} = m_2 \begin{pmatrix} \sin\theta \\ \cos\theta \\ 0 \\ 0 \end{pmatrix}. \tag{2.38}$$

Multiplying out gives

$$\begin{pmatrix} \mathcal{R}(m_L)\sin\theta + m_D\cos\theta \\ m_D\sin\theta + m_R\cos\theta \end{pmatrix} = m_2 \begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix}, \tag{2.39}$$

which gives

$$\tan 2\theta = \frac{2m_D}{m_R - \mathcal{R}(m_L)}. \tag{2.40}$$

Hence, the mixing angle depends on the values of all the real mass-terms.

The value of the phase factor can also be determined in a similar way, by using $j = 1$ in the eigenvalue equation leading to [1]

$$\tan 2\lambda = -\frac{2\mathcal{I}(m_L)}{\mathcal{R}(m_L) + m_R - \sqrt{(\mathcal{R}(m_L) - m_R)^2 + 4m_D^2}}. \tag{2.41}$$

Thus also the phase factor depends on the values of the mass terms. If CP symmetry is conserved, $e^{i\lambda}$ takes on the values $i, -i, 1, -1$. The phase and mixing angle must be chosen so that we obtain positive eigenvalues for the mass.

2.3.3 The seesaw mechanism

If

$$\begin{aligned} m_D &\ll m_R \\ m_L &= 0 \end{aligned} \tag{2.42}$$

we obtain from equation (2.17) or (2.33) that the mass eigenstates are given by

$$\begin{aligned} m_{2,1} &= \frac{1}{2}\left(m_L + m_R \pm \sqrt{(m_L - m_R)^2 + 4m_D^2}\right) \\ &\approx \frac{1}{2}\left(m_R \pm \sqrt{m_R^2 + 4m_D^2}\right) \\ &= \frac{1}{2}\left(m_R \pm m_R\sqrt{1 + 4\frac{m_D^2}{m_R^2}}\right) \\ &\approx \frac{1}{2}\left(m_R \pm m_R(1 + 2\frac{m_D^2}{m_R^2})\right) \\ &\approx \frac{1}{2}\left(m_R \pm m_R \pm 2\frac{m_D^2}{m_R}\right) \\ &\approx \begin{cases} m_R \\ -\frac{m_D^2}{m_R} \end{cases} \end{aligned} \tag{2.43}$$

As we have chosen the positive values, we take the absolute value of the expressions above. The minus sign is due to phase factors in the P matrix, relating the eigenvalue to the physical mass by a factor $e^{2i\lambda}$. This shows that the two mass eigenvalues are very different in size, one is very light and the other one heavy. Looking at equation (2.40), the mixing angle is

$$\tan 2\theta = \frac{2m_D}{m_R} \ll 1 \Rightarrow \theta \approx \frac{n\pi}{2}, n = 0, 1 \tag{2.44}$$

and the phase factor is

$$\tan 2\lambda \approx 0 \Rightarrow \lambda \approx \frac{n\pi}{2}, n = 0, 1, 2, 3, 4 \tag{2.45}$$

As discussed earlier, we want the positive eigenvalues, which we get from $\theta = 0, \lambda = \pm\frac{\pi}{2}$. We then get

$$\begin{aligned} N_L &= U n_L \\ \begin{pmatrix} \nu_L \\ \nu_R^C \end{pmatrix} &= \begin{pmatrix} \cos\theta e^{i\lambda} & sin\theta \\ -\sin\theta e^{i\lambda} & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_{1L} \\ \nu_{2L} \end{pmatrix} \\ &\approx \begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \nu_{1L} \\ \nu_{2L} \end{pmatrix} \\ &= \begin{pmatrix} i\nu_{1L} \\ \nu_{2L} \end{pmatrix}. \end{aligned} \tag{2.46}$$

This shows that ν_{1L} consists mostly of $i\nu_L$ and ν_{2L} consists mostly of ν_R^C. Thus, the ν_{2L} consists mostly of a sterile neutrino, as the righthanded neutrino is known to not interact with the particles of the standard model.

Having calculated the two mass eigenstates, one of the great properties of the seesaw mechanism appears; it explains the smallness of the mass of the active neurinos. Experiments show that the mass of the neutrinos are only a small fraction of the masses of the other fermions in the standard model. If the neutrinos acquire mass through the Higgs mechanism like the other fermions, why should the neutrino be so much lighter? There is no reason why he Dirac mass term for the neutrino should not be of the same order as the Dirac mass terms for the charged fermions. The up quark and the electron both have masses of order 1 MeV, the electron neutrino is much lighter. In any case, the Dirac mass term appears in the symmetry-breaking Higgs mechanism, and thus the magnitude of the mass must be not be much larger than the electroweak scale.

The Dirac mass term is a consequence of breaking the electroweak symmetry, and thus it is protected by the electroweak symmetries. Since only the electroweak symmetry is broken in the Higgs mechanism, the mass can only be of the same order as the electroweak scale. A greater mass would have broken more than just the electroweak symmetry. The electroweak scale is of the order 10^2 GeV. Since the righthanded neutrino is a singlet in the standard model, its mass is not protected by the symmetries of the standard model. Thus it could be as big as it likes. But righthanded neutrinos is a sign of our not complete understanding of particle physics, and is thus a hint of physics behind the standard model. A common way to address this inadequate description, is to incorporate the standard model into a high energy theory, introducing grand unified theories. In these theories, the righthanded neutrino is not a singlet, but part of a multiplet under a higher symmetry. Then the righthanded Majorana mass term is only restricted by the energy scale of the new more general theory, as the righthanded neutrino acquire mass when the symmetry of the more general theory is broken. The breaking scale of the grand unified theories are of the order $10^{14} - 10^{16}$ GeV.

Thus, assuming $m_D \approx 10^2$ GeV and $m_R \approx 10^{15}$ GeV, the Seesaw mechanism claims that

$$m_1 = \frac{m_D^2}{m_R} \approx 10^{-11} GeV = 10^{-2} eV. \tag{2.47}$$

We have now seen that a heavy righthanded neutrino could induce very small neutrino-masses!

2.4 Three Generation Seesaw Mechanism

In this subsection, I will generalize the discussion made in the previous section to the case of three generations. We introduce three flavour neutrinos, and allows for N_s righthanded (or sterile) neutrinos. We will see that we end up with two seesaw scenarios, called the type I and the type II seesaw mechanism. The first gives no firm restrictions on the hierarchy, the second is able to explain degenerate masses. In this section, I have, in addition to the references listed in the beginning of this chapter, made use of the references [2, 18, 19, 20, 21].

With three generations of active neutrinos, we can introduce a array of left-handed flavour neutrino fields

$$\boldsymbol{\nu_L} \equiv \begin{pmatrix} \nu_{eL} \\ \nu_{\mu L} \\ \nu_{\tau L} \end{pmatrix}. \tag{2.48}$$

The Majorana left mass term then becomes

$$\mathcal{L}^{M,L}_{mass} = -\frac{1}{2}\overline{\boldsymbol{\nu_L^C}} M_L \boldsymbol{\nu_L} + h.c. = -\frac{1}{2}\sum_{\alpha,\beta=e,\mu,\tau} \overline{\nu^C_{\alpha L}} M_{L\alpha\beta}\nu_{\beta L} + h.c. \tag{2.49}$$

where M_L can be shown to be symmetric by using the properties of the change conjugation operator $\mathcal{C}$ [1].

The number of righthanded sterile neutrinos are not limited, so we say that we have N_s righthanded neutrinos. To find the Dirac mass term and the righthanded Majorana mass term, we must introduce a vector of righthanded flavour neutrino fields

$$\boldsymbol{\nu_R} \equiv \begin{pmatrix} \nu_1 \\ \nu_2 \\ \vdots \\ \nu_{N_s} \end{pmatrix}. \tag{2.50}$$

The Majorana right mass term then becomes

$$\mathcal{L}^{M,R}_{mass} = -\frac{1}{2}\overline{\boldsymbol{\nu_R^C}} M_R \boldsymbol{\nu_R} + h.c. = -\frac{1}{2}\sum_{s,s'=1,2,..,N_s} \overline{\nu^C_{sR}} M_{Rss'}\nu_{s'R} + h.c. \tag{2.51}$$

and the mass matrix M_R is $N_s \times N_s$ dimensional and symmetric.

The Dirac mass term is

$$\mathcal{L}^{D}_{mass} = -\sum_{s=1,2,..,N_s}\sum_{\alpha=e,\mu,\tau} \overline{\nu_{sR}} M_{Ds\alpha}\nu_{\alpha L} \tag{2.52}$$

where M_D is a $N_S \times 3$ dimensional matrix. All the mass matrices are complex.

To obtain the fields with definite masses, we need to combine the mass terms to one matrix equation at the same form as equation (2.13). First, we define the column matrix of a full lefthanded field

$$\boldsymbol{N_L} \equiv \begin{pmatrix} \boldsymbol{\nu_L} \\ \boldsymbol{\nu_R^C} \end{pmatrix} \tag{2.53}$$

where

$$\boldsymbol{\nu_R^C} \equiv \begin{pmatrix} \nu^C_1 \\ \nu^C_2 \\ \vdots \\ \nu^C_{N_s} \end{pmatrix}. \tag{2.54}$$

Then the total mass Lagrange density is

$$\mathcal{L}_{mass}^{D+M} = -\frac{1}{2}\overline{\boldsymbol{N_L^C}} M_{D+M} \boldsymbol{N_L} + h.c. \tag{2.55}$$

where the mass matrix M_{D+M} is given by

$$M_{D+M} = \begin{pmatrix} M_L & M_D^T \\ M_D & M_R \end{pmatrix}. \tag{2.56}$$

This matrix is symmetric, but it is not generally real, and we have to diagonalize it to find the mass eigenvalues. The main idea is to diagonalize M_{D+M} in such a way that the heavy and light neutrino fields decouple. In analogy with the case of only one generation, we aim to describe the active neutrinos as light, and the sterile neutrinos as very heavy.

2.4.1 Diagonalizing the mass matrix

Once again, the aim is to diagonalize M_{D+M} by introducing a unitary matrix W, such that

$$\mathbf{N_L} = W\mathbf{n_L} \qquad\qquad \mathbf{n_L} = \begin{pmatrix} \nu_{1L} \\ \nu_{1L} \\ \vdots \\ \nu_{N_t L} \end{pmatrix} \tag{2.57}$$

where W diagonalizes M_{D+M}. $N_t = 3 + N_s$ is the total number of states in the lefthanded array $\mathbf{N_L}$. I make the ansatz [18] that the matrix diagonalizing the mass matrix is

$$W = \begin{pmatrix} \sqrt{\mathbf{1} - BB^\dagger} & B \\ -B^\dagger & \sqrt{\mathbf{1} - B^\dagger B} \end{pmatrix}, \tag{2.58}$$

where B is a general $3 \times N_s$ matrix and W is unitary by construction, so that

$$W^T M_{D+M} W = \begin{pmatrix} M_{light} & 0 \\ 0 & M_{heavy} \end{pmatrix}, \tag{2.59}$$

where M_{light} is 3×3 dimensional and M_{heavy} is $N_s \times N_s$ dimensional, and both matrices are symmetric. The mass eigenvalues of the light neutrinos are the eigenvalues of M_{light}, and likewise for the heavy neutrinos. The square roots in equation (2.58) is to be understood by powerseries, such that

$$\sqrt{1-x} = 1 - \frac{1}{2}x - \frac{1}{8}x^2 - \frac{1}{16}x^3 - \dots \tag{2.60}$$

giving

$$\sqrt{\mathbf{1} - BB^\dagger} = 1 - \frac{1}{2}BB^\dagger - \frac{1}{8}BB^\dagger BB^\dagger - \frac{1}{16}BB^\dagger BB^\dagger BB^\dagger - \dots \tag{2.61}$$

B is assumed to be expandable in a powerseries in $1/m_R$, which is possible if $m_R \gg 1$ such that

$$B = B_1 + B_2 + B_3 + \cdots = \sum_{j=1}^{\infty} B_j \tag{2.62}$$

where B_j is proportional to $(m_R)^{-j}$. m_R is the order of magnitude of the eigenvalues of $\sqrt{M_R^\dagger M_R}$. Of course the eigenvalues of $\sqrt{M_R^\dagger M_R}$ might not all be of the same order, but the important idea is that the eigenvalues of M_R is assumed to be much greater than the eigenvalues of M_L and M_D.

Calculating $M = W^T M_{D+M} W$ we get

$$\begin{aligned} W^T M_{D+M} W &= \begin{pmatrix} \sqrt{\mathbf{1} - BB^\dagger} & B \\ -B^\dagger & \sqrt{\mathbf{1} - B^\dagger B} \end{pmatrix}^T \begin{pmatrix} M_L & M_D^T \\ M_D & M_R \end{pmatrix} \\ &\quad \times \begin{pmatrix} \sqrt{\mathbf{1} - BB^\dagger} & B \\ -B^\dagger & \sqrt{\mathbf{1} - B^\dagger B} \end{pmatrix} \\ &= \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}, \end{aligned} \tag{2.63}$$

We get

$$\begin{aligned} M_{11} &= (\sqrt{\mathbf{1} - B^* B^T}) \left(M_L \sqrt{\mathbf{1} - BB^\dagger} - M_D^T B^\dagger \right) - \\ &\quad -B^* \left(M_D \sqrt{\mathbf{1} - BB^\dagger} - M_R B^\dagger \right) \\ &= \sqrt{\mathbf{1} - B^* B^T} M_L \sqrt{\mathbf{1} - BB^\dagger} - \sqrt{\mathbf{1} - B^* B^T} M_D^T B^\dagger \\ &\quad -B^* M_D \sqrt{\mathbf{1} - BB^\dagger} + B^* M_R B^\dagger \end{aligned} \tag{2.64}$$

and

$$\begin{aligned} M_{22} &= B^T \left(M_L B + M_D^T \sqrt{\mathbf{1} - B^\dagger B} \right) + \sqrt{\mathbf{1} - B^T B^*} \left(M_D B + M_R \sqrt{\mathbf{1} - B^\dagger B} \right) \\ &= B^T M_L B + B^T M_D^T \sqrt{\mathbf{1} - B^\dagger B} + \sqrt{\mathbf{1} - B^T B^*} M_D B \\ &\quad + \sqrt{\mathbf{1} - B^T B^*} M_R \sqrt{\mathbf{1} - B^\dagger B} \end{aligned} \tag{2.65}$$

and

$$\begin{aligned} M_{12} &= \sqrt{\mathbf{1} - B^* B^T} \left(M_L B + M_D^T \sqrt{\mathbf{1} - B^\dagger B} \right) - B^* \left(M_D B + M_R \sqrt{\mathbf{1} - B^\dagger B} \right) \\ &= \sqrt{\mathbf{1} - B^* B^T} M_L B + \sqrt{\mathbf{1} - B^* B^T} M_D^T \sqrt{\mathbf{1} - B^\dagger B} - B^* M_D B \\ &\quad -B^* M_R \sqrt{\mathbf{1} - B^\dagger B} \end{aligned} \tag{2.66}$$

and

$$\begin{aligned} M_{21} &= B^T \left(M_L \sqrt{\mathbf{1} - BB^\dagger} - M_D^T B^\dagger \right) + \sqrt{\mathbf{1} - B^T B^*} \left(M_D \sqrt{\mathbf{1} - BB^\dagger} - M_R B^\dagger \right) \\ &= B^T M_L \sqrt{\mathbf{1} - BB^\dagger} - B^T M_D^T B^\dagger + \sqrt{\mathbf{1} - B^T B^*} M_D \sqrt{\mathbf{1} - BB^\dagger} - \\ &\quad \sqrt{\mathbf{1} - B^T B^*} M_R B^\dagger. \end{aligned} \tag{2.67}$$

We see that $M_{21} = M_{12}^T$ as is should be, as $M = W^T M_{D+M} W$ similar by unitary transformation to a symmetric matrix. In order to diagonalize M_{D+M}, we need

$$M = W^T M_{D+M} W = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} = \begin{pmatrix} M_{light} & 0 \\ 0 & M_{heavy} \end{pmatrix}. \quad (2.68)$$

Thus I require

$$B^T M_L \sqrt{\mathbf{1} - BB^\dagger} - B^T M_D^T B^\dagger + \sqrt{\mathbf{1} - B^T B^*} M_D \sqrt{\mathbf{1} - BB^\dagger} - \sqrt{\mathbf{1} - B^T B^*} M_R B^\dagger = 0. \quad (2.69)$$

To solve the equation above, I expand it according to equation (2.61). This gives

$$\begin{aligned}
\sqrt{\mathbf{1} - BB^\dagger} &= 1 - \frac{1}{2} BB^\dagger - \frac{1}{8} BB^\dagger BB^\dagger - \frac{1}{16} BB^\dagger BB^\dagger BB^\dagger - \dots \\
&= 1 - \frac{1}{2} (B_1 + B_2 + B_3 + \dots)(B_1 + B_2 + B_3 + \dots)^\dagger \\
&\quad - \frac{1}{8} (B_1 + B_2 + B_3 + \dots)(B_1 + B_2 + B_3 + \dots)^\dagger (B_1 + B_2 + B_3 + \dots) \\
&\quad \times (B_1 + B_2 + B_3 + \dots)^\dagger \\
&\quad - \frac{1}{16} (B_1 + B_2 + B_3 + \dots)(B_1 + B_2 + B_3 + \dots)^\dagger (B_1 + B_2 + B_3 + \dots) \\
&\quad \times (B_1 + B_2 + B_3 + \dots)^\dagger (B_1 + B_2 + B_3 + \dots)(B_1 + B_2 + B_3 + \dots)^\dagger - \dots \\
&= 1 - \frac{1}{2} B_1 B_1^\dagger - \frac{1}{2} \left(B_1 B_2^\dagger + B_2 B_1^\dagger \right) \\
&\quad - \frac{1}{2} \left(B_1 B_3^\dagger + B_3 B_1^\dagger + B_2 B_2^\dagger + \frac{1}{4} B_1 B_1^\dagger B_1 B_1^\dagger \right) \\
&\quad - \frac{1}{2} \left(B_1 B_4^\dagger + B_4 B_1^\dagger + B_2 B_3^\dagger + B_3 B_2^\dagger \right) \\
&\quad + \frac{1}{4} \left(B_1 B_1^\dagger B_1 B_2^\dagger + B_1 B_1^\dagger B_2 B_1^\dagger + B_1 B_2^\dagger B_1 B_1^\dagger + B_2 B_1^\dagger B_1 B_1^\dagger \right) + \dots \quad (2.70)
\end{aligned}$$

To zeroth order in m_R, equation (2.69) reads:

$$\begin{aligned}
& M_D - M_R B_1^\dagger = 0 \\
& \Rightarrow B_1 = M_D^\dagger \left(M_R^{-1} \right)^\dagger .
\end{aligned} \quad (2.71)$$

To first order in m_R:

$$\begin{aligned}
& (B_1 + B_2)^T M_L \left(1 - \tfrac{1}{2} B_1 B_1^\dagger \right) \\
& - (B_1 + B_2)^T M_D^T (B_1 + B_2)^\dagger \\
& + \left(1 - \tfrac{1}{2} B_1^T B_1^* \right) M_D \left(1 - \tfrac{1}{2} B_1 B_1^\dagger \right) \\
& - \left(1 - \tfrac{1}{2} B_1^T B_1^* \right) M_R (B_1 + B_2)^\dagger \\
& \qquad \approx B_1^T M_L - M_R B_2^\dagger \\
& \qquad = 0 \\
& \Rightarrow B_2 = M_L^\dagger M_D^T M_R^{-1} \left(M_R^{-1} \right)^\dagger ,
\end{aligned} \quad (2.72)$$

where I have made use of M_R being symmetric.

To second order in m_R the same kind of calculations give:

$$\begin{aligned} B_3 &= M_L^\dagger M_L^T M_D^\dagger \left(M_R^{-1}\right)^\dagger M_R^{-1} \left(M_R^{-1}\right)^\dagger \\ &\quad -M_D^\dagger \left(M_R^{-1}\right)^\dagger M_D^* M_D^T M_R^{-1} \left(M_R^{-1}\right)^\dagger \\ &\quad -\frac{1}{2} M_D^\dagger \left(M_R^{-1}\right)^\dagger \left(M_R^{-1}\right) M_D M_D^\dagger \left(M_R^{-1}\right)^\dagger \end{aligned} \tag{2.73}$$

We could also calculate B_4 and so on in the same way, but our main goal is to find the mass terms, so we stop here. The important feature about the Bs is that they are recursive calculable.

2.4.2 Seesaw mechanism type I

We go back to M_{light} and M_{heavy}, and calculate (using equation (2.64)) to second order in $1/m_R$

$$\begin{aligned} M_{light} &= M_{11} \\ &= \sqrt{\mathbf{1} - B^* B^T} M_L \sqrt{\mathbf{1} - B B^\dagger} - \sqrt{\mathbf{1} - B^* B^T} M_D^T B^\dagger - B^* M_D \sqrt{\mathbf{1} - B B^\dagger} \\ &\quad + B^* M_R B^\dagger \\ &\approx M_L - M_D^T M_R^{-1} M_D \\ &\quad -\frac{1}{2}\left(M_D^T M_R^{-1} \left(M_R^{-1}\right)^* M_D^* M_L + M_L M_D^\dagger \left(M_R^{-1}\right)^* M_R^{-1} M_D \right) \end{aligned} \tag{2.74}$$

where the calculation is preformed by counting terms to second order and inserting for the Bs. Similarly

$$\begin{aligned} M_{heavy} &= M_{22} \\ &= B^T M_L B + B^T M_D^T \sqrt{\mathbf{1} - B^\dagger B} + \sqrt{\mathbf{1} - B^T B^*} M_D B \\ &\quad + \sqrt{\mathbf{1} - B^T B^*} M_R \sqrt{\mathbf{1} - B^\dagger B} \\ &\quad \approx M_R + \frac{1}{2}\left(M_D M_D^\dagger \left(M_R^{-1}\right)^* + \left(M_R^{-1}\right)^* M_D^* M_D^T \right) \\ &\quad + \frac{1}{2}\left(M_D M_L^* M_D^T M_R^{-1} \left(M_R^{-1}\right)^* + \left(M_R^{-1}\right)^* M_R^{-1} M_L^* M_D M_D^T \right). \end{aligned} \tag{2.75}$$

In seesaw mechanism type I, M_L is zero, and $M_D \ll M_R$, meaning that the eigenvalues of $\sqrt{M_D^\dagger M_D}$ is much smaller than the eigenvalues of $\sqrt{M_R^\dagger M_R}$. Looking at equation (2.72), his gives $B_2 = 0$. Going to higher order, one can show [18]

$$M^L = 0 \Rightarrow B_{2N} = 0 \qquad\qquad N = 1, 2, 3, \ldots \tag{2.76}$$

For M_{light} and M_{heavy} we then get

$$M_{light} = -M_D^T M_R^{-1} M_D \tag{2.77}$$

and

$$M_{heavy} = M_R + \frac{1}{2}\left(M_D M_D^\dagger \left(M_R^{-1}\right)^* + \left(M_R^{-1}\right)^* M_D^* M_D^T\right) \tag{2.78}$$

to second order in $1/m_R$. To leading order, we find

$$\begin{aligned} M_{light} &\approx -M_D^T M_R^{-1} M_D \\ M_{heavy} &\approx M_R. \end{aligned} \tag{2.79}$$

This is essentially the same relations as we found for the one generation scenario.

2.4.3 Seesaw type II

If the condition $M_L = 0$ is not met, equations (2.75) and (2.74) will look a bit different to first order, as they will not simplify:

$$\begin{aligned} M_{light} &\approx M_L - M_D^T M_R^{-1} M_D \\ M_{heavy} &\approx M_R + \frac{1}{2}\left(M_D M_D^\dagger \left(M_R^{-1}\right)^* + \left(M_R^{-1}\right)^* M_D^* M_D^T\right). \end{aligned} \tag{2.80}$$

The mass-term for the righthanded neutrinos is rather complex, but as we don't see them, we tend to care more about the light left-handed neutrinos, so we'll begin the discussion there.

If the second term in M_{light} dominates, we pretty much have the same situation as in the Seesaw type I regime. If, however, the first term dominates, the seesaw mechanism as discussed in the previous section no more explains the smallness of the neutrino mass scale. Thus we need to explain why M_L should be small, in order to explain the smallness of the neutrino masses.

One common way to account for the smallness of M_L, is introducing scalar triplets ([19, 20, 2]). The mass terms in section 2.4.2 can be accounted for in the normal standard model of the Higgs mechanism for electroweak theory with the introduction of two Higgs doublets. We want to explain the smallness of M_L with the additional introduction of an $SU(2)_L$ scalar triplet Δ. The matrix representation for this triplet is

$$\Delta = \begin{pmatrix} H^+ & \sqrt{2}H^{++} \\ \sqrt{2}H^0 & -H^+ \end{pmatrix}. \tag{2.81}$$

and the neutral component (H^0) couples to the neutrino. When this term is included in the Lagrange density, and if the neutral component has a non-zero vacuum expectation value, $< H^0 >= v_\Delta$, it breaks the symmetry and gives raise to a mass term for the lefthanded neutrinos. From imposing the conditions of a global stable minimum, one finds [19]

$$v_\Delta = \frac{\mu v^2}{M^2} \tag{2.82}$$

where μ is a coupling, v is the vacuum expectation value of the Higgs doublet, and M is the mass of the Higgs triplet. The mass term of the lefthanded Majorana is proportional to the vacuum expectation value of the triplet times a Yukawa coupling constant. This gives, according to [2],

$$M_L = Y_\nu v_\Delta = Y_\nu \frac{\mu v^2}{M^2}. \tag{2.83}$$

This term is even possible without introducing any righthanded neutrinos at all, introducing a heavy Higgs triplet, as the mass occurs in a lefthanded Majorana mass term, and thus this is a way of addressing the question of neutrino masses without introducing righthanded neutrinos and the seesaw mechanism [2].

With the inclusion of righthanded neutrinos, in a left-right symmetric seesaw model, the vacuum expectation value of the triplet will take the form [2]

$$v_\Delta \simeq f \frac{v^2}{v_R} \tag{2.84}$$

where v_R is the vacuum expectation value of the righthanded part of the neutral component of the triplet, $< H_R^0 >$. To first order in $1/m_R$ we then obtain

$$M_{light} \simeq f \frac{v^2}{v_R} - M_D^T M_R^{-1} M_D \tag{2.85}$$

where f is a coupling constant.

Seesaw type II gives rise to a good explanation for the possible degenerate behaviour of the neutrino masses. With the seesaw type I, with $M_{light} \approx -M_D^T M_R^{-1} M_D$, there is no reason why M_{light} should give rise to degenerate neutrino masses. The type I mechanism does not prohibit degenerate masses, but it does not encourage it either. In type II, the mass matrix takes the form $M_{light} \approx f\frac{v^2}{v_R} - \left(M^D\right)^T \left(M^R\right)^{-1} M_D$. If the first term dominates over the last term, this gives rise to degenerate neutrino masses, as the first term is a scalar times the identity matrix, while the last term is a matrix that has few restrictions on the relative sizes of the matrix components.

Although the seesaw mechanism provides the framework for describing the neutrino mass generating process, it is not confirmed. It is, however, a popular model, and if neutrinoless double β-decays are confirmed, giving confirmation of Majorana behavior of neutrinos, it might increase the popularity, as the seesaw mechanism is in need of Majorana nature of neutrinos.

2.5 Other Ways to Generate Neutrino Masses

Although the seesaw mechanism is a very popular way to explain the small neutrino masses, there are of course alternative approaches to the matter. In this section, I have made use of the references [22, 23, 2, 19].

One way to explain small neutrino masses, is by introducing universe models with large extra dimensions. In these theories, we live in a universe with extra dimensions. Our "everyday" perceived 3+1 dimensions lives on a brane inside a higher dimensional space. This higher dimensional space is called the bulk, and we need at least two extra dimensions. The extra dimensions must be small, otherwise gravitational effects would have revealed them, they must be smaller than one millimeter by order of magnitude. Also in this model one need righthanded neutrinos, and they live in the bulk, and not necessarily in our brane. There can be interactions between particles in the bulk-world and particles in the brane-world, but the couplings are suppresed by a factor inversely proportional to squareroot of the volume of the extra dimensions. Thus the lefthanded neutrinos in the brane can mix with the righthanded bulk neutrinos, generating a Dirac mass. The coupling is weak doe to the different dimensions occupied by the righthanded and lefthanded neutrinos, causing the mass to be small.

This model, though not close to be confirmed, is particularly interesting because of its ability to generate small Dirac masses for the neutrino. If the neutrinos turn out to be Dirac particles, the seesaw mechanism can not be used to explain the smallness of the neutrinomass.

Also other models aim to explain the smallness of the neutrino mass. The Zee model and the Zee-Babu model add extra singlets and Higgs doublets to the standard model multiplets. These extra particles are assumed to arise from super symmetry. In these models, the neutrino mass arises as the self energy in one-loop and two-loop diagrams. These models are not very predictive of the neutrino mass scale, but on the other hand, a confirmation of the neutrino mass model and a better measurement of the neutrino mass could give more information about supersymmetric models.

2.6 Using Mass Models to Predict Mass Scale

The mechanisms generating neutrino masses described in this section are all too little predictive to give any good information. For example, we showed that the lightest neutrino mass, in the case of type II seesaw mechanism, was given by equation (2.83):

$$Y_\nu \frac{\mu v^2}{M^2} \tag{2.86}$$

or ():

$$\frac{f v^2}{v_R} \tag{2.87}$$

This would be predictive if one something about the parameters. But as v is the vacuum expectation value of the Higgs, which we do know something about, M is the mas of the Higgs triplet, which is, in GUT theory, limited to the GUT scale, but it can also be a lot smaller. μ is a coupling about which we know very little, and Y_ν is not determined. The same problem arises from the second expression with f

and v_R. Thus it is very difficult to use the model to give information to restrain the neutrino mass, as there is a huge degeneracy in the parameters. But neutrino mass experiments can give restrictions to the relation between the unknown parameters, which can give very useful information both when it comes to determining which model is right and when it comes to exploring the model that lays behind the mass generating process. Thus neutrino mass is a key to yet unexposed physics!

Part II

Neutrinos from the cosmologist's point of view

Chapter 3

Cosmology

The big bang model accounts very well for almost the entire history of the universe. It explains how the universe starts out containing hot gas, which cools down as the universe expands. Starting out dominated by radiation, eventually overtaken by nonrelativistic matter, and now possibly dominated by dark energy, the big bang model of the universe gives a good description of the evolution of the universe from very early on and until today. There are, however, questions left open by the model, and among the most interesteing parts left out is the initial conditions.

In this section, I will outline the cosmological standard model. I will introduce the Robertson-Walker universe and the Friedman-Robertson-Walker line element, the Einstein field equation and the Friedman equations. I will then discuss the dynamics of the universe and how it varies with the type of energy that dominates the energy density. Some of the shortcomings of the model will be discussed in section 3.3, and the proposed solution to many of the open questions is to introduce early *inflation*; a short period of accelerated cosmic expansion in the early universe. This is discussed in section 3.4.

The main references used in this chapter is [24, 25, 26, 27]. Other references are given when used.

3.1 The Friedmann-Robertson-Walker line element

Some of the most useful equations in cosmology are derived from Einstein's field equation ($c = 1$ and $\hbar = 1$ is used in this and the reminding chapters):

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} - \Lambda g_{\mu\nu}. \tag{3.1}$$

$G_{\mu\nu}$ is the Einstein tensor, which contains the curvature of space-time and is a complicated non-linear function of the metric $g_{\mu\nu}$ and its derivatives. Λ is a cosmological constant. G is Newton's gravitational constant. $T_{\mu\nu}$ is the energy-momentum tensor, which contains the energy- and momentum- densities, the pressure, the stress and the shear forces:

$$T_{\mu\nu} = \begin{pmatrix} T_{00} & T_{01} & T_{02} & T_{03} \\ T_{10} & T_{11} & T_{12} & T_{13} \\ T_{20} & T_{21} & T_{22} & T_{23} \\ T_{30} & T_{31} & T_{32} & T_{33} \end{pmatrix} \tag{3.2}$$

The components have the following phsyical interpretations:

- T_{00}: energy density
- T_{i0}: momentum density
- T_{0i}: energy flux
- T_{ii}: pressure
- $T_{ij|_{j\neq i}}$: shear forces

$T_{\mu\nu}$ is thus only dependent on the content of the universe. The left hand side of the Einstein field equation is geometrical, while the right hand side is describing the physical content of the universe, and the Einstein's field equation relates the two. Thus, as John A. Wheeler quite famously is supposed to have said it: "*Matter tells space how to curve. Space tells matter how to move.*"

Local conservation of mass and momentum can be expressed using the enegry-momentum tensor

$$T^{\mu\nu}_{;\nu} = 0, \tag{3.3}$$

where "$;\nu$" indicates the covariant derivative with respect to ν. One often apply a perfect-fluid description of the universe. By doing so, one assumes that there are no shear stress nor energy or momentum flux. The expression for the energy-momentum tensor is then

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu + p g_{\mu\nu}, \tag{3.4}$$

where ρ is the mass density, p is the pressure of the perfect fluid, $u_\mu = x^\mu_{;0}$ is the 4-velocity of the fluid, and $g_{\mu\nu}$ is the metric. In a comoving orthonormal basis, the energy-momentum tensor becomes

$$T_{\hat{\mu}\hat{\nu}} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}, \tag{3.5}$$

where the hat indicates orthonormal comoving basis. The relation between ρ and p is often expressed as an equation of state

$$p = \omega\rho, \tag{3.6}$$

where ω is called the equation of state parameter.

