O. Koop

Detecting Extragalactic Double White Dwarfs

With the Laser Interferometer Space Antenna

Bachelor thesis

29 july 2018

Thesis supervisors: Dr. E.M. Rossi
Dr. R.I. van der Veen
V. Korol

Leiden University
Sterrewacht and Mathematical Institute

Abstract

While many articles have been published about characterizing extragalactic double systems of white dwarfs as source of noise for observations with LISA, it would also be of interest to actually observe these targets. We will set up the theory of Gravitational Waves from scratch, to provide a more mathematically accurate insight behind the theory of curvature and the Riemann tensor than is provided in general reading material for physicists. This also allows for a more concise handling of the calculations we perform for binary systems. We find the maximal distance at which we can still observe a binary system with the properties of $J065133.34 + 284423.4$ (shortened to $J06$), and show its dependence on location. By restricting our search to the nearby galaxy Andromeda, we predict characteristics of visible double white dwarfs in that galaxy. This resulted in a prediction of 4 visible double white dwarfs in the Andromeda after observing for 4 years, which grows to 30 after 10 years. In the future, predictions could also be made for populations in other galaxies or for individual double white dwarfs in the Local Group.

Contents

Introduction

After Henri Poincaré proposed, and Albert Einstein predicted Gravitational Waves (GWs) on basis of the general theory of relativity, many scientists have been considering a possible way to detect this gravitational radiation. By studying the exact effect theoretical GWs would have on the surrounding space, one way of observing these waves was found to be laser interferometry. This has been used in building the many GW observatories across the globe (for example LIGO or VIRGO). In september of 2015, the first GW (due to a Black Hole merger) was detected with LIGO [1], and using the availability of more observatories, even more events have been recorded since then [6].
Since 1980, studies for designs for a GW detector in space were performed. LISA was first proposed to ESA in 1990, and was pushed forward since then. In 2013, ESA announced that the theme 'the Graviataional Universe' would be selected for a 2034 launch. A mission proposal for LISA was submitted january 2017[2].
More than 2000 papers have been written describing the theory about detections with LISA. Many of these (for example [7], [19] and [15]) focus on the primary scientific objectives named in the proposal [2], and thus consider extragalactic binary systems as source of noise. This thesis will explore the possibility of actually detecting GWs emitted by binary systems of White Dwarfs (double white dwarfs (DWD)) in satellite galaxies of the Milky Way (like the Small and Large Magellanic Clouds), or even the nearby galaxy Andromeda (M31).
Being able to observe GWs from sources in the Andromeda galaxy would provide more information about a possible star formation history of the galaxy. We would also be able to determine the distance to the parts of Andromeda where the systems are situated more precisely. By also observing and mapping intergalactic DWDs, we can gain a better understanding of the formation and history of the structures in our Local Group.
To theoretically determine the observation limit, we have to explore the mathematics behind the curvature of spacetime due to GWs and use that theory to predict the signal received from extragalactic DWD systems.

The LISA-mission will detect GWs with an interferometric measurement of a difference in optical pathlength (due to geodesic deviation (see chapter 1)) along three sides of a triangular configuration defined by three free-falling test masses, contained inside co-orbiting drag-free spacecrafts. The variations caused by GWs are of orders of pico- or nanometers, while variations due to solar system celestial dynamics are of the order of tens of thousands of kilometers. Still the GWs can be distinguished because GWs have periods at mHz frequencies, while the latter variations have periods of months, so they are quiet at mHz frequencies [2]. All three components will send out laser beams to their neighbours, such that they send a signal back (phase-locked with the incoming signal) along with a fixed offset frequency. This set-up allows LISA to be viewed as two independent virtual Michelson interferometers (making it possible to view both polarizations of GWs), along with a third 'Sagnac' configuration to characterize instrumental noise background. This adds up to six active laser links.

Due to the rotation of the constellation about itself and the sun, the location of a source observed for several weeks can be reconstructed.
Per spacecraft, two test masses (TMs) are used, each one belonging to one interferometric arm. To limit the relative accelerations of these TMs and the spacecrafts, the spacecrafts will be endowed with thrusters to follow the TM along the interferometry arm, without forces applied to the TM along these axes.
Due to instrumental noise and other sources, the strain sensitivity in the frequency domain where we want LISA to observe GWs (10^{-5}Hz$\leq f \leq$ 1Hz) is seen in figure 1. The dots in this figure represent the signal from $J06$, the double white dwarf we will use as an example througout. In figure 2 of [2] we see the various theoretical frequencies with their strain sensitivity of possible objects we want to study with LISA, including LIGO-type black holes, Galactic Binaries and the stochastic background, both Cosmological and Astrophysical. As seen, due to the ability of characterizing the noise from foregrounds, the strain due to astrophysical sources will decay while time advances.
To now find the orbital average, we need to multiply with $\sqrt{3/20}$. This will be used to find the SNR for an observation of a binary system.
In the proposal it can be read that it is proposed to reach a nominal mission lifetime of 4 years. It is predicted that it is also possible that LISA will observe for 10 years or longer. These time-scales will thus be considered mainly throughout this thesis.

Double White Dwarfs (DWD) are systems where two white dwarfs orbit each other. Because these are both compact objects, and the configuration is asymmetric, they produce Gravitational Waves during this process. Because these systems are very common (about $2.6 \cdot 10^7$ expected in the Milky Way alone [14],[13]), they make for great targets to verify Einstein's theory of Gravitation and to verify the feasibility of LISA when launched. Furthermore, we can easily predict physical effects in DWD-systems because the system remains predictable for most of its life. Only when the inspiral phase has progressed such that tidal effects have an impact on the system and the orbital speed is not negligible to the speed of light anymore, the predicted theories do not hold. These events are thus very interesting to try to observe, for new information could be found. Several assumptions have to be made before we simplified the theory enough to use it. These involve the *slow motion approximation* and the *weak field approximation*, as well as the assumption that the orbit is circular and the masses can be viewed as point masses.
These assumptions will not hold during the inspiral phase, so constraints on parameters can be found to limit our search. When for example the orbital separation is too short, tidal forces will play a role, so the point-mass approximation does not hold anymore. For heavier white dwarfs, the radius is smaller, thus the orbital period can become much shorter. More massive objects with a shorter period send out signal with bigger amplitudes, so they become more easily visible.
It should be noted that white dwarfs can not have a mass higher than $1.4M_\odot$ ($M_\odot$ is the solar mass), known as the Chandrasekhar limit[4].

The first chapter of this thesis will focus on the mathematics of manifolds and curvature. By defining what we mean with a spacetime and using tensor fields, we define the Riemann tensor, which is the main characterization of curvature of manifolds.
In the second chapter, it will be shown how the Riemann tensor of our spacetime is connected to the presence of matter and energy in the universe by exploring the principles of General Relativity.
The third chapter will then limit the theory to Gravitational Waves (GWs) as solution of Einstein equations in vacuum. The quadrupole formalism will be explored to derive the form of GWs emitted by binary systems. Finally, a discussion of the theory of signals and systems will be performed to

analyse the technique of matched filtering to determine the optimal signal to noise ratio (SNR) of a detection of a GW.
The theory will be used to analyse the maximal distance for observing binary systems of white dwarfs in chapter 4. The same techniques will be applied to a simulated population of binary white dwarfs in the Milky Way, moved computationally to M31, to analyse the expected properties of visible systems in M31.
A conclusion with possible follow-up research will be presented at the end.

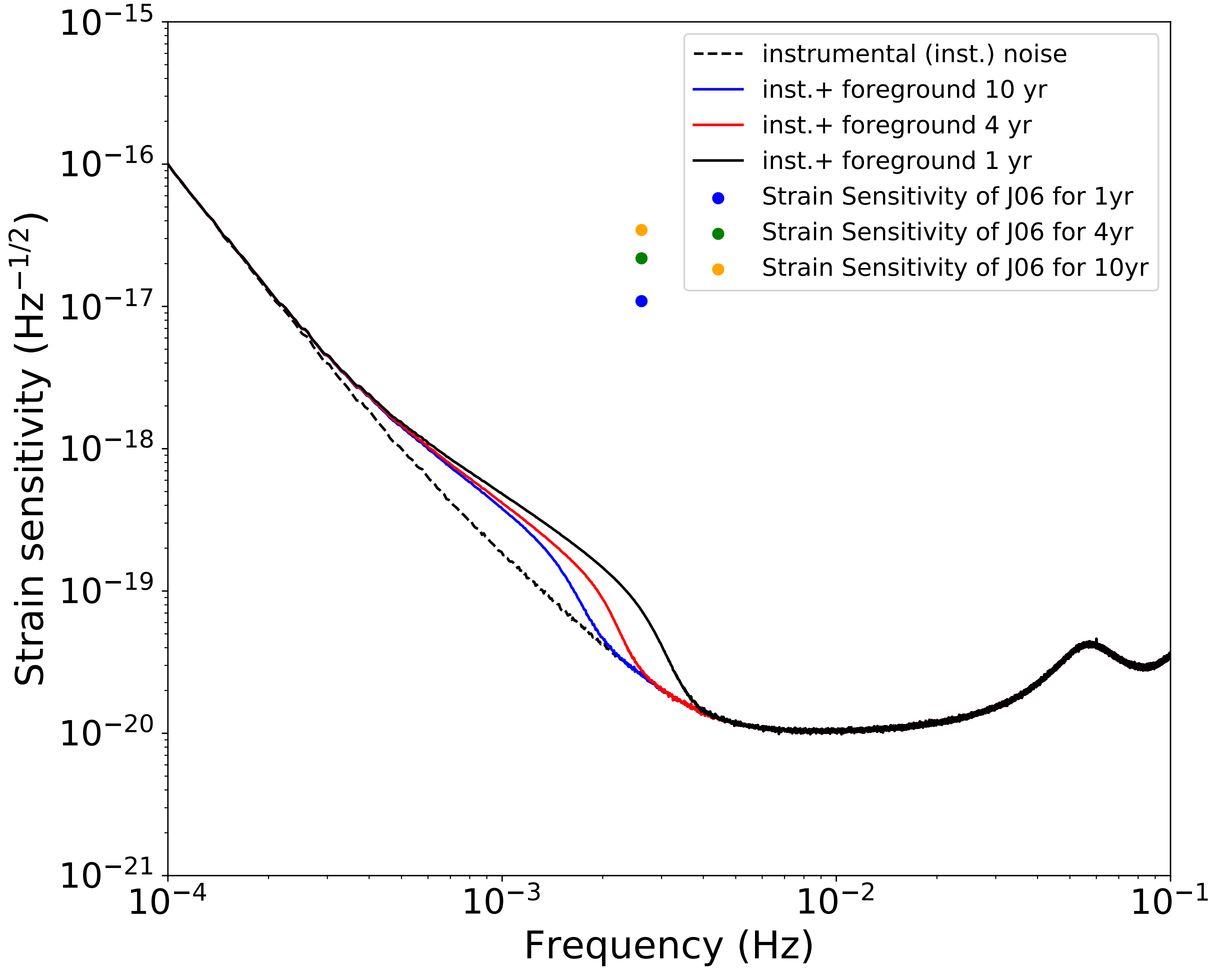


Figure 1: Strain Sensitivity for LISA. The dashed line signifies the effects of instrumental (abbreviated to inst.) noise, while the solid lines are the instrumental strain with addition of foreground strain due to astrophysical sources. These are shown for 1, 4 and 10 years of observing time. Also shown for 1, 4 and 10 year of observing time are the strain sensitivities of the Double White Dwarf J06. The data was taken from [2].

Chapter 1

Curvature of Manifolds

Before we can start our discussion on Gravitational Waves, we need to rigorously describe the relevant elements of the theory of General Relativity. This we will do by studying the mathematical definition of curvature, where the Riemann tensor will appear to be the most important quantity. To do this, let us first start by describing what we mean by a manifold.

1.1 Manifolds

The definition of a manifold used throughout will be a generalisation of the notion of a submanifold of $\mathbb{R}^n$, dependent on an atlas for a topological space X:

Definition 1.1.1. An n**-dimensional topological atlas** $\mathcal{A}$ for X is a set $\mathcal{A} = \{(U_i, h_i, V_i) | i \in I\}$ with I an index set, such that for each $i \in I$ the following statements hold: U_i is open in X and $X = \bigcup_{i \in I} U_i$, V_i is open in $\mathbb{R}^n$ and $h_i : U_i \to V_i$ is a homeomorphism.
This atlas is **differentiable** if for all $i, j \in I$ the map

$$h_j \circ h_i^{-1} : h_i(U_i \cap U_j) \to h_j(U_i \cap U_j) \tag{1.1}$$

is differentiable[1] as function from $\mathbb{R}^n$ to $\mathbb{R}^n$.
It is **maximal** if for any other atlas $\mathcal{B}$ for which $\mathcal{A} \cup \mathcal{B}$ is another atlas it holds that $\mathcal{B} \subseteq \mathcal{A}$.