To get the metric, it is common to assume a Friedmann-Robertson-Walker (FWR) model, in which space is homogeneous and isotropic on large scales:

- *Homogeneous* means of the same kind or nature, or uniform. In a cosmological setting, a homogeneous universe means that there are no special place in the universe. All locations have the same properties.
- *Isotropic* means having the same properties inn all directions, or having equal properties along all axes. In cosmological terms, an isotropic universe will look the same independent on the direction in which you are observing.

This assumption of isotropy and homeogeneity is referred to as *the cosmological principle*. The universe is not homogeneous and isotropic on small scales. The density that I observe is not the same if I look to the ground as if I look to the sky. But on large scales, homogeneity and isotropy seems to be a good zeroth-order approximation. Only on scales of about 100 Mpc or more, the cosmological principle is a good approximation. The cosmological principle is assumed to be valid at all times, thus the evolution of the universe has to be the same everywhere.

The assumptions above gives rise to a line element in comoving coordinates:

$$ds^2 = -dt^2 + a^2(t)\left(d\chi^2 + r^2(\chi)d\Omega^2\right), \tag{3.7}$$

where t is *cosmic time*; the time meassured by an observer moving with the expansion of space. χ is the comoving radial coordinate. $a(t)$ is the scale factor, or expansion factor, which relates the comoving coordinate distance to the physical distance: $dl = \sqrt{g_{\chi\chi}}d\chi = a(t)d\chi$. By demanding isotropy of curvature, one obtains

$$\frac{dr}{d\chi} = \sqrt{1 - kr^2}, \tag{3.8}$$

where k is a constant which characterizes the curvature of the universe. The line element is then

$$ds^2 = -dt^2 + a^2(t)\left(\frac{dr^2}{1 - kr^2} + r^2(\chi)d\Omega^2\right), \tag{3.9}$$

which is the Friedmann-Robertsom-Walker line element. k is taken to be a parameter that describes the spatial curvature of the universe described by the line element, and has dimension m^{-2}. $k = 0$ gives a spatially flat universe, and we have Euclidean spatial geometry. $k > 0$ gives positiv spatial curvature, and we have a "closed model" with spherical spatial geometry. $k < 0$ gives negative curvature and a open model with hyperbolic spatial geometry.

3.2 The Friedmann equations

When the metric is established, according to equation (3.1), the equations of motion of the universe can be calculated. By applying the Cartan formalism to the

line element of equation (3.9), the Einstein's field equation can be solved. Solving for the diagonal components ($G_{tt} = G_{00}$ and $G_{mm} = G_{ii}$, $i = 1, 2, 3$) in an local orthogonal basis (indicated by a hat on the indices), one finds [27]

$$\begin{aligned} G_{\hat{t}\hat{t}} &= 3\frac{\dot{a}^2 + k}{a^2}, \\ G_{\hat{m}\hat{m}} &= -2\frac{\ddot{a}}{a} - \frac{\dot{a}^2 + k}{a^2}. \end{aligned} \tag{3.10}$$

Looking back at the Einstein's field equation in equation (3.1) and the expression for the energy monentum tensor of a perfect fluid in equation (3.5), one obtains

$$\begin{aligned} 8\pi G\rho &= 3\frac{\dot{a}^2 + k}{a^2} - \Lambda, \\ 8\pi G p &= -2\frac{\ddot{a}}{a} - \frac{\dot{a}^2 + k}{a^2} + \Lambda, \end{aligned} \tag{3.11}$$

which are the Friedmann equations. In a more familiar way of writing, the equation above reads:

$$\begin{aligned} \left(\frac{\dot{a}}{a}\right)^2 \equiv H^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{k}{a^2} \qquad (a), \\ \frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3} \qquad (b), \end{aligned} \tag{3.12}$$

where the Hubble parameter $H = \frac{\dot{a}}{a}$ is introduced. This parameter relates to the dynamics of the universe, as it is a fraction between the rate of change of the scale factor and the scale factor. From the last Friedmann equation, one can see that if the cosmological constant is set to zero, $\Lambda = 0$, and the universe is spatially flat, $k = 0$, the acceleration or deceleration of the universe is determined by the relation between the mass density and the pressure. Applying the equation of state, equation (3.6), one can easily see that

- $3p < -\rho \rightarrow \omega < -1/3$ implies an accelerated expansion of the universe
- $3p > -\rho \rightarrow \omega > -1/3$ implies a deceleration of the universe.

There is one more useful equation that we will need; the equation of adiabatic expansion:

$$\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + p). \tag{3.13}$$

This equation is derived by assuming that the universe expands adiabatically, which is quite reasonable given that there could only be heat transfer if the heat had somewhere to go outside our universe, which seems rather unlikely.

From this point on, I will assume that the universe is spatially flat, meaning that $k = 0$.

3.2.1 Perfect fluid components

There are three types of perfect fluids that are well worth noticing: dust, radiation and a cosmological constant. Dust has $\omega = 0$, as it has no pressure. Radiation has $\omega = 1/3$, as the energy momentum tensor of an ultrarelativistic field (like the electromagnetic field) has trace equal to zero. The most used form of vacuum energy in cosmology is Lorentz invariant vacuum energy (LIVE), where it is assumed that it is not possible to measure any velocities relative to vacuum, which implies $\omega = -1$ [27] (when referring to vacuum energy, I mean LIVE unless other choices are stated). One thus see that dust and radiation slow down the expansion of the universe, while vacuum energy speeds up the expansion. Energy components with $\omega < -1/3$ is referred to as dark energy. The vacuum energy is thus a kind of dark energy.

Looking at the equation of adiabatic expansion, equation (3.13), we find

$$\rho = \rho_0 \frac{a_0}{a}^{3(1+\omega)}, \tag{3.14}$$

where I have used $\rho(t_0) = \rho_0$, $a(t_0) = a_0$. $a(t) = a$ and $\rho(t) = \rho$.

By solving the first Friedmann equation, (3.12 a), for one component only, one can find the dependence of the scalefactor on time:

$$a(t) \propto \begin{cases} t^{\frac{2}{3(1+\omega)}} & \omega \neq -1 \\ e^{\sqrt{\frac{\Lambda c^2}{3}}t} & \omega = -1. \end{cases} \tag{3.15}$$

We then see that

- Dust, which has $\omega = 0$, gives $\rho \propto a^{-3}$. This is quite logical: $\rho = M/V$ and the volume V increases like a^3.
- Radiation, which has $\omega = 1/3$, gives $\rho \propto a^{-4}$. This is logical too, it only requires us to think more. The density of "radiation particles" decreases like a^{-3} as the same time as the energy decreases with a^{-1} because of the wave behaviour. The wavelengths are redshifted with the expansion, and the energy scales inversely proportional to the wavelength, which increases with a.
- Vacuum, which has $\omega = -1$, gives $\rho = \rho_0$. Thus the energy density is constant! Thus LIVE is often referred to as a cosmological constant. The cosmological constant is not a necessity in the Friedmann equations, it turns up when solving the Einstein equation and might well be equal to zero. Throughout history, the presence of a nonzero cosmological constant has been debated. Einstein originally demanded a nonzero cosmological constant to get a stable static universe, but later regretted this when he realized that the universe is not stable but expanding. He later remarked that introducing a cosmological constant to get a static universe was the greatest blunder of his career, and the cosmological constant became disfavoured among cosmologists for quite some time, when it was set to zero. It has, however,

become more popular to assume a nonzero cosmological constant lately, and it seems to be a necessity rather than just an option to include a nonzero cosmological constant of some sort when accounting for the dynamics of the universe.

From the discussion above, it is clear that if the universe ever was radiation dominated, it must have been so in the beginning. Matter would eventually become dominant over radiation, if there is a cosmological constant, however small, an expanding universe will eventually be dominated by the vacuum energy.

3.3 Puzzles of the Big Bang model

As stated at the beginning of this chapter, the standard model as presented above has some limitations. There are some fundamental problems which it does not address, like why the universe became so close to flat and homogeneous on large scales, and where did the perturbations come from? One could say that it just happened, that our universe is very close to flat by coincidence, and the perturbations was just created by chance. And no doubt, if someone one day sat down and decided to make infinitely many universes, some of them would be almost flat with small perturbations, but that would be only a tiny fraction. Thus the probability of creating our particular universe, by chance, is very little, and this scenario is not very satisfying.

It would be more gratifying to explain the flatness and homogeneity, and the small perturbations, as something that happened after the creation of the universe, by a process which would have created an almost flat, homogeneous universe with small perturbations independently of the nature of the universe before this process. Luckily, someone came up with a process solving all the problems above, and the process is called inflation. But before learning about inflation, we will look into the listed problems in a more detailed manner.

3.3.1 The horizon problem/large scale smoothness

Observations shows that the universe is rather homogeneous and isotropic on large scales. Why is that? The cosmological principle states that it should be so, but that is not a good explanation. Objects on opposite sides of the sky have the same properties. And if the two sides of the sky had different properties in the early phases of the Big Bang, they have not necessarily been able to even out the differences during the history of the universe.

The particle horizon sets the causal scale of the universe, and the scale of how far one can “see” in the universe, restricted by the speed of light and the expansion of the universe. It is given by

$$d_{PH}(t) = a(t) \int_{t_0}^{t} \frac{dt'}{a(t')}, \tag{3.16}$$

which is the proper distance confining the area in causal contact at time t with an event at time t_0. t_0 is typically the beginning of time (the beginning of the universe). If the distance today between two patches of the sky is greater than the particle horizon, they have never been in causal contact. We can think of our particle horizon as the radius of the volume of the patches of the sky that we are in contact with.

The universe is thought to have been dominated by radiation in early times. At the time of recombination, when the most of the electrons bind to nuclei forming neutral atoms, the co-moving particle horizon had a size that today extends about one degree on the sky in the CMB. This means that the photons coming from patches of the sky separated by more than about two degrees, cannot have been in causal contact before recombination. And after recombination the interaction rate dropped substantially. How can it be that the universe looks so smooth and isotropic on large scales? This is the *horizon problem*.

A useful scale in cosmology is the *Hubble length*

$$d_H(t) = \frac{1}{H(t)}. \tag{3.17}$$

Hubble's law states that the velocity at which a particle, moving only with the expansion of the universe, travels away from a reference point is given by the Hubble constant times the proper distance from the point to the particle:

$$v = \frac{dr}{dt} = Hr. \tag{3.18}$$

Thus, the Hubble length given by equation (3.17) is the proper distance radius of the sphere which moves away from an observer at the speed of light. This implies that a light-signal emitted from the distance of a Hubble length will not be able to reach the observer if the universe keeps expanding at the same rate. The *Hubble horizon* is often referred to as the causal scale; particles separated by more than one Hubble length cannot communicate in the sense that a light-signal emitted now by one particle will not be observed by the other given that the expansion-properties of the universe does not change. The particle horizon if often approximated by the Hubble length.

A common timescale is the *Hubble time*. It is given by

$$t_H = \frac{1}{H}. \tag{3.19}$$

Two particles separated by a distance R is moving appart with the speed $v = HR$ according to equation (3.18). If they were together at $t = 0$, the time taken them to separate by a distance R is $t = \frac{R}{v} = \frac{R}{HR} = H^{-1}$ given that the Hubble parameter has been constant throughout the expansion time. Thus the Hubble time is an estimate of the age of the universe in a Big Bang model. The Hubble parameter has not been constant throughout the history of the expanding universe, but taking $H_0 = 72 kms^{-1}Mpc^{-1}$ gives $t_H = 13.6$ Gyr, which is a fairly good approximation to the age of the universe.

3.3.2 Why is the universe so flat?

Another problem that is the universe appears to be very close to spatially flat.

From the first Friedman equation, (3.12 a), we have

$$H^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{k}{a^2} \tag{3.20}$$

where ρ contains the energy density of matter and radiation, $\rho = \rho_m + \rho_r$. We could define ρ to also include the energy density from the cosmological constant: $\rho = \rho_m + \rho_r + \rho_\Lambda$ where $\rho_\Lambda = \Lambda/8\pi G$. The redefined Friedman equation then reads

$$H^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}. \tag{3.21}$$

When $k = 0$, which corresponds to a spatially flat universe, we get

$$\rho_c = \frac{3H^2}{8\pi G}. \tag{3.22}$$

This is the mass density that gives a flat universe, and we call it the critical density. It is common to define the critical density today by

$$\rho_{c0} = \frac{3H_0^2}{8\pi G}. \tag{3.23}$$

This quantity is used to measure densities relative to the critical density:

$$\Omega \equiv \frac{\rho}{\rho_c} = \frac{\sum_i \rho_i}{\rho_c} \tag{3.24}$$

where i runs over all components of energy.

The Friedmann equation, equation (3.21), can then be rewritten as

$$\Omega(t) - 1 = \frac{kc^2}{a^2H^2}. \tag{3.25}$$

This means that $\Omega(t) - 1$ is a measure of how much the density deviates from the critical density as a function of time. Looking back at equation (3.15), which shows the time-dependence of the scalefactor, we calculate the time-dependence of aH for different models:

- Dust dominated model. $\omega = 0 \rightarrow a(t) \propto t^{\frac{2}{3}} \rightarrow aH \propto t^{-1/3}$, thus aH decreases with time, meaning that $|\Omega(t) - 1| \propto t^{2/3}$ increases with time. Thus, if the universe once was close to flat (i.e. $|\Omega(t) - 1| \approx 0$), it will become less flat as time goes by.

- Radiation dominated model. $\omega = 1/3 \rightarrow a(t) \propto t^{\frac{1}{2}} \rightarrow aH \propto t^{-1/2}$, thus aH decreases with time, meaning that $|\Omega(t) - 1| \propto t$ increases with time.

- Vaccum dominated model. $\omega = -1 \rightarrow a(t) \propto e^{\kappa t} \rightarrow aH \propto e^{\kappa t}$ where κ is a constant that is positive if Λ is positive, which I have assumed. Thus $|\Omega - 1|$ decreases as time goes by! A positive cosmological constant would make the universe flatter! We also see that the Hubble parameter is constant.

For both models dominated by dust and models dominated by radiation the universe gets less flat with time. Then how come our universe is so close to flat today? It must have been very flat long back, and that is possible but not very likely! Adding a cosmological constant could possibly solve the problem.

3.3.3 Where does the small scale perturbations come from?

The universe is very homogeneous and isotropic on large scales. The picture is not the same on small scales, where there are galaxies, stars, planets and humans. There are plenty of structure on small scales, and it is not evenly distributed. This structure would not have formed if the universe had no structure in the beginning.

From the cosmic microwave background radiation, one can see that there are fluctuations, and those fluctuations are a reflection of the universe at the time of the last scattering, when photons and matter stopped exchanging information. The background radiation tels us that the initial density fluctuations were of the order 10^{-5}. What caused these fluctuations?

3.4 Inflation to the Rescue

Fortunately, there exists a model that can account for all the puzzles explained above. Inflation. The idea is that an early epoch of accelerated expansion in the early history of the universe can explain both how the universe became as flat as it must have been in order to be close to flat today, and why all the parts of the sky that we observe today to a very good approximation is isotropic. By introducing a scalar field to “drive” the inflation, it is also possible to explain the initial density perturbations as quantum fluctuations of this scalar field.

So, how can an early epoch of rapid expansion solve the horizon problem? Today, we are only observing back to the time of the last scattering. To explain the observed isotropy on large scales, the part of the universe that we are observing today must have been causally connected at that time (or before). If the universe underwent a rapid accelerated expansion before that time, it would mean that patches of the universe that once were very close, would rapidly move apart from each other. The volume of the universe that once were in causal contact would then grow dramatically. Thus, even if the parts of the sky that are now separated by 180 degrees, could have been in contact before the time of the last scattering.

As mentioned earlier on, the particle horizon can be approximated by the Hubble radius $1/H(t)$. The co-moving Hubble radius is given by $1/a(t)H(t)$, and this quantity is often used to describe the co-moving radius of the volume which is causally connected.

From the discussion of the evolution of $a(t)H(t)$ we had in section 3.3.2, we see that the comoving Hubble radius is increasing with time in a universe dominated by either dust or radiation, and decreases in a universe dominated by a positive cosmological constant. Thus, in a universe existing of "ordinary matter", i.e. dust and radiation, the Hubble radius will increase during the history of the universe, meaning that places further appart on the sky recognize eachoter as time goes by. A universe dominated by vaccum energy would behave oppositely; patches of the sky that once were in contact will loose contact as time goes by. And most interestingly: A universe dominated by vaccum energy in the form of a cosmological constant early on and then dominated by dust or radiation would behave in a familiar way. Parts of the universe which are not causually connected today could have been causally connected a long time ago! Early domination by vaccum energy could solve the horizon problem!

What then about the flatness problem? We saw allready back in section 3.3.2 that the universe become less flat with time in a dust or matter dominated model. In a model dominated by a cosmological constant, however, the universe becomes flatter as time goes by. An early period of vacuum energy domination could "drive" the universe towards a flatter geometry, since the radius of curvature of the universe would grow. Thus, a model dominated by vacuum energy in the form of a cosmological constant early in the history of the universe could explain both the horizon problem and the flatness problem. It is, however, not so simple as to say that the universe early on was dominated by a cosmological constant, and then developed into a radiation and dust dominated universe. As we saw in section 3.2.1, the dependence of the energy density on time ensures that if a cosmological constant dominates at one time, it will not be "overtaken" by dust or radiation in an expanding universe (unless one all of a sudden increases the total content of dust or radiation of course). How did the phase of vacuum energy end?

3.4.1 Scalar field inflation

It turns out that a scalar field with certain properties can supply us with a kind of inflation that ends. A scalar field with potential $V(\phi)$ will have energy density and pressure given by [24]

$$\begin{aligned} \rho c^2 &= \frac{1}{2}\left(\frac{d\phi}{dt}\right)^2 + V(\phi), \\ p &= \frac{1}{2}\left(\frac{d\phi}{dt}\right)^2 - V(\phi), \end{aligned} \tag{3.26}$$

where ϕ is taken to be the zeroth order part of the field ϕ, which is assumed to be well approximated by a homogeneous part $\phi(t)$ and a perturbation $\delta\phi(\vec{x}, t)$. The dimension of ϕ is energy. A reasonable question to ask is why such a scalar field could give rise to cosmological inflation.

If the scalar field is the dominant energy component of the universe, and

$$\frac{1}{2}\left(\frac{d\phi}{dt}\right)^2 \ll |V(\phi)| \tag{3.27}$$

we obtain $\rho \approx -p$ which gives $\omega \approx -1$, giving a proximation to a cosmological constant. This implies that a scalar field behaving as a cosmological constant can help us solve all our problems.

In the following it is assumed that the scalar field dominates a flat universe. Early in the inflationary epoch this is not necessarily the case, but during inflation the universe gets flatter and the other energy-components gets diluted by the expansion of space, so not long after the onset of inflation this will be true.

Inserting equation (3.26) into equation (3.12 a), we obtain

$$H^2 = \frac{8\pi G}{3}\rho = \frac{8\pi G}{3}\left(\frac{1}{2}\left(\frac{d\phi}{dt}\right)^2 + V(\phi)\right). \tag{3.28}$$

The adiabatic expansion equation (equation (3.13)) gives

$$\begin{aligned} &\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + p) \\ &\rightarrow \frac{d\phi}{dt}\frac{d^2\phi}{dt^2} + \frac{dV}{d\phi}\frac{d\phi}{dt} = -3\frac{\dot{a}}{a}(\rho + p). \end{aligned} \tag{3.29}$$

Applying equation (3.26) gives

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0, \tag{3.30}$$

where $\dot{\phi}$ is the time derivative of ϕ and $V'(\phi)$ the derivative of the potential with respect to ϕ. It is assumed that ϕ is not constant ($\dot{\phi} \neq 0$). It is also common to introduce a term $\Gamma_\phi\dot{\phi}$ where Γ is the decay width of the ϕ particles [28]. This decay width is assumed to be zero until inflation ends, so we will neglect it.

Equation (3.30) is analogous to the equation of motion of a particle moving in a potential well $V(\phi)$ experiencing a frictional force proportional to $\dot{\phi}$. When $\ddot{\phi} = 0$, the particle moves with a constant velocity given by

$$\dot{\phi} = -\frac{V'(\phi)}{3H}, \tag{3.31}$$

which is the terminal velocity of the particle. In what is called the slow-roll regime, $\ddot{\phi}$ is negligible, meaning that terminal velocity is approximately reached. We then have equation (3.31). To get inflation like we have seen in the section above, we want to fulfill equation (3.27), and applying equation (3.31) we find:

$$\begin{aligned} &\frac{1}{2}\dot{\phi}^2 \ll |V(\phi)| \\ &\rightarrow \big(V'(\phi)\big)^2 \ll 18H^2|V(\phi)|. \end{aligned} \tag{3.32}$$

Applying equation (3.27) to equation (3.28) and inserting for H in the equation above, we find

$$\begin{aligned} \big(V'(\phi)\big)^2 &\ll \frac{48\pi V(\phi)^2}{E_{Pl}^2} \\ \rightarrow \frac{1}{3}\frac{E_{Pl}^2}{16\pi}\left(\frac{V'(\phi)}{V(\phi)}\right)^2 &\ll 1 \end{aligned} \tag{3.33}$$

where E_{Pl} is the Planck energy given by $E_{Pl}^2 = \hbar c^5/G$. It is common to define a *slow roll parameter* ϵ

$$\epsilon = \frac{E_{Pl}^2}{16\pi}\left(\frac{V'(\phi)}{V(\phi)}\right)^2. \tag{3.34}$$

The first slow roll condition then reads

$$\epsilon = \frac{E_{Pl}^2}{16\pi}\left(\frac{V'(\phi)}{V(\phi)}\right)^2 \ll 1. \tag{3.35}$$

There is one more slow roll parameter, which relates to the curvature of the potential. It was earlier assumed that $\ddot{\phi} = 0$. This will not always be the case, but in equation (3.30) it is an approximation which is valid for

$$\ddot{\phi} \ll V'(\phi). \tag{3.36}$$

From equation (3.31) we then find (first take the derivative of the equation, then apply the equation again):

$$\frac{V''(\phi)}{9H^2} \ll 1, \tag{3.37}$$

where the observation of a constant Hubble parameter during inflation from section (3.3.2) has been applied. It has been assumed that $V'(\phi) \neq 0$. In a universe dominated by the scalar field , $H^2 = 8\pi G V(\phi)/3c^2$. Inserting this for the Hubble parameter, one obtains

$$\frac{1}{3}\frac{V''(\phi)}{8\pi E_{Pl}^2 V(\phi)} \ll 1, \tag{3.38}$$

which gives the second slow roll parameter

$$\eta = \frac{V''(\phi)}{8\pi E_{Pl}^2 V(\phi)}, \tag{3.39}$$

and the second slow roll condition is (when taking into account that the curvature might be negative)

$$|\eta| = \left|\frac{V''(\phi)}{8\pi E_{Pl}^2 V(\phi)}\right| \ll 1. \tag{3.40}$$

The equation for the Hubble parameter (equation 3.28) and the equation of motion for the scalar field (equation 3.30) then becomes

$$H^2 \approx \frac{8\pi G}{3} V(\phi) \tag{3.41}$$

$$3H\dot{\phi} \approx -V'(\phi) \tag{3.42}$$

which are the *slow roll equations*.

The slow roll parameters and the slow roll equations are very useful. Inflation takes place if $\ddot{a} > 0$. The scale factor is always positive, so $\frac{\ddot{a}}{a}$ is then also positive. We get

$$\frac{\ddot{a}}{a} = \dot{H} + H^2 > 0. \tag{3.43}$$

If $\dot{H}$ is positive, then this is easily satisfied, and we get inflation although not by a scalar field. If $\dot{H}$ is negative, we require

$$-\frac{\dot{H}}{H^2} < 1 \tag{3.44}$$

to have inflation. Looking at the first slow roll approximation we find

$$\dot{H} = \frac{4\pi G}{3c^2} V'(\phi) \frac{\dot{\phi}}{H} \tag{3.45}$$

and by also applying the second slow roll equation also we get

$$\dot{H} = -\frac{4\pi G}{3c^2} \frac{E_{Pl}^2}{8\pi} \frac{V'(\phi)}{V(\phi)}. \tag{3.46}$$

Then, using the first slow roll equation one final time, one finds

$$-\frac{\dot{H}}{H^2} = \frac{E_{Pl}^2}{16\pi} \left(\frac{V'(\phi)}{V(\phi)} \right)^2 = \epsilon < 1 \tag{3.47}$$

as the condition for inflation. And this is exactly the first slow roll equation. Thus, $\epsilon \ll 1$ for a scalar field gives inflation given that the slow roll equations are valid. Hence, under the condition that the slow roll approximation holds, meaning that $\epsilon \ll 1$ and $|\eta| \ll 1$, a scalar field gives rise to inflation.

It is common to say that inflation ends at $\epsilon(\phi_e) = 1$. During inflation, the scale-factor grows as

$$a(t) = a(t_i) e^{H_i (t - t_i)} \tag{3.48}$$

where t_i is the time when inflation begins, and $H_i = \sqrt{\frac{\Lambda c^2}{3}}$ is the Hubble parameter which is constant during the period dominating by a cosmological constant.

During inflation, the scale factor grows by a factor

$$\frac{a(t_e)}{a(t_i)} = e^N \tag{3.49}$$

where N is the *N-number* (the number of e-foldings by which the scale factor grows) given by

$$N = H_i(t_e - t_i). \tag{3.50}$$

Thus we see that if the Hubble time of inflation $t_H = H_i^{-1}$ is small compared to the duration of inflation, the N-number will be large, meaning that the universe increases by a large factor. The N-number can also be expressed in terms of the scalar field potential:

$$N = \ln\left(\frac{a(t_e)}{a(t_i)}\right) \tag{3.51}$$

and

$$\ln(a) = \int \frac{da}{a} = \int \frac{\dot{a}dt}{a} \tag{3.52}$$

thus

$$N = \int_{t_i}^{t_e} \frac{\dot{a}dt}{a} = \int_{t_i}^{t_e} H_i dt. \tag{3.53}$$

By applying the slow roll equations (3.41) and (3.42) one obtains

$$N = \int_{t_i}^{t_e} H_i dt = -\frac{8\pi G}{\hbar c^5}\int_{t_i}^{t_e} \frac{V(\phi)}{V'(\phi)}\dot{\phi}dt = \frac{8\pi}{E_{Pl}^2}\int_{\phi_e}^{\phi_i} \frac{V(\phi)}{V'(\phi)}d\phi. \tag{3.54}$$

An N-number of about $N \approx 60$ is needed in order to solve the problems of homogeneity and isotropy.

3.4.2 What causes inflation

The inflation criterion given by the slow roll equations and the slow roll conditions can be met by many different scalar fields. The potential of the field, $V(\phi)$ is unknown. Power-law potentials are generally assumed, possibly with a constant term, but there is no consensus about what caused the inflation. Explaining inflation by a scalar field is only fully satisfactory as long as one come up with a decent explanation for how the scalar field comes into existence. This section is based on [28] as well as the references listed at the beginning of this chapter.

Spontaneous symmetry breaking is a familiar concept in quantum field theory, by giving masses to the weak bosons and the fermions through the Higgs mechanism. It occurs if the symmetry of a system is not shared by its vacuum state. In the standard model of quantum field theory, the underlying gauge symmetry is $SU(3)_C \times SU(2)_L \times U(1)_Y$, while the symmetry of the vacuum state is "only" $SU(3)_C \times U(1)_{EM}$.

In cosmology, spontaneous symmetry breaking is interesting because ever since the big bang, the universe has changed from being extremely hot and dense to rather cold and "dilute". In the true spirit of William of Ockham, it is believed among many physicist that the symmetries today might be the broken relics of symmetries

that would be restored at higher temperatures, and that the high-temperature symmetries might be unifying the physical forces of nature. The spontaneous breakdown of the gauge or global symmetries might have had phase transitions associated with the symmetry breaking process. And the phase transitions might give rise to scalar field inflation!

Grand Unified Theories (GUT) states that the electroweak force and the strong force unifies into some large symmetry at high energies. The scale of this unification, the GUT scale, is of the order 10^{14} GeV. Thus the phase transition causing the minimum of the scalar potential to move and creating a slight slope in the potenitial on which the scalar field "rolls" into the new minimum occurs at temperatures at about the GUT scale, and the new minimum of the potential is at the GUT scale to.

Inflation caused by a phase transition due to a spontaneous symmetry breaking event at the GUT scale is not favoured today. It was suggested early in the development of inflationary ideas, but has been ruled out, and new theories does not to the same extent bother to explain the creation of the scalar field.

One "theory" of inflation, which describes inflation rather than going into what really causes it, is chaotic inflation. Is states that inflation has a rather simple potential, i.e. $V(\phi) = \lambda\phi^4$, without giving any explanation as to what starts the inflation.The minimum of the potential is at $\phi = 0$, and to get inflation, the scalar field must have been displaced from its minimum. The displacement from the minimum (ϕ_i) can take on different values in different parts of the universe. The potential of a chaotic inflation does not have a constant term, so that

$$V(\phi) = M\phi^n, \tag{3.55}$$

where M and n are constants.

The cause of inflation is not established, and inflation is not established as a commonly accepted valid theory, although it is applied assumed in most cosmological models.

3.4.3 Structure formation by inflation

Another very useful property of inflation is its ability to explain the density fluctuations observed in the universe today. The universe might have been very smooth just after the big bang, but in order to allow for matter to contract to form structure, there must have been some initial fluctuations. Large scale structures have formed during the history of the universe, and there are fluctuations in the cosmic microwave background radiation, hence there must have been density fluctuations in the early universe to give birth to the present fluctuations. Inflation gives a possible explanation to this problem; vacuum fluctuations in the scalar field that causes inflation gives rise to density fluctuations.

The density fluctuations of the early universe can be determined from observations, and this have been done by COBE and WMAP. The density perturbation $\frac{\Delta\rho}{\rho}$

at the time of horizon-crossing for the scales observed today are [26]

$$\frac{\Delta\rho}{\rho} \sim 10^{-5}. \tag{3.56}$$

Thus there must have been a mechanism generating such fluctuations in the early universe, as it is not given that there are any fluctuations at all initially.

One simplified, but useful, way of calculating the size of density fluctuations caused by a scalar field slowly rolling towards the minimum of the potential is to apply Heisenberg's uncertainty principle, stating that the uncertainty in time multiplied by the uncertainty in the energy is limmited by

$$\Delta t \Delta E \sim \hbar. \tag{3.57}$$

Thus in a time interval Δt the precision to which the energy can be measured is limited by the equation above. If inflation takes place at the energy scale m, where m is the mass of the scalar field, then the typical energy per particle is given by

$$E = k_B T = m. \tag{3.58}$$

In the early universe, when radiation dominates, one can show that [26]

$$k_B T \sim E_{Pl}\sqrt{\frac{t_{Pl}}{t}} \tag{3.59}$$

within a factor of some small number. This gives a order of magnitude of the energy fluctuation of

$$\Delta E \sim \frac{\hbar}{t} \sim \frac{m^2}{E_{Pl}} \tag{3.60}$$

and thus

$$\frac{\Delta E}{E} \sim \frac{m}{E_{Pl}}. \tag{3.61}$$

This can be related to the density-perturbations, since $\rho \propto T^4$ for ultra relativistic particles, which most of the universe consisted of during inflation. We also know that $E \propto T$, and thus

$$\frac{\Delta\rho}{\rho} \sim \frac{d\rho}{\rho} = \frac{dE}{E} \sim \frac{\Delta E}{E} \tag{3.62}$$

meaning that the energy fluctuations and the density fluctuations are of the same order of magnitude.