Now the definition we will use for a manifold is as follows:

Definition 1.1.2. An n**-dimensional differentiable manifold** is a pair $(X, \mathcal{M})$ where X is such that every open can be written as union of a selection of opens in a countable collection of open sets and such that every two points admit disjoint open neighbourhoods (Hausdorff). $\mathcal{M}$ needs to be a maximal n-dimensional differentiable atlas for X.

The Hausdorff condition is made such that we can always represent the manifold locally by finding a mapping to a Euclidean space $\mathbb{R}^n$.

Example 1.1.1.

[1] With "differentiable" we mean that it is an element of $C^\infty(\mathbb{R}^n)$, where $C^\infty(X)$ denotes the set of functions on X that are n times differentiable for all $n \in \mathbb{N}$.

- The Euclidean space $\mathbb{R}^n$ combined with the atlas $\mathcal{M} = \{(U, h, V) : U, V \subset \mathbb{R}^n$ open and $h : U \to V$ a homeomorphism$\}$ is an n-dimensional differentiable manifold.

- The 2-dimensional sphere $S^2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\} \subset \mathbb{R}^3$ is another example of a 2-dimensional manifold. This is seen by letting $(U_i^\pm, h_i^\pm, V_i)$ be the charts given with $U_i^\mp = \{(x_1, x_2, x_3) \in S^2 : x^i \lesseqgtr 0\}$, $V_i = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 < 1\}$ and $h_i^\pm = (x_1, ..., \hat{x_i}, ..., x_3)$, meaning projection on the i-th coordinate (the element with the hat is omitted). The inverse of $h_i^\pm$ are:

$$(h_i^\pm)^{-1}(x_1, x_2) = \left(x_1, ..., x_{i-1}, \pm\sqrt{1 - \sum_{j=1}^{n} x_j^2}, x_i, ..., x_3 \right). \tag{1.2}$$

  The functions $h_j^\pm \circ (h_i^\pm)^{-1}$ are given as:

$$h_j^\pm \circ (h_i^\pm)^{-1}(x_1, ..., x_3) = \left(x_1, ..., \hat{x_j}, ..., x_{i-1}, \pm\sqrt{1 - \sum_{k=1}^{n} x_k^2}, x_i, ..., x_3 \right), \tag{1.3}$$

  when $j < i$. Of course a similar formula holds for $i > j$, and when $i = j$ the map is just id_{V_i}, the identity mapping on V_i. These are all differentiable components, so the atlas consisting of these charts is differentiable. ■

1.2 Tensors

We aim to define the notion of a tensor field for our manifold. Because of the way a tensor field is defined, it will be independent of chosen coordinates, and thus we can describe the curvature of a manifold in terms of a tensor field, called the Riemann tensor (where we refer to the tensor field with the term 'tensor'). This will make for an intrinsic definition of curvature, which we need for General Relativity, for we can not collect extrinsic data about spacetime, because we can not make observations from an embedding space of higher dimension. We need to introduce some concepts before we can talk about tensor fields, but let us start with the definition of a tensor. The concepts named in that definition will be explored in this chapter.
The definition that we will use for a tensor is:

Definition 1.2.1. A **tensor** of type (k, l) on a manifold M in a point $p \in M$ is a multilinear map $T : T_pM^* \times ... \times T_pM^* \times T_pM \times ... \times T_pM \to \mathbb{R}$ (the first is k times, the second l times). Here T_pM is the **tangent space** at one point $p \in M$ and T_pM^* its **dual space**.

This means that a tensor takes as arguments k *one-forms* and l *vectors*. We will now introduce these concepts more clearly.

An important notation that will be used is the **Einstein summation convention**. This means that an index that is present in an equation both as upper and as lower index implies a sum, for example for two vectors in $\mathbb{R}^n$, $\vec{V} = (V^1, V^2, ..., V^n)$ and $\vec{W} = (W_1, W_2, ..., W_n)$, their standard inner product can be written as:

$$\langle \vec{V}, \vec{W} \rangle = V^\mu W_\mu = \sum_\mu V^\mu W_\mu, \tag{1.4}$$

where $\langle -, - \rangle$ denotes the standard inner product on an n-dimensional Euclidean space.

1.2.1 Tangent Space

For a general manifold M, we can define the tangent space to that manifold in a point $p \in M$. The most intuitive way is dependent on parametrised curves in the manifold. We will follow the lines of [17].

Definition 1.2.2. For M an n-dimensional differentiable manifold a **differentiable curve in** M **through** p is a mapping $\gamma : I := (-\epsilon, \epsilon) \to M$ for $\epsilon > 0$ with $\gamma(0) = p$. The set of all these curves is called $\mathcal{K}_p$.

Let (U, h, V) be a chart for M around p. For $\gamma \in \mathcal{K}_p$ and suitable $0 < \delta \leq \epsilon$ we obtain a differentiable curve $h \circ \gamma : (\delta, \delta) \to V$ and define:

$$\dot{\gamma}(0)_h := \frac{d}{dt}(h \circ \gamma)(0) \in \mathbb{R}^n. \tag{1.5}$$

Definition 1.2.3. Two curves $\gamma_1, \gamma_2 \in \mathcal{K}_p$ are **equivalent** ($\gamma_1 \sim \gamma_2$) if $\dot{\gamma_1}(0)_h = \dot{\gamma_2}(0)_h$.

This equivalence is proven to be well-defined in [17] (Lemma 2.3.1). Now we can define:

Definition 1.2.4. The **tangent space** to a manifold M at a point $p \in M$ is the quotient[2]:

$$T_pM = \mathcal{K}_p / \sim .$$

The equivalence class of $\gamma \in \mathcal{K}_p$ is denoted as $[\gamma] \in T_pM$, and is called a **tangent vector** to M at p.

As we know from Linear Algebra, the tangent space in every point of the manifold $\mathbb{R}^n$ is homeomorphic to $\mathbb{R}^n$ again. The tangent space to a point $p \in S^2$ can be identified as the subspace of $\mathbb{R}^3$ that is orthogonal to the radial unit vector through p with respect to the normal Euclidean inner product.
The tangent space has a natural structure of an n-dimensional vector space, as can be seen by the fact that the map $\Phi_h : \mathbb{R}^n \to T_pM$ is bijective and that the structure on T_pM is independent of the chosen chart. This is shown in [17] (Theorem 2.3.3). We define Φ_h for $v \in \mathbb{R}^n$ to be:

$$\Phi_h(v) := [\gamma_v] \in T_pM, \tag{1.6}$$

where $\gamma_v : (-\epsilon, \epsilon) \to M$ is defined as $\gamma_v(t) := h^{-1}(h(p) + t \cdot v)$, for $t \in (-\epsilon, \epsilon)$. Such an ϵ always exists because V is open, so indeed $h(p) + t \cdot v \in V$ for all $t \in (-\epsilon, \epsilon)$.
We now define:

Definition 1.2.5. For the unit basis $(e_i)_{i=1,\ldots,n}$[3] of $\mathbb{R}^n$ the **partial derivative** is defined as:

$$\partial_i := \frac{\partial}{\partial x^i}(p) := \Phi_h(e_i), \tag{1.7}$$

where Φ_h is as in (1.6).

[2]The elements of the set $\mathcal{K}_p / \sim$ are the **equivalence classes** of curves. These equivalence classes contain all curves that are equivalent to each other, and the equivalence class of γ is denoted by $[\gamma]$. If γ is equivalent to τ, we can also denote the class by $[\tau]$, so all elements of a class can be a **representative** of the class.

[3]From here on forth, the indices for a basis are implied by the dimensionality, so we do not write them explicitly.

The notation we will use is thus that a vector can be written as:

$$\vec{V} := V^\alpha \partial_\alpha \tag{1.8}$$

where (∂_α) is the natural basis associated with a chosen basis (x^α). Here the V^α are the components of $\vec{V}$ w.r.t. the chosen basis. These components can be found by finding the derivative of the respective coordinates of a curve of which $\vec{V}$ is the tangent vector in $t = 0$ after mapping it to an appropriate coordinate system. With use of these vectors, we can introduce the concept of a partial derivative of a mapping:

Definition 1.2.6. The **partial derivative** in $p \in M$ of a map $f : M \to N$ between manifolds is:

$$\frac{\partial f}{\partial x^\mu} := \frac{\partial}{\partial x^\mu}(\psi \circ f \circ \phi^{-1}), \tag{1.9}$$

where ϕ is a chart on M containing p, and ψ a chart on N containing $f(p)$.

It is of course of more general importance to observe mappings between manifolds. Let us define:

Definition 1.2.7. The **differential** f_* of a map $f : M \to N$, where M is an n-dimensional differentiable manifold and N an m-dimensional one, is the linear transformation $f_* : T_pM \to T_{F(p)}N$ defined as follows. For $x \in T_pM$ let $\gamma = \gamma(t)$ be a curve on M with $\gamma(0) = p$ and $\dot{\gamma}(0) = x$. Then $f_*[\gamma] = [f \circ \gamma]$.

Note that this is independent of the chosen curve, as long as it satisfies the given constraints. For more details on differentials and intuition behind vectors and their components, see [9]. The matrix of this linear transformation in terms of the bases (∂^x_α) at p and (∂^y_α) at $f(p)$ is exactly the **Jacobian matrix** $(f_*)^j_i = \dfrac{\partial f^j}{\partial x^i}(p) = \dfrac{\partial y^j}{\partial x^i}(p)$ known from analysis.

In order now to define the dual space, we first define the elements of such a vector space:

Definition 1.2.8. A **one-form** in a point $p \in M$ is a mapping $\omega : T_pM \to \mathbb{R}$ that is linear, so $\omega(\lambda v + u) = \lambda\omega(v) + \omega(u)$ for $u, v \in T_pM$ and $\lambda \in \mathbb{R}$.
The collection of all one-forms is the **dual or cotangent space** to T_pM, denoted by T_pM^*, with operations (for $\alpha, \beta \in T_pM^*$, $c \in \mathbb{R}$ and $v \in T_pM$):

$$(\alpha + \beta)(v) := \alpha(v) + \beta(v); \quad (c\alpha)(v) = c\alpha(v). \tag{1.10}$$

If a basis $\partial_1, ..., \partial_n$ for T_pM is given, the basis $\mathbf{d}x^1, ..., \mathbf{d}x^n$ of T_pM^* is called the **dual base** if $\mathbf{d}x^i(\partial_j) = \delta^i_j$[4].

Remark.

- The dual basis exists and is unique. (Proposition 6.3 from [18]).

- The $\mathbf{d}x^i$ act on a vector $v \in T_pM$ by giving the i-th component. This is seen by writing:

$$\mathbf{d}x^i(\partial_j v^j) = \mathbf{d}x^i(\partial_j)v^j = \delta^i_j v^j = v^i. \tag{1.11}$$

[4]This is the Kronecker-delta, for which it holds that:

$$\delta^i_j = \begin{cases} 1 & \text{if } i = j \\ 0 & if i \neq j \end{cases}$$

- We can write for any linear form that $\omega = \omega_\alpha \mathbf{d}x^\alpha$. This then means that $\omega(v) = \omega_\alpha v^\alpha$ for $\omega \in T_pM^*$ and $v \in T_pM$. ▲

Now remember the definition of tensors given at the start of this section. The integer $k+l$ is called the **rank** of a tensor. A few examples are a vector, which is a type $(1,0)$ tensor, a one-form, which is a tensor of type $(0,1)$ and a **bilinear form**, a tensor of type $(0,2)$. A bilinear form maps couples of vectors to real numbers in a linear way for each vector. One example of a bilinear form is the common inner product in $\mathbb{R}^n$.
Note that the space of all (k,l) tensors defined for a $p \in M$ is a vector-space connected to that p, with addition and scalar multiplication component-wise, comparable to vector spaces of matrices.

Definition 1.2.9. A smooth k-dimensional **vector bundle** of a manifold M is a pair of smooth manifolds E and M, along with a surjective mapping $\pi : E \to M$, for which the following conditions hold:

(i) For every $p \in M$, the set $E_p = \pi^{-1}(p)$ (the **fiber** of E over p) has the structure of a vector space.

(ii) For every $p \in M$, there exists an open set $p \in U$ and a diffeomorphism $\phi : \pi^{-1}(U) \to U \times \mathbb{R}^k$, a **local trivialization**, such that for the projection on the first coordinate $\pi_1 : U \times \mathbb{R}^k \to U$ it holds that $\pi_1 \circ \phi = \pi$.

(iii) The restriction of ϕ to each fiber should be a linear isomorphism.

Example 1.2.1.