The fluctuations caused by inflation occurs because the inflation is not simultaneous all over space. It begins and ends at different times in different parts of the universe. If two volume elements of space both have the same energy, E, but are stretched differently during inflation, there will be a difference in the energy density after inflation. If one volume element is stretched by a scale factor a_1 and

the other by a scale factor $a_2 = a_1 + \Delta a$ caused by different duration of inflation, Δt, then the energy density after inflation differs by

$$\begin{aligned}
\Delta\rho &= E\left(\frac{1}{a_1^3} - \frac{1}{a_2^3}\right) \\
&= E\left(\frac{1}{a_1^3} - \frac{1}{(a_1+\Delta a)^3}\right) \\
&\approx \frac{E}{a_1^3}\left(1 - \frac{1}{(1+\frac{\dot{a}}{a_1}\Delta t)^3}\right) \\
&\approx \frac{E}{a_1^3}\left(1 - \left(1 - 3\frac{\dot{a}}{a_1}\Delta t\right)\right) \\
&= 3H\Delta t\rho \\
&\rightarrow \frac{\Delta\rho}{\rho} \approx 3H\Delta t \sim H\Delta t,
\end{aligned} \tag{3.63}$$

where I have assumed that a differentable and Δt is small.

The difference in ending time of the inflation is expressed as

$$\Delta t \approx \left|\frac{\Delta\phi}{\dot{\phi}}\right|. \tag{3.64}$$

Thus

$$\frac{\Delta\rho}{\rho} \sim H\left|\frac{\Delta\phi}{\dot{\phi}}\right|, \tag{3.65}$$

and an expression for $\Delta\phi$ must be found. By applying the Heisenberg uncertainty principle with the Hubble time as the time interval to the scalar field, one finds

$$|\Delta\phi| \sim \hbar H, \tag{3.66}$$

and the final expression is

$$\frac{\Delta\rho}{\rho} \sim \frac{\hbar H^2}{|\dot{\phi}|}. \tag{3.67}$$

By applying the slow roll equations (3.41) and (3.42), one get

$$\frac{\Delta\rho}{\rho} \sim \frac{(\hbar c)^{3/2}}{E_{Pl}^3}\frac{V(\phi)^{3/2}}{V'(\phi)}. \tag{3.68}$$

Thus we can restrict our potential by applying the observational result that $\frac{\Delta\rho}{\rho} \sim 10^{-5}$ (although in doing so one should be more accurate in the derivation of the above). Unfortunately this is not as simple as it might look, as the potential depends on ϕ. $\frac{\Delta\rho}{\rho}$ is measured today, and is proportional to the temperature fluctuations in the cosmic microwave background radiation, which has been frozen in since the

last scattering. Thus the value of ϕ needed is the value of the scalar field when the mode of interest crosses the horizon during inflation. This makes it hard to find a theoretical value for $\frac{\Delta\rho}{\rho}$, as the potential at the time of horizon crossing is highly model dependent, and the number of possible models are large.

Even though it is hard to predict a concrete model of inflation, most physicists still assume it has taken place. In the rest of this thesis, I am assuming a cosmological model with early inflation.

3.5 Summary

In this chapter, I have introduced the homogeneous and isotropic FRW universe model. The Friedman equations have been discussed, and the puzzles of the standard model was introduced. It is shown that the flatness problem, the horizon problem and the origin of initial fluctuations can all be solved by an early phase of accelerated cosmic expansion called inflation. We have looked at scalar field inflation and the slow roll approximation in some detail.

Chapter 4

Cosmological Perturbations and The Matter Power Spectrum

In this chapter, I will introduce cosmological perturbation theory.

In the previous chapter, the emphasis was placed on isotropic and homogeneous universe models and the Friedman equations. This is a very useful approach to cosmology, and provides a lot of information about the dynamics of the universe as a whole, as the universe is homogeneous and isotropic on large scales. But on cosmologically small scales (i.e. the size of a galaxy), the universe is far from homogeneous and isotropic. The universe is filled with structure, and we should be thankful for that.

The large scale structure of the universe has, as we saw in the last chapter, a history that dates back to the epoch of inflation, which occurred long before even the smallest atoms started to form, in the quantum fluctuations of the scalar field. As time went by, these small initial density fluctuations grew due to gravity, and the result is visible to us today in galaxies, our solar system and in the tiny temperature fluctuations in the cosmic microwave background radiation.

In order to understand the nature of the cosmic structure formation, we must understand the initial fluctuations and their development from inflation and until today. In this section, we will look at how the density perturbations can be described as perturbations to a homogeneous and isotropic background. I will start to perturb the metric, and then go on to perturb the energy density. Then I will also introduce the matter power spectrum, which is the correlation function of the matter density distribution today.

This chapter is, when otherwise is not stated, based on the references [24, 29, 30, 31].

4.1 The Metric Perturbations

In this section, I will introduce the perturbations to the metric. The perturbations to the metric is a consequence of inflation, and the metric depends on the energy

content of the universe. Thus the metric evolves as the universe evolves.

The behaviour of the universe with small perturbations is often calculated using perturbation theory. The perturbed states are assumed to be perturbations to the Friedman-Robertson-Walker universe described in chapter 3. In such models, the unperturbed metric reads

$$g^0_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & a^2(t) & 0 & 0 \\ 0 & 0 & a^2(t) & 0 \\ 0 & 0 & 0 & a^2(t) \end{pmatrix}, \tag{4.1}$$

and the line element is

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -dt^2 + dx^2 + dy^2 + dz^2. \tag{4.2}$$

It is useful to introduce conformal time. The conformal time η is given by

$$a^2(\eta)d\eta^2 = dt^2, \tag{4.3}$$

thus

$$\eta = \int_0^t \frac{dt'}{a(t')}. \tag{4.4}$$

This looks a lot like the particle horizon defined in (3.16), only short of the scale factor. In fact, the conformal time is nothing but the comoving particle horizon. Hence conformal time is an important parameter as it determines the comoving volume of causal contact. Using conformal time, the line element becomes

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = a^2(\eta)(-d\eta^2 + dx^2 + dy^2 + dz^2), \tag{4.5}$$

where

$$g^0_{\mu\nu} = a^2(\eta) \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.6}$$

We can then introduce physical inhomogeneities to the universe by adding a first order perturbation to the FRW metric:

$$g_{\mu\nu} = g^0_{\mu\nu} + \delta g_{\mu\nu}, \tag{4.7}$$

which allows us to write out the line element as

$$ds^2 = \left(g^0_{\mu\nu} + \delta g_{\mu\nu}\right) dx^\mu dx^\nu. \tag{4.8}$$

The perturbation to the metric consists of three parts, the scalar, vector and tensor perturbations. The three types of perturbations evolves independently of each other, and can thus be treated individually. The scalar perturbations couples to the density of matter and radiation, giving rise to inhomogeneities and anisotropies in

the universe today. The tensor perturbations are called gravitational waves, and in first order perturbation theory they do not couple to matter. The vector perturbations are not very interesting in a cosmological context, as they do not couple to matter and are not generated by the simplest inflationary models. They also decay rapidly. I will mainly care for the scalar perturbations in my further study, thus we need to get into some detail about these perturbations. The scalar perturbations are important in neutrino cosmology as neutrinos affect the gravitational collapse and the creation of structure in the universe.

4.1.1 Scalar perturbations

The form of the scalar perturbation to the metric is dependent upon choice of gauge. By choice of gauge I mean the choice of coordinates used to describe the system. The choice of gauge does not affect the physical results obtained. In conformal Newtonian gauge, the perturbation to the metric is

$$\delta g_{\mu\nu} = a^2(\eta) \begin{pmatrix} -2\Psi & 0 & 0 & 0 \\ 0 & 2\Phi & 0 & 0 \\ 0 & 0 & 2\Phi & 0 \\ 0 & 0 & 0 & 2\Phi \end{pmatrix}, \tag{4.9}$$

such that the line element becomes

$$ds^2 = a^2(\eta) \left(-(1+2\Psi)d\eta^2 + (1+2\Phi)\left(dx^2 + dy^2 + dz^2\right)\right). \tag{4.10}$$

The functions Ψ and Φ depend on both time and space.

This perturbed metric influences the energy distribution, and the energy distribution influences the metric. Solving for the time evolution of the perturbations, we need to look at the Boltzmann equations and the Einstein equations for the perturbed metric. Doing so to great detail is out of the scope for this thesis, but a more detailed discussion is given in [24]. In such a calculation, one has to take into consideration that not only does the metric and the gravitational forces affect all energy components of the universe, and that the constituents of the universe affects the metric, but also that the particles of the universe affects *each other*, see figure 4.1. There is Coulomb scattering between photons and protons, and Compton scattering between electrons and photons. Thus the equations for all the components need to be solved simultaneously.

The particle distribution is given by the Boltzmann equation

$$\frac{df_i(x,t,p)}{dt} = C[f], \tag{4.11}$$

where f_i is the distribution function for particle type i and $C[f]$ contains all possible collision terms which depends on the interaction with other energy components. The particles also respond to perturbations in the metric. The metric perturbations react to the perturbations in the energy density in the universe, and the relation between the two is given by the Einstein equation (3.1).

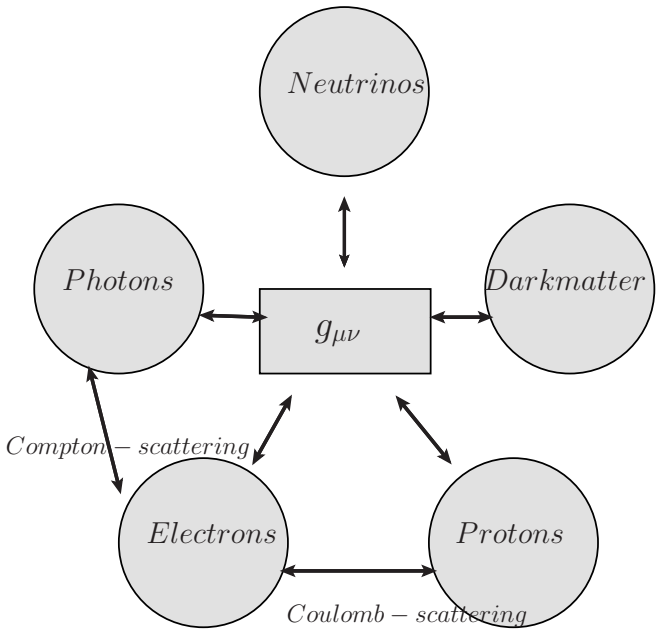


Figure 4.1: All the energy components interact with the metric. This means that the Einstein equations and the Boltzmann equations has to be solved simultaneously for all species. This figure is inspired by [24]

4.2 The Power Spectrum

The power spectrum of the matter fluctuations, or the matter power spectrum for short, is a measure of the fluctuations of matter density. It is essentially a statistical measure of the amplitude of fluctuations, or matter clustering, at different scales. In this section, I will introduce power spectra in general, and the matter power spectrum in particular. I go into the level of details needed to understand the effect of neutrinos on the structure formation in the universe.

In linear perturbation theory, the matter density is given by

$$\rho_m = \rho_m^{(0)}(1 + \delta_m) \tag{4.12}$$

where $\rho_m^{(0)} =< \rho_m >$ is the spatial average density, and

$$\delta_m(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \rho_m^{(0)}}{\rho_m^{(0)}} \tag{4.13}$$

is the matter overdensity. To simplify the equations used in cosmological perturbation theory, one tends to work in Fourier space. This simplifies the equations derived from the Einstein and Boltzmann equations a lot, as it transforms the partial differential equations into decoupled ordinary differential equations which can

be solved for each Fourier mode individually. Also, transforming into *k-space*, the scale dependence comes out nicer. $\mathbf{k}$ is the wave-vector, and $k = |\mathbf{k}|$ is the wavenumber, or to be stricly correct, the comoving wavenumber,

$$k = \frac{2\pi}{\lambda^C}, \tag{4.14}$$

such that a small value of k indicates perturbations on a large comoving scale, λ^C. The physical wavelength relates to the comoving wave-number by

$$\lambda = a(t)\frac{2\pi}{k}. \tag{4.15}$$

The Fourier transform

$$F(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{x}\cdot\mathbf{k}} F'(\mathbf{k}), \tag{4.16}$$

where $F'(\mathbf{k})$ is the Fourier transform of $F(\mathbf{x})$, is used. The overdensity in interest is then $\delta_m(\mathbf{k})$. In Fourier space, two overdensities or perturbations at different values of $\mathbf{k}$ is independent, that is $\delta_m(\mathbf{k})$ is uncorrelated with $\delta_m(\mathbf{k}')$.

4.2.1 The primordial power spectrum

The overdensity is, obviously, zero on average, but the variance

$$\sigma^2 \equiv < \delta_m^2 > \tag{4.17}$$

over the whole distribution can be non-zero. The matter power spectrum is defined as the two-point correlation function of the matter fluctuations in Fourier space:

$$< \delta_m(\mathbf{k})\delta_m(\mathbf{k}') >= (2\pi)^3 P(k)\delta^3(\mathbf{k} - \mathbf{k}') \tag{4.18}$$

where the brackets denote the average over the whole distribution, in a (on large scales) homogeneous and isotropic universe. The delta function is due to the uncorrelation of overdensities with different values of $\mathbf{k}$, a nice feature of Fourier transformations and linear regime. It is worth noting that the matter power spectrum $P(k)$ is not dependent on the direction of $\mathbf{k}$. This is because it is assumed that the universe is homogeneous and isotropic.

The matter power spectrum essentially tells us about the variance in the distribution of matter fluctuation. If the value of $P(k)$ is large, it indicates that at the scale k the overdensities and/or underdensities are large. A small value of $P(k)$ tells us that the universe is smooth at scale k.

As we saw in the previous section, the metric is perturbed, and so is the density contrast. The inflationary scalar field will also have perturbations given by:

$$\phi(\mathbf{x}, t) = \phi^{(0)}(t) + \delta\phi(\mathbf{x}, t), \tag{4.19}$$

and the perturbation in the scalar field and the perturbation to the metric plays together in such a way that they couple. Since others, like [24], already has calculated the power spectrum of the scalar perturbations, I will only state the result:

$$P_\Psi(k) = P_\Phi(k) = \frac{2}{9k^3}\left(\frac{aH^2}{\dot{\phi}^{(0)}}\right)^2\Bigg|_{aH=k}. \tag{4.20}$$

Without going into too much detail, this equation needs an explanation. First of all; the first equality is made assuming that anisotropic stress is negligible[1]. The initial conditions for the Boltzmann and Einstein equations used in cosmological perturbation theory is set when a scale leaves the horizon during inflation. But what does this mean?

The ratio of k to η is a good measure of the size of a scale. The wavenumber k scales inversely to the comoving wavelength. η is the comoving horizon. Thus $k\eta$ is the ratio of the comoving horizon to the comoving wavelength of the perturbation. If this ratio is less than one, then the wavelength of the perturbation is greater than the horizon, and no causal physics could affect it. If $k\eta$ is greater than one, then causal physics can affect the perturbation. The comoving horizon is decreasing during inflation and increasing after inflation, but the comoving wavelength is constant. Thus perturbations will move out and in of the horizon during the history of the universe. Initial conditions for the cosmological perturbations are set when the perturbation wavelength crosses the horizon during inflation. During inflation the Hubble parameter is by a good approximation constant, and from equation (4.4) we find that during inflation (but not too close to the end of it) $\eta \approx 1/aH$. Thus $aH = k$ in the equation above indicates that $P_\Psi(k)$ is found by evaluating the equation at horizon crossing. It can be shown [24, 30] that the equation above can be expressed using the slow roll parameter ϵ like

$$P_\Psi(k) = P_\Phi(k) = \frac{8\pi G}{9k^3}\frac{H^2}{\epsilon}\Bigg|_{aH=k}. \tag{4.21}$$

This is the equation we use later.

The power spectrum in equation (4.21) has an explicit k^{-3} dependence. A spectrum in which $k^3 P_\Phi(k)$ is independent of k is gives rise to a scale-invariant matter power spectrum, as we shall see later. This is referred to as a Harrison-Zel'dovich-Peebles spectrum. The power spectrum is often written in such a way that one easily can spot any deviations from a scale-invariant spectrum:

$$P_\Phi(k) = \frac{50\pi^2}{9k^3}\left(\frac{k}{H_0}\right)^{n_s-1}\delta_H^2\left(\frac{\Omega_m}{D_1(a=1)}\right). \tag{4.22}$$

[1] Anisotropic stress is the non-diagonal components to the spatial part of the energy-momentum tensor. Only relativistic particles contributes, and it relates to the quadrupole moments of the relativistic particles. Compton scattering effectively reduces the contribution from photons in the early universe while the photon contribution to the energy density is dominant. The neutrino contribution is not negligible for massive neutrinos, and gives a small contribution to the anisotropic stress [32], but for now we neglect neutrinos. When no anisotropic stress is assumed, $\Psi = \Phi$ [24].

δ_H is the scalar amplitude, the amplitude at horizon crossing. D_1 is a growth factor that describes the growth of the matter perturbations at late times, and n_s is the scalar *spectral index*, which tells us about the deviation from a scale-invariant spectrum. I will return to the spectral index later.

4.2.2 The matter power spectrum

Having found the scalar perturbation power spectrum, we still need to do some work to find the matter power spectrum. The matter power spectrum today is derived using the scalar perturbation spectrum generated by inflation as an initial condition.

Throughout the evolutionary history of the universe, the scalar perturbations evolve differently depending on the substance dominating the energy density of the universe at the time. The perturbation modes also evolves differently depending on whether or not they are inside the horizon. Both at very early times, when all modes are outside the horizon and well before matter-radiation-equality, and at very late times, after horizon crossing and matter-radiation equality, all modes behaves in the same way. But sometime in between, when modes starts to cross the horizon and around the time of matter-radiation-equality, modes on different scales behave very differently. When relating the potential today to the primordial potential, the potential set up during inflation, it is common to factorize it as

$$\Phi(\mathbf{k}, a) = \frac{9}{10}\Phi_p(\mathbf{k}) \times \{TransferFunction(k)\} \times \{GrowthFunction(a)\}\,. \tag{4.23}$$

$\Phi_p(\mathbf{k})$ is the premordial value of the potential (the scalar perturbation to the metric) set during inflation. The transfer function, $T(k)$ describes how the k-mode evolves through the epoch of horizon crossing and matter-radiation equality. The transfer function is defined in such a way that it is one or large scales, causing the 9/10 factor on the right hand side. The growth function describes how the mode evolves well after that epoch, and is defined by

$$GrowthFunction = \frac{\Phi(a)}{\Phi(a_{late})} = \frac{D_1(a)}{a} \qquad\qquad a > a_{late}. \tag{4.24}$$

The evolution of late times is independent of the scale of the mode, hence the growth function only depends on the scale factor a. a_{late} is the scale factor when a mode is well after the regime of the transfer function, when the evolution can be described by the scale factor alone.

The potential Φ can be related to the density perturbation today using a combination of components of the Einstein equations on large scales, called Poisson's equation

$$\Phi = \frac{4\pi G\rho_m a^2 \delta}{k^2} \qquad\qquad a > a_{late} \tag{4.25}$$

giving (after some simple substitutions)

$$\delta(\mathbf{k}, a) = \frac{k^2 \Phi(\mathbf{k}, a) a}{(3/2)\Omega_m H_0^2} \qquad a > a_{late}. \tag{4.26}$$

Then, using equations (4.23) and (4.24), we find

$$\delta(\mathbf{k}, a) = \frac{3}{5} \frac{k^2}{\Omega_m H_0^2} \Phi_p(\mathbf{k}) T(k) D_1(a) \qquad a > a_{late}, \tag{4.27}$$

showing how the matter overdensity is related to the scalar perturbation to the metric. This causes the matter power spectrum to be expressed by the scalar perturbation power spectrum calculated in equation (4.22). The matter power spectrum can then be calculated, giving:

$$P(k, a) = 2\pi^2 \delta_H^2 \frac{k^{n_s}}{H_0^{n_s+3}} T^2(k) \left(\frac{D_1(a)}{D_1(a=1)} \right)^2 \qquad a > a_{late}. \tag{4.28}$$

The power spectrum today simplifies to

$$P(k) = 2\pi^2 \delta_H^2 \frac{k^{n_s}}{H_0^{n_s+3}} T^2(k). \tag{4.29}$$

4.2.3 The spectral index

It can be shown [24, 30] that the spectral index can be expressed using the slow roll parameters

$$n_s = 1 - 4\epsilon - 2\delta \tag{4.30}$$

in the case of scalar field slow roll inflation. The slow roll parameters are time dependent, but we are interested in n_s at a particular time. As the spectral index first appears in the primordial power spectrum, in equation (4.22), the scale factor of interest is measured when a mode crosses the horizon during inflation, thus the slow roll approximation is valid, and $\epsilon \ll 1, \delta \ll 1$. Thus $n_s \lesssim 1$. A more physical way of understanding the above is that the Hubble parameter changes very slowly during inflation, and the slow roll parameters are related to the rate of change of the Hubble parameter.

The fact that slow roll scalar field inflation gives a specific spectral index is of great help, as a detection of a spectral index that deviates greatly from one would imply that slow roll scalar field inflation did not happen. Unfortunately, the scalar index can be one even without slow roll scalar field inflation, so detecting $n_s = 1$ does not provide sufficient proof of inflation.

4.2.4 The turnover in the matter power spectrum

The transfer function gives information on how a mode changes during the period of horizon crossing and radiation-matter equality. These two events are important

to the evolution of structure formation; when a mode crosses the horizon, it can collapse due to gravity as it is within the causal scale. When the dominant energy component of the universe changes, then so does the rate of perturbation growth, thus the transition period from radiation to matter also affects the evolution of the modes.

At early time in the history of structure formation, radiation dominates the energy content of the universe. All the way during radiation domination, and until photon decoupling, radiation pressure works against the force of gravity. Thus gravitational collapse is less efficient than during matter domination after photon decoupling, where gravitational collapse is very efficient. This means that the small scale modes, entering the horizon well inside the radiation dominated era, evolves very differently from the large scale modes that crosses inside the horizon during matter domination. When a scale enters the horizon during the radiation dominated era, the potential (Φ) decays, retarding the growth of the density contrast δ. Thus, the earlier a scale enters the horizon during radiation domination, the more retarded the density perturbation growth gets. When a mode enters the horizon during matter domination, there is no retardation. The primordial power spectrum of the denity contrast can be found from equations (4.22) and (4.26) to be

$$P(k)_{primordial} \propto k^{n_s}, \tag{4.31}$$

with $n_s \lesssim 1$. Thus, on large scales, the modes that cross the horizon well inside matter domination was proportional to k^{n_s} before horizon crossing and before matter domination. When matter dominates, all scales inside the horizon grows at the same rate. This implies that the power spectrum has a turnover for the mode that crosses the horizon at matter-radiation equality. The power spectrum looks pretty much like figure 4.2. We can easily see the expected turnover. The logarithmic plot also displays the linear behaviour anticipated on large scales.

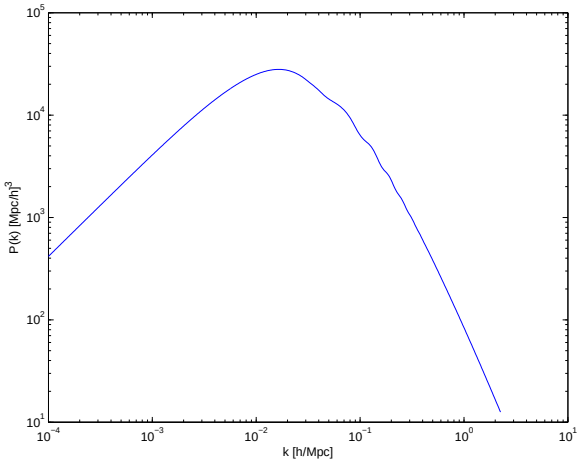


Figure 4.2: The matter power spectrum for typical values of the cosmological parameters. One can easily see the behaviour described in section 4.2.4, with a turnover for the scales that crossed the horizon about the time of matter radiation equality. The power spectrum is made using CAMB [33].

Chapter 5

Cosmological structure formation

Structures has been formed in the universe ever since the big bang. In this section, I will look at the evolution of structure in the universe, and how neutrinos affect this picture. I will start with a short history of the very early years of the universe, and go onto discussing how pressure and neutrino free streaming influences structure formation by gravitational collapse.

I will also discuss two observables, the baryon acoustic oscillations (BAO) caused by the counteraction exerted by pressure on the gravitational forces, and the mass variation in a sphere of radius $8hMpc^{-1}$, σ_8, which is the measure of the over-densities on scales of the size of galaxy clusters. These two parameters are observed in the matter and CMB power spectra, and are sensitive to the neutrino mass.

5.1 A Very Rushed Tour through the History of the Universe

In this subsection, I will very briefly outline the evolutionary history of our universe. The description will not go into very much detail, but an outline is needed to understand the formation of the *cosmic microwave background radiation* (CMB) and the other observables in cosmology. In this section, I have mainly used the references [34, 29, 24].

5.1.1 At the very beginning of time

The history of the universe is thought to have begun some 14 Gy ago, by a very big bang. The *bang* started in a singularity, and as singularities are not in the scope if this thesis, we will not dwell upon what created the universe in the first place. The important thing to know is that very early on, the universe was rapidly expanding, incredibly hot and dense and consisting of ionized plasma with rapid interaction between the particles. As the universe expanded, the plasma cooled

down, eventually forming nuclei, atoms and the larger scale structure as galaxies, and smaller scale structures like us and the starts.

At very early times, the universe is thought to have undergone a phase of very rapid expansion, known as inflation. Before the universe had existed for as much as a second, its size underwent a very short period of expansion in which its size increased by a factor greater than e^{60} times in less than 10^{-30} seconds. The reason why it is believed that the universe underwent inflation, is that such an epoch in the history of the universe can both explain why the universe seems to be so flat, isotropic and homogeneous, and why small perturbations are formed in such a way that we observe the power spectrum of matter to be scale invariant

$$P(k) \propto k^{n_s}, \tag{5.1}$$

where n_s is the spectral index. Inflation is discussed in section 3.4 and the power spectrum is explained in section 4.2.2, and we will now just accept that an early accelerated expansion probably took place. After inflation, the universe was dominated by the inflationary non-relativistic scalar field particles for a while. This era is referred to as the cold big bang. These particles decayed into relativistic particles, leaving a universe dominated by radiation. This process is called reheating, as massive non-relativistic particles decays into relativistic light particles.

5.1.2 Formation of very small structure; nucleosynthesis and recombination

In the radiation dominated post-inflationary universe, nuclei starts to form when the temperature lowers. First, quarks combine making hadrons. Then, at about 1-0.1 MeV, the hadrons combine forming light nuclei; hydrogen and helium. The nuclei and the electrons interact very strongly through the Coulomb scattering, and photons and electrons are strongly coupled via Compton scattering. As the universe continues to cool down, the interaction rates decrease.

When the temperature drops below the ionization energy of hydrogen, *recombination* begins. The process

$$H + \gamma \leftrightarrow p + e^- \tag{5.2}$$

drops out of equilibrium, and the same goes for the equivalent process for helium. This causes the free electrons to be captured by the nuclei, and the universe goes from being ionized to neutral.

5.1.3 Decoupling of photons and the formation of cosmic microwave background radiation

Since the number of free electrons drops rapidly during the epoch of recombination, the rate of interaction between photons and electrons drops as well. When the scattering rate between electrons and photons drops below the expansion rate

of the universe, the photons are said to be decoupled from the electrons, as their mean free path becomes comparable to the size of the observable universe. Before decoupling, the universe was opaque, as the photons never travelled very far before interacting and scattering. After decoupling, the photons could travel almost undisturbed through the universe, mainly affected by the cosmic expansion. These are the photons that we observe today as the CMB radiation. As the photons were in thermal equilibrium with each other and the baryonic matter of the universe at the time of decoupling, and as they all "last scattered" at basically the same time, they all had basically the same temperature at the time of recombination. Moving through the universe towards us, they have all been subjected to the same redshift due to cosmic expansion. Thus, the CMB has basically the same temperature in all directions. There are, however, small fluctuations of the order $\delta T/T \sim 10^{-5}$. These small fluctuations, already present at the time of recombination, were the seeds that initiated structure formation, forming all the cosmic structure that we see today.

5.2 Jeans Scale and Structure Formation in a ΛCDM Model

In a universe containing nothing but matter, gravitational forces or curvature of space-time, depending on the level of detail, causes matter collapse into structure. This happens in our universe too, but as the energy density of our universe is not purely due to matter, pressure and cosmic expansion complicates the simplified picture of gravitational collapse. In this section, I will focus on the contribution of pressure. I assume a so called flat Λ*CDM model*, a universe containing baryons, photons, cold dark matter and a dark energy in the form of a cosmological constant. I have mainly used the references [34, 29, 24] in this section, other references are stated when used.

When talking about structure formation in the early universe, even before the decoupling of photons and baryons, there are one scale in particular that, together with the horizon, plays an important part, and that is the *Jeans scale*, which sets the bounds to the size of structure formation when matter and radiation are still coupled. The picture is really rather simple: baryonic over-densities like to collapse due to gravitational forces, and they will, if not prevented by pressure exerted by the photons (or other forces), which are coupled to the baryons. Thus, quite simplified, the pressure of the photons hinder the baryons to collapse due to gravity.

If a general sphere of gas is compressed, a pressure gradient will build up that works in the opposite direction of gravity, see figure 5.1. If there were no pressure, an overdense sphere would collapse in a time given by

$$t_{collapse} \sim \frac{1}{\sqrt{G\rho^0}}, \tag{5.3}$$

as one can see from the Friedman equations.

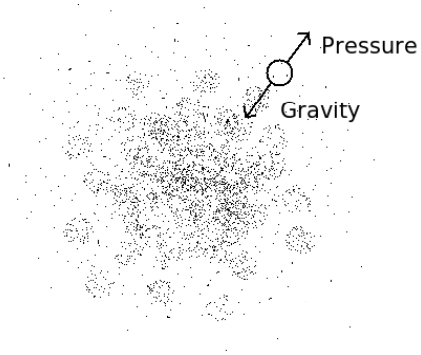


Figure 5.1: The pressure in the gas will exert an outward force, while the gravity exerts an inward force.

If the pressure is nonzero, the pressure will increase as the sphere of gas collapses. Pressure cannot, however, build up instantaneously, it only changes with the sound speed c_s. The speed of sound is found from

$$c_s = \sqrt{\frac{\partial P}{\partial \rho}} = \sqrt{\omega}. \tag{5.4}$$

When both matter and photons contribute, the sound speed is given by

$$c_s = \sqrt{\frac{1}{3(1+R)}} \tag{5.5}$$

where R is the ratio of baryons to photons

$$R = \frac{3\rho_b}{4\rho_\gamma}. \tag{5.6}$$

The dark matter does not influence the sound speed, as it does not interact, and the weak interaction of the neutrinos makes their contribution negligible.

With the introduction of a sound speed, we can introduce a sound horizon, analogous to the particle horizon.The particle horizon is given by

$$d_{PH} = a(t) \int_{t_0}^{t} \frac{c dt'}{a(t')}. \tag{5.7}$$

To find the sound horizon, we simply substitute the speed of light by the speed of sound:

$$d_{SH} = a(t) \int_{t_0}^{t} \frac{c_s dt'}{a(t')}. \tag{5.8}$$

The sound speed sets the scale reachable by pressure waves, analogous to the speed of light sets the scale which is in causal contact.

The time taken to change the pressure in a region of radius λ is then

$$t_{pressure} \sim \frac{\lambda}{c_s}. \tag{5.9}$$

If the sphere of gas takes longer to collapse than the pressure take to react to the increased density, then pressure have time to counter the gravitational attack. This occurs if

$$t_{pressure} < t_{collapse}, \tag{5.10}$$

which gives

$$\lambda < c_s \frac{1}{\sqrt{G\rho^0}}, \tag{5.11}$$

thus, if the size of the gas sphere is smaller than a quantity decided by the density of the gas and the speed of sound, then the sphere of gas cannot collapse by gravity due to the induced pressure force. This means that small perturbations will not grow. The radius λ is the *Jeans length*, denoted by λ_J.

A more precise derivation of the Jeans length (using the continuity, Euler and Poisson equations), gives

$$\lambda_J = c_s \sqrt{\frac{\pi}{G\rho^0}}, \tag{5.12}$$

from solving

$$\frac{d^2\delta}{dt^2} + \frac{2}{a}\frac{da}{dt}\frac{d\delta}{dt} + \left(c_s^2 k^2 - 4\pi G\rho_0\right)\delta = 0 \tag{5.13}$$

for the oscillating solutions. Scales smaller than λ_J is said to be Jeans stable. In a flat Friedmann-Robertson-Walker universe, the Jeans length can be reduced to (using the Friedmann equations)

$$\lambda_J = 2\pi \sqrt{\frac{2}{3}} \frac{c_s}{H}, \tag{5.14}$$

On scales shorter than the Jeans scale, not only will gravitational collapse of baryons not occur, but the modes will tend to oscillate acoustically. This gives rise to acoustic peaks in the CMB and matter power spectra today, which I will discuss later on.