- The so-called **tangent bundle** TM of a manifold M (the disjoint union of the tangent space T_pM for all $p \in M$) and the **cotangent bundle** T^*M, the disjoint union of all cotangent spaces T_pM^* for all $p \in M$ are vector bundles.
- By taking the union for all $p \in M$ of $\left(T^k_l\right)_p M$, the **type** (k,l) **tensor space**, is another example of a vector bundle, denoted by $T^k_l M$. ■

There is a natural way to define the product between two tensors such that the resulting object is a tensor again, but with other dimensions. For two tensors $F \in T^k_l M$ and $G \in T^p_q M$ the tensor $T \otimes G \in T^{k+p}_{l+q} M$ is defined such that for one-forms ω^i and vectors v_i:

$$F \otimes G(\omega^1, ..., \omega^{l+q}, v_1, ..., v_{k+p}) = F(\omega^1, ..., \omega^l, v_1, ..., v_k) G(\omega^{l+1}, ..., \omega^{l+q}, v_{k+1}, ..., v_{k+p}). \tag{1.12}$$

Given a basis (∂_α) of T_pM with its dual basis $(\mathbf{d}x^\alpha)$ of T_pM^* we can construct a basis for the vector space of (k,l)-tensors in the point p by using this tensor product defined as the type (k,l)-tensor for which the image of $(\omega^1, ..., \omega^k, v_1, ..., v_l)$ is equal to:

$$(\partial_{\alpha_1} \otimes ... \otimes \partial_{\alpha_k} \otimes \mathbf{d}x^{\beta_1} \otimes ... \otimes \mathbf{d}x^{\beta_l})(\omega^1, ..., \omega^k, v_1, ..., v_l) = \prod_{i=1}^{k} \prod_{j=1}^{l} \omega_i(\partial_{\alpha_i}) \cdot \mathbf{d}x^{\beta_j}(v_j). \tag{1.13}$$

This means that we can write for a tensor T of type (k,l):

$$T = T^{\alpha_1 ... \alpha_k}_{\beta_1 ... \beta_l} \partial_{\alpha_1} \otimes ... \otimes \partial_{\alpha_k} \otimes \mathbf{d}x^{\beta_1} \otimes ... \otimes \mathbf{d}x^{\beta_l}. \tag{1.14}$$

Here the $T^{\alpha_1 ... \alpha_k}_{\beta_1 ... \beta_l}$ are the **components of the tensor T w.r.t. the basis** (∂_α). These are unique and fully characterize the tensor, which is just a generalization of the fact that all vectors can uniquely be represented w.r.t. a basis of the vector space they are an element of.

Remark. Tensors transform under a very specific law when performing coordinate transformations in the form of $\partial'_l = \left(\dfrac{\partial x^s}{\partial x'^l}\right)\partial_s$ and $\mathbf{d}x'^i = \left(\dfrac{\partial x'^i}{\partial x^c}\right)\mathbf{d}x^c$. Here we denote all objects defined w.r.t. the new coordinates primed.
This transformation law follows from the multilinearity of the tensor, so if:

$$W^{i...j}_{k...l} = W(\mathbf{d}x^i, ..., \mathbf{d}x^j, \partial_k, ..., \partial_l), \tag{1.15}$$

then:

$$\begin{aligned} W'^{i...j}_{k...l} &= W(\mathbf{d}x'^i, ..., \mathbf{d}x'^j, \partial'_k, ..., \partial'_l) \\ &= \left(\frac{\partial x'^i}{\partial x^c}\right) ... \left(\frac{\partial x'^j}{\partial x^d}\right)\left(\frac{\partial x^r}{\partial x'^k}\right) ... \left(\frac{\partial x^s}{\partial x'^l}\right) W^{c...d}_{r...s}. \end{aligned} \tag{1.16}$$

Note that thus, tensors are by construction invariants with respect to coordinates, which is one of the main reasons Einstein developed his theory of relativity using tensor analysis, see chapter 2. ▲

1.2.2 Tensor fields and the Metric tensor

We can describe a general inner product on a vector space V with a symmetric and non-degenerate type (0,2) tensor, by writing $\langle v, w\rangle$ for the inner product between $v, w \in V$, we see with a basis (e_i):

$$\langle v, w\rangle = \langle \sum_i e_i v^i, \sum_j e_j w^j\rangle = \sum_i \sum_j v^i \langle e_i, e_j\rangle w^j := \sum_{i,j} v^i g_{ij} w^j := vGw. \tag{1.17}$$

Here $G = (g_{ij})_{i,j=1,...,n}$ is a matrix, also called the **metric tensor**.

We can now describe the inner product on a manifold with a so-called metric tensor field, which generalizes the notion above. Before we talk about a tensor field, we need to define a section:

Definition 1.2.10. If $\pi : E \to M$ is a vector bundle over M, a **smooth section** of E is a smooth map $\sigma : M \to E$ (smooth as map between manifolds) such that $\pi \circ \sigma = \mathrm{id}_M$, thus $\sigma(p) \in E_p$ for all p.

Remark. If we define a **smooth section along a smooth curve** $\gamma : I \to M$, with I an open interval, we need that $\pi(\sigma(t)) = \gamma(t)$ for all $t \in I$ holds as well. ▲

Notation: The space of all smooth sections of a vector bundle E will be denoted by $\mathcal{E}(M)$. Because it is so common, the space of all smooth sections of the tensor bundle $T^k_l M$ is denoted by $\mathcal{T}^k_l(M)$.

Definition 1.2.11. A **tensor field of type** (k, l) on a manifold M is a map that assigns a tensor to each point $p \in M$ in a differentiable manner, meaning that as functions from $\mathbb{R}^n \to \mathbb{R}$, the components are differentiable.
In other words, it is a smooth section of a type (k, l) tensor bundle $T^k_l M$, or an element of $\mathcal{T}^k_l(M)$.

Note that vector fields can be described the same way as these tensor fields, because vectors are type $(1, 0)$ tensors.
For vector fields we can define the **commutator of two vector fields** by how it acts on a smooth function between manifolds $f : M \to N$:

$$[X, Y](f) = X(Y(f)) - Y(X(f)). \tag{1.18}$$

How then the commutator acts on other objects is defined component-wise.
Now we can introduce the notion of the metric tensor:

Definition 1.2.12. A **pseudo-Riemannian metric tensor** is a tensor field of type $(0, 2)$ such that for $p \in M$ and $v, u \in T_pM$ it is symmetric, so $g(u, v) = g(v, u)$, and non-degenerate, so $g(u, v) = 0$ for all v iff[5] u is the null vector.

A **pseudo-Riemannian manifold** is now a couple (M, g) where M is a differentiable manifold and g is a pseudo-Riemannian metric tensor on M.

Example 1.2.2. The Minkowski metric on $\mathbb{R}^4$ is the metric defined by the metric tensor:

$$\eta = (\eta_{ij}) = \begin{pmatrix} -c^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} =: \text{diag}(-c^2, 1, 1, 1). \tag{1.19}$$

Here c is the speed of light, but often units of $c = 1$ are used. This example will be explored in later chapters. ■

The Minkowski metric is an example of a **Lorentzian metric**, so $(\mathbb{R}^4, \eta)$ is one example of a **Lorentzian manifold**, which is a metric g with **signature** $\text{sign}(g) = (-, +, +, +)$. The signature is well-defined, for if there are s negative components for the metric in one base of T_pM where it is diagonal, there are necessarily also s negative components in all other bases where g is diagonal. This is a result of Sylvester's law of inertia (paragraph 2.3.2 from [12]).

Example 1.2.3. The standard inner product on $\mathbb{R}^n$ has $s = 0$, and is an example of a **Riemannian metric**. If $s = 0$, the metric g is also called **positive definite**, meaning that $g(v, v) \geq 0$, with equality iff $v = 0$. ■

Definition 1.2.13. A vector v in the tangent space T_pM of a Lorentzian manifold (M, g) is called **timelike** if $g(v, v) < 0$, **null** if $g(v, v) = 0$ and **spacelike** if $g(v, v) > 0$. The subset of T_pM consisting of all null vectors is called the **timecone** of g at p.
A basis (∂_α) for the tangent space T_pM is an **orthonormal basis** if $g(\partial_0, \partial_0) = -1$ and $g(\partial_i, \partial_i) = 1$ for $i = 1, 2, 3$ and $g(\partial_i, \partial_j) = 0$ for $i \neq j$.

The **line element** $ds^2 := g(dl, dl) = g_{\alpha\beta} dx^\alpha dx^\beta$ is the square of an infinitesimal displacement vector from $p = (x^\alpha)$ to $q = (x^\alpha + dx^\alpha)$ where $p, q \in M$, $dl = \mathbf{d}x^\alpha \partial_\alpha$ ([12]). This line element is used in most texts to describe the metric.

Example 1.2.4. The **de Sitter space** of dimension $n > 1$ is the subset $\Sigma^n \subset (\mathcal{M}^{n+1}, g_{\mathcal{M}^{n+1}})$, where $\mathcal{M}^{n+1}$ is the **Minkowski space** of dimension $n+1$ with the Minkowski metric $g_{\mathcal{M}^{n+1}}$ belonging to that space. This subset is characterized by the constraint that, for every element of Σ^n, it should hold that, for $H \in \mathbb{R} \setminus \{0\}$, $-(X^0)^2 + \sum_{i=1}^n (X^i)^2 = H^{-2}$. Here the X^i are coordinates in $\mathcal{M}^{n+1}$ for which the metric can be represented as $g =$diag$(-c^2, 1, ..., 1)$. We can find one possible metric on this space by introducing hyperspherical coordinates (we do this because simpler coordinates do not cover the

[5] if and only if

entire space). Introduce coordinates $t, \tau_1, ..., \tau_{n-1}$ such that:

$$\begin{aligned}
X^0(t, \tau_1, ..., \tau_{n-1}) &= \frac{\sinh(Ht)}{H}; \\
X^1(t, \tau_1, ..., \tau_{n-1}) &= \frac{\cosh(Ht)}{H}\cos(\tau_1); \\
X^2(t, \tau_1, ..., \tau_{n-1}) &= \frac{\cosh(Ht)}{H}\sin(\tau_1)\cos(\tau_2); \\
&... \\
X^{n-1}(t, \tau_1, ..., \tau_{n-1}) &= \frac{\cosh(Ht)}{H}\sin(\tau_1)\sin(\tau_2)...\sin(\tau_{n-2})\cos(\tau_{n-1}); \\
X^n(t, \tau_1, ..., \tau_{n-1}) &= \frac{\cosh(Ht)}{H}\sin(\tau_1)\sin(\tau_2)...\sin(\tau_{n-2})\sin(\tau_{n-1}).
\end{aligned} \tag{1.20}$$

Here $\tau_i \in (-\pi/2, \pi/2)$ for $i = 1, ..., n-1$ and $\tau_n \in (-\pi, \pi)$. Thus we have a map $\phi : \mathbb{R}^n \to \mathcal{M}^{n+1}$ given by:

$$\phi(t, \tau_1, ..., \tau_{n-1}) = \left(\frac{\sinh(Ht)}{H}, \frac{\cosh(Ht)}{H}\cos(\tau_1), ..., X^n(t, \tau_1, ..., \tau_{n-1})\right). \tag{1.21}$$

Now we can find the induced metric of the de Sitter space by finding:

$$g_{\Sigma^n, ij}(p) = g_{\mathcal{M}^{n+1}}\left(\frac{\partial\phi}{\partial x^i}(p), \frac{\partial\phi}{\partial x^j}(p)\right). \tag{1.22}$$

This way we find that the line-element in the de Sitter space is given as:

$$ds^2 = -c^2dt^2 + \frac{\cosh^2(Ht)}{H^2}\sum_{j=1}^{n-1}\left(\prod_{i=1}^{j-1}\sin^2\tau_i\right)d\tau_j^2. \tag{1.23}$$

Here the sum over j arises from the line element of a $n-1$ dimensional sphere when using spherical coordinates and the induced metric.
We can thus write the metric tensor g_{Σ^n} w.r.t. the introduced coordinates as:

$$\operatorname{diag}\left(-c^2, \frac{\cosh^2(Ht)}{H^2}, \frac{\cosh^2(Ht)}{H^2}\sin(\tau_1), ..., \frac{\cosh^2(Ht)}{H^2}\left(\prod_{i=1}^{n-2}\sin^2\tau_i\right)\right) \blacksquare \tag{1.24}$$

1.3 Curvature

Having the definition of a tensor, we can search for a way to express the curvature of a manifold in terms of such a tensor. To do this, we need to look at how we define derivatives in manifolds, because if we compare, curvature of a curve as known from analysis is a property linked to second derivatives. Furthermore, we need a way to link tangent spaces to different points in a manifold to each other, so we can objectively compare two tangent vectors to a manifold, while they are defined with respect to different base-points. This means that the notion of a manifold itself is not enough to define the concept of curvature. We need to define additional structures. One of these is the 'connection'. The other structure, which will be proven to be linked to the connection in special cases, is the metric defined in the last section. Second derivatives of this metric are present in the definition of the curvature tensor we will use throughout the rest of this thesis.

1.3.1 Covariant Derivative

Firstly, because we want to talk about curvature and linking vectors, we observe the partial derivative of a vector $\vec{V} = V^\alpha \partial_\alpha$:

$$\frac{\partial \vec{V}}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta}\partial_\alpha + V^\alpha \frac{\partial \partial_\alpha}{\partial x^\beta}. \tag{1.25}$$

The first term here is a vector again. The second term, however, involves a derivative of a basisvector along another one. This is a quantity that involves two objects in different tangent spaces, so we need to define the way to connect these tangent spaces by a thus so-called 'connection'.
One important example is the flat spacetime, where we can impose the connection to be the one where one basis vector in p is the basis vector pointing in the same direction in every other point $p' \in M$. This then means that $\frac{\partial \partial_a}{\partial x^\beta} := 0$.

We see that, to add the two parts above together, we need the $\frac{\partial \partial_\alpha}{\partial x^\beta}$ to be vectors, so we need to express them in terms of the basis-vectors. This we will do later on, because we will first need some more definitions, based on [16].

Definition 1.3.1. For a vector bundle $\pi : E \to M$ over a manifold M, a **connection** in E is a map $\nabla : \mathcal{T}(M) \times \mathcal{E}(M) \to \mathcal{E}(M)$, sending a pair (X, Y) to $\nabla_X Y$, where $\mathcal{T}(M)$ is the space of smooth sections of TM. The map has to satisfy the following properties:

(i). $\nabla_X Y$ is linear over $C^\infty(M)$ in X, thus for $f, g \in C^\infty(M)$ and $X_1, X_2 \in \mathcal{T}(M)$ it holds that:

$$\nabla_{fX_1+gX_2} Y = f\nabla_{X_1} Y + g\nabla_{X_2} Y. \tag{1.26}$$

(ii). $\nabla_X Y$ is linear over $\mathbb{R}$ in Y, so for $a, b \in \mathbb{R}$ and $Y_1, Y_2 \in \mathcal{E}(M)$ it holds that:

$$\nabla_X(aY_1 + bY_2) = a\nabla_X Y_1 + b\nabla_X Y_2. \tag{1.27}$$

(iii). ∇ satisfies the Leibniz rule, which states for $f \in C^\infty(M)$ it holds that:

$$\nabla_X(fY) = f\nabla_X Y + (X(f))Y \tag{1.28}$$

$\nabla_X Y$ is also called the **covariant derivative** of Y in the direction of X.
A connection in TM is called a **linear connection**.

Now we can indeed find a way to express the partial derivative of a basis vector as seen above. This is done by defining:

Definition 1.3.2. The **connection coefficients** of a linear connection ∇, Γ^i_{jk} with respect to the frame (∂_α) are defined such that:

$$\frac{\partial \partial_k}{\partial x^j} := \nabla_{\partial_j}\partial_k := \Gamma^i_{jk}\partial_i. \tag{1.29}$$

Remark. Given a coordinate basis, the connection coefficients form a smooth function from the coordinate chart around a point to $\mathbb{R}$. ▲

Lemma 1.3.1. *Let ∇ be a linear connection on M, then there exists a unique connection in each tensor bundle $T^k_l M$, denoted with ∇, such that the following conditions hold:*

(i). *On TM, ∇ and the given connection agree.*

(ii). *On T^0M, the space of smooth real-valued functions on M, ∇ is given as $\nabla_X f = X(f)$.*

(iii). *The following product rule holds for tensor products of tensors F and G (with types such that the tensor product is defined):*

$$\nabla_X(F \otimes G) = (\nabla_X F) \otimes G + F \otimes (\nabla_X G). \tag{1.30}$$

Furthermore, this connection satisfies the properties:

(a) *For a tensor field of type $(1,0)$ ω and a vector field Y, it holds that:*

$$\nabla_X(\omega(Y)) = (\nabla_X \omega)(Y) + \omega(\nabla_X Y). \tag{1.31}$$

(b) *For any $F \in \mathcal{T}_l^k(M)$, vector fields Y_i and 1-forms ω^j it holds that:*

$$(\nabla_X F)(\omega^1, ..., \omega^l, Y_1, ..., Y_k) = X(F(\omega^1, ..., \omega^l, Y_1, ..., Y_k)) \tag{1.32}$$

$$- \sum_{j=1}^{l} F(\omega^1, ..., \nabla_X \omega^j, ..., \omega^l, Y_1, ..., Y_k) \tag{1.33}$$

$$- \sum_{i=1}^{k} F(\omega^1, ..., \omega^l, Y_1, ..., \nabla_X Y_i, ..., Y_k). \tag{1.34}$$

Remark.

- We call a function from $\mathbb{R}$ to $\mathbb{R}$ **smooth** if it is C^∞. Functions of more components such as vector fields are smooth if their components are.

- While the partial derivative of a type (k,l) tensor is generally not a tensor again, the covariant derivative projects the partial derivative onto the appropriate vector space containing all tensors of type $(k, l+1)$ defined for a point infinitesimally close to p.

- The connection coefficients, as shown in eq.(1.29), give the components of the vector you get when parallel transporting (see 'Parallel Transport and geodesics' later on in this chapter) the k-th basis vector along the coordinate curve with constant j-th coordinate. ▲

Since we will only study coordinate frames, meaning that our basisvectors are always described as partial derivatives along coordinate curves, as the ∂_α are, we do not introduce the notation of ω^i_{jk} for the connection coefficients as in [9] or [20].

Due to the product rule for derivatives concerning not only products but also tensor products, we see that the form for the general covariant derivative of a tensor $T^{\alpha_1 ... \alpha_k}_{\beta_1 ... \beta_l}$ of type (k,l) is given with respect to a chosen system of coordinates as:

$$\nabla_\rho T^{\alpha_1 ... \alpha_k}_{\beta_1 ... \beta_l} = \partial_\rho T^{\alpha_1 ... \alpha_k}_{\beta_1 ... \beta_l} + \sum_{i=1}^{k} \Gamma^{\alpha_i}_{\sigma\rho} T^{\alpha_1 ... \alpha_{i-1}\sigma\alpha_{i+1} ... \alpha_k}_{\beta_1 ... \beta_l} - \sum_{j=1}^{l} \Gamma^{\sigma}_{\beta_j \rho} T^{\alpha_1 ... \alpha_k}_{\beta_1 ... \beta_{j-1}\sigma\beta_{j+1} ... \beta_l}. \tag{1.35}$$

This is a tensor of type $(k, l+1)$, as follows from the Lemma above.

Note that the connection coefficients have to transform under coordinate transformations in a specific way for the covariant derivative of a tensor to indeed be a tensor again as stated in the remark above. If the new coordinates are indicated by Greek letters and the old ones by Latin, the transformation law needs to be:

$$\Gamma^{\nu}_{\mu\lambda} = \frac{\partial x^m}{\partial x^\mu}\frac{\partial x^\nu}{\partial x^n}\frac{\partial x^l}{\partial x^\lambda}\Gamma^{n}_{ml} - \frac{\partial x^m}{\partial x^\mu}\frac{\partial x^l}{\partial x^\lambda}\frac{\partial^2 x^\nu}{\partial x^m \partial x^l}. \tag{1.36}$$

This follows from writing for an arbitrary vector that the transformation of its covariant derivative needs to be:

$$\nabla_\mu V^\nu = \frac{\partial x^m}{\partial x^\mu}\frac{\partial x^\nu}{\partial x^n}\nabla_m V^n. \tag{1.37}$$

The left hand side can be expanded with the chain rule and known coordinate transformations as:

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^{\nu}_{\mu\lambda}V^\lambda = \frac{\partial x^m}{\partial x^\mu}\frac{\partial x^\nu}{\partial x^n}\partial_m V^n + \frac{\partial x^m}{\partial x^\mu}V^n\frac{\partial}{\partial x^m}\frac{\partial x^\nu}{\partial x^n} + \Gamma^{\nu}_{\mu\lambda}\frac{\partial x^\lambda}{\partial x^l}V^l. \tag{1.38}$$

The right hand side can be expanded as:

$$\frac{\partial x^m}{\partial x^\mu}\frac{\partial x^\nu}{\partial x^n}\partial_m V^n + \frac{\partial x^m}{\partial x^\mu}\frac{\partial x^\nu}{\partial x^n}\Gamma^{n}_{ml}V^l. \tag{1.39}$$

Because these expressions need to be equal for any vector $\vec{V}$ we can eliminate the vector on both sides, and by multiplying with $\dfrac{\partial x^l}{\partial x^\lambda}$ on both sides we find the expression used above.
This transformation law implies that the connection coefficients are not components of a tensor.
The most important example of a connection is explored in the next subsection.

1.3.2 The Levi-Civita Connection

There is one connection that is of special interest in our discussion of Lorentzian manifolds. This is the so-called **Levi-Civita connection**, also called the **Riemann connection** or the **Christoffel connection**. This is the connection defined by the following theorem:

Theorem 1.3.2. *On an n-dimensional manifold with a metric $g_{\mu\nu}$, there exists a unique symmetric connection that is compatible with the metric, meaning that $\nabla_\rho g_{\mu\nu} = 0$.*

The proof of this Theorem rests on the equality:

$$\begin{aligned} &\partial_\rho g_{\mu\nu} - \partial_\mu g_{\nu\rho} - \partial_\nu g_{\rho\mu} + 2\Gamma^{\lambda}_{\mu\nu} g_{\lambda\rho} = 0 \\ &\Rightarrow \Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}(g_{\nu\rho,\mu} + g_{\rho\mu,\nu} - g_{\mu\nu,\rho}), \end{aligned} \tag{1.40}$$

which holds after introducing a coordinate system. The complete proof is in the appendix. We will only use the Levi-Civita connection from now on. The connection coefficients of this connection are called the **Christoffel symbols** and are given by eq.(1.40) for a given coordinate system.

Remark. The symmetry follows from the fact that we deal with coordinate frames, where $[\partial_i, \partial_j] = 0$. Otherwise the connection found by switching the lower two indices is another metric compatible one, and it would not be unique anymore.
The metric compatibility means that two vectors having a certain inner product at one point p in the manifold conserve this inner product when parallel transporting them to another point in the manifold. The notion of parallel transport is explored in the next section. ▲

In the remaining part of this thesis, we will mainly focus on subsets of a spacetime, which is defined as follows:

Definition 1.3.3. A **spacetime** (M, g, ∇) is a connected four-dimensional Lorentzian manifold (M, g), paired with the unique Levi-Civita connection ∇ compatible with a Lorentzian metric tensor g on M.

Example 1.3.1.

- The four-dimensional Minkowski-space $\mathcal{M}^4$ with metric $g_{M,\mu\nu}$ =diag$(-c^2, 1, 1, 1)$ has no non-zero Christoffel symbols, because the metric is constant. It is one example of a spacetime, and the covariant derivative here coincides with the partial derivative.

- The n-dimensional sphere S^n is not a spacetime because it has a Riemannian metric. We can however compute the Christoffel symbols for the sphere, because the notion of the uniqueness and existence of the Levi-Civita connection also holds for Riemannian metrics, and the same expression can be derived.
  The non-zero Christoffel symbols that come from the induced metric on the sphere when using spherical coordinates $\tau \in (0, \pi)$ and $\phi \in (0, 2\pi)$ are:
  $$\Gamma^{\tau}_{\phi\phi} = -\cos(\tau)\sin(\tau); \quad \Gamma^{\phi}_{\tau\phi} = \cos(\tau)/\sin(\tau). \tag{1.41}$$
  The induced metric is given by:
  $$g_{S^2,ij} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2(\tau) \end{pmatrix}. \tag{1.42}$$
  We will show that the sphere is the object with constant positive curvature, for which thus the de Sitter space is the analogue for the Minkowskian metric with respect to the standard Riemannian metric on $\mathbb{R}^n$.