On scales larger that the Jeans scale but smaller than the horizon (approximated by the Hubble length), gravitationally collapse is not fully supressed by pressure. Thus one could think that the perturbations would grow on scales smaller than the Hubble scale, and in a static universe, that would no doubt be the case. But in an expanding universe, the scale factor must also be considered. In a radiation

dominated universe, solving an equation very similar to equation (5.13) for large scales with some help of the Friedman equations and the equation of state, one finds that the growing mode develops as [35]

$$\delta \propto t \propto a^2. \tag{5.15}$$

Thus, on scales larger than the Jeans scale, the radiation perturbations seems to grow. The matter perturbations does not grow, as the matter density is too small at this stage to gravitationally conquer the rapid expansion.

Solving for the dark matter perturbations inside the Jenas length during the radiation dominated era, one finds ([24, 36]) that the perturbations grow logarithmically on scales smaller than the Hubble horizon. On scales larger than the Hubble horizon, the modes are constant. Thus the cold dark part of the matter distribution gets a head start on the baryonic matter when it comes to density perturbations and structure growth on small scales during the radiation dominated epoch.

During radiation domination, the photons are responsible of the pressure force that prevents gravitational collapse of baryons on scales smaller than the Jeans length. At photon decoupling, the photons and baryons forms two decoupled gases, and the radiation pressure no longer prevents the baryonic matter to gravitationally collapse on all scales. There is, however, a pressure due to the high temperature of the baryons, but this pressure is lower than the pressure associated with the photons, and the Jeans length for baryons drops substantially at recombination. As the baryons cool down by the expansion of the universe, the Jeans scale drops even more. But most importantly, at decoupling baryonic matter is allowed to start forming structures on scales small compared to the Hubble length. Thus, baryonic structure starts growing, and although the dark matter gets a head start, the baryonic matter grows faster after decoupling than it would have if there were no cold dark matter. The cold dark matter perturbations on small scales has grown all through the radiation dominated era, and thus have created gravitational structure, which traps the baryonic matter in its gravitational potntials.

When matter dominates, all matter grows as

$$\delta_m \propto a \tag{5.16}$$

on causal scales (inside the Hubble horizon), and this is the most "productive" period of structure formation. During the latest era, dominated by a cosmological constant, the growth of causal perturbations slows down, due to decay of the metric perturbations. Assuming that the universe is flat, increasing the density parameter of the cosmological constant today would have to be compensated by a reduction in the matter density parameter. Less matter in the universe would mean less material to build structure, and a shorter period of efficient structure formation. Thus, the cosmological constant is not doing much good for structure formation.

It is worth noting that $\lambda_J \propto \rho^{-1/2}$, and thus is evolving not only when changing from one era to another, it also changes within each era. In a radiation dominated universe, we have $\lambda_J \propto a^2$, and in a matter dominated universe, $\lambda_J \propto a^{3/2}$. If a

cosmological constant dominates, then the Jeans scale is constant. The fact that the Jeans scale grows during radiation and matter domination means that larger and larger physical scales will be inside the Jeans length, and thus not able to cluster unhindered.

In this section we have made a number of simplifications. Only linear perturbations have been considered, and when dense structure start to form and $\delta_m \ll 1$ is no more a good approximation, the discussion in this section does not hold. We have also neglected the effects of phase transitions between different epochs. Going into more detail on this is mathematically quite complicated, and is thus omitted.

5.2.1 Baryon Acoustic Oscillations

The Jeans scale at the time of photon decoupling and the baryon acoustic oscillations scale can be extracted from cosmological observations. This provides a useful tool for determining the energy content in the universe. This section is based mainly on the references [37, 38, 39, 40].

As mentioned earlier, at early times the baryonic matter is constrained from gravitational collapse on small scales until the time of photon decoupling. On greater scales, however, the baryons do collapse, only restrained by the size of the Jeans scale, which is a growing function of time. This causes the baryonic matter to "pile up" at the Jeans scale by the time of decoupling. At this time, the Jeans scale is a characteristic scale of the distribution, and a spherical density perturbation will up until decoupling have been propagating radially. After the decoupling, the baryon fluid is no more affected by the photon pressure, and the density perturbations are free to grow on all scales (this is not entirely true, the neutrino free streaming affects the perturbations, and I will return to this later, but for now neutrinos are neglected). Thus, the baryon acoustic oscillation has caused a scale of density excess equal to the Jeans length at decoupling.

The baryon perturbations have developed from decoupling and until today, but structure formation was given an initial condition at the time of decoupling which still affects the matter power spectrum today. Since the dark matter perturbations were allowed to collapse on all scales even before decoupling, these perturbations and the baryon perturbations looks different just after the time of decoupling. The dark matter had clustered on small scales before decoupling, and after decoupling the baryonic matter also clustered on small scales, a process that was accelerated by the gravitational potential already created by the dark matter. But the characteristic scale of the overdensity of the baryonic matter also affected the dark matter. A figure of how a spherical density perturbation would evolve, with the mass profile as a function of the comoving radius is presented in figure 5.2. The radius of the shell of overdensity marks a preferred separation of matter, which later on causes the same scale to be a preferred separation of large scale structure. At recombination, this shell has a radius of roughly 150 Mpc [37]. Since the dark matter both is dominant compared to the baryonic matter and has had time to form structure on all scales, the acoustic scale is not dominating the large scale structure formation. In a correlation function the scale of acoustic oscillation is, however, visible as a peak at about 150 Mpc separation.

The existence of baryonic acoustic oscillations, and the position of the peak in the correlation function, is a confirmation of the presumed evolution of large scale structure in the universe. But more importantly, it provides a standard ruler with witch to measure distances in the universe. There are two measurable parameters of a structure on the sky which can be related to parameters that are harder to measure; the angular separation $\Delta\theta$ and the difference in redshift from the front to

the back of the separation Δz. The comoving distance r_c to a redshift z is

$$r_c(z) = \int c(1+z)dt, \tag{5.17}$$

thus

$$\frac{dr_c}{dz} \simeq \frac{\Delta r_c}{\Delta z} = \frac{c}{H(z)}. \tag{5.18}$$

This gives

$$H(z) = \frac{c\Delta z}{\Delta r_{c||}}, \tag{5.19}$$

where $\Delta r_{c||}$ is the comoving distance between the front and the back of the separation in question. If $\Delta r_{c||}$ is known, the Hubble parameter can be calculated as a function of redshift. This can give important information about the energy content of the universe.

Another quantity that can be measured from objects on the sky is the comoving angular diameter distance d_A. For objects far away

$$\Delta\theta \simeq \sin\Delta\theta = \frac{\Delta r_{s\perp}}{d_A(z)} \tag{5.20}$$

where $\Delta r_{s\perp}$ is the comoving distance perpendicular to the line of sight of the separation in section.

When it comes to baryon acoustic oscillations, the trouble is that we observe it as a statistical property of the matter distribution in the baryonic matter. This implies that measuring $\Delta\theta$ and Δz is hard to do.

What we do observe, is the matter power spectrum. The Fourier transform of the power spectrum is a two-point correlation function, that gives information about the relative excess clustering in a given scale, relative to a uniform distribution [40]. Thus it gives information on the preferred separation distances of cosmic objects.

To compare theory to observations, it is useful to introduce the distance measure [37]

$$D_V(z) = \left[d_A(z)^2 \frac{cz}{H(z)}\right]^{1/3}, \tag{5.21}$$

which relates to the two point correlation function.

The distance measure in equation (5.21) is dependent on the Hubble parameter today, but the observations used to determine the BAOs are not very sensitive to H_0. To get rid of that dependence, another parameter is introduced by Eisenstein [37]:

$$\mathcal{A}(z) = D_V(z)\frac{\sqrt{\Omega_m H_0^2}}{0.35c}. \tag{5.22}$$

This parameter is redshift dependent, and calculable by the integral function

$$\mathcal{A}(z) = \sqrt{\Omega_m}E(z)^{-1/3}\left[\frac{1}{z}\int_0^z \frac{dz'}{E(z')}\right]^{2/3} \tag{5.23}$$

where [40]

$$E(z) \equiv \frac{H(z)}{H_0} = \sqrt{(1+z)^3\Omega_m + f(z)\Omega_{DE} + (1+z)^2\Omega_k + (1+z)^4\Omega_{rad}}. \tag{5.24}$$

In the equation above, $f(z)$ depends on the nature of the dark energy; for a cosmological constant, $f(k) = 1 + z$. The contribution from radiation is dominated by photons, and for low redshifts (late times) this contribution is negligible. For a flat universe, $\Omega_k = 0$.

I have assumed a flat universe, and the values of z used is consistent with a negligible contribution from radiation. Neutrinos are assumed to contribute to the matter density parameter, thus I have assumed

$$E(z) \equiv \frac{H(z)}{H_0} = \sqrt{(1+z)^3\Omega_m + f(z)\Omega_{DE}}. \tag{5.25}$$

The z-dependence of the dark energy contribution can be approximated by

$$\begin{aligned} f(z) &= (1+z)^{3+3\omega} \\ &\approx (1+z)^{3+3\omega_0} \end{aligned} \tag{5.26}$$

where ω_0 is taken to be the value today. This approximation is made assuming that the equation of state parameter of dark energy, ω_X, is not contstant with time, bur rather $\omega_X = \omega_0 + \omega_a \frac{z}{1+z}$. As the universe is assumed to be flat, and only matter and dark energy is assumed to contribute, I have approximated

$$\Omega_{DE} \approx 1 - \Omega_m \tag{5.27}$$

Hence I have used

$$E(z) \equiv \frac{H(z)}{H_0} = \sqrt{(1+z)^3\Omega_m + (1+z)^{3+3\omega_0}(1-\Omega_m)}. \tag{5.28}$$

The BAO parameter $\mathcal{A}(z)$ will be used in chapter 8 when addressing the effect of different neutrino mass hierarchies on cosmology.

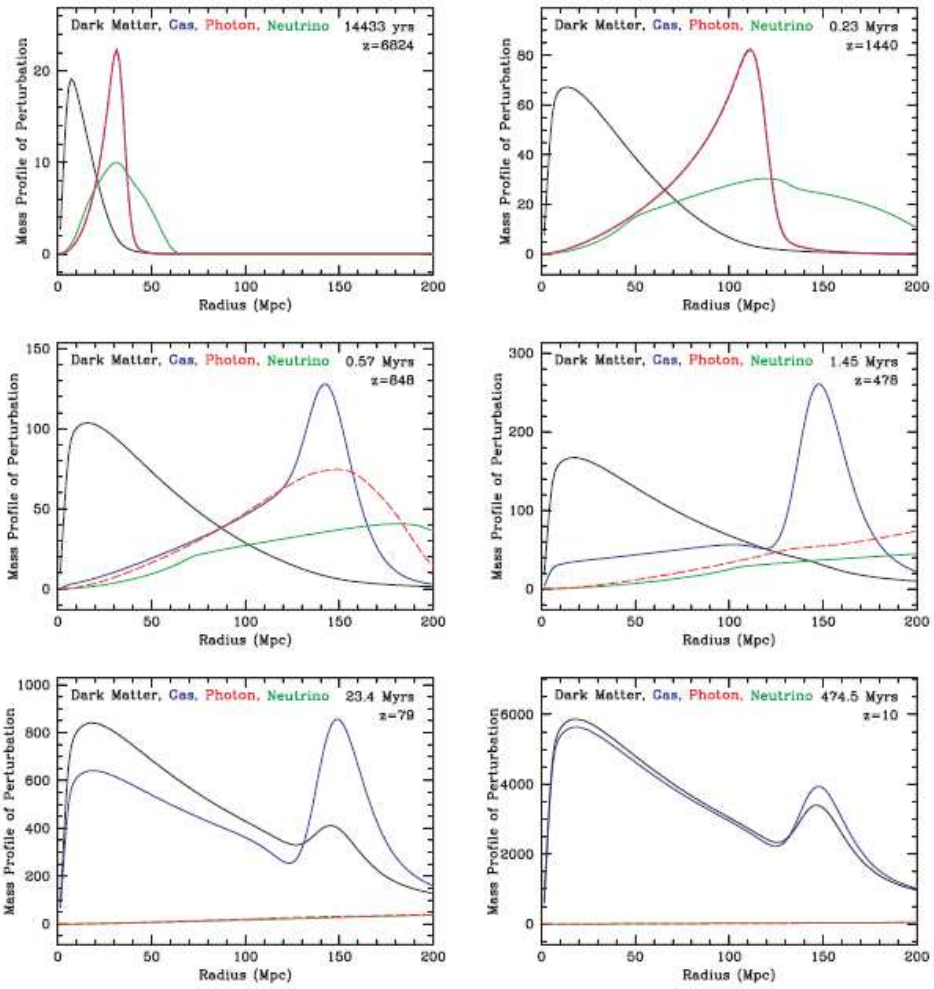


Figure 5.2: Snapshots of the evolution of the radial mass profile vs. comoving radius of an initially pointlike overdensity located at the origin. One can clearly see that the interacting matter and the radiation travels outward from the centre until recombination, when the photons "escape" from the baryonic matter at the time of the recombination, and the baryons are allowed to fall into the gravitational potential well created by the dark matter close to the centre. The scale of the baryon peak at the time of the photon decoupling also affects the dark matter, and approaching the present time of the universe, the baryons and dark matter share the same mass profile, as the baryonic matter catches up with the dark matter due to the potentials created by the dark matter. This figure is from [38].

5.3 Adding Neutrinos to the ΛCDM Model

Adding neutrinos to the ΛCDM model changes the process of structure formation. Knowing that the neutrinos are infact massive, it is tempting to think that neutrinos could make up all the dark matter, and thus demystify the dark matter as neutrinos are well known particles. The problem, if one could say so, is that the neutrinos are not cold. They thus need a special treatment in structure formation. Relativistic neutrinos also changes the time of radiation-matter equality. The main references of this subsection is [41, 42, 36, 34, 1]

5.3.1 Neutrino background

In the standard big bang model of the universe, there exists a sea of relic neutrinos, almost as numerous as the photons. They were kept in equilibrium with the other components of the universe early on, but as they only interact weakly and weak interactions are rather slow, they decoupled before the photon decoupling, when the weak interaction rate drop below the expansion rate of the universe.

After the neutrino decoupling, the relativistic neutrinos keep a temperature $T \propto a^{-1}$. Also the rest of the energy components cools down in the same way as radiation still dominates, leaving the neutrinos with the same temperature as, although decoupled from, the other energy components. When the reaction

$$e^+e^- \leftrightarrow \gamma\gamma \tag{5.29}$$

falls out of equilibrium, at $kT \approx m_e c^2$, free electrons and positrons are eradicated, and the number of photons increase. This leads to an increase of photon temperature, as the effective number of relativistic degrees of freedom decreases. For entropy to not decrease, the temperature must raise. If the entropy remains constant through this process, which is assumed, then one find that the temperature of the photons raise by a factor of $(11/4)^{1/3}$. Assuming that he neutrinos don't take part in this temperature raise because they are decoupled from electrons and photons at this time gives a temperature relation between the photon and neutrino temperature after the annihilation of electron/positrons given by

$$T_\nu = \left(\frac{4}{11}\right)^{1/3} T_\gamma. \tag{5.30}$$

It is not entirely true that the neutrinos are not affected by the electron/positron annihilation, and this gives a slight correction when calculating both the neutrino temperature and other quantities. This is caused by the short time between the neutrino decoupling and the electron/positron annihilation, meaning that there were still some neutrinos left to interact with the charged fermions. This caused a slightly smaller increase of the photon temperature than calculated above, and the neutrinos would gain some energy. Also, the big bang nucleosynthesis is affected by this [43, 44]. Using $N_{eff} = 3.046$ [45] for the effective neutrino number in the

colsmological Boltzmann codes and when calculating the energy density corrects for this. I will keep to equation (5.30) in this section for calculating the parameters which depend on the neutrino temperature.

Being relativistic, the photon temperature drops as $T \propto a^{-1}$. The neutrinos has not been relativistic ever since they decoupled, but their temperature can be assumed to develop the same way, assuming that the number density in a comoving volume is constant. Thus, after the electron/positron annihilation, the photon and neutrino temperature develops in the same way as the photon temperature, and we can thus calculate the temperature of the relic neutrinos today from the present CMB temperature.

The number density of each neutrino species can be calculated from the distribution function, using the approximation in equation (5.30), giving

$$n_{\nu_i} = \frac{6}{11}\frac{\zeta(3)}{\pi^2}T_\gamma^3, \tag{5.31}$$

where I have used that the number of degrees of freedom are two for massive neutrinos, independently of Dirac or Majorana behaviour [36]. Using $T_\gamma = 2.725K$ [46] today and converting to physical units, we find

$$n_{\nu_i} \approx 112cm^{-3}. \tag{5.32}$$

The relic neutrino energy density can then be calculated, and since the neutrinos are non-relativistic today (this is an assumption, but true for at least two of the mass states), we use the relation [43]

$$\Omega_{\nu_i} = \frac{n_{\nu_i} m_{\nu_i}}{\rho_c} = \frac{m_{\nu_i}}{93.14eVh^2}, \tag{5.33}$$

and thus

$$\Omega_\nu = \frac{\sum_i m_{\nu_i}}{93.14eVh^2} = \frac{M_\nu}{93.14eVh^2}. \tag{5.34}$$

If the third neutrino mass state is still relativistic, it will give a small, but negligible, correction to the equation above.

This can be used to make a rather careless limit on the total neutrino-mass. Since observations support a close to flat universe, and $h \approx 0.7$, then

$$\sum_i m_{\nu_i} = M_\nu \leq 93\Omega_\nu eVh^2 = 46\Omega_\nu eV. \tag{5.35}$$

This estimate can be improved by noting that about 30% of the energy density in the universe comes from matter (baryons, chareged leptons, dark matter and non-relativistic neutrinos). The rest is dark energy (radiation does not contribute a great deal today). Thus, there are a maximum 30% content of neutrinos . This improves the limit to

$$M_\nu \leq 0.3 \cdot 46eV = 14eV. \tag{5.36}$$

But can neutrinos really solve the dark matter problem? No, we will see in the next section that neutrino free streaming prohibits neutrinos to make up all the dark matter.

5.3.2 Hot dark matter and free streaming

Neutrinos are *hot dark matter*, and by hot it is meant that neutrinos were non-relativistic at the time of decoupling from the thermal background. This causes the neutrino to *free stream*, and thus to suppress structure formation on small scales. And the scale of which density contrasts are erased is mass dependent.

What does free streaming mean? The simplified answer is that massive neutrinos does contribute to the mass density in the universe, and thus contribute to the gravitational collapse discussed in the section above. A spherical mass distribution has a collapse-time of $t_{collapse} \propto 1/\sqrt{G\rho}$. If the radius of the sphere is small and the neutrinos move fast, then all the neutrinos might escaped the spherical overdensity before it collapses. And if the neutrinos makes up most of the original matter density within the sphere, the sphere might no more contain an overdensity. This suppresses the growth of structure which has a radius that is smaller than the distance travelled by the neutrinos in the collapse time.

The velocity of the neutrinos are mass dependent, and thus the scale in which mass clustering is suppressed by free streaming depends on the neutrino mass as well, which gives us a hint as to why neutrino mass limits can be made from cosmological observations. The free streaming scale can be found by simply substituting the speed of sound (which is the speed of travel of the pressure gradient) with the speed of the neutrinos in the definition of the Jeans length in equation (5.14):

$$\lambda_{FS} = 2\pi\sqrt{\frac{2}{3}}\frac{v_\nu}{H}. \tag{5.37}$$

The characteristic speed of the neutrinos is their thermal velocity. When ultra-relativistic, the speed of the neutrinos basically equals the speed of light, and the free streaming scale equals the Hubble horizon, which is an increasing function of time during radiation and matter domination. When they become non-relativistic, the neutrinos thermal velocity is calculated from $|p| = m_\nu v_{th}$, thus [36]

$$v_\nu \approx v_{th} = \frac{|p|}{m_\nu} = \frac{3k_BT}{m_\nu} = \frac{3k_BT_\nu^0 a_0}{a(t)m_\nu} = \frac{1.51\times10^5}{a(t)}\left(\frac{eV}{m_\nu}\right)m/s, \tag{5.38}$$

where the assumption has been made that $|p| \approx E = 3k_BT$. Thus, if the mass scale of the neutrinos are ~ 0.1 eV, then the typical thermal velocity today would be $\nu_{th,\nu} \sim 10^6$m/s.

Using the Friedman equations for a flat universe with negligible photon contribution, the free streaming scale evolves like

$$\begin{aligned}\lambda_{FS} &= 2\pi\sqrt{\frac{2}{3}}\frac{v_\nu}{H_o\sqrt{\Omega_\Lambda + \Omega_m a^{-3}(t)}} \\ &= 2\pi\sqrt{\frac{2}{3}}\frac{7.75}{a(t)h\sqrt{\Omega_\Lambda + \Omega_m a^{-3}(t)}}\left(\frac{eV}{m_\nu}\right)Mpc,\end{aligned} \tag{5.39}$$

where Ω_m and Ω_Λ are the density parameters today.

From equations (5.37) and (5.38) we see that $\lambda_{FS} \propto 1/aH$. During the matter dominated epoch, $1/aH \propto t^{1/3}$, meaning that the physical free streaming scale is increasing. But the scale factor grows faster than the free streaming scale, $a \propto t^{2/3}$, and hence the comoving free streaming scale decreases like

$$\lambda_{FS}^{C} = \frac{\lambda_{FS}}{a} \sim \frac{1}{a^2 H^2} \sim t^{-1/3}. \tag{5.40}$$

Hence the comoving free streaming scale of non-relativistic neutrinos decreases during matter domination, causing the comoving wavenumber at which free-streaming is occurring to decrease. Thus, if neutrinos become non-relativistic during matter domination, the comoving free-streaming wavelength λ_{FS}^{C} passes through a maximum at the time of the non-relativistic transition, and the comoving wavenumber passes through a minimum. The minimal value for the comoving wavenumber is given by [42]

$$k_{nr} = 0.0026 \left(\frac{m_\nu}{eV}\right)^{1/2} \sqrt{\Omega_m} h Mpc^{-1} \tag{5.41}$$

in the case of tree equal neutrino-masses. The Fourier modes with $k > k_{nr}$ will be affected by free streaming, while the smaller modes will not.

After the time of transition of at least the heaviest neutrinos, the matter power spectrum looks like

$$\begin{aligned} P(k) &= \left\langle \left(\frac{\delta\rho_{cdm} + \delta\rho_b + \delta_\nu}{\rho_{cdm} + \rho_b + \rho_\nu} \right)^2 \right\rangle \\ &= \left\langle \left(\frac{\Omega_{cdm}\delta_{cdm} + \Omega_b\delta_b + \Omega_\nu\delta_\nu}{\Omega_{cdm} + \Omega_b + \Omega_\nu} \right)^2 \right\rangle . \end{aligned} \tag{5.42}$$

It can be shown [36] that for at the end of matter domination and all the way through Λ domination it can be expected that $\delta_\nu = \delta_{cdm} = \delta_b$ for $a_0 H_0 < k < k_{nr}$, and for $k > k_{nr}$ the relation is that $\delta_\nu/\delta_{cdm} = \delta_\nu/\delta_b < 1$. Thus the matter power spectrum today can be express012ed as

$$P(k) = \begin{cases} \left\langle \delta_{cdm}^2 \right\rangle & k < k_{nr} \\ (1 - \Omega_\nu/\Omega_m)^2 \left\langle \delta_{cdm}^2 \right\rangle & k \gg k_{nr} \end{cases} \tag{5.43}$$

if one do not take into consideration that the neutrinos affects the metric, and thus the evolution of the power spectrum is even greater than shown in the formulae above.

The presence of neutrinos, which does not contribute to gravitational clustering on small scales but which do contribute to Hubble expansion, causes the metric perturbations Ψ and Φ to decay. This causes a slower growth of the matter perturbation contrast δ_m than if neutrinos were not present. On large scales, in the matter dominate era, the matter perturbations still grows proportional to a, but on scales with $k > k_{nr}$, the density contrast only grows as [36] $\delta_{cdm} = \delta_b \simeq a^{1-\frac{3}{5}\frac{\Omega_\nu}{\Omega_{cdm}}}$, in the limit $\Omega_\nu/\Omega_{cdm} \ll 1$, where the effects of a cosmological constant also is

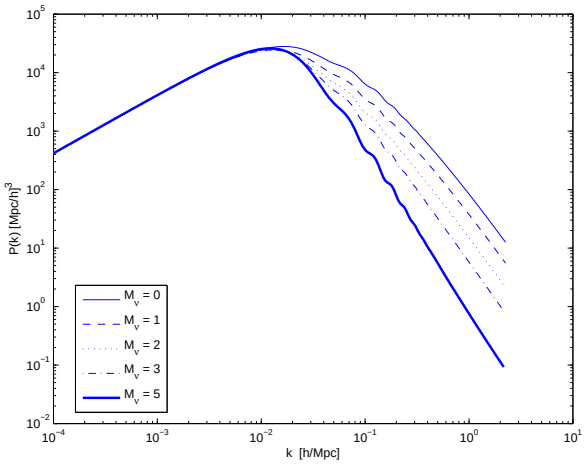


Figure 5.3: The matter power specter P(k) for different total neutrinomasses M_ν. A flat universe and a constant baryon density parameter has been assumed, and other parameters has been set to typical values in a ΛCDM model. When the neutrino density contribution is increased, the cold dark matter contribution has been reduced accordingly. On large scales (small values of k), the matter power spectrum is noticeably affected by the neutrino mass. On smaller scales (larger k), the power spectrum is suppressed by large neutrino masses. This behaviour is caused by the neutrino free streaming. On large scales, the neutrinos contribute to the matter poser spectrum in the same way as dark cold matter. On small scales, the neutrino velocity cannot be neglected, causing free streaming. This figure is produces using CAMB.

taken care of. The neutrinos also affect the time of matter-radiation equality, which affects the scale of the turn over in the matter power spectrum. Assuming that the neutrinos are relativistic at the time of equality, they contribute to the radiation and thus an increased neutrino energy density postpones the time of equality relative to in a model where all the dark matter is cold. A later time of equality means that the matter fluctuations will grow less, as they develop most during the matter dominated era.

It can be shown that when adding these extra effects, the relation between powerspectra with and without massive neutrinos is [42]

$$\frac{P(k)|_{(\Omega_\nu/\Omega_{cdm})} - P(k)|_{(\Omega_\nu/\Omega_{cdm}=0)}}{P(k)|_{(\Omega_\nu/\Omega_{cdm}=0)}} \approx -8\frac{\Omega_\nu}{\Omega_{cdm}} \tag{5.44}$$

for $k \gg k_{nr}$ and small $\frac{\Omega_\nu}{\Omega_{cdm}}$ [36].

The effect of massive neutrinos on the matter power spectrum is shown in figure 5.3 for different neutrino masses.

5.4 The Fluctuation Amplitude on Cluster Scales

In this subsection, I will introduce the σ_8 parameter; a measure on the mass density variance inside a sphere with a radius of $8h^{-1}Mpc$. This parameter is sensitive to the neutrino mass scale, and is thus potentially useful for determining the neutrino mass hierarchy. This section is based on references [47, 48, 29, 24, 35].

The rms overdensity in a sphere of radius R is theroreticaly given by

$$\sigma_R^2 \equiv \left\langle \delta_R^2(x) \right\rangle = \frac{1}{2\pi^2} \int_0^\infty W^2(kR) P(k) k^2 dk. \tag{5.45}$$

The function $W(kR)$ is a window function, given by

$$W(k) = 3\frac{\sin k - k\cos k}{k^3} \tag{5.46}$$

$W(kR)$ is the Fourier transform of $W(x/R)$ divided by V where $W(kR)$ is the top hat function defined by $W = 1$ for $x/R \leq 1$ and $W = 0$ for $x/R > 1$. V is the volume defined by the top hat function $V = \int W(x/R) dx^3$.

The top hat function makes sure that we only count the contribution from $P(k)$ for values where $x/R \leq 1$. Although we use the top hat position function transformed into Fourier space, we only count the contribution coming from the matter power spectrum with separation less than our chosen value of R.

It is common to use $R = 8h^{-1}Mpc$. Then the integral in equation (5.45) is a measure of the mass variance inside a sphere of radius $8h^{-}1Mpc$. This scale is about the size of a cluster of galaxies. The mass variance can be theoretically calculated and is directly observable (counting only the visible matter of course), and is thus a good tool to constrain cosmological models and parameters. As the power spectrum, as we have seen, is affected by the neutrino mass on small scales, the σ_8 is sensitive to mass variations, implying that we can get information about the mass scale from this parameter. In 2006, [47] found constrains on σ_8 using weak gravitational lensing observations directly. His constrain were later used by [48] to find new robust upper limits on the total neutrino mass.

For most experiments, the matter power spectrum observed, and thus also the fluctuation amplitude on cluster scales observed, is only the contribution from the visible matter. We really observe the galaxy power spectrum, not the matter power spectrum, which also contains dark matter. It is common to introduce a bias parameter b relating the observed galaxy power spectrum to the matter power spectrum, $P_g(k) = b^2 P_m(k)$. This is an approximation, as there is no apriori reason why the galaxy structure formation should reflect exactly the total mass distribution in the universe on all scales. For large scales, however, b seems to be approximated well by a constant. This approximation reduces the validity of the power spectrum used in determining cosmological parameters.

Experiments utilizing gravitational lensing, like Euclid, will in the future provide better data on the total matter power spectrum, improving the constraints given on the cosmological parameters.

Chapter 6

The Cosmic Microwave Background Radiation

The cosmic microwave background radiation was discovered in 1964 by Penzias and Wilson. It was discovered as an isotropic background radiation from all over the sky. The fact that the universe is filled with low temperature radiation, almost uniform in all directions, is predicted by the big bang theory. It is, however, not the energy of this radiation that is in the interest of most cosmologists, but the fluctuations in the temperature and the angular distribution of these. The temperature fluctuations on different scales tells us a lot about the universe, and essentially gives us a way of observing the universe at the time of the last photon scattering. In this section, I will discuss the temperature power spectrum, building on what we already know from the matter power spectrum from previous discussion. I will also take a brief look at how different cosmological parameters affects the CMB, and in particular how the neutrino mass is visible in the angular power spectrum. This section is inspired by [24, 31, 36].