- The 4-dimensional **de Sitter space** Σ^4 defined above is another example of a spacetime. Due to the squares of the coefficients it indeed has the needed signature. The non-zero Christoffel symbols for the given coordinates above are found to be[6]:
  $$\begin{aligned}
  \Gamma^{t}_{\tau_1\tau_1} &= \frac{\cosh(Ht)\sinh(Ht)}{H}; \quad \Gamma^{t}_{\tau_2\tau_2} = \frac{\cosh(Ht)\sinh(Ht)\sin^2(\tau_1)}{H};\\
  \Gamma^{t}_{\tau_3\tau_3} &= \frac{\cosh(Ht)\sinh(Ht)\sin^2(\tau_1)\sin^2(\tau_2)}{H}; \quad \Gamma^{\tau_1}_{t\tau_1} = H\tanh(Ht);\\
  \Gamma^{\tau_1}_{\tau_2\tau_2} &= -\cos(\tau_1)\sin(\tau_1); \quad \Gamma^{\tau_1}_{\tau_2,\tau_2} = -\cos(\tau_1)\sin(\tau_1)\sin^2(\tau_2);\\
  \Gamma^{\tau_2}_{t\tau_2} &= H\tanh(Ht); \quad \Gamma^{\tau_2}_{\tau_1\tau_2} = \cos(\tau_1)/\sin(\tau_1); \quad \Gamma^{\tau_2}_{\tau_3\tau_3} = -\cos(\tau_2)\sin(\tau_2);\\
  \Gamma^{\tau_3}_{t\tau_3} &= H\tanh(Ht) \quad \Gamma^{\tau_3}_{\tau_1\tau_2} = \cos(\tau_1)/\sin(\tau_1); \quad \Gamma^{\tau_3}_{\tau_2\tau_3} = \cos(\tau_2)/\sin(\tau_2).\blacksquare
  \end{aligned} \tag{1.43}$$

1.3.3 Parallel Transport and Geodesics

As mentioned before, we need a way to actually express the effect of curvature on functions or objects defined on the manifold, for we need to perform physics on the Lorentzian manifold describing the spacetime in order to discuss gravitational waves. Therefore we need to find a way to "move a vector from one point to another while keeping it invariant". Naturally, we will need to use the

[6] I used SageMath[22] with the built-in package SageManifolds[11] to calculate these.

covariant derivative for this, for parallel transport can intuitively be described by stating that for each infinitesimal displacement, the displaced vector must be parallel to the original one, and must have the same length.
Note that in flat space, this is independent of path, for the connection coefficients of the Levi-Civita connection vanish, because the metric is constant in every basis, so its first derivatives are all zero in eq.(1.29). On a sphere, however, the parallel transport of a vector from the north pole to the equator, along the equator around a part of the equator and back to the north pole again is dependent on the length of the part of the path that is along the equator, and will only be equal to the starting vector again if the path goes along the whole equator.
This means that on a curved manifold it is generally impossible to define a globally parallel vector field to a vector. The parallel transport is dependent on the path along which it is performed.
Let $\gamma : I \to M$ be a (piecewise) smooth curve and let σ be a section of the tangent bundle of M along γ. This section is called **parallel** iff $\nabla_{\dot\gamma(t)}\sigma = 0$ for all $t \in I$.

Let V be an element of $T_{\gamma(0)}M$ and let $\dot\gamma(t)$ be the tangent vector to the curve. We define the **parallel transport** of V along the curve as the parallel section $\mathbb{V}$ along γ of the tangent bundle such that $\mathbb{V}(0) = V$.
This $\mathbb{V}$ is unique due to the Theorem stating unicity of solutions for Ordinary Differential Equations given boundary conditions.([3], Section 1.10).

Following paragraph 1.8 from [8] we can choose coordinates in every point along the curve such that the components of the parallel transport of V are the same as those of the original in every point of the curve[7]. Let (e_i) be the coordinates in such a frame, and $x^\mu(t)$ the coordinates of the curve, such that $\dot\gamma^\mu = dx^\mu/dt$.
If we now consider a so-called **autoparallel** curve, for which the tangent vector $\dot\gamma$ is transported parallel, so $\nabla_{\dot\gamma}\dot\gamma = 0$. Introducing the coordinates, we see:

$$\dot\gamma^\nu(\partial_\nu\dot\gamma^\mu + \Gamma^\mu_{\rho\nu}\dot\gamma^\rho) = 0. \tag{1.44}$$

When we use that $\dot\gamma^\mu = dx^\mu/dt$, we find that:

$$\frac{dx^\nu}{dt}\left[\frac{\partial}{\partial x^\nu}\frac{dx^\mu}{dt} + \Gamma^\mu_{\rho\nu}\frac{dx^\rho}{dt}\right] = 0 \Rightarrow \frac{d^2x^\mu}{dt^2} + \Gamma^\mu_{\nu\rho}\frac{dx^\nu}{dt}\frac{dx^\rho}{dt} = 0. \tag{1.45}$$

This expression is also called the **geodesic equation**, and autoparallel curves are called **geodesics**. A clearer explanation of geodesics can also be found in [23]. Geodesics are of special interest, for they are the curves along which particles move through spacetime, as seen from a physical perspective.

Example 1.3.2.

- In Euclidean spaces, all straight lines are geodesics. In the n-dimensional sphere, all circles with constant distance to the center of the sphere are geodesics.
- In Minkowski space, the Christoffel symbols are all zero, so the geodesic equation becomes: $d^2x^\mu/dt^2 = 0$, so again all straight lines are geodesics.
- In the sphere $S^2 \subset \mathbb{R}^3$, the geodesics are exactly curves that are part of a **great circle**, which is given as $\gamma(t) = \cos(at)\epsilon_1 + \sin(at)\epsilon_2$, where ϵ_1 and ϵ_2 are a pair of orthogonal unit vectors of $\mathbb{R}^3$.■

[7] If the connection coefficients vanish in those frames as well, it is called a Local Inertial Frame (LIF), which thus is a coordinate frame moving along a geodesic such that the tangent space is homeomorphic to four-dimensional Minkowski space, thus the metric in these coordinates is the Minkowski-metric.

1.3.4 Riemann curvature tensor

We are now able to define the Riemann tensor field of type $(1,3)$:

Definition 1.3.4. Let x_p, y_p and v_p be vectors at a point $p \in M$ where M is a Lorentzian manifold and let X, Y and V be smooth vector fields defined in a neighbourhood of p such that they attain the values x_p and y_p in the point p.
The tensor field, expressed in terms of a coordinate basis (∂_i):

$$R(X,Y)V := \nabla_X(\nabla_Y V) - \nabla_Y(\nabla_X V) - \nabla_{[X,Y]}V = (R^i_{jkl}X^k Y^k V^j)\partial_i, \tag{1.46}$$

where the **Riemann tensor** $R(X,Y)$ has components:

$$R^i_{jkl} := \partial_k \Gamma^i_{lj} - \partial_l \Gamma^i_{kj} + \Gamma^i_{kr}\Gamma^r_{lj} - \Gamma^i_{lr}\Gamma^r_{kj}, \tag{1.47}$$

This is dependent on the connection, which is in turn again dependent on the metric of the manifold.
This is essentially a type $(1,3)$ tensor field. It is antisymmetric with respect to its last two indices and satisfies the identity:

$$R_{klmn} = g_{kr}R^r_{lmn} = \frac{1}{2}(\partial_l\partial_n g_{km} - \partial_l\partial_m g_{kn} + \partial_k\partial_m g_{ln} - \partial_k\partial_n g_{lm}). \tag{1.48}$$

This follows from the expression of the Christoffel symbols in eq. 1.40.
This leads to more properties:

$$\begin{aligned} R_{klmn} &= -R_{lkmn}; \\ R_{klmn} &= R_{mnkl}; \\ 3R_{k[lmn]} &:= R_{klmn} + R_{kmnl} + R_{knlm} = 0. \end{aligned} \tag{1.49}$$

Remark. The Riemann tensor is also interpretable as the commutator of covariant derivatives. This is seen as follows:

$$\nabla_a\nabla_b V^m = \partial_a\nabla_b V^m + \Gamma^m_{sa}\nabla_b V^s - \Gamma^s_{ba}\nabla_s V^m. \tag{1.50}$$

In a local inertial frame (LIF) (which we can always find along a geodesic through the point as stated in paragraph 1.8 from [8]), where the connection coefficients vanish, this becomes:

$$\nabla_a\nabla_b V^m = \partial_b\partial_a V^m + \partial_a\Gamma^m_{nb}V^n; \tag{1.51}$$

and

$$\nabla_b\nabla_a V^m = \partial_a\partial_b V^m + \partial_b\Gamma^m_{na}V^n, \tag{1.52}$$

so:

$$[\nabla_a, \nabla_b]V^m = (\partial_a\Gamma^m_{nb} - \partial_b\Gamma^m_{na})V^n. \tag{1.53}$$

But in a LIF $R^m_{nab} = \partial_a\Gamma^m_{nb} - \partial_b\Gamma^m_{na}$, so because it is a tensorial equation it holds in every frame that $R^m_{nab}V^n = [\nabla_a, \nabla_b]V^m$. ▲

In a four-dimensional Lorentzian manifold, the antisymmetry in the last two indices implies that there are only 6 independent combinations of the components, which is also true for the first couple of indices.
Then the second condition above implies that $R_{AB} = R_{BA}$ if A is the first pair and B the second, so we have a 6×6 symmetric matrix, which has 21 independent components, so of the 4^4 components of

R^i_{jkl} in relativistic spacetime of four dimensions, only 20 independent components are left, for the last identity above implies one extra constraint.
Intuitively, the Riemann tensor gives the components i of the vector resulting from parallel transporting (see above) a vector v^j_p along an infinitesimal curve spanned in the directions of vectors x^k_p and y^l_p.
Now define the **Ricci tensor** as a contraction of the Riemann tensor, $R_{ij} = R^k_{ikj} = -R^k_{ijk}$ (so it is symmetric). This is the only contraction of interest, for $R^k_{kij} = 0 = g^{rk}R_{ijrk}$ due to the symmetries of the Riemann tensor.
It represents a type $(0,2)$ tensor field, and we note that the Ricci tensor can thus be written as:

$$R_{\mu\nu} = \partial_\alpha \Gamma^\alpha_{\mu\nu} - \partial_\nu \Gamma^\alpha_{\mu\alpha} + \Gamma^\alpha_{\sigma\alpha}\Gamma^\sigma_{\mu\nu} - \Gamma^\alpha_{\sigma\nu}\Gamma^\sigma_{\mu\alpha}. \tag{1.54}$$

The Ricci tensor is a symmetric tensor, so it has 10 independent components.
We can take the trace of the Ricci tensor to find: $R = R^i_i$, the **Ricci scalar** or **scalar curvature**, a smooth function (type $(0,0)$ tensor field).

Example 1.3.3.

- In the Minkowski space, all Christoffel symbols were equal to zero, so the components of the Riemann tensor are zero as well, meaning the Ricci tensor and curvature are both also trivial.

- To show that the n-dimensional de Sitter space Σ^n is the analogue of the n-dimensional sphere S^n we show that, characteristically, they are the objects with constant positive curvature due to their respective embeddings in $\mathbb{R}^{n+1}$ with the Minkowskian and standard Riemannian metric. The non-zero components of the Riemann tensor for spherical coordinates on the S^2 are:

$$\begin{aligned} R^\tau_{\phi\tau\phi} &= \sin^2(\tau); \quad R^\tau_{\phi\phi\tau} = -\sin^2(\tau); \\ R^\phi_{\tau\tau\phi} &= \frac{\cos^2(\tau)-1}{\sin^2(\tau)}; \quad R^\phi_{\tau\phi\tau} = 1. \end{aligned} \tag{1.55}$$

  This has as a consequence that the Ricci tensor becomes equal to the metric tensor in the two-dimensional case. The Ricci scalar thus is constant and equal to 2.
  When altering the radius to be r instead of 1, the scalar curvature becomes $2/r^2$.

- The four-dimensional de Sitter space has Christoffel symbols found in equation (1.43), for which the non-zero Riemann tensor components are equal to:

$$\begin{aligned} R^t_{\tau_1 t\tau_1} &= \cosh^2(Ht); \\ R^t_{\tau_2 t\tau_2} = R^{\tau_1}_{\tau_2\tau_1\tau_2} = -R^{\tau_3}_{\tau_2\tau_2\tau_3} &= \cosh^2(Ht)\sin^2(\tau_1) \\ R^t_{\tau_3 t\tau_3} = R^{\tau_1}_{\tau_3\tau_1\tau_3} = R^{\tau_2}_{\tau_3,\tau_2,\tau_3} &= \cosh^2(Ht)\sin^2(\tau_1)\sin^2(\tau_2); \\ R^{\tau_1}_{tt\tau_1} = R^{\tau_2}_{tt\tau_2} = R^{\tau_3}_{tt\tau_3} &= H^2; \\ R^{\tau_2}_{\tau_1\tau_1\tau_2} = R^{\tau_3}_{\tau_1\tau_1\tau_3} &= -\frac{\sinh^2(Ht)\sin^2(\tau_1) - \cos^2(\tau_1)+1}{\sin^2(\tau_1)}. \end{aligned} \tag{1.56}$$

  The Ricci tensor then becomes:

$$R_{\mu\nu} = \mathrm{diag}\left[-3H^2, 3\cosh^2(Ht), 3\cosh^2(Ht)\sin^2(\tau_1), 3\cosh^2(Ht)\sin^2(\tau_1)\sin^2(\tau_2)\right]. \tag{1.57}$$

  This has as a consequence that the Ricci scalar is equal to $R = 12H^2$, because the inverse metric $g^{\mu\nu}_{\Sigma^4}$ is just defined by $g^{\mu\mu}_{\Sigma^4} := \dfrac{1}{g_{\Sigma^4,\mu\mu}}$, and the off-diagonal elements are zero.

- In the general n-dimensional de Sitter space, the Riemann tensor is given by that of a maximally symmetric space (see [25] for details):

$$R_{ijkl} = H^2(g_{ik}g_{jl} - g_{il}g_{jk}). \tag{1.58}$$

This implies that, for the Ricci tensor:

$$R_{ij} = (n-1)H^2 g_{ij}. \tag{1.59}$$

And thus the Ricci scalar becomes:

$$R = n(n-1)H^2. \tag{1.60}$$

This is, when $r = 1/H$, exactly the expression for the Ricci scalar of the n-sphere with radius r. Thus indeed we can speak about the de Sitter space as the analogue of the n-sphere in Minkowskian space. ■

Lemma 1.3.3. *(The Bianchi Identities): It holds that, for the type* $(0,5)$ *tensors given (where we lowered the index of the type* $(1,3)$ *tensor with the metric):*

$$\nabla_\lambda R_{\alpha\beta\mu\nu} + \nabla_\nu R_{\alpha\beta\lambda\mu} + \nabla_\mu R_{\alpha\beta\nu\lambda} = 0. \tag{1.61}$$

Proof: Taking the partial derivative of the Riemann tensor we see that:

$$\partial_l R_{abmn} = \frac{1}{2}(\partial_b\partial_m\partial_l g_{an} - \partial_b\partial_n\partial_l g_{am} + \partial_a\partial_n\partial_l g_{bm} - \partial_a\partial_m\partial_l g_{bn}). \tag{1.62}$$

And due to the fact that g_{ab} is symmetric we see then that:

$$\partial_l R_{abmn} + \partial_n R_{ablm} + \partial_m R_{abnl} = 0. \tag{1.63}$$

But in a local inertial frame this also holds for covariant derivatives, and because it is a tensorial equation, the Bianchi identities indeed hold. □

Using the Bianchi identities, we see that:

$$0 = g^{\nu\sigma}g^{\mu\lambda}(\nabla_\lambda R_{\rho\sigma\mu\nu} + \nabla_\rho R_{\sigma\lambda\mu\nu} + \nabla_\sigma R_{\lambda\rho\mu\nu}) = \nabla^\mu R_{\rho\mu} - \nabla_\rho R + \nabla^\nu R_{\rho\nu}$$

$$\Rightarrow \nabla^\mu R_{\rho\mu} = \frac{1}{2}\nabla_\rho R. \tag{1.64}$$

Defining the **Einstein tensor** (type (0,2) tensor field) as $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}$, we see that $\nabla^\mu G_{\mu\nu} = 0$. This is the property that had to be satisfied by a tensor describing the curvature of spacetime to relate it to the Stress-Energy tensor (describing the matter and energy in the space) to be defined in the next chapter.

Interpretations

To give a more intuitive understanding of the Ricci tensor and scalar, we have indicated that this essentially represents the difference in volume between a sphere of radius ϵ in Euclidean space with respect to the embedding of a sphere defined as the surface with geodesic distance ϵ (arclength along a geodesic) to a point in the manifold (the origin of the local coordinates) everywhere. Intuitively it represents the change of volume of a test element when moving it along a geodesic (see above). This idea is worked out in more detail in [21]. One can gain more insight in the effect of the change of the metric by looking at the Geodesic Deviation induced by the curvature of the space.

1.3.5 Geodesic Deviation

One of the ways curvature is manifested is by the geodesic deviation, describing the relative acceleration of neighbouring geodesics. This is of importance when observing the effect of curvature, especially a change in curvature, on particles at certain coordinates in spacetime.
To define the geodesic deviation, we will consider a one-parameter family of geodesics $x^\mu(v)$, where v labels the geodesic and t is the affine parameter for a given geodesic. Let $U^\mu = \dfrac{\partial x^\mu}{\partial t}\partial_t$ be the tangent vector. Connecting two points on neighbouring geodesics by an infinitesimal displacement vector $\eta^\mu := \dfrac{\partial x^\mu}{\partial v}\partial_v$, where ∂_v is the vector in the direction of the geodesic labelled by $v + \delta v$, where $\delta v \ll 1$. This implies that:

$$\begin{aligned}
(\nabla_U \eta)^\mu = \nabla_\nu \eta^\mu U^\nu &= \left[\frac{\partial \eta^\mu}{\partial x^\nu} + \Gamma^\mu_{\rho\nu}\eta^\rho\right]\frac{\partial x^\nu}{\partial t} \\
&= \frac{\partial \eta^\mu}{\partial t} + \Gamma^\mu_{\rho\nu}\frac{\partial x^\rho}{\partial v}\frac{\partial x^\nu}{\partial t}\partial_v \\
&= \partial_v\left[\frac{\partial^2 x^\mu}{\partial t \partial v} + \Gamma^\mu_{\rho\nu}\frac{\partial x^\rho}{\partial v}\frac{\partial x^\nu}{\partial t}\right] \\
&= \partial_v\left[\frac{\partial^2 x^\mu}{\partial t \partial v} + \Gamma^\mu_{\rho\nu}\frac{\partial x^\rho}{\partial t}\frac{\partial x^\nu}{\partial v}\right] \\
&= \left[\frac{\partial U^\mu}{\partial t} + \Gamma^\mu_{\rho\nu}\frac{\partial x^\rho}{\partial t}\frac{\partial x^\nu}{\partial v}\right]\partial_v \\
&= \left[\frac{\partial U^\mu}{\partial x^\nu} + \Gamma^\mu_{\rho\nu}U^\rho\right]\frac{\partial x^\nu}{\partial v}\partial_v = (\nabla_\nu U^\mu)\eta^\nu = (\nabla_\eta U)^\mu.
\end{aligned} \tag{1.65}$$

The relative acceleration is now given by $\nabla_U(\nabla_U \eta)$, but the geodesic equation implies $\nabla_U U = 0$ so $\nabla_\eta \nabla_U U = 0$, thus $\nabla_U \nabla_\eta U + [\nabla_\eta, \nabla_U]U = 0$.
This means that:

$$\nabla_U \nabla_U \eta = -[\nabla_\eta, \nabla_U]U = [\nabla_U, \nabla_\eta]U. \tag{1.66}$$

But the commutator of these covariant derivatives is the Riemann curvature tensor, so we may define the **geodesic deviation**:

$$\frac{D^2 \eta^\sigma}{dt^2} = R^\sigma_{\mu\lambda\kappa}U^\mu U^\lambda \eta^\kappa. \tag{1.67}$$

Here we see the Riemann tensor again. We can also apply this theory to surface instead, leading to an expression with the Ricci tensor, or with volumes, leading to an expression with the Ricci scalar.

1.4 Summary

In this chapter we defined a manifold, additional structures on a manifold like a metric and a connection and we derived the intrinsic property of curvature of a manifold. This way, we have set up the basic mathematical concepts needed to describe curvature in spacetime. In the next chapter, we will use this to define the law relating curvature to presence of matter. This will be necessary to describe Gravitational Waves in chapter 3.

Chapter 2

Principles of General Relativity

To describe Gravitational Waves (GW) mathematically, we need to discuss some of the principles of General Relativity (GR), which came forth from Einstein's Special Relativity (SR). GR is the physical theory of gravity formulated by Einstein in 1915. Two principles are of central importance in Einstein's formulation, the Equivalence Principle of Gravitation and Inertia (EP), and the Principle of General Covariance (PGC). But why do we need a generalisation of the Newtonian theory of gravity? Why do we need differential geometry to formulate the laws? What is the role of these equivalence principles? In this introductory chapter we will explore the origin of the Einstein equations:

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{2.1}$$

where G is the gravitational constant, c the speed of light, $G_{\mu\nu}$ the Einstein tensor introduced in the previous chapter and $T_{\mu\nu}$ the stress-energy tensor. Both $G_{\mu\nu}$ and $T_{\mu\nu}$ are elements of $\mathcal{T}^0_2 M$, smooth sections of the vector bundle with type $(0,2)$-tensors.

2.1 Principles

Before we can start our discussion of the various components present in Einstein equations, we need to introduce the fundamental principles of the theory of General Relativity, because these impose rules that do not directly follow from the mathematical theory.
Newton published his theory of gravity in 1685 stated in two laws. The first is Newton's law: $F = m_I a$, where F is the force on a particle, m_I the inertial mass of a particle, and a its acceleration. The second is Newton's law of gravitation: $F_G = m_G g$, where m_G is the gravitational mass, F_G the gravitational force on a particle and g the gravitational acceleration, dependent on the position of the particle with respect to other particles with mass. It is given as:

$$g = -\frac{G \sum_i M_{Gi}(r - r_i)}{|r - r_i|^3}. \tag{2.2}$$

Here M_{Gi} are the masses of other particles and r_i their positions, while r is the position of the particle with mass m_G. Observing a particle freely falling in a gravitational field, experiencing a force $F = m_G g$, so accelerating with $a = F/m_I = g m_G/m_I$. Experiments have shown to a high level of accuracy that m_G/m_I is the same for all materials, by scaling the Gravitational Constant G present in the expression for g above such that $m_I = m_G$ we impose that the EP states that $m_I = m_G$, so an experimenter releasing objects and timing their fall will not be able to tell whether he is in a gravitational field or

being accelerated through empty space.
This generalizes to the *strong Principle of Equivalence*:
"In an arbitrary gravitational field, at any given spacetime point, we can choose a locally inertial[1] reference frame (LIF) such that, in a sufficiently small region surrounding that point, all physical laws reduce to the form they would take in the absence of gravity, the form prescribed by SR."

The *Principle of General Covariance* states: "A physical law is "true" if it preserves its form under arbitrary coordinate transformations, meaning that it has to be expressible as a tensorial equation".

This principle is the reason we want to express the laws of GR in a tensorial form with the tensors defined on a curved manifold with the concept of distance defined in terms of the metric of this manifold.
We want this to hold because we already know from Newton's formulation that laws are invariant under rotations of the reference frame and from SR that they are invariant under Lorentz transformations. In addition now, the strong Equivalence Principle states that an observer can not feel the difference between acceleration in free space or a gravitational field. Because the coordinate transformation for an acceleration is non-linear, we indeed want our principle to extend to arbitrary coordinate transformations to account for other degrees of symmetry.

2.1.1 Special Relativity

Einstein proposed the laws of Special Relativity (SR) and Minkowski showed that the effects of time dilation and length contraction can be explained by a four-dimensional spacetime with the Minkowskian line element[2] $ds^2 = -c^2dt^2 + \sum_{i=1,2,3}(dx^i)^2$ [5]. This was generalized by Einstein to a form where, in any local coordinates $x^0 = t, x^1, x^2, x^3$, the line element of the metric takes the form:

$$ds^2 = g_{00}(t,\vec{x})dt^2 + 2g_{0\beta}(t,\vec{x})dtdx^\beta + g_{\alpha\beta}(t,\vec{x})dx^\alpha dx^\beta. \tag{2.3}$$

Here g_{00} has to be negative for it to be a generalisation of Minkowski's flat spacetime.

Assuming that the universe is stationary so we can choose the local coordinates such that the coefficients are independent of coordinate time t and that the mixed temporal-spatial terms vanish (because a rotating body might produce a stationary metric), we find a **static metric**:

$$ds^2 = g_{00}(\vec{x})dt^2 + g_{\alpha\beta}(\vec{x})dx^\alpha dx^\beta. \tag{2.4}$$

We can now try to find expressions for these metric-components that rely on the distribution of energy and matter through the universe. The g_{00}-component can be calculated when knowing the Newtonian gravitational potential Φ, as is shown in a later section.
In flat spacetime, the Christoffel symbols are zero, so we get $G_{\mu\nu} = 0$. This is the simplest case for Einstein's equations.

[1]Locally Inertial means that the tangent space is locally homeomorphic to four-dimensional Minkowski-space, thus we can find a basis for the tangent space such that the metric is the Minkowski-metric.

[2]Note that many other books using special or general relativity use the convention that the Minkowski metric is given as $\eta = \mathrm{diag}(c^2, -1, -1, -1)$. This convention implies in particle physics that $p^2 = m^2$ instead of $p^2 = -m^2$, when m is the mass, and p the four-momentum of a particle. Because we will be mostly concerned with distances, we will use the first convention, because then a positive length in 'normal' three-dimensional Euclidean space is still positive in spacetime, while the other sign would mean that it becomes negative.

To precisely calculate all quantities needed we need to introduce the concept of a **proper time** along the **world line** (the geodesic along which the particle moves through space-time) of a particle, which is just the time measured by a clock moving along the geodesic with the particle. This proper time is denoted by $d\tau^2$, and is generally used in calculations throughout SR.
Important quantities describing energy and matter are collected in the **momentum four-vector**, the four-vector $(E/c, m\gamma\mathbf{v})$, where $\mathbf{v} = d\mathbf{x}/dt$ and $\gamma = dt/d\tau$, creating a smooth section of the tangent bundle along the geodesic of a particle.

2.2 Einstein equations

As stated in the last section, Φ and g_{00} are related, but what can we say about the other metric coefficients?
First compare Einstein's gravitation with 10 independent metric components with Electromagnetism, governed by a 4-vector potential A, and Newtonian gravitation, governed by a single potential Φ. Newtonian gravitation is a scalar theory and electromagnetism is a vector theory. Extending this, we need Einstein's gravitation to be a symmetric covariant second-rank tensor theory to connect the metric with curvature and the presence of matter and energy. The components g_{ij} can be interpreted as 'metric potentials'. But which are the field equations satisfied by the g_{ij}?
In this section we will define a tensorial way to describe matter and fields. Then we will search for the appropriate tensor describing curvature that can be related to the other quantity.