6.1 The Cosmic Microwave Background Radiation

The temperature fluctuations in a direction $\mathbf{x}$ in the sky is defined as

$$\theta(\mathbf{x}, \hat{n}, \eta) = \frac{1}{T}\delta T(\mathbf{x}, \hat{n}, \eta), \tag{6.1}$$

where η is the conformal time and $\hat{n}$ is the direction of propagation of the photon. The parameter θ is essentially the photon equivalent of the matter density contrast δ. Since we are interested in how the temperature perturbations behave on different scales, we transform into Fourier space, just like we did when dealing with matter perturbations:

$$\theta(\mathbf{x}, \hat{n}, \eta) = \int \frac{d^3k}{(2\pi)^3}\theta(\mathbf{k}, \hat{n}, \eta)e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{6.2}$$

A definition that makes our notation more convenient is to introduce a new quantity, μ, given by $\mu = \hat{\mathbf{k}} \cdot \hat{n} = \cos\theta$, such that θ is the angle between the photon direction $\hat{n}$ and the direction in which the temperature is changing, $\hat{\mathbf{k}}$. Then θ can be written as a sum of Legendre polynomials [31]

$$\theta(\mathbf{k}, \mu, \eta) = \sum_{l=0}^{l=\infty} (-i)^l (2l+1) \theta_l(\mathbf{k}, \eta) P_l(\mu) \tag{6.3}$$

where P_l is the Legendre polynomial or order l, and θ_l is the lth multipole moment of of the temperature field, defined as [24]

$$\theta_l(\mathbf{k}, \eta) \equiv \frac{1}{(-i)^l} \int_{-1}^{1} \frac{d\mu}{2} P_l(\mathbf{k}, \mu, \eta) \theta(\mu). \tag{6.4}$$

The orthogonality relation for Legendre polynomials comes in handy:

$$\int_{-1}^{1} d\mu P_l(\mu) P_{l'}(\mu) = \frac{2\delta_{ll'}}{2l+1}. \tag{6.5}$$

As were the case for the matter power spectrum, we want to look at the correlation between different points at the sky. The CMB that we observe today, was all emitted at the same time, and we observe the radiation all in one place, so we look at the correlation between radiation from different angles today:

$$C(\beta) = \left\langle \theta(\mathbf{x}, \hat{n}, \eta_0), \theta(\mathbf{x}, \hat{n'}, \eta_0) \right\rangle \tag{6.6}$$

where $\beta = \hat{n} \cdot \hat{n'}$. Inserting equation (6.2) into equation (6.6), and applying equation (6.3), we get

$$\begin{aligned} C(\beta) &= \left\langle \theta(\mathbf{x}, \hat{n}, \eta_0), \theta(\mathbf{x}, \hat{n'}, \eta_0) \right\rangle \\ &= \int \frac{d^3k}{(2\pi)^3} \int \frac{d^3k'}{(2\pi)^3} \left\langle \theta(\mathbf{k}, \hat{n}, \eta_0) \theta(\mathbf{k'}, \hat{n'}, \eta_0) \right\rangle P_l(\mu) P_{l'}(\mu') e^{i(\mathbf{k}+\mathbf{k'})\cdot\mathbf{x}} \\ &= \int \frac{d^3k}{(2\pi)^3} \int \frac{d^3k'}{(2\pi)^3} \sum_{l,l'=0}^{l=\infty} (-i)^{l+l'} (2l+1)(2l'+1) \left\langle \theta_l(\mathbf{k}, \eta_0) \theta_{l'}(\mathbf{k'}, \eta_0) \right\rangle \\ &\quad \times P_l(\mu) P_{l'}(\mu') e^{i(\mathbf{k}+\mathbf{k'})\cdot\mathbf{x}}. \end{aligned} \tag{6.7}$$

Assuming that the temperature fluctuations are Gaussian, an assumption that seems to fit the physical world quite well, we expect the different θ_l modes to be orthogonal, such that

$$\left\langle \theta_l(\mathbf{k}, \eta_0) \theta_{l'}(\mathbf{k'}, \eta_0) \right\rangle = (2\pi)^3 \delta(\mathbf{k} + \mathbf{k'}) \delta_{ll'} \left\langle |\theta_l(\mathbf{k}, \eta_0)|^2 \right\rangle. \tag{6.8}$$

This is very helpful when we carry out the integral in equation (6.7), since then only $\mathbf{k'} = -\mathbf{k}$ and $l = l'$ contribute. We also assume that the universe is statsistically

isotropic, meaning that averaging over all directions, no direction is special. Thus

$$\begin{aligned} C(\beta) &\approx \int \frac{d^3k}{(2\pi)^3}\int \frac{d^3k'}{(2\pi)^3}\sum_{l,l'=0}^{l=\infty}(-i)^{l+l'}(2l+1)(2l'+1)\left\langle\theta_l(k,\eta_0)\theta_{l'}(k',\eta_0)\right\rangle \\ &\quad P_l(\mu)P_{l'}(\mu')e^{i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}} \\ &= \int \frac{d^3k}{(2\pi)^3}\int \frac{d^3k'}{(2\pi)^3}\sum_{l=0}^{l=\infty}(-i)^{2l}(2l+1)^2(2\pi)^3\delta(\mathbf{k}+\mathbf{k}')\left\langle|\theta_l(k,\eta_0)|^2\right\rangle \\ &\quad P_l(\mu)P_l(\mu')e^{i(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}} \\ &= \sum_{l=0}^{l=\infty}\int \frac{d^3k}{(2\pi)^3}(2l+1)^2\left\langle|\theta_l(k,\eta_0)|^2\right\rangle P_l(\hat{\mathbf{k}}\cdot\hat{n})P_l(\hat{\mathbf{k}}\cdot\hat{n'}). \end{aligned} \tag{6.9}$$

It is, however, not this quantity that we want to en up with. In the same way as we wrote $\theta(\mathbf{k},\hat{n},\eta)$ in terms of Legendre polynomials in equation (6.3), we write $C(\beta)$ in terms of a sum of Legendre polynomials:

$$C(\beta)=\frac{1}{4\pi}\sum_{l=0}^{\infty}(2l+1)C_lP_l(\cos\beta). \tag{6.10}$$

To do this, we need the identity

$$\int d\Omega_k P_l(\hat{\mathbf{k}}\cdot\hat{n})P_l(\hat{\mathbf{k}}\cdot\hat{n'})=\frac{4\pi}{2l+1}P_l(\hat{\mathbf{n}}\cdot\hat{\mathbf{n}'}), \tag{6.11}$$

where $d\Omega_k$ is the solid angle element. We then find

$$\begin{aligned}
C(\beta) &= \sum_{l=0}^{l=\infty} \int \frac{d^3k}{(2\pi)^3}(2l+1)^2 \left\langle |\theta_l(k,\eta_0)|^2 \right\rangle P_l(\hat{\mathbf{k}}\cdot\hat{n})P_l(\hat{\mathbf{k}}\cdot\hat{n'}) \\
&= \sum_{l=0}^{l=\infty} \int\int\int \frac{k^2\sin\theta dk d\phi d\theta}{(2\pi)^3}(2l+1)^2 \left\langle |\theta_l(k,\eta_0)|^2 \right\rangle P_l(\hat{\mathbf{k}}\cdot\hat{n})P_l(\hat{\mathbf{k}}\cdot\hat{n'}) \\
&= \sum_{l=0}^{l=\infty} \int \frac{k^2dk}{(2\pi)^3}(2l+1)^2 \left\langle |\theta_l(k,\eta_0)|^2 \right\rangle \int d\Omega P_l(\hat{\mathbf{k}}\cdot\hat{n})P_l(\hat{\mathbf{k}}\cdot\hat{n'}) \\
&= \sum_{l=0}^{l=\infty} \int \frac{k^2dk}{(2\pi)^3}(2l+1)^2 \left\langle |\theta_l(k,\eta_0)|^2 \right\rangle \frac{4\pi}{2l+1} P_l(\hat{\mathbf{n}}\cdot\hat{\mathbf{n}'}) \\
&= \sum_{l=0}^{l=\infty} \int \frac{k^2dk}{(2\pi)^3}(2l+1)^2 \left\langle |\theta_l(k,\eta_0)|^2 \right\rangle \frac{4\pi}{2l+1} P_l(\cos\beta) \\
&= \sum_{l=0}^{l=\infty} \int \frac{4\pi k^2dk}{(2\pi)^3}(2l+1) \left\langle |\theta_l(k,\eta_0)|^2 \right\rangle P_l(\cos\beta) \\
&= \sum_{l=0}^{l=\infty} \int \frac{d^3k}{(2\pi)^3}(2l+1) \left\langle |\theta_l(k,\eta_0)|^2 \right\rangle P_l(\cos\beta).
\end{aligned} \tag{6.12}$$

We then see, from equation (6.10), that

$$C_l = 4\pi \int \frac{d^3k}{(2\pi)^3} \left\langle |\theta_l(k,\eta_0)|^2 \right\rangle . \tag{6.13}$$

When C_l is defined in this way, it is always positive. A large C_l implies large temperature fluctuations on scale l. The l indicates the multipole moment, if $l = 0$, it refers to the monopole, a uniform background, and $l = 1$ is the dipole. If C_l is large for a given value of l, it means that there are large fluctuations in the temperature on scales associated with the "l-pole", which really is a word that I just invented. If C_l is large for large values of l, it indicates that the temperature fluctuations are large on small scales, an a large value of C_l for small values of l indicates large fluctuations on large scales.

Another way of understanding the quantity C_l is found by expanding the temperature fluctuations in terms of spherical harmonics:

$$\theta(\mathbf{x},\hat{\mathbf{n}},\eta) = \sum_{l=1}^{\infty}\sum_{m=-l}^{l} a_{lm}(\mathbf{x},\eta)Y_{lm}(\hat{\mathbf{n}}). \tag{6.14}$$

Since the spherical harmonics are orthogonal in l and m, the a_{lm} (related to the temperature fluctuations in Fourier space), can be expressed as [24]

$$a_{lm}(\mathbf{x},\eta) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \int d\Omega Y^*_{lm}(\hat{\mathbf{n}})\theta(\mathbf{k},\hat{\mathbf{n}},\eta). \tag{6.15}$$

This quantity is dependent on direction and position. We cannot predict individual a_{lm}s by theory, but we can predict the distribution from which it is drawn, thus we are interested in the variance of the a_{lm}s. It can then be shown [24] that C_l is the variance of the a_{lm}s is:

$$\langle a_{lm} a^*_{l'm'} \rangle = \delta_{ll'} \delta_{mm'} C_l. \tag{6.16}$$

This again gives the same interpretation as before; as a_{lm} tells us something about the amount of variation on a specific scale and in a specific direction, the variance of the a_{lm} is averaged over space and tells us something about the temperature fluctuations on a scale l. C_l is just a statistical property, it relates to the fluctuations on a scale, but tells us nothing about where the fluctuations are. Thus C_l only describes the background radiation well in an isotropic universe. If anisotropies comes into the picture, more information is needed to describe the CMB in a good way.

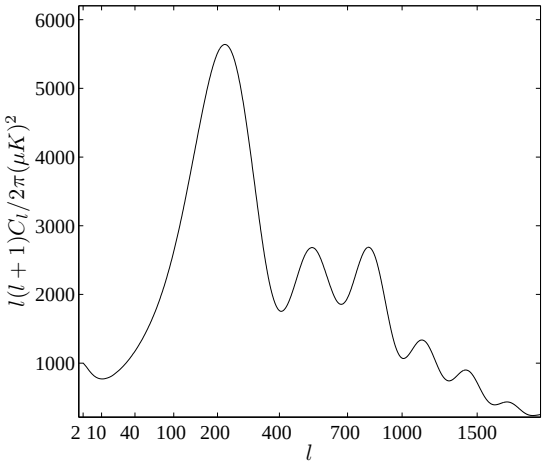


Figure 6.1: A CMB power spectrum for typical parameter values, with $M_\nu = 0$. The plot is made using CAMB.

On large scales, in a universe with scale factor $n_s = 1$, it can be shown that [24] that the quantity $l(l+1)C_l$ is constant:

$$l(l+1)C_l = \frac{\pi}{2} \left(\frac{\Omega_m}{D_1(a=1)} \right)^2 \delta_H^2. \tag{6.17}$$

Thus C_l is proportional to $\frac{1}{l(l+1)}$ for small values of l. For this reason, one typically plot $l(l+1)C_l$ when representing the angular temperature fluctuations in the CMB. A typical theoretical prediction of the angular power spectrum is displayed in figure 6.1.

6.2 The Effect of Different Cosmological Parameters

I will now provide a very quick explanation of why the CMB angular power spectrum looks the way it does. The peaks in the power spectrum are caused by the same physics as baryon acoustic oscillations. As the radiation and baryons were strongly coupled to each other until photon decoupling, which happened at approximately the same time as the photon last scattering, the photons oscillated in pace with the baryonic matter. The photons thus had basically the same distribution as the baryons at the time of photon decoupling on small scales. After recombination, the photons could not move totally independent of the baryonic matter, as the baryonic matter affected the gravitational potential at the time. The photons seen today has travelled out of the potentials set up during the baryon acoustic oscillations. This affects the energy of the photons.

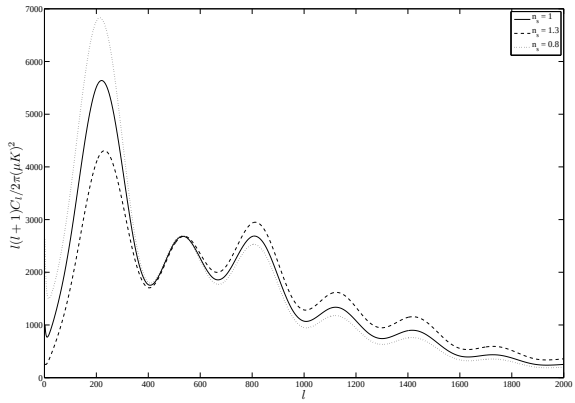


Figure 6.2: A CMB power spectrum for typical parameter values, for different values of the spectral index n_s.

We saw earlier that the Jeans scale, which controls the acoustic oscillations, grows with time. Thus larger and larger scales starts to oscillate. This means that different scales have had the time to complete a different number of oscillations before recombination. As a mode enters the horizon, it starts to grow, and as it enters the Jeans horizon, it starts to oscillate. The first peak in the CMB spectrum, indicates the mode that has reached its first peak at the time of recombination. The second peak is at the mode that has reached the first decompression at the time of recombination.

The flat part for small l relates to the Sachs-Wolfe effect, predicting a constant C_l for small l in case of $n_s = 1$. If n_s deviates from 1, then the power spectrum will be tilted. Thus n_s is often referred to as the tilt parameter. The effect is shown in figure 6.2.

The baryon density also highly affects the CMB angular power spectrum. Adding

more baryons decreases the sound speed of the universe. This decreases the Jeans scale, increasing the distance between the peaks, but more importantly, it also affects the height of the acoustic peaks. The odd peaks corresponds to overdensities. If the total baryon density in the universe is increased, then more baryons takes part in the oscillation. Increased baryon density in the overdensity increases the gravitational potential, and an increased gravitational potential makes the compressions stronger than the decompressions. Thus the odd peaks grows more than the even peaks when increasing the baryon density. The effect is shown in figure 6.3.

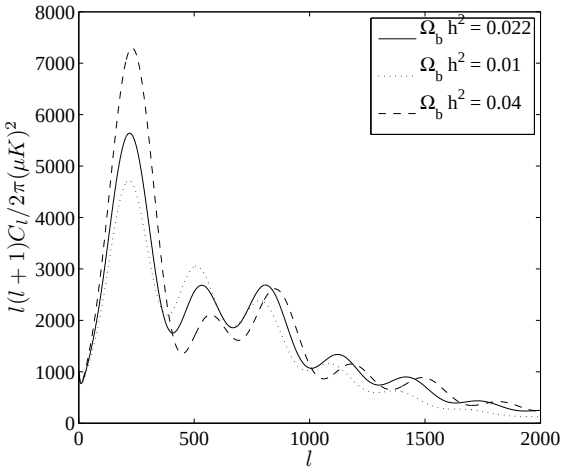


Figure 6.3: A CMB power spectrum for typical parameter values, for different amount of baryons, $\Omega_b h^2$.

6.3 Massive Neutrinos and the CMB

In this section, I will look at some of the effects massive neutrinos have on the CMB angular power spectrum. The CMB power spectrum is often used in combination with other observations (in particular observations of large scale structure) to place limits on the total neutrino mass M_ν. One can also, however, place upper limits on the total neutrino mass from CMB alone, and in 2005 Ichikawa et. al. [14] found $M_\nu < 2$ from the Wilkinson Mikrowave Anisotropy Probe observations for a flat ΛCDM model. This section is based on the references [14, 36, 31].

6.3.1 Effects of massive neutrinos on the CMB

Neutrinos affects the acoustic peaks in the CMB power spectrum. The way they influence the power spectrum depends on their mass, and in particular the time when they became non-relativistic. There are two main effects from the neutrinos on the CMB angular power spectrum; a increase in the height of the acoustic peaks, and a shift in the position of the peaks.

The height of the peaks

If the neutrinos were still relativistic (and hence counting as radiation) at the time of matter-radiation equality, keeping $\Omega_{dm} = \Omega_{cdm} + \Omega_\nu$ constant, this implies a later time of equality than if neutrinos were non-relativistic well before the time of equality. The change of time of equality is the main effect on the CMB power spectrum of massive neutrinos. A later equality decreases the small scale perturbations in the matter power spectrum due a prolonged radiation era where perturbations decrease. The effect on the CMB power spectrum is in fact opposite, due to the early integrates Sachs-Wolfe effect.

The free streaming of massive neutrinos causes a decay in the gravitational potential, as it smooths out the gravitational wells and suppresses perturbation growth on scales smaller than the free streaming scale. This causes larger temperature fluctuations inside the neutrino free streaming scale [14, 36]. The relation between the multipole corresponding to the free streaming scale l_{nr}, the conformal time of the non-relativistic transition $\eta_{nr} = \eta(n_{nr})$ and the comoving angular diameter distance to the surface of recombination $r_\theta(\eta_{rec})$ is given by

$$l_{nr} \simeq \frac{2\pi r_\theta(\eta_{rec})}{\eta_{nr}}. \tag{6.18}$$

Thus scales corresponding to multipoles with $l > l_{nr}$ is affected by the streaming. The neutrinos became non-relativistic when their momentum was comparable to their mass, $p_\nu \sim m_\nu$, meaning $3k_B T_{\nu,nr} = m_\nu$. The corresponding redshift is

found from

$$\begin{aligned} 1 + z_{nr} &= \frac{a_{nr}}{a_0} \\ &= \frac{T_{\nu,nr}}{T_{\nu_0}} \\ &= \frac{\frac{m_\nu}{3k_b}}{\left(\frac{4}{11}\right)^{1/3} * T_{CMB} * k_b} \\ &= \frac{m_\nu}{\left(\frac{4}{11}\right)^{1/3} * 3 * 2.728K * 1.38 \times 10^{-23} J/K} \\ &= 1.99 * 10^4 \left(\frac{m_\nu}{eV}\right). \end{aligned} \tag{6.19}$$

The neutrinos becoming non-relativistic at the time of recombination, ie. $z = 1088$, has a mass of $m_\nu = 0.55$eV, i.e. indicating degenerate masses. The time of recombination is insensitive to the neutrino mass, as neutrinos decoupled long before this time. If $m_\nu < 0.55$, then they are relativistic at recombination.

According to [14], this tells us that only the multipole amplitudes with $l > 300$ is affected by the free streaming with enhanced temperature fluctuations, given $m_\nu = 0.55$ eV. For smaller neutrino masses, smaller values of l are affected by an increased height of the peaks is lower.

The position of the acoustic peaks

Massive neutrinos also shifts the CMB power spectrum horizontally. This is due to the effect of the neutrinos on the cosmic expansion, and the change in the time of radiation matter equality.

The acoustic scale of the CMB power spectrum is given by

$$l_A = \pi \frac{r_\theta(\eta_{rec})}{r_s(\eta_{rec})} \tag{6.20}$$

where $r_s(\eta_{rec})$ is the comoving sound horizon at recombination, and $r_\theta(\eta_{rec})$ is the angular diameter distance to the last scattering surface. The sound horizon is defined by

$$r_s(a) = \int_0^{\eta(a)} c_s d\eta = \int_0^a c_s(a') \frac{da'}{a'^2 H}. \tag{6.21}$$

The Hubble parameter is affected by the neutrino mass, and thus the neutrino mass is affecting the acoustic scale. As seen before in equation (5.5) the sound speed is not affected by the neutrino mass, as it only depends on the ration of baryon density to photon density. It is shown in [31] that the acoustic scale decays with an increasing $\Omega_\nu h^2$.

The position of the mth peak of the power spectrum, l_m, can be parametrized by

$$l_m = l_A(m - \phi_m) \tag{6.22}$$

where ϕ_m is a phase factor arising from the integrated Sachs Wolfe effect, due to the decay of gravitational potential before matter domination. The Sachs Wolfe effect is the red shift effect of photons travelling out of gravitational potential wells. As the gravitational potential wells are caused by density perturbations, and density perturbations are affected by neutrinos due to free streaming, ϕ_m is a measure of neutrino mass. If some neutrinos were relativistic and some non-relativistic at the time, the relativistic neutrinos count as radiation and the non-relativistic as matter when calculating energy densities.

The phase factor ϕ_m can be approximated by the fitting formula [49]

$$\phi_m \approx b_m \left(\frac{\rho_{r,rec}/\rho_{m,rec}}{0.3} \right)^{0.1}, \tag{6.23}$$

showing that the ratio of radiation to matter density at recombination is important, and hence the neutrino mass is important. The fitting formula above is calculated for massless neutrinos, but as the effect of massive neutrinos is mainly changing the ratio of radiation to matter energy density, it is still a useful approximation. By defining

$$\frac{\rho_r}{\rho_m} = \frac{\rho_\gamma + \rho_{\nu,r}}{\rho_{CDM} + \rho_b + \rho_{\nu,m}}, \tag{6.24}$$

where $\rho_{\nu,r}$ is the neutrinos' relativistic contribution to the energy density and $\rho_{\nu,m}$ is the non-relativistic contribution, and applying $b_1 = 0.267$ [49], one can then calculate the position of the first peak as a function of the neutrino mass. It is shown in [31] that the phase factor ϕ is close to constant for small neutrino masses, and then is a decreasing function of $\Omega_\nu h^2$. It is shown in [14] that the effect of the neutrino mass on the phase factor ϕ_1 is very small for neutrinos that were non-relativistic at recombination, ie. $m_\nu < 0.55$ eV, and that the effect is significant if neutrinos became non-relativistic before recombination, $m_\nu > 0.55$ eV. The nett effect of massive neutrinos on the shift of the first peak is [14, 31] to move the first peak to the left.

6.3.2 Using CMB to obtain limits on M_ν

The net effect of massive neutrinos on the CMB power spectrum is, as we have seen, to increase the heights of the acoustic peaks, and to shift the peak positions. The increased heights of the peaks is due to postponed radiation matter equality and due to the early integrated Sachs-Wolfe effect. The shift in the powerspectrum to the left is mainly due to the changed angular scale caused by the altered expansion properties of the universe, and it is also affected by the altered Sachs-Wolfe effect due to the neutrino contribution.

It is, however, hard to use the CMB matter power spectrum alone to get good limits on the neutrino mass for small neutrino masses. The reason for this is that the CMB power spectrum is mostly sensitive to neutrinos that became relativistic before recombination, $M_\nu > 3 \times 0.55$ eV $= 1.65$ eV, which we shall see in the next

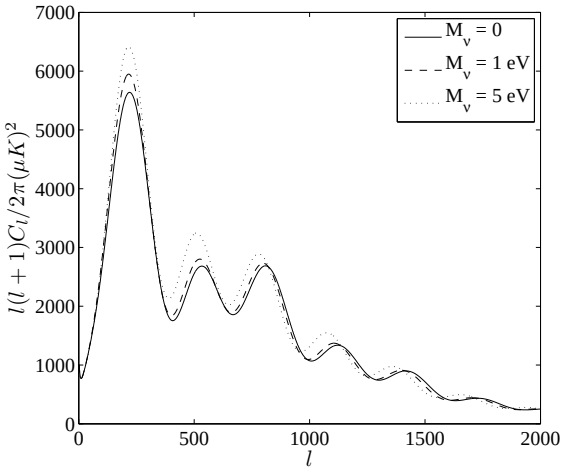


Figure 6.4: A CMB power spectrum for typical parameter values, for different total neutrino masses, M_ν. In order to keep the curvature constant, Ω_{CDM} has been altered to compensate for the change in Ω_ν.

chapter indicates degenerate neutrino masses. We have already seen that the shift in the acoustic peaks is small if the neutrinos were relativistic at decoupling. It also turns out that although the position and height of the first peak is sensitive to small neutrino masses, the height of the second and third peak are mostly sensitive to neutrinos that were non-relativistic at recombination. Thus, having only two good observables to determine the neutrino mass, there is too much degeneracies with other parameters to provide sufficiently good limits.

In figure 6.4, the CMB angular power spectrum is plotted for $m_\nu = 0$ eV, $M_\nu = 1$ eV and $M_\nu = 5$ eV. For $M_\nu = 5$ eV we can see that the height of all the peaks is increased, and that the spectrum is shifted towards smaller values of l. For $M_\nu = 1$ eV, only the first peak is increased notably, and the shift is very small. [14] places an upper limit on the ability of CMB alone to give good neutrino mass limits at $M_\nu \lesssim 1.5$ eV. In combination with other cosmological observations, or priors from experiments in laboratories on earth, however, CMB is a powerful tool to use when finding neutrino mass limits. It provides some constraints on the neutrino mass, and it is very useful for determining other cosmological parameters that need to be determined in order to get good limits on M_ν.

Part III

Results

Chapter 7

Method

In this section, I will discuss the methods used in my thesis. I will start by reviewing the background for the Markov Chain Monte Carlo approach used by the Fortran code CosmoMC, which I have used in order to estimate cosmological parameters. This code utilizes the likelihood function, and I will discuss the likelihood function and likelihood analysis in section 7.1, and then go on to discussing the CosmoMC code in section 7.2.

I have made my own hierarchy dependent prior on the total neutrino mass, using the best values of the neutrino parameters today. In section 7.3 I will discuss how this prior was made.

7.1 The Likelihood Function

In this thesis, I have made use of the CosmoMC code [50], which applies a Markov chain Monte Caro (MCMC) method on cosmological parameters and observations which uses multiple runs of the Boltzmann code CAMB [33] to estimate cosmological parameters. Before I review some of the properties of the Markov Chain Monte Carlo method, I will look at the likelihood function which on which the statistical method is based. This section is mainly based on the references [24, 51].

The likelihood function is the main building block of the analysis of cosmological parameters. It is defined as *"the probability that a given experiment would get the data it did get given a theory"* [24], or the probability of a data set given a chosen model. This is a very powerful tool for parameter estimation, as we tend to have the observations and want to find the theoretical model which best describes the data. We will see in this section that if the likelihood of one model is higher than the likelihood of another model, the latter gives a better theoretical description of what we can observe. The likelihood function is used to determine the best fit cosmological parameters of the theory; the parameters that together gives the largest likelihood within the given model.

The goal of likelihood analysis is to take us from a set of observable quantities $d = \{d_i\}$ to a set of unobservable cosmological parameters $\theta = \{\theta_j\}$. The

likelihood function of the defined situation is given by

$$P(\theta|d) = \frac{P(d|\theta)P(\theta)}{P(d)} = \frac{\mathcal{L}(\theta)P(\theta)}{P(d)}. \tag{7.1}$$

(which really is Bayes theorem) where $P(\theta)$ contains information about priors on the theoretical model (this is called the prior) and $P(d)$ gives information about the prior on the data set (called the evidence). $P(\theta|d)$ is the probability of the data given the theoretical model, which is called posterior. $P(d|\theta)$ is the probability of the data given the theory, and is called the likelihood.

As $P(\theta|d)$ is a probability, it should integrate to 1. Hence

$$P(d) = \int P(d|\theta)P(\theta)d\theta. \tag{7.2}$$

This is called the *evidence*, and is independent of the theory θ. When working with the likelihood, we have one set of data d and we want to find the peak of the likelihood when varying the parameters θ, and for such purposes the evidence is not important, as it neither affects the peak nor the width of $P(\theta|d)$. Thus

$$P(\theta|d) \propto \mathcal{L}(\theta)P(\theta). \tag{7.3}$$

If the prior is flat, then the posterior is proportional to the likelihood, hence maximizing the posterior and find the most probable parameters given the data that we have is equivalent to maximizing the likelihood.

In most cases, and in cosmology in particular, we deal with more than one parameter. For N parameters, to find the probability distribution of each parameter individually, one have to marginalize over the other parameters:

$$P(\theta_i|d) = \int \dots \int P(\theta|d)d\theta_1 \dots d\theta_{i-1}d\theta_{i+1} \dots d\theta_N. \tag{7.4}$$

This integral gives the one dimensional probability distribution. It is also often useful to see how two parameters depends on each other, to make contour-plots. To find this relation, one have to marginalize over all the parameters but the two parameters of interest.

The normal distribution with mean μ and variance σ is given by

$$p(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{7.5}$$

We normally refer to $\frac{(x-\mu)^2}{\sigma^2}$ as χ^2. For multiple variables,

$$\chi^2 = \sum_i -\frac{(x_i - \mu_i)^2}{\sigma_i^2} = \sum_i \chi_i^2. \tag{7.6}$$

If the likelihood function is Gaussian, the relation between the likelihood and χ^2 is then

$$\chi^2 = -2 \ln \mathcal{L}. \tag{7.7}$$

Maximizing the likelihood thus means minimizing χ^2 for a Gaussian distribution.

7.2 CosmoMC

CosmoMc is a public available Fortran code, which applies a Markov Chain Monte Carlo approach to cosmology, using the Boltzmann code CAMB. CAMB produces powerspectras (both for the CMB and for matter), and CosmoMC uses the likelihood function to compare data with theoretical models.

CosmoMC explores the cosmological parameterspace using a Markov Chain Monte Carlo method (MCMC). A Markov chain is a tool for statistical modelling. The goal of a Markov Chain Monte Carlo sampler is to sample over all of parameter space in a manner that mirrors the probability distribution of the parameter space. With the MCMC method, each step in the chain is only dependent on the previous step, but not on the steps before. Thus it behaves as a random walker. The next step in the chain is based on the current position in parameter space, with a transitional probability dependent on the current position and the proposed position alone, ie. the random walker is not completely ramdom. To ensure that all of parameter space will be visited by the random walker if it gets enough time, the sampler must fulfil the requirements of detailed balance and be ergodic: the probability of going from point a to point b must equal the probability of the opposite and the sampler must be able to sample all over parameter space.

The great advantage of the MCMC approach, is that it samples data where the probability is large. It does not necessarily sample all the points in parameter space, although it is important that it can sample all the points, and this saves us a lot of time, compared to a grid-based computation of the likelihood.

7.2.1 The Metropolis-Hastings algorithm

In this subsection, I will explore the CosmoMC code, and the Metropolis-Hastings algorithm which is used in the CosmoMC code. I will not go into particularities in the Fortran code of CosmoMC, but I will present the background needed in order to understand how a program like CosmoMC can estimate cosmological parameters from observations. The relevant references used are [50, 52].

When using CosmoMC, we want to explore the probability distribution $D(\theta)$, which is the stationary asymptotic distribution. In CosmoMC, this distribution is given by the likelihood function. In general, the random walker will move from position θ_i in parameter space to another position θ_{i+1} with the transition probability $W(\theta_i, \theta_{i+1})$. CosmoMC uses the Metropolis-Hastings algorithm to generate this transition probability. This is done by introducing a proposal density $q(\theta_i, \theta_{i+1})$, which ideally is close to the stationary distribution that we want our Markov chain to converge to. The proposal density $q(\theta_i, \theta_{i+1})$ is used to propose a new point to which the random walker can move. Then, an acceptance test is performed, deciding whether or not the random walker is allowed to move. This is done by calculating the acceptance probability $p(\theta_i, \theta_{i+1})$. Thus, the transitional probability is the product of the probability of choosing point θ_{i+1}, $q(\theta_i, \theta_{i+1})$, and probability

of accepting the choice, $p(\theta_i, \theta_{i+1})$:

$$W(\theta_i, \theta_{i+1}) = q(\theta_i, \theta_{i+1}) p(\theta_i, \theta_{i+1}). \tag{7.8}$$

The acceptance probability must be less than one, and is given by

$$p(\theta_i, \theta_{i+1}) = \min\left(1, \frac{D(\theta_{i+1}) q(\theta_{i+1}, \theta_i)}{D(\theta_i) q(\theta_i, \theta_{i+1})}\right). \tag{7.9}$$

The probability distribution used when performing the acceptance test is the likelihood. The acceptance probability in equation (7.9) is a product between the ratio of likelihoods of the two parameter points and the ratio of the proposal densities of going from the one point to the other. If the two proposal densities are equal, $q(\theta_{i+1}, \theta_i) = q(\theta_i, \theta_{i+1})$, and the proposed parameter space position has a greater likelihood function than the present parameter space position, then the acceptance probability will be one, and the random walker will move to the new point. If, however, the acceptance probability is less than one, there must still be a chance that the random walker moves in order to attain detailed balance and the ergodic principle. In such cases, the acceptance probability is compared to a random number, and the new state is accepted given that some criteria between the acceptance probability and the random number is met.

The proposal density has nothing to say for the outcome of the result of the computation, but choosing a smart proposal density efficiently decreases the time spent on sampelling the probability distribution. The initial proposal density is set in the input file, and as we do have some idea to where our parameters should be in parameter space it is not too hard to suggest a reasonable starting proposal density. CosmoMC allows for a continuous update of the proposal density during the iterations, which also accounts for the correlation between different parameters.

In CosmoMC, the random walker starts in a position in parameter space θ_1 which the user decide in the input file. Then CosmoMc calls on CAMB to calculate the power spectra for the parameters θ_1. CosmoMC then compares the theoretical power spectra with observations, and calculates the likelihood for the parameters θ_1. Then CosmoMC picks a new point in parameter space, θ_2, from the parameter distribution decided by the user in the input file. CAMB finds the power spectra, and CosmoMC calculates the likelihood of the parameters θ_2. Knowing now the likelihood of both positions in parameter space, CosmoMC decides whether or not to accept the new point and move to position θ_2. If the new parameter point is accepted, CosmoMC proposes a new point θ_3, and the same procedure is repeated. If the Metropolis-Hastings algorithm rejects point θ_2, a new point, θ_{2b} has to be chosen from θ_1. The parameter points that gets accepted is stored in an output file. If the random walker gets stuck in the same point for several iterations, the position in parameter space is given a weight counting the number of times a movement was rejected.

7.2.2 Convergence statistics

CosmoMC has a burn-in time. During the first part of the MCMC cycle, the chain has not yet converged to its stationary distribution which is used for calculation the parameters, and the points are affected by the starting point. Thus the first part of the output should be omitted. One typically uses a burn-in time of order 100-1000. Also, ideally, all points in the parameter chain should be independent of the point before. This is not the case, as every point to a certain degree depends on the point before. To rid oneself of that correlation, one can thin the chain, such that only every 4th or 1000th element in the chain are used in the final analysis. When the stationary distribution is reached, it is however not necessary to thin the chain, as although two subsequent positions in parameter space are correlated, they are both chosen from the stationary distribution.