2.2.1 Stress-Energy tensor

The tensor we will use to describe matter and fields in GR is the
Stress-Energy tensor. It is defined as a type $(0,2)$ tensor field where the components T^{ij} describe the flux of the i-th component of the four-momentum through the surface orthogonal to the j-th coordinate:

1. T^{00} is the flux of $p^0 = E/c$ through the surface orthogonal to the coordinate time, so it is the **energy density**.
2. $T^{0i}/c = T^{i0}/c$ for $i \neq 0$ is in the same sense the **momentum density**.
3. T^{ij} for $i \neq 0$ and $j \neq 0$ is the **current of the momentum**, meaning that the components T^{ii} are the i-th components of the pressure and the off-diagonal elements represent the shearing terms.

This dimensionality comparison means we can express the components of the stress-energy tensor as:

$$T^{\alpha\beta} = c\sum_n \int p_n^\alpha \frac{dx_n^\beta}{d\tau_n} \delta^4(\vec{x} - \vec{x}_n(\tau_n))d\tau_n = \sum_n p_n^\alpha \frac{dx_n^\beta(t)}{dt} \delta^3(\mathbf{x} - \mathbf{x}_n(t)), \tag{2.5}$$

where we consider a system of n non-interacting particles at points $x_n(t)$ with energy-momentum four-vector p_n^α and proper time τ_n, the δ represents the Dirac δ-function and the boldfaced vectors represent spatial parts.
If now $v \ll c$, we have $p_n^0 \sim m_n c$ and $T^{00} \sim \sum\limits_n m_n c^2 \delta^3(\mathbf{x} - \mathbf{x}_n(t))$, which reduces to the density of matter ρc^2. This same dimensionality check can be performed for the other components.

Remark. Note that these expressions are only defined for discrete distributions of particles. We can extend this definition to a continuous distribution, characterized by densities. Also note that the delta functions are not smooth, but we can replace these by narrow Gaussians to again ensure that $T^{\mu\nu}$ is a tensor. A more general definition, in terms of the Lagrangian density of the fields and matter $\mathcal{L}$, depending on all potentials $\Phi^{(i)}$ and their covariant derivatives $\nabla_\alpha \Phi^{(i)}$, is described in the appendix. It results in:

$$T_{\mu\nu} = -2c\left[\frac{\partial \mathcal{L}}{\partial g^{\mu\nu}} - \frac{1}{2}\mathcal{L}g_{\mu\nu}\right]. \tag{2.6}$$

This expression contains a derivative to a metric component, which we can calculate using the chain-rule for partial derivatives. ▲

Example 2.2.1. To illustrate the connection between the two definitions, we will study the case of free particles through space-time the action of such a particle is given by:

$$S = -mc\int\sqrt{-g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}}dt. \tag{2.7}$$

If we transform to proper time τ, this becomes $S = -mc^2\int d\tau$, because τ is an affine parameter thus the quantity within the square root is actually constant. Expanding $d\tau$ in the low-velocity limit, we recover $L = \frac{1}{2}mv^2$.

The energy-momentum tensor density $\mathcal{T}^{\mu\nu} = \sqrt{-g}T^{\mu\nu}$ (g is the determinant of the metric tensor) can be found from:

$$\frac{1}{2}\sqrt{-g}T^{\mu\nu} = -\frac{\partial \mathcal{L}}{\partial g_{\mu\nu}} \Rightarrow \mathcal{T}^{\mu\nu} = -2\frac{\partial \mathcal{L}}{\partial g_{\mu\nu}}. \tag{2.8}$$

Here we define:

$$\mathcal{L} = -\int mc\sqrt{-g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}}\delta^4(x^\mu - x^\mu(t))dt. \tag{2.9}$$

so we find that the tensor density is equal to:

$$\mathcal{T}^{\mu\nu} = \int \frac{mc\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}}{\sqrt{-g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}}}\delta(\tau - \tau(t))\delta^3(\mathbf{x} - \mathbf{x}(\tau)) \tag{2.10}$$

$$= m\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}\frac{dt}{d\tau}\delta^3(\mathbf{x} - \mathbf{x}(t(\tau))), \tag{2.11}$$

where we transformed the integral over t to one over τ and back again. Also we used the invariance of the quantity in the root when describing this with an affine parameter.

In the rest frame of the particle, the energy-momentum tensor density has only the term $\mathcal{T}^{00} = mc^2$. Note that $dt/d\tau = \sqrt{1 - v^2/c^2}$, and that now:

$$\mathcal{T}^{0j} = mc\frac{d\tau}{dt}\frac{dx^j}{dt}\frac{dt}{d\tau} = \frac{mc\mathbf{v}}{\sqrt{1 - v^2/c^2}}. \tag{2.12}$$

This results in $\mathcal{T}^{0\mu} = cp^\mu$, where p is the four-momentum again[10]. ■

Example 2.2.2. Consider an isotropic (without directional preference) fluid, viewed from the rest frame of that fluid. Let ρ be the rest mass-energy density of the fluid, p the pressure in each direction

and u the velocity of the particles. The operator $g^i_j + u^i u_j$ projects a four-vector orthogonally into the three-space orthogonal to u.
The stress-energy tensor can now be expressed as:

$$T^{ij} = \rho u^i u^j + p(g^{ij} + u^i u^j) = (\rho + p)u^i u^j + pg^{ij}. \tag{2.13}$$

This can also be seen by letting $\Lambda(\mathbf{v})$ represent a Lorentz transformation into a frame with velocity $\mathbf{v}$, then [24]:

$$\Lambda(\mathbf{v}) = \begin{pmatrix} 0 & 0 \\ 0 & \delta_{ij} - \dfrac{v_i v_j}{\mathbf{v}^2} \end{pmatrix} + \gamma \begin{pmatrix} 1 & -v_j \\ -v_i & \dfrac{v_i v_j}{\mathbf{v}^2} \end{pmatrix}. \tag{2.14}$$

Then:

$$T_{\mu\nu}(\mathbf{v}) = \Lambda^\alpha_\mu(\mathbf{v}) T_{\alpha\beta}(0) \bar{\Lambda}^\beta_\nu(\mathbf{v}) = pg_{\mu\nu} + (\rho + p)u_\mu u_\nu. \tag{2.15}$$

Here $\bar{\Lambda}$ is the inverse of Λ. ■

Now we will be able to derive the tensor that has to be on the left-hand side of Einstein's equations. We will first do this in a physically intuitive way and derive the same equations using the variational principle in the appendix, giving a more mathematically rigorous explanation.

2.2.2 Einstein tensor

We need a connection between curvature of spacetime and presence of matter. To give some insight in this proposition consider an idealized fluid without pressure, meaning the individual particles are freely falling under gravitational influence. Each thus moves along a geodesic. Let γ be the geodesic describing the world-line of one particular particle, and $\delta x_t = (\delta x^0, ..., \delta x^3)$ the variation vector between this particle and another at time t. In [9] paragraph 4.1b it is stated that δ and d/dt commute, so we find with Newtons law:

$$\frac{d^2}{dt^2}(\delta x^\alpha) = \delta \frac{d^2 x^\alpha}{dt^2} = -\delta \frac{\partial \Phi}{\partial x^\alpha} = -\frac{\partial^2 \Phi}{\partial x^\alpha \partial x^\beta} \delta x^\beta. \tag{2.16}$$

But the geodesic deviation states in the Newtonian limit with a weak field that also:

$$\frac{D^2 \delta x^\alpha}{d\tau^2} = R^\alpha_{\beta\mu\nu} U^\beta U^\mu \delta x^\nu \Rightarrow \frac{d^2 \delta x^\alpha}{dt^2} \sim R^\alpha_{0\beta 0} \delta x^\beta. \tag{2.17}$$

Since $R^j_{k00} = 0$ we obtain:

$$\nabla^2 \Phi = \sum_{\alpha=1,2,3} \frac{\partial^2 \Phi}{\partial x^\alpha \partial x^\alpha} \sim -\sum_{\alpha=1,2,3} R^\alpha_{0\alpha 0} \sim -R_{00}. \tag{2.18}$$

Here indeed the curvature is connected to the matter density ρ for $\nabla^2 \Phi = 4\pi G \rho$ (Poissons' law).

To give a connection between the stress-energy tensor and the curvature, we need to find another tensor describing the curvature, which we call $\bar{G}_{\mu\nu}$ for now. In order for this connection to take the form $\bar{G}_{\mu\nu} = KT_{\mu\nu}$, the following properties need to be satisfied:

1. $\bar{G}_{\mu\nu} = KT_{\mu\nu}$ has to hold for a constant K that we will have to determine.
2. Therefore $\bar{G}_{\mu\nu}$ has to be symmetric and it has to satisfy the law that $\nabla_\nu \bar{G}^{\mu\nu} = 0$, because $T_{\mu\nu}$ has these properties as well.

3. In the weak field it has to reduce to $\bar{G}_{00} \sim -\nabla^2 g_{00}$.
4. It has to be linear in second derivatives of $g_{\mu\nu}$ and contain products of first derivatives of $g_{\mu\nu}$. This is a reasonable assumption for if it would contain terms with higher or lower orders of derivatives, we would need to multiply with a constant having dimensions of a suitable power of a length, making the equations scale-dependent, which is unacceptable in view of the PGC.

The need for it to be a tensor and the two properties 2. and 4. are met if we define:

$$\bar{G}_{\mu\nu} = C_1 R_{\mu\nu} + C_2 g_{\mu\nu} R, \tag{2.19}$$

where C_1 and C_2 are constants such that $C_2/C_1 = -1/2$. This is because we know that the Ricci tensor and the scalar curvature are the two general objects containing the necessary derivatives of $g_{\mu\nu}$, and a linear combination of this kind is automatically another tensor. The ratio is constrained by comparing this expression with the expression of the Einstein tensor, for which it indeed holds that $\nabla_\nu G^{\mu\nu} = 0$.

When we let Latin indices indicate spatial parts and Greek indices all four components, we see that in the weak field and Newtonian limit $|T_{ij}| \ll |T_{00}|$, which should consequently also hold for $\bar{G}_{ij}$ and $\bar{G}_{00}$. Therefore we see that:

$$|C_1(R_{ij} - \frac{1}{2} g_{ij} R)| \ll |\bar{G}_{00}|. \tag{2.20}$$

Hence $R_{ij} \simeq \frac{1}{2} g_{ij} R$ and since $g_{ij} \simeq \delta_{ij}$ we see that $R_{kk} \simeq \frac{1}{2} R$. Consequently:

$$R = g^{\mu\nu} R_{\mu\nu} \simeq \eta^{\mu\nu} R_{\mu\nu} = -R_{00} + \sum_k R_{kk} = -R_{00} + \frac{3}{2} R. \tag{2.21}$$

From this it is found that $R \simeq 2R_{00}$, so because $\bar{G}_{\mu\nu} = C_1(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R)$ it holds that $\bar{G}_{00} \simeq C_1 2 R_{00}$. Computing R_{00} in the weak field limit we find that the non linear part is second order and retaining only the first order terms and imposing stationarity we get:

$$R_{00} \simeq -\frac{1}{2} \eta^{ij} \frac{\partial^2 g_{00}}{\partial x^i \partial x^j} = -\frac{1}{2} \nabla^2 g_{00}. \tag{2.22}$$

This then implies that:

$$G_{00} \simeq -C_1 \nabla^2 g_{00} \Rightarrow C_1 = 1. \tag{2.23}$$

This implication uses assumption 3. above. In conclusion, $\bar{G}_{\mu\nu} = G_{\mu\nu}$, the Einstein tensor already introduced.