In my work, I have used 10 MPIs, such that 10 chains are run simultaneously. When all the 10 chains have converged to approximately the same distribution, I assume that the correct stationary probability distribution is sampled. The chains are run in a parallell matter, where they communicate. They exchange information about their distributions and the chain means, and from the Gelman and Rubin "variance of chain means"/"mean of chain variances" criterion it is decided whether or not all the chains have converged. This is done by defining a parameter R = $\frac{\text{variance of chain means}}{\text{mean of chain variances}}$. When $R-1$ becomes smaller than a predetermined limit, the iterations are stopped. I have used $R-1=0.03$ as my limit.

When chain thinning is performed and burn-in subtracted, we are left with the data points to use, which are uncorrelated and sampled after the stationary distribution is reached. To get from an array of positions in parameter space and the likelihood and number of repeated events in each parameter point to the actual parameter distribution, we make histograms. The probability of a neutrino mass M_ν in an interval $M_\nu \epsilon [i,j]$ equals the number of visits by the random walker in that particular interval, divided by the total number of sampled values of neutrino masses.

7.2.3 Adding a prior

CosmoMC samples the probability distribution of the cosmological parameters. For most purposes, the "histogram" described above is used; the number of samples in a parameter point is directly proportional to the probability of that combination of parameters. But it CosmoMC also calculates the likelihood, and it gives the likelihood for each point visited as an output. As the chain converges, the number of visits in each bin echoes the likelihood in that bin. We could have explored all of parameter space by a grid calculation of the likelihood to find the probability distribution of the parameters, but the MCMC method is a lot quicker (grid calculation scales exponentially with the number of parameters, a MCMC-method does not increase the number of samples needed by much when adding an extra parameter). One can also use the likelihood of every point in parameter space, and for a

Gaussian distribution it is then simple to relate the likelihood to a mean value and a standard deviation, or a two-dimensional contour plot showing contours containing 95% or 68% of the distribution.

Adding a prior on one or more parameters can be done by multiplying the likelihood found in cosmoMC with the likelihood given by the prior

$$\mathcal{L} = \mathcal{L}_{CosmoMC} \times \mathcal{L}_{prior}. \tag{7.10}$$

CosmoMC has the $-ln(\mathcal{L})$ as an output, and GetDist (a code that makes Matlab files for plotting the likelihood distributions from the CosmoMC output), which is a code that comes along with CosmoMC takes $-ln(\mathcal{L})$ as an input. Hence, the most convenient way to add a prior after the CosmoMC run is to add a negative log likelihood function. I have done this to add a hierarchy-dependent prior on the neutrino mass.

When there are more than one parameter point that fits into the grid used to find probability distributions and contour plots, an average over the likelihoods of all the points inside the grid bin has been used.

The marginalized distribution is given by

$$\mathcal{L}(\theta_i) = \int \cdots \int \mathcal{L}(\theta) d\theta_1 \ldots d\theta_{i-1} d\theta_{i+1} \ldots d\theta_N. \tag{7.11}$$

In the limit of a small step size, this can be approximated by

$$\mathcal{L}(\theta_i) \approx \sum_{\theta_{1,max}}^{\theta_{1,max}} \cdots \sum_{\theta_{N,max}}^{\theta_{N,max}} \mathcal{L}(\theta_{1\ldots i\ldots N}) \Delta\theta_1 \ldots \Delta\theta_{i-1} \Delta\theta_{i+1} \ldots \Delta\theta_N. \tag{7.12}$$

For this to be a good approximation, the sum over θ_1 is performed with equal step size $\Delta\theta_1$. In that case, we can take the step size out of the sum, and just sum over $\mathcal{L}(\theta_{1\ldots i\ldots N})$:

$$\begin{aligned} \mathcal{L}(\theta_i) \;\approx\; & \Delta\theta_1 \ldots \Delta\theta_{i-1} \Delta\theta_{i+1} \ldots \Delta\theta_N \\ & \times \sum_{\theta_{1,max}}^{\theta_{1,max}} \cdots \sum_{\theta_{i-1,max}}^{\theta_{1-1,max}} \cdots \sum_{\theta_{i+1,max}}^{\theta_{i+1,max}} \cdots \sum_{\theta_{N,max}}^{\theta_{N,max}} \mathcal{L}(\theta_{1\ldots i\ldots N}) \end{aligned} \tag{7.13}$$

In this case, the marginalized likelihood of parameter $\mathcal{L}(\theta_i)$ is proportional to the mean of likelihoods distribution. It is, however, not so that the step size from one sample to another is the same for each parameter for every sample. Thus using the mean likelihood distribution as an approximation to the real likelihood distribution is not accurate.

When calculating the mean likelihood, I have only used the individual parameter points and ignored the fact that some points are counted multiple times due to a low acceptance probability of the next proposed point in the MCMC cycle.

Ideally, I would rather have applied the prior on the neutrino mass as an input parameter to CosmoMC, and use the number-of-visits histogram as the probability distribution. I was, hoverer, not able to manipulate the CosmoMC code to include such a prior, and hence the approximation above has been applied.

7.3 From Hierarchy and Experimental Results to Mass Distribution

From the neutrino oscillation experiments, we can extract mass square differences, Δm^2_{ij}, and constraints on the mixing angles, $\sin^2 2\theta_{ij}$. And from tritium β-decay, we find an upper limit on the effective electron neutrino mass, $m_{e,max}$. This gives us enough information to calculate a probability distribution of the total neutrino-mass. This quantity is useful when calculating the neutrino-mass from cosmological observations, as we can use it as a prior on the total neutrino mass.

The effective neutrino mass of a flavour is given by [11]:

$$m^2_{\alpha,eff} = \sum_i |U_{\alpha i}|^2 m_i^2, \tag{7.14}$$

thus the effective neutrino mass is given by (according to equation (1.40))

$$\begin{aligned} m^2_{e,eff} &= |U_{e1}|^2 m_1^2 + |U_{e2}|^2 m_2^2 + |U_{e3}|^2 m_3^2 \\ &= c_{12}^2 c_{13}^2 m_1^2 + s_{12} c_{13}^2 m_2 + s_{13}^2 m_3^2 \\ &\leq m^2_{e,max}. \end{aligned} \tag{7.15}$$

The total neutrino mass

$$M_\nu = \sum_i m_i = m_1 + m_2 + m_3 \tag{7.16}$$

can be parametrized by m_1 and Δm^2_{ij}. To find the probability distribution of the total neutrino mass, we also impose the parameters $\sin^2 2\theta_{ij}$ and the experimental limit on $m_{e,max}$. Knowing the distribution of Δm^2_{ij} and $\sin^2 2\theta_{ij}$, and demanding $m_{e,eff} \leq m_{e,max}$, one only need to know the mass hierarchy to find a probability distribution of the total neutrino-mass. Procedure:

1. Draw the parameters Δm^2_{ij}, $\sin^2 2\theta_{ij}$ and m_1 from a distribution based on experimental results.

2. Calculate $m^2_{e,eff}$. If this value exceeds $m^2_{e,max}$: go back to *step 1*.

3. If $m_{e,eff} \leq m_{e,max}$, calculate M_ν, given the hierarchy chosen.

4. Save this value of M_ν, and repeat step 1-4 for as long as you like.

Step 1 needs no further explanation. The only part requiring a bit of consideration is that you have to calculate the interval from which to draw m_1. This is done by finding the interval that allows for $m_{e,eff}$ to cover the entire interval $[0 : m_{e,max}]$.

To perform *step 2*, a transformation from $\sin^2 2\theta_{ij}$ to s_{ij} and c_{ij} is needed:

$$\begin{aligned} \sin\theta_{ij} &= \sqrt{\frac{1-\sqrt{1-\sin^2 2\theta_{ij}}}{2}} \\ \cos\theta_{ij} &= \sqrt{1-sin^2\theta_{ij}}. \end{aligned} \tag{7.17}$$

Then, $m_{e,eff}$ is easily calculated from equation (7.15).

In *step 3*, M_ν is calculated using equation (7.16). For normal and inverted hierarchy the formulas for calculating the masses is given in table 7.1 and 7.2. The minimum value of M_ν is calculated from the demand that the smallest mass is positive.

Hierarchy	m_2	m_3
Normal	$\sqrt{m_1^2+\Delta m_{21}^2}$	$\sqrt{m_1^2+\Delta m_{21}^2+\lvert\Delta m_{32}^2\rvert}$
Inverted	$\sqrt{m_1^2+\Delta m_{21}^2}$	$\sqrt{m_1^2+\Delta m_{21}^2-\lvert\Delta m_{32}^2\rvert}$

Table 7.1: Calculating m_2 and m_3 for different hierarchies.

Hierarchy	M_{min}
Normal	$\sqrt{\Delta m_{21}^2}+\sqrt{\Delta m_{21}^2+\lvert\Delta m_{32}^2\rvert}$
Inverted	$\sqrt{\lvert\Delta m_{32}^2\rvert-\Delta m_{21}^2}+\sqrt{\lvert\Delta m_{32}^2\rvert}$

Table 7.2: Calculating M_{min} for different hierarchies.

7.3.1 At what scale is the neutrino masses degenerate?

If the neutrino masses degenerate? I stated in the previous chapter that if the neutrinos were all non-relativistic before decoupling, then the total neutrino mass would be about $M_\nu = 1.65$ eV, and that then the neutrino masses would be degenerate. In this section, I will take a closer look at how big the neutrino masses must be before they are considered to be degenerate.

The masses are said to be degenerate if $M_\nu \gg \Delta m_{ij}$ for all combinations of i and j. The largest neutrino mass difference is about $\Delta m_{23}^2 \approx 3\times10^{-3}$ eV, giving $|\Delta m_{23}| \approx 0.05$ eV. Thus the neutrino masses are degenerate for some value $m \gg 0.05$ eV. To find a more exact mass at which to start assuming a degenerate hierarchy, I have plotted the neutrino mass eigenvalues as a function of the total neutrino mass in figure 7.1 for both normal and inverted hierarchy. From these plots, one can make a rather rough assumption of the degeneration scale. For a total M_ν of more than about 0.4 eV, the individual neutrino masses seems to merge to a degenerate value.

To make a better estimate, I have plotted the relative difference between m_1 and m_3 in figure 7.2. In this plot we see that at $M_\nu \approx 0.11$ eV, the relative difference between the two eigenvalues drops below 0.1. The estimate we gave earlier of degeneracy when $M_\nu \approx 0.4$ eV gives a relative difference of about 0.01. This fulfills $\Delta m_{ij} \ll M_\nu$, and I have chosen to use this limit in the further work of this thesis.

When calculating M_ν for a degenerate hierarchy, I have used the same formula as for the normal hierarchy, but the lower limit is set so that $M_\nu > 0.2$ eV.

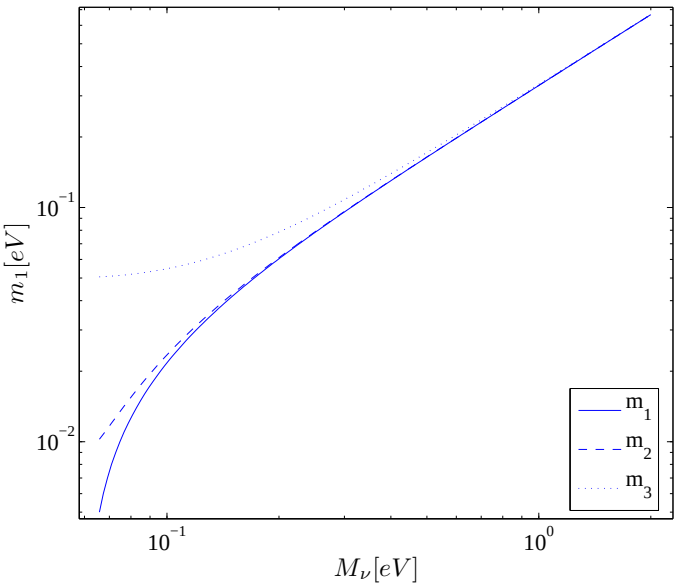
7.3.2 Distribution of the Mass Parameters

The limits on the mass and mixing parameters used in this thesis are:

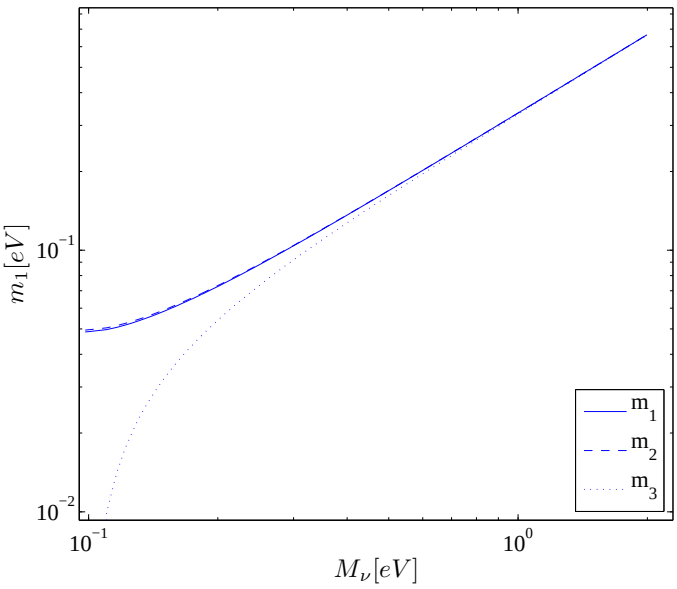
$$\begin{aligned} &\Delta m^2_{21} = (8.0 \pm 0.3) \times 10^{-5} eV^2 \\ &|\Delta m^2_{32}| = (1.9 - 3.0) \times 10^{-3} eV^2 \\ &\sin^2(2\theta_{12}) = 0.8^{+0.03}_{-0.04} \\ &\sin^2(2\theta_{13}) \leq 0.19. \\ &m_{e,max} = 2.0 eV \end{aligned} \tag{7.18}$$

These values are taken from [6], and combines results from various experiments, and are supposed to be the best values by the summer of 2008. In my calculations, I have used flat priors on $\sin^2(2\theta_{13})$ and $|\Delta m^2_{32}|$ in the allowed interval. On Δm^2_{21} I have used a Gaussian distribution with $mean = 8.0 \times 10^{-5} eV^2$ and $std = 0.3 \times 10^{-5} eV^2$, and on $\sin^2(2\theta_{12})$ I used $mean = 0.86$ and $std = 0.03$. The prior used on m_1 is flat, and for all the hierarchies, the only upper restriction placed on the maximum of the mass is given by $m_{e,max}$. The probability distribution for the different hierarchies then come out as shown in figure 7.3. This is the prior that has been applied to the CosmoMC output.

From figure 7.3 we see that the prior looks similar for all the hierarchies at high values of M_ν. At low values of M_ν, there are differences which are mainly due to the different values of M_{min}.



(a) Normal hierarchy



(b) Inverted hierarchy

Figure 7.1: The relation between total neutrino mass M_ν and mass eigenvalues m_i for normal (a) and inverted (b) hierarchies.

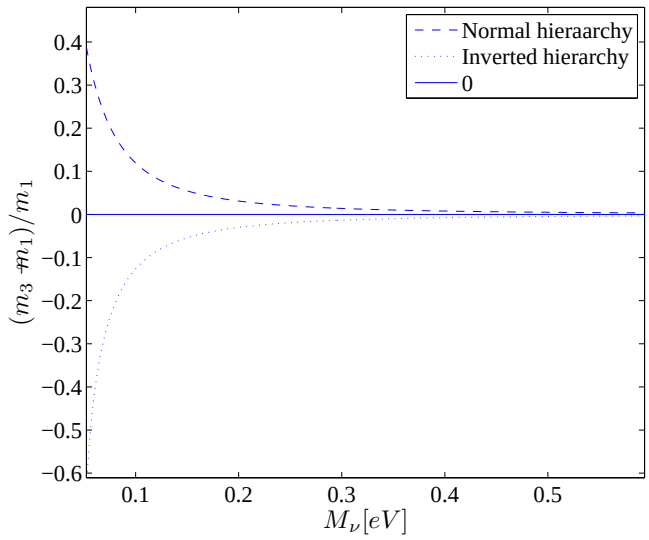


Figure 7.2: The relative difference between m_1 and m_3 for normal and inverted hierarchies.

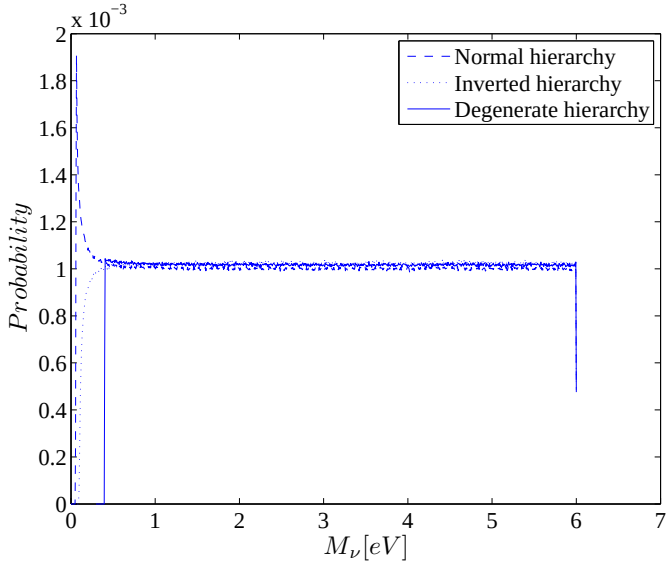


Figure 7.3: The hierarchy dependent prior found using the values parameter values of equation 7.18.

Chapter 8

Results: Hierarchy and Cosmology

So far in this work, I have looked at neutrinos in the standard model, and possible mechanisms for generating neutrino masses. I have established that as the mass square differences are known, or at least the absolute value of the mass square differences, and the absolute mass scale is not very well restricted, there are several possibilities for the ordering of the mass eigenvalues called mass hierarchies.

In chapter 5 and chapter 6 I discussed the effects of massive neutrinos on cosmology, and how the total neutrino mass influences observables and the matter and CMB power spectra in particular. It is now time to look at the effects of the neutrino mass hierarchies on three of the cosmological parameters discussed earlier; the spectral index n_s discussed in section 4.2.3 and 6.2, the baryon acoustic oscillations parameter $\mathcal{A}(z)$ discussed in section 5.2.1 and the fluctuation amplitude on cluster scales σ_8 introduced in section 5.4.

I will start to look at how the neutrino mass hierarchies an the total neutrino mass affects these parameters when CMB data alone is used for a nice flat ΛCDM $+\nu$ model in section 8.3. Then I will include also large scale structure data in section 8.4. Eventually I will look at what happens if one permits the dark energy to differ from a cosmological constant.

8.1 Cosmology and Neutrino Mass Hierarchies

As discussed in chapter 6, CMB data alone is not very likely to be precise about the total neutrino mass, unless the total mass is greater than about 1.6 eV. In chapter 5 we saw that the effect of neutrino masses on the matter power spectrum was dependent on k_{nr}, which is the comoving wavelength at the time when the neutrinos became non-relativistic. This is the case even if the neutrinos are very light, and thus observations of the matter power spectrum through observations of the large scale structure in the universe are potentially more powerful when it comes to determining the total neutrino mass. In cosmology, it is common to assume

three equal massed neutrinos, and then the non-relativistic transition occurs at the same time for all the neutrinos. For a degenerate neutrino hierarchy, this is a good approximation. If, however, the neutrinos are very light and the masses are non-degenerate, this would happen at different times for the different species. This will lead to different k_{nr} for the different species.

If the total neutrino mass is found to be larger than about 1 eV, then the neutrinos are degenerate, and although we still would like to know which mass is the larger, it is not so very important from the astrophysical point of view, as the assumption of equal masses then is a good approximation. Thus the mass scale can determine the hierarchy, if the total mass is found to be large. Also, in the other end of the scale, a normal hierarchy allows for a smaller total neutrino mass than an inverted hierarchy ($M_{nu} \lesssim 0.1$ eV is only allowed for a normal hierarchy, given the present neutrino parameters in equation (7.18)). But if the total mass is somewhere in between, we want to be able to deduce the neutrino hierarchy without being able to use the mass scale.

8.2 The Model Used

In section 8.3 and 8.4 I have used a flat ΛMDM model; a model consisting mostly of dark energy in the form of a cosmological constant, and a mix of cold and warm dark matter, of which neutrinos is a part.

I have used a cosmological parameter with eight parameters, $\Omega_b h^2$, $\Omega_{dm} h^2$, θ, τ, M_ν, n_s, A_s and a_{SZ}. These parameters are assumed to be free, and are varied in the run of CosmoMC. In order to run with M_ν as a parameter rather than $f_\nu = \Omega_\nu / \Omega_{dm}$, the CosmoMC code has been slightly modified. As a flat universe is assumed, Ω_Λ is varied indirectly. The θ parameter is the ratio of the sound horizon to the angular diameter distance, $\theta = r_\theta / r_{SH}$. τ is the optical depth at reionization. A_s is an amplitude parameter for the primordial scalar perturbations. a_{SZ} relates to the Sunyaev-Zeldovich distortion of photons from the CMB being scattering of hot gas within clusters. This effect is very small, as other matter components outweigh the clusters as dominate energy components [29]. In section 8.7, the equation of state parameter of dark energy, ω_X, is used as a free parameter in addition to the above mentioned parameters.

8.3 Numerical Results from CMB Alone

In this section, I present the results of the CosmoMC code using CMB data alone. I have used the 7 year WMAP data [53] available from LAMBDA [54], the ACBAR experiment [55, 56], the CBI experiment and BOOMERANG experiment [57].

The marginalised likelihoods are presented in figure 8.1, for the total neutrino mass M_ν, the the spectral index n_s, the fluctuations amplitude on cluster scales σ_8 and the baryon acoustic oscillation (BAO) parameter $\mathcal{A}(z)$ for different values of the redshift z. Using only CMB data, the results for various parameters are listed

in table 8.1 to 8.8. I have listed the parameter value with the highest likelihood, and the 95% confidence limits. For the neutrino mass, I have included the one-tail central credible interval 95% C.L where no lower limit can be obtained. For the other parameters, which all have a two-tail probability distribution, I have included the central credible interval limits of 95% C.L.

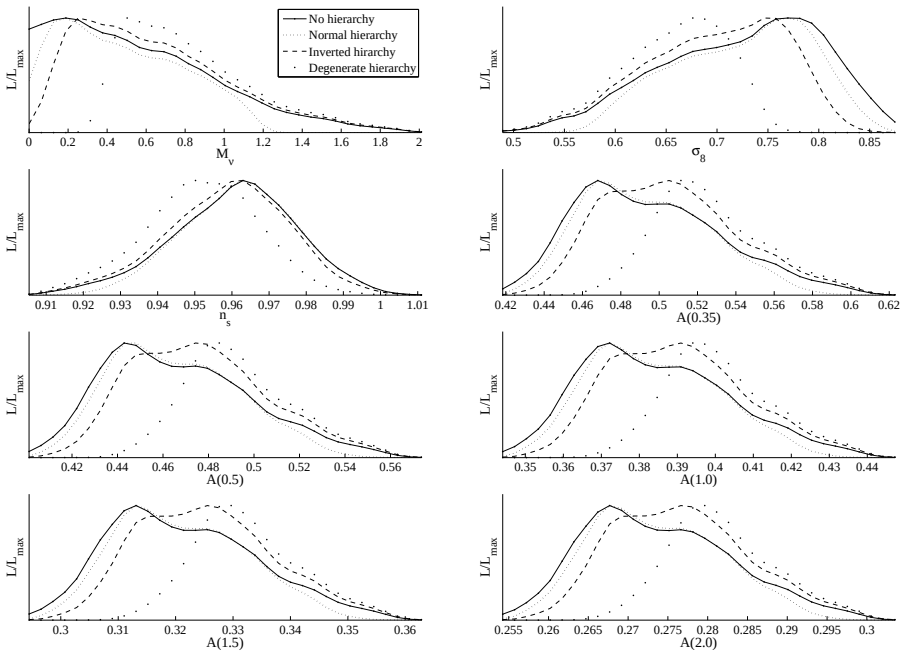


Figure 8.1: The marginalized parameters, using CMB data alone. There is not much deviation between the normal and inverted hierarch and the no-hierarchy likelihood functions except for the BAO parameter $\mathcal{A}(z)$. The confidence limits corresponding to the figures are found in table 8.1 to 8.8

In general, there is very little deviation between results before the hierarchy prior is added and the results after adding a normal hierarchy prior, as seen from the marginalized likelihood distributions in figure 8.1. The same trend is found in the confidence intervals and the maximal likelihood parameter values obtained. The inverted hierarchy prior affects some of the likelihood functions more than others. For the total mass, the spectral index and the fluctuation amplitude, the inverted hierarchy likelihood function is similar to the function for the normal hierarchy and no hierarchy. For the BAO parameter, the inverted hierarchy maximal likelihood position and the likelihood function looks more like the degenerate hierarchy functions. The degenerate hierarchy prior affects the likelihood-distributions more than the other priors for all the observables, and the confidence interval is generally shifted together with the point of maximal likelihood.

Hierarchy	Max $\mathcal{L}$ $M_\nu[eV]$	95% lower limit	95% upper limit $[eV]$
None	0.1855	none	1.3823
Normal	0.1885	none	1.0681
Inverted	0.2513	0.1257	1.5708
Degenerate	0.5027	0.3770	1.6336

Table 8.1: Confidence limits on the total neutrino mass, M_ν, obtained from CMB alone.

Hierarchy	Max $\mathcal{L}$ σ_8	95% lower limit	95% upper limit
None	0.7697	0.5594	0.8514
Normal	0.7580	0.5545	0.8397
Inverted	0.7463	0.5477	0.8047
Degenerate	0.6762	0.5361	0.7463

Table 8.2: Confidence limits on the fluctuation amplitude on cluster scales, σ_8, obtained from CMB alone.

Hierarchy	Max $\mathcal{L}$ n_s	95% lower limit	95% upper limit
None	0.9630	0.9246	0.9918
Normal	0.9630	0.9310	0.9918
Inverted	0.9630	0.9214	0.9918
Degenerate	0.9502	0.9182	0.9822

Table 8.3: Confidence limits on the spectral index, n_s, obtained from CMB alone.

Hierarchy	Max $\mathcal{L}$ $A(0.35)$	95% lower limit	95% upper limit
None	0.4680	0.4370	0.5795
Normal	0.4680	0.4370	0.5695
Inverted	0.5052	0.4494	0.5857
Degenerate	0.5114	0.4742	0.5919

Table 8.4: Confidence limits on the BAO parameter $\mathcal{A}(z)$ for $z = 0.35$ obtained from CMB alone.

Hierarchy	Max $\mathcal{L}$ $A(0.5)$	95% lower limit	95% upper limit
None	0.4429	0.4168	0.5370
Normal	0.4482	0.4168	0.5213
Inverted	0.4743	0.4273	0.5422
Degenerate	0.4847	0.4534	0.5474

Table 8.5: Confidence limits on the BAO parameter $\mathcal{A}(z)$ for $z = 0.5$ obtained from CMB alone.

Hierarchy	Max $\mathcal{L}$ $A(1.0)$	95% lower limit	95% upper limit
None	0.3722	0.3534	0.4285
Normal	0.3722	0.3566	0.4191
Inverted	0.3910	0.3597	0.4285
Degenerate	0.3941	0.3753	0.4316

Table 8.6: Confidence limits on the BAO parameter $\mathcal{A}(z)$ for $z = 1.0$ obtained from CMB alone.

Hierarchy	Max $\mathcal{L}$ $A(1.5)$	95% lower limit	95% upper limit
None	0.3131	0.3007	0.3504
Normal	0.3131	0.3028	0.3442
Inverted	0.3255	0.3048	0.3504
Degenerate	0.3297	0.3152	0.3524

Table 8.7: Confidence limits on the BAO parameter $\mathcal{A}(z)$ for $z = 1.5$ obtained from CMB alone.

Hierarchy	Max $\mathcal{L}$ $A(2.0)$	95% lower limit	95% upper limit
None	0.2677	0.2587	0.2946
Normal	0.2677	0.2602	0.2901
Inverted	0.2766	0.2617	0.2946
Degenerate	0.2796	0.2707	0.2961

Table 8.8: Confidence limits on the BAO parameter $\mathcal{A}(z)$ for $z = 2.0$ obtained from CMB alone.

8.4 Adding Large Scale Structure

In this section, I will add large scale structure (LSS) observations to improve the limits on M_ν. As discussed previously, the CMB is not very sensitive to neutrino masses smaller than $M_\nu \approx 1.6$. As the upper limit on the total neutrino mass from CMB alone is found to be much lower than this limit, we should not trust our results as much as we would like. As seen from earlier discussion, the matter power spectrum is able to provide limits on the total neutrino mass for smaller mass scales. Thus, adding observations on large scale structure can improve the total neutrino mass limits, and hopefully also give better information about the hierarchy of the masses. Hence, I have added the SDSS LRG data set [58] to the CMB data used in the previous section in the CosmoMC run.

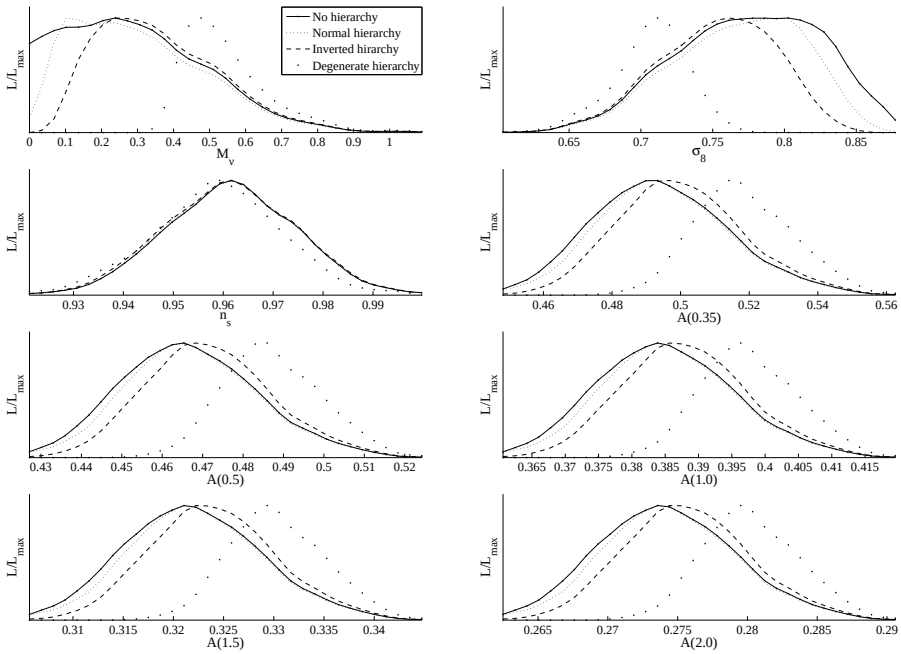


Figure 8.2: The marginalized parameters, found by using CMB data and large scale structure observations. The confidence limits obtained from the same data are found in tables 8.9-8.16. Similar plots for CMB alone is found in figure 8.1.

The one-dimensional marginalised likelihood distributions with CMB and large scale structure data is showed in figure 8.2, and the confidence-intervals are displayed in tables 8.9-8.16.

Introducing LSS data effectively decreases the upper limits on the neutrino masses relative to the CMB alone scenario of the last two sections. The upper limit is approximately halved by the introduction of large scale structure information, and an upper confidence limit in the range $M_\nu < 0.65 - 0.85$ eV is in agreement

Hierarchy	Max $\mathcal{L}$ M_ν [eV]	95% lower limit	95% upper limit
None	0.2382	none	0.6466
Normal	0.1021	0.0340	0.7147
Inverted	0.2382	0.1021	0.7487
Degenerate	0.4764	0.3744	0.8508

Table 8.9: Confidence limits on the total neutrino mass, M_ν, obtained from CMB in combination with large scale structure data.

Hierarchy	Max $\mathcal{L}$ σ_8	95% lower limit	95% upper limit
None	0.8027	0.6702	0.8606
Normal	0.7944	0.6702	0.8440
Inverted	0.7613	0.6919	0.8275
Degenerate	0.7116	0.6454	0.7530

Table 8.10: Confidence limits on the fluctuation amplitude on cluster scales, σ_8, obtained from CMB in combination with large scale structure data.

with the "conservative estimates" given in section 1.4.2. Also, adding large scale structure observations seems to make the likelihood functions for the normal and inverted hierarchies more similar. They are both quite similar to the no-hierarchy likelihood functions, while the degenerate hierarchy likelihood function tends to differ from the "low mass" scenarios.

Hierarchy	Max $\mathcal{L}$ n_s	95% lower limit	95% upper limit
None	0.9617	0.9355	0.9879
Normal	0.9617	0.9355	0.9879
Inverted	0.9617	0.9355	0.9879
Degenerate	0.9593	0.9331	0.9831

Table 8.11: Confidence limits on the spectral index, n_s, obtained from CMB in combination with large scale structure data.

Hierarchy	Max $\mathcal{L}$ $A(0.35)$	95% lower limit	95% upper limit
None	0.4933	0.4586	0.5383
Normal	0.4933	0.4586	0.5383
Inverted	0.4967	0.4656	0.5383
Degenerate	0.5141	0.4933	0.5487

Table 8.12: Confidence limits on the BAO parameter $\mathcal{A}(z)$ for $z = 0.35$ obtained from CMB in combination with large scale structure data.

Hierarchy	Max $\mathcal{L}$ $A(0.5)$	95% lower limit	95% upper limit
None	0.4654	0.4360	0.5035
Normal	0.4654	0.4360	0.5036
Inverted	0.4683	0.4419	0.5036
Degenerate	0.4860	0.4654	0.5124

Table 8.13: Confidence limits on the BAO parameter $\mathcal{A}(z)$ for $z = 0.5$ obtained from CMB in combination with large scale structure data.

Hierarchy	Max $\mathcal{L}$ $A(1.0)$	95% lower limit	95% upper limit
None	0.3839	0.3660	0.4070
Normal	0.3839	0.3678	0.4070
Inverted	0.3856	0.3714	0.4088
Degenerate	0.3963	0.3839	0.4124

Table 8.14: Confidence limits on the BAO parameter $\mathcal{A}(z)$ for $z = 1.0$ obtained from CMB in combination with large scale structure data.

Hierarchy	Max $\mathcal{L}$ $A(1.5)$	95% lower limit	95% upper limit
None	0.3211	0.3092	0.3364
Normal	0.3211	0.3104	0.3364
Inverted	0.3222	0.2128	0.3376
Degenerate	0.3293	0.3211	0.3400

Table 8.15: Confidence limits on the BAO parameter $\mathcal{A}(z)$ for $z = 1.5$ obtained from CMB in combination with large scale structure data.

Hierarchy	Max $\mathcal{L}$ $A(2.0)$	95% lower limit	95% upper limit
None	0.2736	0.2650	0.2846
Normal	0.2736	0.5659	0.2846
Inverted	0.2744	0.2676	0.2855
Degenerate	0.2795	0.2736	0.2872

Table 8.16: Confidence limits on the BAO parameter $\mathcal{A}(z)$ for $z = 2.0$ obtained from CMB in combination with large scale structure data.

8.5 Comparing CMB Alone and CMB + LSS

As seen from figures 8.1 and 8.2 and the corresponding confidence limits, the case of normal and inverted hierarchy has only minor deviations from the case of no hierarchy prior applied, particularly when both CMB and LSS observations has been applied.

The upper limit on the neutrino mass is approximately halved by adding large scale observations. This is due to the low sensitivity of CMB observations to small neutrino masses. The upper limit on the total neutrino mass is lower for the normal hierarchy is lower than for the inverted hierarchy, which again is lower than for the degenerate hierarchy.

The mass fluctuation amplitude on cluster scales, σ_8 is found to be lower for a degenerate hierarchy than for normal or inverted hierarchy. This is due to suppression of the matter power spectrum by massive neutrinos. As a degenerate hierarchy favours heavier neutrinos, it gives rise to a power spectrum that is suppressed for large values of k, and thus equation (5.45) leads to a smaller value of σ_8. The point of highest likelihood is moved towards higher values of σ_8 when the large scale structures are added, and the confidence interval becomes smaller for all the hierarchies. The narrower confidence interval is due to the greater ability of the large scale structure to determine σ_8, as it is directly related to the matter power spectrum. The spectral index n_s is lower for a degenerate hierarchy than for the normal and inverted hierarchies.

In the previous sections, we saw that the Inverted hierarchy highest likelihood positions for $\mathcal{A}(z)$ resembled the degenerate hierarchy more than the normal hierarchy for CMB alone, while resembled the normal hierarchy for CMB + LSS. This is probably due to the better values obtained for the likelihood function when large scale structure observations are added.

For the BAO parameter $\mathcal{A}(z)$, is it easy to see that it is decreasing for higher values of z. A degenerate hierarchy leads to a larger value of $\mathcal{A}(z)$ for all values of z than a normal hierarchy, and also the inverted hierarchy increases the best fit value of $\mathcal{A}(z)$ slightly. This is due to the effect of increased neutrino mass. $\mathcal{A}(z)$ is a function of z, Ω_m and Ω_{DE}. As the universe is kept spatially flat and radiation is assumed to have negligible contribution, Ω_m grows and Ω_Λ decreases as the neutrino mass is increased. This is shown in figures 8.3 and 8.4 when the dark energy is a cosmological constant.

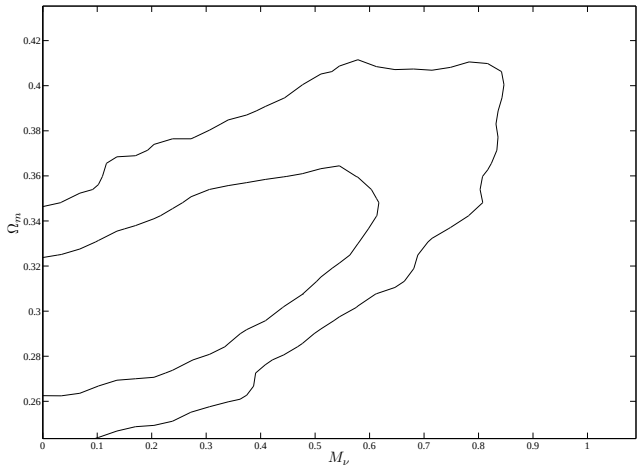


Figure 8.3: Contour-plot of Ω_m and the total neutrino mass M_{nu}. Both CMB data and large scale structure observations are used. We see that an increased neutrino mass gives an increased value of Ω_m.

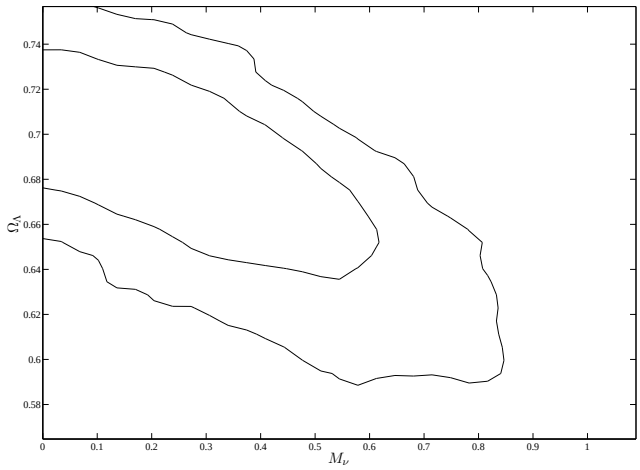


Figure 8.4: Contour-plot of Ω_Λ and the total neutrino mass M_{nu}. Both CMB data and large scale structure observations are used. We see that an increased neutrino mass gives an decreased value of Ω_Λ.

8.6 The Effects of Neutrino Hierarchy

As commented on before, the normal and inverted hierarchy likelihood distributions are quite similar to the no-hierarchy distribution for all the parameters. The degenerate hierarchy likelihood distributions show some deviations from the other hierarchies. My goal is to find out if we can use n_s, σ_8 or $\mathcal{A}(z)$ for the chosen values of z to say something about the mass hierarchy.

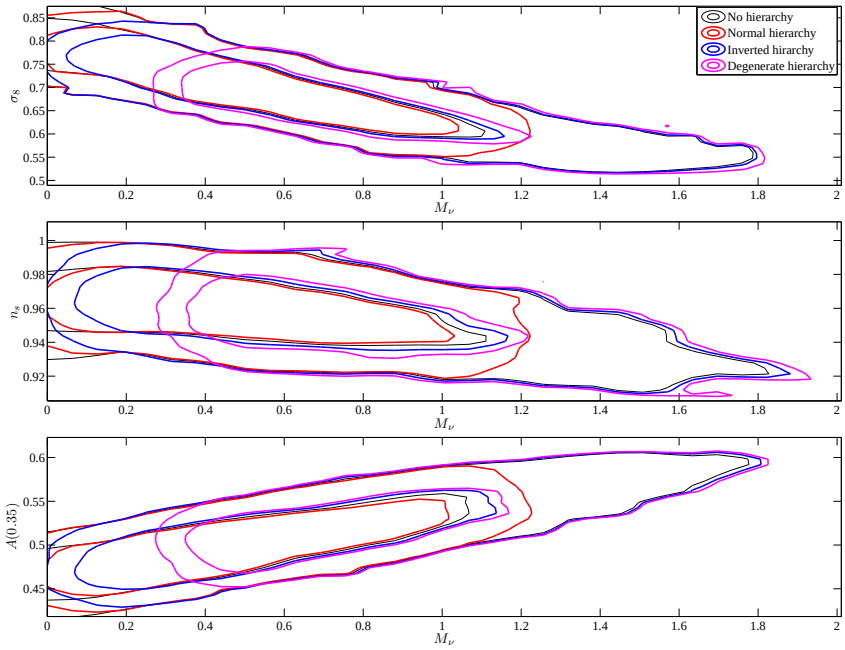


Figure 8.5: Contour-plots of the matter fluctuations, σ_8, the scale factor, n_s and the BAO parameter $\mathcal{A}(0.5)$ against the total neutrino mass M_ν. The plots are produced using GetDist on the results of CMB data alone. The contours are drawn from 68% and 95% C.L.

8.6.1 What is mass, and what is hierarchy?

As there is a general trend in all the marginalised one-dimensional plots that the normal and inverted hierarchies seems to have more in common with each other than with the degenerate hierarchy, the question arises whether the discrepancy between the low-mass hierarchies and the high-mass hierarchy is only due to the difference in total mass. For this purpose, the contour-plots of the investigated parameters and the total neutrino mass are useful. Looking at the two-dimensional contour-plots in figure 8.5 and 8.6, we can see the relation between the investigated parameters and the total neutrino mass. The first thing to notice, is that for the

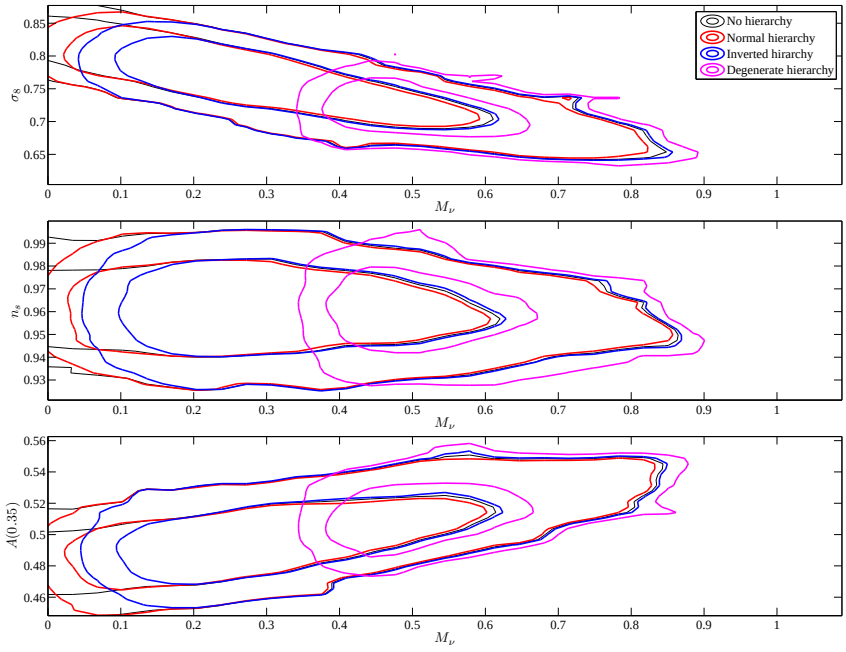


Figure 8.6: Contour-plots of the fluctuation amplitude on cluster scales, σ_8, the spectral index n_s and the BAO parameter $\mathcal{A}(z)$ for $z = 0.35$ against the total neutrino mass $M\nu$. the plots are produced using the GetDist on the results obtained using CMB data in combination with large scale structure observations. The contours drawn encircles 68% and 95% of the distribution.

"mid-range" values of total neutrino mass, ie. $M_\nu \in (0.4 - 0.9)$ for CMB alone and $M_\nu \in (0.4 - 0.6)$ for CMB + LSS, the contour lines of all the hierarchies behaves in the same way. They also behave similarly to the no-hierarchy contour lines. The degenerate hierarchy contour lines are, particularly when bot CMB and LSS are used, a bit wider in the paremeter direction than the other hierarchies, but this is due to the narrower M_ν interval. For the CMB alone contours, the inverted, degenerate and no hierarchy contours come nicely together in the upper end of the M_ν interval. The normal hierarchy has a lower upper limit on the total neutrino mass, as its likelihood for low values of the total mass is enhanced by the hierarchy prior. For small values, the contour lines closes at different values, due to the different lower bounds on the total neutrino mass.

For CMB + LSS, the stronger constraints on the total neutrino mass imposed by the large scale structure observations causes the normal hierarchy to get a more similar behaviour for large values of the total neutrino mass to that of the other hierarchies. The degenerate hierarchy has higher upper limits on the neutrino mass, due do the higher lower limit. There is nothing in the contour plots indicating that the different hierarchy contours does not all come from the same distribution, only

modified by the different allowed lower limits on M_ν.

We see the same trend as we did in the marginalised likelihood functions when it comes to the similar behaviour of the low mass hierarchies, where they only really differ for low neutrino masses. The low-mass hierarchies also has the same behaviour at high masses as the data that does not depend on hierarchy.

The contour-plot is also a good tool to investigate how the parameters n_s, σ_8 and $\mathcal{A}(z)$ depend upon the total neutrino mass. From the behaviour of the confidence lines, we can see that the three parameters depend in different ways on M_ν. σ_8 decays when the neutrino mass increases, and thus a good measurement of σ_8 would be very helpful in determining M_ν. n_s also decreases when M_ν increases when using CMB alone, but the dependence is weaker, and an accurate measurement of the scale factor can still leave the total neutrino mass with a wide interval of possible values. For CMB + LSS there seems to be no dependence of n_s on M_ν. The BAO parameter $\mathcal{A}(z)$ has, in the same way as σ_8, a quite strong dependence of M_ν, and a larger neutrino mass seems to increase the value of $\mathcal{A}(z)$. I have only plotted $\mathcal{A}(z)$ for $z = 0.35$, as the behaviour is the same for all the used values of z.

Our main interest, however, is to tell how the parameters n_s, σ_8 and $A(z)$ are affected by the hierarchy of the neutrino masses. Unfortunately, our data does not seem to be very helpful. The tendency of the low-mass hierarchies to behave very similarly leaves us with not much hope to be able to distinguish between the normal and inverted hierarchies from observing n_s, σ_8 and $A(z)$. Only for very small neutrino masses is there a difference between the two hierarchies. This different behaviour is, however, expected, as the normal hierarchy allows for smaller neutrino masses than the inverted hierarchy. The contour-lines for the inverted hierarchy close for higher masses than observed for the normal hierarchy; an effect that is due to the difference of the mass scale, and not the hierarchy alone. The contour-lines for the degenerate hierarchy is shorter in the M_ν direction, and must therefore be wider in the parameter-direction. This is only to encircle the same confidence interval as the normal and inverted hierarchy contour-lines, and has thus no real physical significance. Thus, from the marginalised likelihood distributions, the highest likelihood values and the contour-plots it seems like any attempt to determine the neutrino mass hierarchy from the parameters in question using CMB data and the SDSS LRG data is fruitless.

It is, however, possible to say something about the total neutrino mass from the parameters investigated, and thus to say something about whether the neutrino masses obeys the degenerate hierarchy or not. A value of $\sigma_8 \approx 0.8$ would strongly suggest that the neutrinos have too small masses to be degenerate, and a very low value of $\mathcal{A}(z)$ would suggest that the neutrinos are not degenerate.

It is not very surprising that the effect of the prior on separating the different hierarchies is small. The prior likelihood distribution, as seen in figure 7.3, only properly separates the hierarchies at low masses, and this is an effect due to different allowed minimal values for M_ν. Using the raw data rather than the best fit values when producing the prior may improve the results, but most likely not sig-

nificantly. The prior only accounts for the distribution of the total neutrino mass, and does not account for the effects of each individual neutrino mass on cosmology.

8.6.2 Is this study still useful?

So far, we have seen that the three chosen cosmological parameters/observables are not much good for determining the neutrino hierarchy with the applied methods. Although a bit disappointing, some conclusions can be drawn from this study.

Cosmologist almost always assume that neutrino masses are degenerate. And by degenerate, they do not mean that the neutrino mass is so large that the mass splitting is small compared to the individual masses, but that the mass splitting can be neglected. Three neutrinos with the same mass, such that $m_\nu = M_\nu/3$ is used for simplicity. And it seems like this is a very decent approximation. From my study of n_s, σ_8 and $A(z)$, it seems like the mistake done by assuming that the mass splitting does not affect cosmology is negligible, as all the results obtained resembles that of no hierarchy, only adjusted for the allowed interval of M_ν.

The maximal likelihood point of each parameter, and the confidence interval, seems only to be constrained by the total neutrino mass, an not the hierarchy. If neglecting the low-mass end of each contour-line, which closes at the scale of the lowest permitted total mass, the contour-lines follows the no-hierarchy lines almost perfectly, and any deviation can be explained by the mass scale. The maximum likelihood positions and confidence levels of all the parameters are very close for the normal, inverted and no hierarchy cases, making it hard to distinguish between the non-degenerate hierarchies. Thus, for the data used in this study, assuming that the neutrino mass hierarchy is negligible is appropriate.

8.7 Allowing for Dark Energy with $\omega \neq -1$

So far in this thesis, we have used LIVE as a model for the dark energy, which is equivalent to a cosmological constant. This is often used in cosmology, but it is not necessary that the dark energy comes in the form of a cosmological constant. We know that there is an "extra" energy component that has an equation of state parameter $\omega_X < -1/3$, but exactly how this dark energy behaves is not established. The reasons for assuming a cosmological constant are several. First of all, it is beautiful with a cosmological constant. It behaves nicely, it does not cluster and it makes the equations nice and simple. And "it is there" in the Einstein and Friedman equations. It really is the simple solution.

Secondly, when allowing ω_X to vary, we get $\omega_X \approx -1$. Thus it is an approximation to use $\omega_X = -1$.

It is, however, not so that we *know* that the dark energy is a cosmological constant. And for all we know, ω_X could vary with time. The present constraints on ω_X varies with the data used to obtain them, but [59] has found the limit $\omega_X = -1.12^{+0.42}_{-0.43}$ from WMAP7 data, and $\omega_x = 0.980 \pm 0.053$ at 68% C.L.

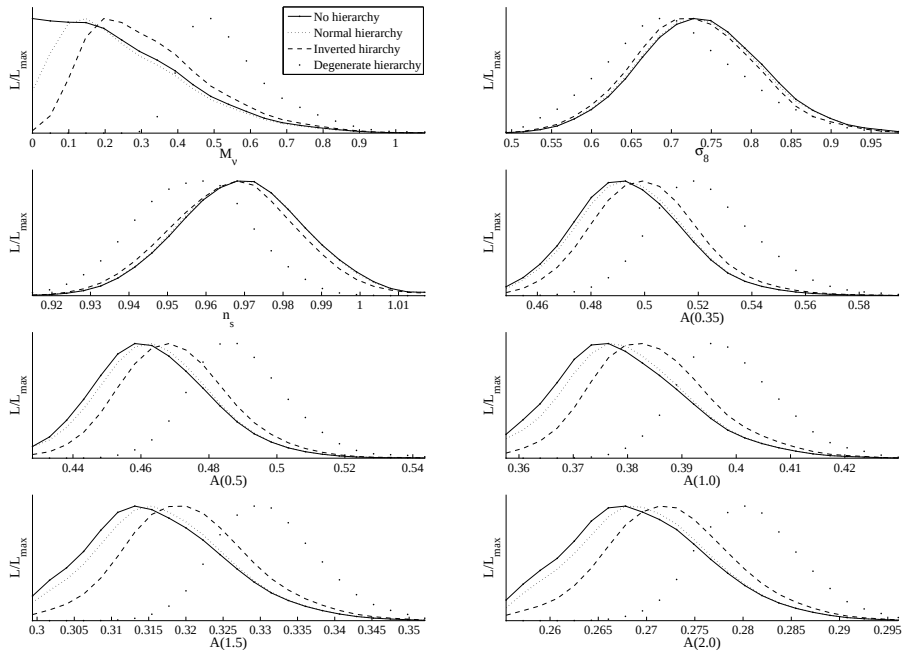


Figure 8.7: The marginalized parameters obtained using CMB data and large scale structure observations when allowing the dark energy to have $\omega_X \neq 1$. The confidence limits obtained from the same results as shown in the this figure is presented in tables 8.17 - 8.24

when combining WMAP7 with observations of the baryon acoustic oscillations and a prior on the present Hubble parameter. The limits are found assuming that ω_X is constant in time, that the universe is spatially flat, and that the neutrino mass is zero. With a non-zero neutrino mass, [60] found $\omega_x = -1.07 \pm 0.12$ at 68% C.L.

I have included ω_X as a parameter in my model, to see how this effects my results. The marginalised parameters are plotted in figure 8.7 and the corresponding confidence limits are shown in tables 8.17 to 8.25, and the contour plot for the parameters used in the sections above is presented in figure 8.8. The plots for the equation of state parameter is presented in figure 8.9 (contour-plot) and 8.10 (marginalised likelihood function).

Adding an additional parameter lowers σ_8, increases n_s and changes the maximal likelihood M_ν value slightly. It does not, however, seem to change the dependency of the chosen cosmological parameters in the hierarchy. This makes sense, adding an extra free parameter should not alter the ability of cosmology to determine the neutrino mass hierarchy.

From the contour-plot of ω_X and M_ν in figure 8.10, the degeneracy between the two parameters seems to be small, but a increased total neutrino mass does

Hierarchy	Max $\mathcal{L}$ M_ν	95% lower limit	95% upper limit
None	0	none	0.5985
Normal	0.1472	0.0332	0.5890
Inverted	0.1963	0.0982	0.763
Degenerate	0.4908	0.3436	0.8344

Table 8.17: Confidence limits on the total neutrino mass, M_ν, obtained from CMB and large scale structure combined. The dark energy is not restrained to a cosmological constant, $\omega_X \neq -1$.

Hierarchy	Max $\mathcal{L}$ σ_8	95% lower limit	95% upper limit
None	0.7457	0.6011	0.8904
Normal	0.7278	0.5784	0.8986
Inverted	0.7278	0.5784	0.8773
Degenerate	0.6851	0.5570	0.8773

Table 8.18: Confidence limits on the fluctuation amplitude on cluster scale, σ_8, obtained from CMB and large scale structure combined. The dark energy is not restrained to a cosmological constant, $\omega_X \neq -1$.

seem to induce a decreased equation of state parameter slightly. But by assessing the marginalised likelihood function of the equation of state parameter in figure 8.9, we see one interesting feature. The normal and no hierarchy priors seems to favour an equation of state parameter $\omega_X > -1$. The degenerate hierarchy favours an equation of state parameter $\omega_X < -1$. This is consistent with what is seen in figure 8.10, and seems to be a pure total mass effect, and not caused by hierarchy.

Hierarchy	Max $\mathcal{L}$ n_s	95% lower limit	95% upper limit
None	0.9704	0.9374	1.0004
Normal	0.9681	0.9371	1.0035
Inverted	0.9681	0.9371	0.9991
Degenerate	0.9592	0.9282	0.9858

Table 8.19: Confidence limits on the spectral index, n_s, obtained from CMB and large scale structure combined. The dark energy is not restrained to a cosmological constant, $\omega_X \neq -1$.

Hierarchy	Max $\mathcal{L}$ $\mathcal{A}(0.35)$	95% lower limit	95% upper limit
None	0.4891	0.4590	0.5408
Normal	0.4928	0.4610	0.5437
Inverted	0.4992	0.4610	0.5437
Degenerate	0.5183	0.4865	0.5691

Table 8.20: Confidence limits on the BAO parameter $\mathcal{A}(z)$, for $z = 0.35$, obtained from CMB and large scale structure combined. The dark energy is not restrained to a cosmological constant, $\omega_X \neq -1$.

Hierarchy	Max $\mathcal{L}$ $\mathcal{A}(0.5)$	95% lower limit	95% upper limit
None	0.4603	0.4365	0.5011
Normal	0.4632	0.4331	0.5034
Inverted	0.4671	0.4381	0.5084
Degenerate	0.4883	0.4632	0.5184

Table 8.21: Confidence limits on the BAO parameter $\mathcal{A}(z)$, for $z = 0.5$, obtained from CMB and large scale structure combined. The dark energy is not restrained to a cosmological constant, $\omega_X \neq -1$.

Hierarchy	Max $\mathcal{L}$ $\mathcal{A}(1.0)$	95% lower limit	95% upper limit
None	0.3757	0.3608	0.4054
Normal	0.3765	0.3606	0.4079
Inverted	0.3827	0.3639	0.4110
Degenerate	0.3953	0.3827	0.4204

Table 8.22: Confidence limits on the BAO parameter $\mathcal{A}(z)$, for $z = 1.0$, obtained from CMB and large scale structure combined. The dark energy is not restrained to a cosmological constant, $\omega_X \neq -1$.

Hierarchy	Max $\mathcal{L}$ $\mathcal{A}(1.5)$	95% lower limit	95% upper limit
None	0.3141	0.3017	0.3358
Normal	0.3154	0.3017	0.3360
Inverted	0.3200	0.3040	0.3383
Degenerate	0.3292	0.3177	0.3452

Table 8.23: Confidence limits on the BAO parameter $\mathcal{A}(z)$, for $z = 1.5$, obtained from CMB and large scale structure combined. The dark energy is not restrained to a cosmological constant, $\omega_X \neq -1$.

Hierarchy	Max $\mathcal{L}$ $\mathcal{A}(2.0)$	95% lower limit	95% upper limit
None	0.2668	0.2572	0.2847
Normal	0.2678	0.2572	0.2855
Inverted	0.2713	0.2590	0.2872
Degenerate	0.2802	0.2713	0.2925

Table 8.24: Confidence limits on the BAO parameter $\mathcal{A}(z)$, for $z = 2.0$, obtained from CMB and large scale structure combined. The dark energy is not restrained to a cosmological constant, $\omega_X \neq -1$.

Hierarchy	Max $\mathcal{L}$ ω_X	95% lower limit	95% upper limit
None	-0.8104	-1.4795	-0.5130
Normal	-0.8577	-1.5096	-0.5318
Inverted	-0.8577	-1.5096	-0.5783
Degenerate	-1.1078	-1.7769	-0.5873

Table 8.25: Confidence limits on the dark energy equation of state parameter ω_x, obtained from CMB and large scale structure combined.

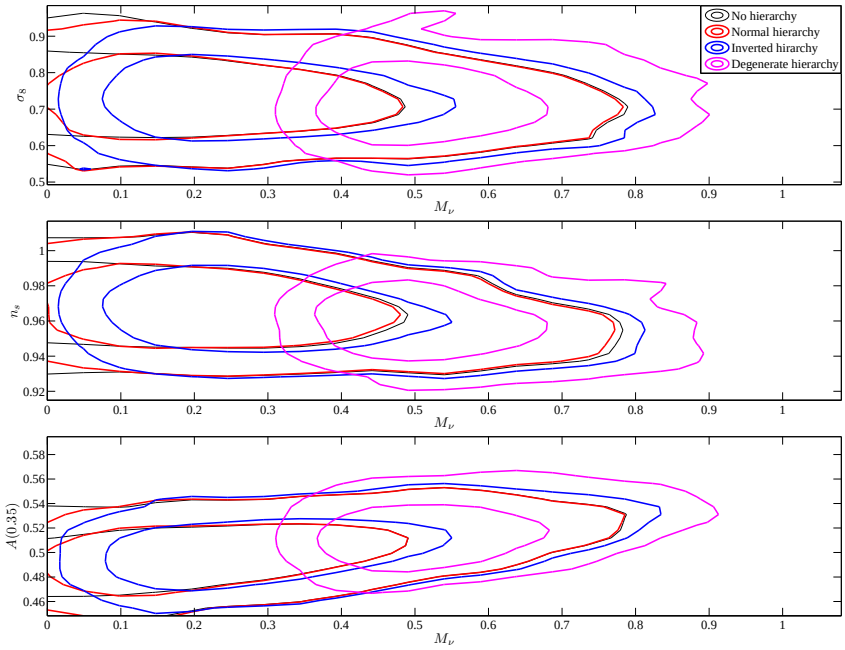


Figure 8.8: The two dimensional contour-plots for the fluctuation amplitude on large scales, σ_8, the spectral index, n_s and the BAO parameter $\dashv(z)$ for $z = 0.35$ against the total neutrino mass, M_ν obtained using CMB data and large scale structure observations when allowing the dark energy to have $\omega_X \neq 1$.

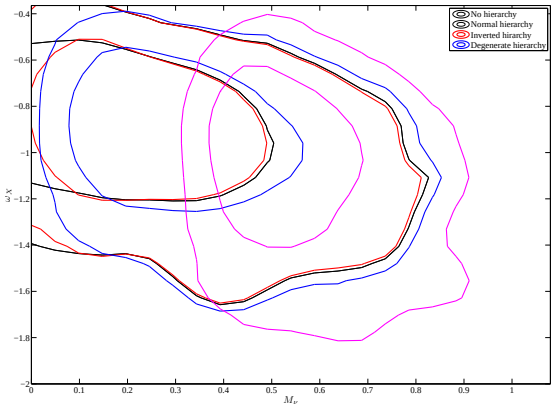


Figure 8.9: Contour-plot of the equation of state parameter of dark energy, ω_X and the total neutrino mass, M_ν. The contour-plot is obtained using CMB data and large scale structure observations.

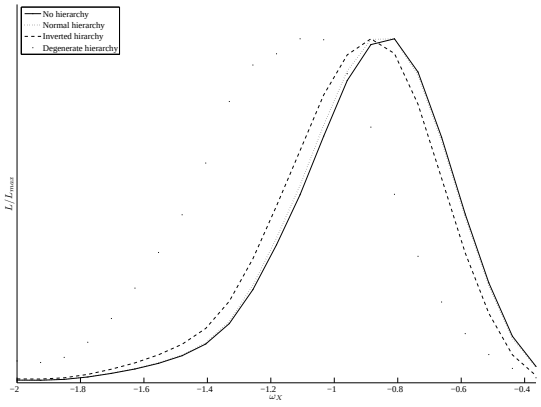


Figure 8.10: Marginalised likelihood function for the equation of state parameter of dark energy, ω_X. The contour-plot is obtained using CMB data and large scale structure observations.

Chapter 9

Summary and Discussion

In this thesis, I have studied the massive neutrino, and its effect on cosmology. I started out by reviewing the neutrino in the standard model of particle physics. I looked at how it is described by particle physicists, and why the neutrino mass is a way of gaining insight to physics beyond the standard model. In the first chapter, I also discussed experimental evidence of a non-zero neutrino mass, the experimental constraints on the mass parameters, and the unknown ordering of the masses.

In chapter 2 I discussed neutrino mass generating mechanisms, with emphasis on the seesaw mechanism. The seesaw mechanism is a popular way to generate neutrino masses, as it can both explain the mystery of the smallness of the mass and account for both degenerate and small neutrino masses.

In part II I looked at cosmology. I started by introducing the FRW universe, and went on to discussing structure formation, the matter power spectrum and the cosmic microwave background radiation. The effects of massive neutrinos were discussed, and it turned out that although the neutrinos are very light and not very keen on interacting with other particles, they leave visible imprints of their existence on cosmological scales which relates to their mass scale. They suppress structure formation on small scales, due to their free streaming. The amount of suppression is increasing with increasing neutrino mass, and the scale at which the suppression begins decreases with increasing neutrino mass. Thus the amount of suppression observed in the matter power spectrum can give a limit on the total neutrino mass. Also the CMB power spectrum can be used to obtain neutrino mass limits, as the power spectrum is sensitive to the total mass of the neutrino. It is, however, not very sensitive to small neutrino masses (masses that was still relativistic at the time of recombination), and thus the CMB power spectrum is not giving very tight limits on the neutrino mass. It is, however, measured with great accuracy and provides strong limits on other cosmological parameters, and is therefore crucial in determining the neutrino mass.

I have used CMB alone and CMB in combination with the large scale structure data from SDSS LRG to see if the neutrino mass hierarchy can be observed

in the fluctuation amplitude on cluster scales σ_8, the spectral index n_s and the baryon acoustic oscillation parameter $\mathcal{A}(z)$. To do so, a mass hierarchy prior was imposed on likelihood distribution functions provided by the CosmoMC code. I found that the effect of the neutrino mass hierarchies in the above mentioned parameters seems to come from the total neutrino mass. The mass hierarchy did not affect the distribution of the chosen parameters to any significant extent that could not be explained by the shift in the preferred total neutrino mass. This shows us that the commonly used approximation in cosmology of neglecting the mass splitting, even at small masses, is valid. The deviation from a no-hierarchy model is only due to lower limits on the total neutrino mass for the different hierarchies.

I also studied the effect of allowing the equation of state parameter for dark energy, ω_X, to be a free parameter. This gave very little new information.

The prior I have used on the total neutrino mass based on the mass hierarchy, is obtained using experimentally obtained limits on the mixing parameters and the mass square differences. Using the raw data rather than best fit values and confidence intervals may improve my results slightly. It seems, however, like using a mass hierarchy prior is not the way to go to obtain a cosmological detection of the neutrino mass hierarchy. The prior only accounts for the distribution of the total neutrino mass, and does not account for the effects of each individual neutrino mass on cosmology.

The results of my work is consistent with results from the literature. On scales $k \gg k_{nr}$ the effect of varying the fraction of the total neutrino mass carried by the individual mass eigenstates is negligible according to [61], and outside of the free streaming scale there is no effect. The effect, even on scales where a deviation from a low-mass degenerate scenario is theoretically calculable and significant, is found in [61] to be small, both for the CMB and the matter power spectrum, compared to the accuracy in the cosmological observations. It is claimed that the degeneracy between the total neutrino mass and other cosmological parameters will make a determination of hierarchy from cosmological observations impossible. Others, like [62] supports the view that the precision obtained and the experiments performed so far is insufficient in deciding on the mass splitting.

Later studies, like [63] claims that detection of mass hierarchy is possible, but not with the present data. In this paper it is claimed that the data from future Euclid survey and the already data collecting Planck experiments might provide strong evidence for either normal or inverted hierarchies. Also [64] supports this result. In the latter study it is found that the matter power spectra in principle distinguishes between the hierarchies, but that it is much less sensitive to hierarchy than to the total neutrino mass and the amplitude of the splitting. This makes it possible, although very challenging, for an ideal experiment to distinguish between normal and inverted hierarchies.

Euclid might be an important experiment when it comes to determining the neutrino mass hierarchy, as it will give good information on the matter power spectrum. Observations like the SDSS are made by counting galaxies. Euclid will utilize weak gravitational lensing and thus trace also the dark matter distribution.

This will provide more accurate constraints on the neutrino mass, and possible the mass splitting and hierarchy.

Appendix A

Calculating equation 1.7

The Dirac equation states that

$$(i\gamma_\mu \partial^\mu - m)\,\psi(x) = 0, \tag{A.1}$$

which can also be expressed as

$$\left(i\gamma_0\partial^0 + i\gamma_i\partial^i - m\right)\psi(x) = 0. \tag{A.2}$$

Multiplying from the right by γ_0:

$$\begin{aligned}
\gamma_0|\left(i\gamma_0\partial^0 + i\gamma_i\partial^i - m\right)\psi(x) &= \left(i\gamma_0^2\partial^0 + i\gamma_0\gamma_i\partial^i - \gamma_0 m\right)\psi(x) \\
&= \left(i\partial^0 + i\gamma_0\gamma_0\gamma_5\sigma_i\partial^i - \gamma_0 m\right)\psi(x) \\
&= \left(i\partial^0 + i\gamma_5\sigma^i\partial^i - \gamma_0 m\right)\psi(x) \\
&= \left(i\partial^0 + i\sigma_i\partial^i\gamma_5 - \gamma_0 m\right)\psi(x) \\
\Rightarrow \left(i\partial^0 + i\sigma_i\partial^i\gamma_5 - \gamma_0 m\right)\psi(x) &= 0.
\end{aligned} \tag{A.3}$$

In the above calculation, I have made use of the identities

$$\begin{aligned}
\gamma_i &= \gamma_0\gamma_5\sigma_i \\
\gamma_o\gamma_0 &= 1 \\
\gamma_5\sigma_i &= \sigma_i\gamma_5.
\end{aligned} \tag{A.4}$$

I then multiply (A.3) by γ_5 from the right:

$$\begin{aligned}
\gamma_5|\left(i\partial^0 + i\sigma_i\partial^i\gamma_5 - \gamma_0 m\right)\psi(x) &= \left(i\gamma_5\partial^0 + i\gamma_5\sigma_i\partial^i\gamma_5 - \gamma_5\gamma_0 m\right)\psi(x) \\
&= \left(i\partial^0\gamma_5 + i\sigma_i\partial^i\gamma_5\gamma_5 - \gamma_5\gamma_0 m\right)\psi(x) \\
&= \left(i\partial^0\gamma_5 + i\sigma_i\partial^i + \gamma_0\gamma_5 m\right)\psi(x) \\
\Rightarrow \left(i\partial^0\gamma_5 + i\sigma_i\partial^i + \gamma_0\gamma_5 m\right)\psi(x) &= 0,
\end{aligned} \tag{A.5}$$

where I also needed the identity

$$\gamma_5\gamma_5 = 1. \tag{A.6}$$

Combining the two equations: (A.3) + (A.5):

$$\begin{aligned} \left(i\partial^0 + i\sigma_i\partial^i\gamma_5 - \gamma_0 m\right)\psi(x) + \left(i\partial^0\gamma_5 + i\sigma_i\partial^i + \gamma_0\gamma_5 m\right)\psi(x) = 0 \\ i\partial^0(1+\gamma_5)\psi(x) + i\sigma_i\partial^i(\gamma_5+1)\psi(x) + m\gamma_0(1-\gamma_5)\psi(x) = 0 \\ \Rightarrow i\partial^0\psi(x)_R + i\sigma_i\partial^i\psi(x)_R = m\gamma_0\psi(x)_L. \end{aligned} \tag{A.7}$$

Calculating (A.3) - (A.5):

$$\begin{aligned} \left(i\partial^0 + i\sigma_i\partial^i\gamma_5 - \gamma_0 m\right)\psi(x) - \left(i\partial^0\gamma_5 + i\sigma_i\partial^i + \gamma_0\gamma_5 m\right)\psi(x) = 0 \\ i\partial^0(1-\gamma_5)\psi(x) + i\sigma_i\partial^i(\gamma_5-1)\psi(x) - m\gamma_0(1+\gamma_5)\psi(x) = 0 \\ \Rightarrow i\partial^0\psi(x)_L - i\sigma_i\partial^i\psi(x)_L = m\gamma_0\psi(x)_R. \end{aligned} \tag{A.8}$$

Equations (A.7) and (A.8) are the results stated in equation (1.7).

In doing these calculations, [3] and [65] have been helpful.

Appendix B

Beta Decay - A Measure of the Mass

Beta decays played an important role in the discovery of the neutrino, and it is still used as a way to measure the mass of the neutrinos, as we saw in the section above. In this section I will show how one can use beta decay to show that the neutrino has a mass, and to find the scale of this mass. I have used some approximations to simplify the calculations, but the results should still give a picture of how the neutrinomass effects the energy spectrum of the electron. The reference [66] has been to great help writing this appendix.

Several approximations can be made to simplify the calculations.

1. The neutron and proton are treated as elementary particles, not composite quark particles. This is a huge oversimplification, but since I'm only really interested in the effect on the result leptons, I avoid the trouble of structure functions and strong interactions in this way.

2. The momentum of the W-boson is much smaller than the mass of the W
$$\rightarrow \frac{-ig_1g_2}{k^2 - m_W^2} \approx \frac{ig_1g_2}{m_W^2} \tag{B.1}$$

3. The released energy is very small compared to the rest energy of the proton.
→ The spinors of the nucleon particles (n and p) can be considered non-relativistic.

4. In the case of non-relativistic spinors, $\overline{u}_s\gamma^\mu u_{s'} = 2m\delta(s - s')$.
→ Assuming that the neutron and the proton has the same spin, we then get

$$\begin{aligned} \left[\overline{u}_{s_3}(\mathbf{p}_3)\gamma^\mu(1-\gamma^5)u_{s_1}(\mathbf{p}_1)\right] & \approx \left[\overline{u}_{s_3}(\mathbf{p}_3)\gamma^0 u_{s_1}(\mathbf{p}_1)\right] \\ \times \left[\overline{u}_{s_4}(\mathbf{p}_4)\gamma_\mu(1-\gamma^5)v_{s_2}(\mathbf{p}_2)\right] & \left[\overline{u}_{s_4}(\mathbf{p}_4)\gamma_0(1-\gamma^5)v_{s_2}(\mathbf{p}_2)\right] \\ & = 2\sqrt{m_n m_p}\left[\overline{u}_{s_4}(\mathbf{p}_4)\gamma_0(1-\gamma^5)v_{s_2}(\mathbf{p}_2)\right] \end{aligned} \tag{B.2}$$

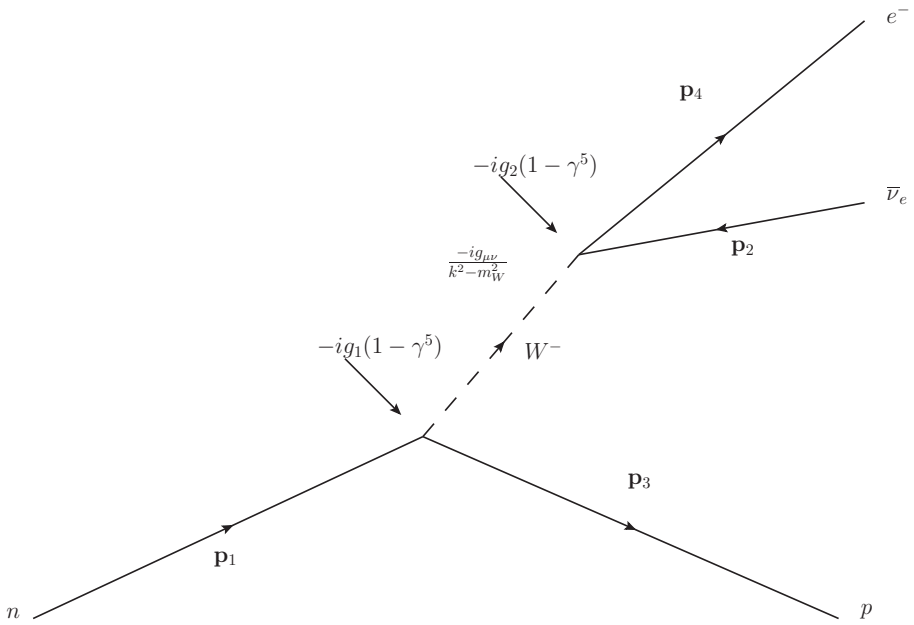


Figure B.1: Feynman diagram for β-decay.

where I have also assumed that the axial current don't contribute (parity conserved for the nucleus).

5. Eventually I will also assume that the electron and the neutrino are highly relativistic, so that $E \approx |\mathbf{p}|$, but this will only be used in the very end of the calculations.

The amplitude, as read from the Feynman diagram in figure B.1 is:

$$\begin{aligned}\mathcal{M} &= \frac{-ig_1 g_2}{k^2 - m_W^2} \left[\overline{u}_{s_3}(\mathbf{p}_3)\gamma^\mu(1-\gamma^5)u_{s_1}(\mathbf{p}_1)\right] g_{\mu\nu} \left[\overline{u}_{s_4}(\mathbf{p}_4)\gamma^\nu(1-\gamma^5)v_{s_2}(\mathbf{p}_2)\right] \\ &= \frac{-ig_1 g_2}{k^2 - m_W^2} \left[\overline{u}_{s_3}(\mathbf{p}_3)\gamma^\mu(1-\gamma^5)u_{s_1}(\mathbf{p}_1)\right] \left[\overline{u}_{s_4}(\mathbf{p}_4)\gamma_\mu(1-\gamma^5)v_{s_2}(\mathbf{p}_2)\right] .\end{aligned} \tag{B.3}$$

Imposing condition we get

$$\mathcal{M} = \frac{ig_1 g_2}{m_W^2} 2\sqrt{m_n m_p} \left[\overline{u}_{s_4}(\mathbf{p}_4)\gamma_0(1-\gamma^5)v_{s_2}(\mathbf{p}_2)\right] , \tag{B.4}$$

giving me

$$\mathcal{M}^\dagger = \frac{-ig_1 g_2}{m_W^2} 2\sqrt{m_n m_p} \left[\overline{v}_{s_2}(\mathbf{p}_2)\gamma_0(1-\gamma^5)u_{s_4}(\mathbf{p}_4)\right] \tag{B.5}$$

and finally, summing over final spin states and averaging over initial spin states (remembering that I demanded that the nucleon particles must have the same spin):

$$\begin{aligned}
|\mathcal{M}|^2 &= \frac{1}{2}\sum_{s_1}\sum_{s_2}\sum_{s_4}\mathcal{M}\mathcal{M}^\dagger \\
&= \frac{g_1^2 g_2^2}{m_W^4} 4 m_n m_p \frac{1}{2}\sum_{s_1}\sum_{s_2}\sum_{s_4}\left[\overline{u}_{s_4}(\mathbf{p}_4)\gamma_0(1-\gamma^5)v_{s_2}(\mathbf{p}_2)\right]\left[\overline{v}_{s_2}(\mathbf{p}_2)\gamma_0(1-\gamma^5)u_{s_4}(\mathbf{p}_4)\right] \\
&= \frac{g_1^2 g_2^2}{m_W^4} 4 m_n m_p \sum_{s_2}\sum_{s_4}\left[\overline{v}_{s_2}(\mathbf{p}_2)\gamma_0(1-\gamma^5)\sum_{s_4} u_{s_4}(\mathbf{p}_4)\overline{u}_{s_4}(\mathbf{p}_4)\gamma_0(1-\gamma^5)v_{s_2}(\mathbf{p}_2)\right] \\
&= \frac{g_1^2 g_2^2}{m_W^4} 4 m_n m_p \sum_{s_2}\left[\overline{v}_{s_2}(\mathbf{p}_2)\gamma_0(1-\gamma^5)\left[\gamma\cdot p_4 + m_e\right]\gamma_0(1-\gamma^5)v_{s_2}(\mathbf{p}_2)\right] \\
&= \frac{g_1^2 g_2^2}{m_W^4} 4 m_n m_p Tr\left(\gamma_0(1-\gamma^5)\left[\gamma\cdot p_4 + m_e\right]\gamma_0(1-\gamma^5)\sum_{s_2} v_{s_2}(\mathbf{p}_2)\overline{v}_{s_2}(\mathbf{p}_2)\right) \\
&= \frac{g_1^2 g_2^2}{m_W^4} 4 m_n m_p Tr\left(\gamma_0(1-\gamma^5)\left[\gamma\cdot p_4 + m_e\right]\gamma_0(1-\gamma^5)\left[\gamma\cdot p_2 - m_{\nu_e}\right]\right) \\
&= \frac{g_1^2 g_2^2}{m_W^4} 4 m_n m_p Tr\left(\gamma^0\left[\gamma^\mu p_{4\mu} + m_e\right]\gamma_0(1-\gamma^5)\left[\gamma^\nu p_{2\nu} - m_{\nu_e}\right]\right) \\
&= \frac{g_1^2 g_2^2}{m_W^4} 8 m_n m_p \left[4(p_2^0 p_4^0 - p_2\cdot p_4 + p_2^0 p_4^0) + 4 m_e m_{\nu_e}\right] \\
&= \frac{g_1^2 g_2^2}{m_W^4} 32 m_n m_p \left[E_e E_\nu + |\mathbf{p}_2||\mathbf{p}_4|\cos\theta + m_e m_{\nu_e}\right]
\end{aligned} \tag{B.6}$$

where θ is the angle between the electron and the neutrino. I have no need for the constant in front, as I am not interested in the actual decay-rate, so I let some of the constants including the couplings form one constant K. Assuming that the electron and the neutrino are highly relativistic, I get

$$|\mathcal{M}|^2 \approx K E_e E_{\nu_e} m_n m_p \left[1+\cos\theta\right]. \tag{B.7}$$

Then I make use of the equation for the decay rate:

$$\begin{aligned}
d\Gamma &= \frac{1}{2M}(2\pi)^4\delta^4(p_1 - p_2 - p_3 - p_4)\left(\frac{d^3\mathbf{p}_2}{2E_2(2\pi)^3}\right)\left(\frac{d^3\mathbf{p}_3}{2E_3(2\pi)^3}\right)\left(\frac{d^3\mathbf{p}_4}{2E_4(2\pi)^3}\right)|\mathcal{M}|^2 \\
&= \frac{(2\pi)^4}{2m_n}\delta^4(p_1 - p_2 - p_3 - p_4)\left(\frac{d^3\mathbf{p}_2}{2E_{\nu_e}(2\pi)^3}\right)\left(\frac{d^3\mathbf{p}_3}{2E_p(2\pi)^3}\right)\left(\frac{d^3\mathbf{p}_4}{2E_e(2\pi)^3}\right) \\
&\times K E_e E_{\nu_e} m_n m_p \left[1+\cos\theta\right] \\
&= K'\delta^4(p_1 - p_2 - p_3 - p_4) d^3\mathbf{p}_2 d^3\mathbf{p}_3 d^3\mathbf{p}_4 \left[1+\cos\theta\right]
\end{aligned} \tag{B.8}$$

where I have assumed that the $m_p/E_p \approx 1$. Performing the $\mathbf{p}_3$-integration gives:

$$d\Gamma = K'\delta^4(E_n - E_p - E_e - E_{\nu_e})d^3\mathbf{p}_2 d^3\mathbf{p}_4 \left[1 + \cos\theta\right]. \tag{B.9}$$

Then,it is time for the $\mathbf{p}_2$-integration:

$$\begin{aligned} d\Gamma &= K' \int \delta^4(E_n - E_p - E_e - E_{\nu_e})|\mathbf{p}_2|^2 d|\mathbf{p}_2| d\Omega d^3\mathbf{p}_4 \left[1 + \cos\theta\right]. \\ &= K'' \int \delta^4(E_n - E_p - E_e - E_{\nu_e}) dE_{\nu_e} |\mathbf{p}_2|^2 d|\mathbf{p}_2| d^3\mathbf{p}_4 \left[1 + \cos\theta\right] \\ &= K'' E_{\nu_e} |\mathbf{p}_2| d^3\mathbf{p}_4 \left[1 + \cos\theta\right] \qquad\qquad E_{\nu_e} = E_n - E_p - E_e \\ &= K''(E_n - E_p - E_e)\sqrt{(E_n - E_p - E_e)^2 - m_{\nu_e}^2}\, d^3\mathbf{p}_4 \left[1 + \cos\theta\right]. \end{aligned} \tag{B.10}$$

Then, I look at the $\mathbf{p}_4$-integral, or the integral over the momentum of the electron, and do the angular integral:

$$\begin{aligned} d\Gamma &= K''(E_n - E_p - E_e)\sqrt{(E_n - E_p - E_e)^2 - m_{\nu_e}^2} \int d|\mathbf{p}_e||\mathbf{p}_e|^2 d\cos\theta d\phi \left[1 + \cos\theta\right] \\ &= K''(E_n - E_p - E_e)\sqrt{(E_n - E_p - E_e)^2 - m_{\nu_e}^2}\, d|\mathbf{p}_e||\mathbf{p}_e|^2 \int d\cos\theta \left[1 + \cos\theta\right] \\ &= K''(E_n - E_p - E_e)\sqrt{(E_n - E_p - E_e)^2 - m_{\nu_e}^2}\, d|\mathbf{p}_e||\mathbf{p}_e|^2 \int_{-1}^{1} d\cos\theta \left(1 + \cos\theta\right) \\ &= C(E_n - E_p - E_e)\sqrt{(E_n - E_p - E_e)^2 - m_{\nu_e}^2}\, d|\mathbf{p}_e||\mathbf{p}_e|^2 \end{aligned} \tag{B.11}$$

The coupling constant is not very interesting, what matters is the behaviour (using $E \approx m$ for the heavy particles):

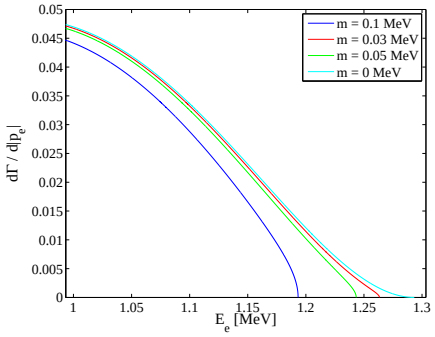
$$\begin{aligned} \frac{d\Gamma}{d|\mathbf{p}_e|} &= C|\mathbf{p}_e|^2(m_n - m_p - E_e)\sqrt{(m_n - m_p - E_e)^2 - m_{\nu_e}^2} \\ &= C|\mathbf{p}_e|^2(m_n - m_p - \sqrt{|\mathbf{p}_e|^2 + m_e^2})\sqrt{(m_n - m_p - \sqrt{|\mathbf{p}_e|^2 + m_e^2})^2 - m_{\nu_e}^2} \end{aligned} \tag{B.12}$$

If the neutrino is massless, we get

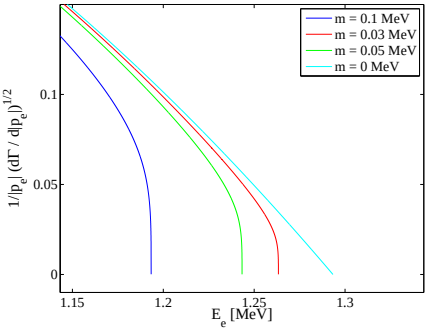
$$\frac{d\Gamma}{d|\mathbf{p}_e|} = C|\mathbf{p}_e|^2(m_n - m_p - \sqrt{|\mathbf{p}_e|^2 + m_e^2})^2. \tag{B.13}$$

From the equation above, we see that

$$\frac{1}{|\mathbf{p}_e|}\sqrt{\frac{d\Gamma}{d|\mathbf{p}_e|}} = \sqrt{C}(m_n - m_p - \sqrt{|\mathbf{p}_e|^2 + m_e^2}) = \sqrt{C}(m_n - m_p - E_e), \tag{B.14}$$



(a) $\frac{d\Gamma}{d|\mathbf{p}_e|}$ as a function of the electron energy, E_e.



(b) $\sqrt{\frac{1}{|\mathbf{p}_e|}\frac{d\Gamma}{d|\mathbf{p}_e|}}$ as a function of the electron energy, E_e.

Figure B.2: Kurie plots of the endpoints of the electron energy spectrum for varying neutrino masses.

for the case of massless neutrinos, which is linear in E_e. For the massive case, however, I have

$$\frac{1}{|\mathbf{p}_e|}\sqrt{\frac{d\Gamma}{d|\mathbf{p}_e|}} = \sqrt{C}\sqrt{(m_n - m_p - E_e)\sqrt{(m_n - m_p - E_e)^2 - m_{\nu_e}^2}} \qquad \text{(B.15)}$$

which is not linear in E_e.

From figure B.2 (a) we see that the endpoint decay-rate is a very good measure of the mass of the neutrino, as its zero point allows for a direct measure of the neutrino mass. Unfortunately, it is not entirely trivial to do such an experiment, as the decay-rate is very low at these energies, independent of the neutrinomass. Also the relation between the square-root of the decay-rate divided by the momentum has a very characteristically behaviour at the endpoint if the neutrinomass is non-zero, it is nonlinear.

Bibliography

[1] C. Giunti and C.W. Kiw. *Fundamentals of Neutrino Physics and Astrophysics*. Oxford University Press, 2007.

[2] David O. Caldwell (Ed.). *Current Aspects of Neutrino Physics*. Springer-Verlag Berlin Heidelberg, 2001.

[3] K. Zuber. *Neutrino Physics*. IoP, 2004.

[4] C. Sutton. *Spaceship Neutrino*. Cambridge University Press, 2002.

[5] D. Griffiths. *Introduction to Elementary Particles*. Wiley-vhc, 2009.

[6] C. Amsler et. al. Review of particle physics. *Physics Letters B, 667:1-6*, 2008.

[7] R. Wendell et. al. Atmospheric neutrino oscillation analysis with sub-leading effects in super-kamiokande i, ii, and iii. *hep-ex/1002.3471*, 2010.

[8] F. Boehm et. al. Results from the palo verde neutrino oscillation experiment. *Physical Review D, 62:072002*, 2000.

[9] Apollonio et. al. Limmits on neutrino oscillation from the chooz experiment. *Physic Letters B 466:415-430*, 1999.

[10] T. Araki et. al. Measurement of neutrino oscillation with kamland: Evidence of spectral disortion. *Physic Review Letters 94*, 2005.

[11] C.Weinheimer. Absolute masses of neutrinos - experimental results and future possibilites. *Physica Scripta Vol. T121*, 2005.

[12] M. Beck. The katrin experiment. *Journal of Physics: Conference series 203*, 2010.

[13] H. V. Klapdor-Kleingrothaus. First evidence for neutrinoless double beta decay - and world status of double beta experiments. *arXiv:hep-ph/0512263*, 2005.

[14] K. Ichikawa et. al. Constraining neutrino masses by cmb experiments alone. *Physiccal Reviev D 71:043001*, 2005.

[15] S. Hannestad. The connection between cosmology and neutrino physics. *arXiv:1003.4119*, 2010.

[16] T. Sekiguchi et. al. Neutrino mass from cosmology: Impact of high-accuracy measurement of the hubble constant. *arXiv:0911.0976v1*, 2009.

[17] K. Ichikawa. Neurino mass contraint from cmb and its degeneracy with other cosmological parameters. *Journal of Physics: Conference Series 120*, 2004.

[18] W. Grimus and L. Lavoura. The seesaw mechanism at arbitrary order: disentangling the small scale from the large scale. *Journal of High Energy Physics 11*, 2000.

[19] W. Grimus. Neutrino physics - model for neutrino masses and lepton mixing. *Proceedings of Science (P2GC) 001*, 2006.

[20] P.F. Perez et. al. Neutrino masses and the lhc: Testing type ii seesaw. *Physical Review D, 78:015018*, 2008.

[21] R. N. Mohapatra and P. Pal, B. *Massive neutrinos in physics and astrophysics, Third edition*. World Scientific Publishing, 2004.

[22] G. Dvali and A. Y. Smirnov. Probing large extra dimensions with neutrinos. *Nuclear Physics B, 563:63-81*, 1999.

[23] I. K. Wehus. Ekstra dimensjoner og kosmologi. Master's thesis, Universtitetet i Oslo, 2001.

[24] S. Dodelson. *Modern Cosmology*. Academic press, 2003.

[25] Ø. Grøn. *Lecture notes on the general theory of relativity*. Springer, 2009.

[26] Ø. Elgarøy. Ast 4220- cosmology i. http://www.uio.no/studier/emner/matnat/astro/AST4220/h09/pensumliste.xml, 2009.

[27] Ø. Grøn and S. Hervik. *Einstein's general theory of relativity, with modern applications in cosmology*. Springer, 2007.

[28] E.W. Kolb and M. S. Turner. *The early universe*. Perseus Publishing, 1994.

[29] A. R. Liddle and D. H. Lyth. *Cosmological Inflation and Large-Scale Structure*. Cambridge University Press, 2006.

[30] Ø. Elgarøy. Skalarfelt og kvantefluktuasjoner. http://www.uio.no/studier/emner/matnat/astro/AST5220/v09/undervisningsplan.xml, 2009.

[31] J. R. Kristiansen. Massive neutrinos and cosmology. Master's thesis, University of Oslo, 2006.

[32] Ø. Elgarøy. Initialbetingelser: I begynnelsen var ϕ. http://www.uio.no/studier/emner/matnat/astro/AST5220/v09/undervisningsplan.xml, 2009.

[33] A. Lewis et. al. Efficient computation of cmb anisotropies in closed frw models. *Astrophysical Journal, 538:473*, 2000.

[34] B. Ryden. *Introduction to cosmology*. Addison Wesley, 2003.

[35] Lucchin F. Coles, P. *Cosmology, The origin and evolution of cosmic structure*. Wiley, 2003.

[36] J. Lesgourgues and S. Pastor. Massive neurinos and cosmology. *Physical reports, 429:307-379*, 2006.

[37] D. J. Eisenstein et al. Detection of the baryon acoustic peak in the large-scale correlation function of sdss luminous red galaxies. *Astrophysical Journal, 633:560-574*, 2005.

[38] D. J. Eisenstein et al. The robustness of the acoustic scale in the low-redshift clustering of matter. *The Astrophysical Journal, 644:660-674*, 2007.

[39] W. J. Percival et al. Measuring the baryon acoustic oscillation scale using the sdss and 2sfgrs. *Monthly Notices of the Royal Astronomical Society, 381:1053-1066*, 2007.

[40] B. A. Bassett and R. Hlozek. Baryon acoustic oscillations. *Dark Energy, Ed. P. Ruiz-Lapuente*, 2009.

[41] S. Pastor. Massive neutrinos and cosmology. *arXiv:hep-ph/0505148*, 2005.

[42] Ø. Elgarøy and O. Lahav. Neutrino masses from cosmological probes. *New Journal of Physics, 7:61*, 2005.

[43] G .Mangano et. al. Relic neutrino decoupling including flavour oscillations. *Nuclear Physics B729:221-234*, 2005.

[44] S. Hannestad and J. Madsen. Neutrino decoupling in the early universe. *Physical Review D, 52:1764*, 1995.

[45] S. Hannestad. Oscillation effects on neutrino decoupling in the early universe. *Physical Review D, 65:083006*, 2002.

[46] J. C. Mather et. al. Calibrator design for the cobe far infrared absolute spectrophotometer (firas). *Astrophysical Journal, 512*, 1999.

[47] H. Dahle. The cluster mass function from weak gravitational lensing. *Astrophysical Journal, 653:954-962*, 2006.

[48] Kristiansen et. al. Using the cluster mass function from weak lensing to constrain neutrino masses. *Physical Review D, 75:083510*, 2007.

[49] W. Hu et. al. Cosmic microwave background observables and their cosmological implications. *The Astrophysical Journal, 549:669*, 2001.

[50] A. Lewis and S. Bridle. Cosmological parameters from cmb and other data: a monte-carlo approach. *Physical Reviw D, 66:103511*, 2002.

[51] A.N. Taylor and T.D. Kitching. Analytical methods for cosmological likelihoods. *astro-ph/10031136*, 2010.

[52] W. H. Press et. al. *Numerical Recipies, The Art and Science of Compunting - Third edition*. Cambridge university press, 2007.

[53] D. Larson et. al. Seven-year wilkinson microwave anisotropy probe (wmap) observations: Power spectra and wmap-derived parameters. *arxiv:1001.4635*, 2010.

[54] Legacy Archive for Mickrowave Bacground Data Analysis. Wmap 7 year data. *http://lambda.gsfc.nasa.gov/*.

[55] C.L. Kuo et. al. High resolution observations of the cmb power spectrum with acbar. *Astrophysical Journal, 600:1200-1219*, 2004.

[56] C.L. Kuo et. al. Improved measurements of the cmb power spectrum with acbar. *Astrophysical Journal, 664:687*, 2006.

[57] B. P. Crill et. al. Boomerang: A balloon-borne millimeter wave telescope and total power receiver for mapping anisotropy in the cosmic microwave background. *Astrophysical Journal Supplements, 148:527-541*, 2003.

[58] B. A. Reid et. al. Cosmological constraints from the clusteriog of the sloan digital sky survey dr7 luminous red galaxies. *arxiv:09071659*, 2009.

[59] N. Jarosik et.al. Seven-year wilkinson microwave anisotropy probe (wmap) observations: Sky maps, systematic errors, and basic results. *arXiv:1001.4744v1*, 2010.

[60] D. N. Spergel et.al. Wilkinson microwave anisotropy probe (wmap) three year results: Implications for cosmology. *Astrophysical Journal Suplement Series, 170:377-408*, 2009.

[61] A. Slosar. Detecting neutrino mass differences with cosmology. *Physical Review D, 73:123501*, 2006.

[62] J. Lesgourgues. Probing neutrino masses with cmb lensing extraction. *Physical Review D, 73:045021*, 2006.

[63] F. De Bernardis et. al. Determining the neutrino mass hierarchy with cosmology. *Physical Review D, 80:123509*, 2009.

[64] R. Jimenez et. al. Can we measure the neutrino mass hierarchy in the sky? *arXiv:1003.5918*, 2010.

[65] W. N. Cottingham and D. A. Greenwood. *An Introduction to the Standard Model of Particle Physics*. Cambridge university press, 2007.

[66] F. Halzen and A. D. Martin. *Quarks & Leptons: An Introductory Course in Modern Particle pHysics*. John Wiley & Sons, 1984.