Newtonian limit of GR

The constant K follows from comparing dimensionality in the weak field, non-relativistic limit.
Assuming the gravitational field is very weak[3], so the metric can be written as $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ where $|h_{\mu\nu}| \ll 1$ is an arbitrary 4×4 matrix and $\eta_{\mu\nu}$ is the Minkowski metric. We call the path a particle follows through spacetime its **world line**. Assuming it travels with a speed very small compared with

[3]This is a realistic assumption because we want to study GWs due to binary dwarf systems. These are far away and, as seen in chapter 3, emit really weak GW-signals.

the speed of light ($v \ll c$) and introducing the proper time again as the time kept by a clock moving with the particle, its unit velocity four vector is:

$$u := \frac{dx}{d\tau} \sim \frac{cdt}{d\tau}. \tag{2.24}$$

Because world lines are geodesics, the geodesic equations state that:

$$\begin{aligned}\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{00}\left(\frac{cdt}{d\tau}\right)^2 = 0 \Rightarrow \frac{d^2x^\mu}{d\tau^2} &= \frac{1}{2}\eta^{\mu\alpha}\frac{\partial h_{00}}{\partial x^\alpha}\left(\frac{cdt}{d\tau}\right)^2 \\ \Rightarrow \frac{d^2\mathbf{x}}{d\tau^2} &= \frac{1}{2}\nabla h_{00}\left(\frac{cdt}{d\tau}\right)^2.\end{aligned} \tag{2.25}$$

Here the first implication follows after only retaining terms of first order in $h_{\mu\nu}$, such that, following from the expression in eq. 1.40:

$$\Gamma^\mu_{00} \sim -\frac{1}{2}\eta^{\mu\sigma}\frac{\partial h_{00}}{\partial x^\sigma}. \tag{2.26}$$

The second implication results from only using the spatial terms. The temporal part is zero assuming the field is stationary $\dfrac{\partial h_{00}}{\partial t} = 0$. We can rescale the time coordinate such that $cdt/d\tau = 1$ to see:

$$\frac{d^2\mathbf{x}}{d\tau^2} = \frac{1}{2}\nabla h_{00}. \tag{2.27}$$

But because $\tau = ct$ and in Newtonian gravity it holds that $d^2\mathbf{x}/dt^2 = -\nabla\Phi$ where Φ is the Newtonian gravitational potential such that $\nabla^2\Phi = 4\pi G\rho$ with ρ the matter density.
The PGC now states that these two expressions have to be equal, so $h_{00} = -2\Phi/c^2 + C$, where C is a constant. Assuming the potential vanishes 'at infinity', we have that $C = 0$ and we find:

$$g_{00} = -1 - 2\frac{\Phi}{c^2}. \tag{2.28}$$

Thus it is shown that a test body freely falling with non-relativistic speeds in a weak gravitational field moves along a geodesic, for which g_{00} then represents the Newtonian gravitational potential.

Now using this result along with constraint 3., we wanted $G_{00} \sim -\nabla^2 g_{00} = 2\nabla^2\Phi/c^2 = 8\pi G\rho/c^2$. We also know that $T_{00} = \rho c^2$, so to get $G_{00} = KT_{00}$ we need $K = 8\pi G/c^4$, finishing the derivation of (2.1):

$$G_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}. \tag{2.29}$$

2.2.3 Gauge Freedom

Note that, since $G_{\mu\nu}$ is symmetric and has two indices ranging over four dimensions, it has 10 independent components. This means that Einstein's equations introduce 10 equations for the 10 independent components of $g_{\mu\nu}$. These are however not independent, because the law that $\nabla_\nu G^{\mu\nu} = 0$ provides four other conditions the Einstein tensor must satisfy. This means that we have six equations and 10 unknown functions. We can choose four constraints to apply to the solutions, making the freedom disappear from the solution.
When observing the derivation of Einstein equations in the appendix, we see that these degrees of freedom do also exist because of the symmetries in the Lagrangian defined for the matter and fields. These symmetries exist because $\mathcal{L}_H \propto \sqrt{-g}R$, and both $\sqrt{-g}$ and R are independent of chosen coordinates.

Harmonic Gauge

One example of an interesting condition to impose is given by:

$$\Gamma^{\lambda} = g^{\mu\nu}\Gamma^{\lambda}_{\mu\nu} = 0. \tag{2.30}$$

A coordinate system where this is satisfied is called the **Harmonic Gauge**. The harmonic gauge is of interest in our case because it allows all waves to be described in a comprehensible way, namely as harmonic functions. A function f is **harmonic** if $\Box f = 0$, where $\Box$ is the **D'Alembertian operator** defined as $\Box = g^{\lambda\kappa}\nabla_{\lambda}\nabla_{\kappa}$[4]. This can be written as:

$$\begin{aligned} g^{\lambda\kappa}\nabla_{\lambda}\nabla_{\kappa}f &= g^{\lambda\kappa}\left(\frac{\partial\nabla_{\lambda}f}{\partial x^{\kappa}} - \Gamma^{\alpha}_{\lambda\kappa}\nabla_{\alpha}f\right) = g^{\lambda\kappa}\left[\frac{\partial^2 f}{\partial x^{\kappa}\partial x^{\lambda}} - \Gamma^{\alpha}_{\lambda\kappa}\frac{\partial f}{\partial x^{\alpha}}\right] \\ &= g^{\lambda\kappa}\frac{\partial^2 f}{\partial x^{\kappa}\partial x^{\lambda}} - \Gamma^{\alpha}\frac{\partial f}{\partial x^{\alpha}}. \end{aligned} \tag{2.31}$$

It should be noted that it is always possible to choose this gauge. This is seen by indicating the indices of the new coordinate system with Greek indices, and those of the old one with Latin indices, we note by relabelling and then contracting eq.(1.36) with $g^{\mu\nu}$:

$$\Gamma^{\lambda} = \frac{\partial x^{\lambda}}{\partial x^{r}}\Gamma^{r} + g^{\mu\nu}\frac{\partial x^{\lambda}}{\partial x^{s}}\frac{\partial^2 x^{s}}{\partial x^{\mu}\partial x^{\nu}} \Rightarrow \Gamma^{\lambda} = \frac{\partial x^{\lambda}}{\partial x^{r}}\Gamma^{r} - g^{rs}\frac{\partial^2 x^{\lambda}}{\partial x^{r}\partial x^{s}}. \tag{2.32}$$

The implication above follows from the fact that:

$$\begin{aligned} g^{\mu\nu}\frac{\partial x^{\lambda}}{\partial x^{s}}\frac{\partial^2 x^{s}}{\partial x^{\mu}\partial x^{\nu}} &= g^{\mu\nu}\left\{\frac{\partial}{\partial x^{\mu}}\frac{\partial x^{\lambda}}{\partial x^{s}}\frac{\partial x^{s}}{\partial x^{\nu}} - \frac{\partial x^{s}}{\partial x^{\nu}}\frac{\partial^2 x^{\lambda}}{\partial x^{\mu}\partial x^{s}}\right\} \\ &= g^{\mu\nu}\left\{\frac{\partial}{\partial x^{\mu}}\delta^{\lambda}_{\nu} - \frac{\partial x^{s}}{\partial x^{\nu}}\frac{\partial x^{r}}{\partial x^{\mu}}\frac{\partial^2 x^{\lambda}}{\partial x^{r}\partial x^{s}}\right\}. \end{aligned} \tag{2.33}$$

Therefore, if Γ^{ρ} is non-zero, we can find a system where $\Gamma^{\lambda'}$ is zero and as such reduce to the harmonic gauge.

Example 2.2.3. As an example, note that the function given by the coordinate $f = x^{\mu}$ satisfies, when $\Gamma^{\lambda} = 0$ in eq.(2.31):

$$\Box x^{\mu} = g^{\lambda\kappa}\frac{\partial^2 x^{\mu}}{\partial x^{\kappa}\partial x^{\lambda}} = g^{\lambda\kappa}\partial_{\kappa}\delta^{\mu}_{\lambda} = 0. \tag{2.34}$$

Thus it behaves as a wave in the harmonic gauge. ■

2.3 Summary

In this chapter we derived Einstein's equations dependent on physical intuition, following along the lines of physicists that perceived this theory. A derivation using more mathematical insight is given in the appendix. These equations will be the tool needed to describe Gravitational Waves in the next chapter.

[4]In coordinates, this can be written (for flat space) as the known: $\Box_F = -\frac{\partial^2}{\partial t^2} + \sum_{i=1,2,3}\frac{\partial^2}{\partial x_i^2}$

Chapter 3

Gravitational Waves

The theory of General Relativity predicts the existence of Gravitational Waves (GWs). These waves can be studied by two different approaches: one based on perturbative methods and a second one, based on an exact solution of the Einstein equations. The second method has so far not led to solutions. This is due to the complexity and non-linearity of the Einstein equations, which makes finding an exact solution almost impossible. The first method will be explored here, and imposing symmetries on the solutions to Einstein's equations makes it easier to solve the equations.
Due to the fact that the metric tensor can also be interpreted as the equivalent of the gravitational potential as described in the previous chapter, GWs can be seen as metric waves. As opposed to electromagnetic waves, which constitute radiation *through* spacetime, GWs are thus deformations *of* said spacetime, behaving as waves. This means that when they propagate the geometry of spacetime changes, thus the proper distance between points in spacetime changes with time. In this chapter, we will mostly follow [8].

3.1 Perturbative solution

Let $g^0_{\mu\nu}$ be a known solution to the Einstein equations and write $g_{\mu\nu} = g^0_{\mu\nu} + h_{\mu\nu}$[1] where $|h_{\mu\nu}| \ll |g^0_{\mu\nu}|$. We have to find the inverse of $g_{\mu\nu}$ to contract some identities, to which end we try $g^{\mu\nu} = g^{0\mu\nu} - h^{\mu\nu}$, where $h^{\mu\nu} = g^{0\mu\alpha}g^{0\nu\beta}h_{\alpha\beta}$. We then see that:

$$(g^{0\mu\nu} - h^{\mu\nu})(g^0_{\nu\alpha} + h_{\nu\alpha}) = g^{0\mu\nu}g^0_{\nu\alpha} + \mathcal{O}(h^2) + g^{0\mu\nu}h_{\nu\alpha} - g^{0\mu\alpha}g^{0\nu\beta}h_{\alpha\beta}g^0_{\nu\alpha} \tag{3.1}$$

$$= \delta^\mu_\alpha + \mathcal{O}(h^2) + g^{0\mu\nu}h_{\nu\alpha} - \delta^\mu_\nu g^{0\nu\beta}h_{\alpha\beta} = \delta^\mu_\alpha + \mathcal{O}(h^2). \tag{3.2}$$

After contracting the Einstein equations (eq.(2.1)) by $g^{\mu\nu}$ and noting that $g^{\mu\nu}g_{\mu\nu} = 4$ we see that (when $T = g^{\mu\nu}T_{\mu\nu}$):

$$R - 2R = \frac{8\pi G}{c^4}T \Rightarrow R = -\frac{8\pi G}{c^4}T \Rightarrow R_{\mu\nu} = \frac{8\pi G}{c^4}(T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T). \tag{3.3}$$

We write from now on $T_{\mu\nu}$ for the complete stress-energy tensor that is associated to $g_{\mu\nu}$. We can solve Einstein equations for $T^0_{\mu\nu}$, finding $g^0_{\mu\nu}$ as solution. Then we solve the equations for known $g_{\mu\nu}$ and find $T_{\mu\nu}$. This means that the quantity $\delta T_{\mu\nu} := T_{\mu\nu} - T^0_{\mu\nu}$ can be associated to $h_{\mu\nu}$.

[1] We implicitly write $h_{\mu\nu} := h \cdot p_{\mu\nu}$, where $p_{\mu\nu}$ is a symmetric 4×4 matrix describing the perturbation and $h \in \mathbb{R}$ small, such that the components of $h_{\mu\nu}$ are small when compared with $g^0_{\mu\nu}$. It thus holds that $g_{\mu\nu} = g^0_{\mu\nu} + hp_{\mu\nu} + \mathcal{O}(h^2)$ when we expand the metric around the background metric g^0.

Remembering the expressions for the Ricci tensor (eq.(1.54)) and the Christoffel symbols (eq.(1.40)), we may write for the Christoffel symbols of the perturbed metric:

$$\begin{aligned}
\Gamma^{\gamma}_{\beta\mu}(g_{\mu\nu}) &= \frac{1}{2}[g^{0\alpha\gamma} - h^{\alpha\gamma}][(\partial_\mu g^0_{\alpha\beta} + \partial_\beta g^0_{\alpha\mu} - \partial_\alpha g^0_{\beta\mu}) + (\partial_\mu h_{\alpha\beta} + \partial_\beta h_{\alpha\mu} - \partial_\alpha h_{\beta\mu})] \\
&= \Gamma^{\gamma}_{\beta\mu}(g^0) + \delta\Gamma^{\gamma}_{\beta\mu}(h) + O(h^2); \\
\delta\Gamma^{\gamma}_{\beta\mu}(h) &= \frac{1}{2}g^{0\alpha\gamma}[\partial_\mu h_{\alpha\beta} + \partial_\beta h_{\alpha\mu} - \partial_\alpha h_{\beta\mu}] - \frac{1}{2}h^{\alpha\gamma}[\partial_\mu g^0_{\alpha\beta} + \partial_\beta g^0_{\alpha\mu} - \partial_\alpha g^0_{\beta\mu}].
\end{aligned} \tag{3.4}$$

This results in the Ricci tensor becoming:

$$\begin{aligned}
R_{\mu\nu}(g_{\mu\nu}) &= R^0_{\mu\nu}(g^0) + \delta R_{\mu\nu}(h) + O(h^2). \\
\delta R_{\mu\nu}(h) &= \frac{\partial}{\partial x^\alpha}\delta\Gamma^{\alpha}_{\mu\nu}(h) - \frac{\partial}{\partial x^\nu}\delta\Gamma^{\alpha}_{\mu\alpha}(h) + \Gamma^{\alpha}_{\sigma\alpha}(g^0)\delta\Gamma^{\sigma}_{\mu\nu}(h) \\
&\quad + \delta\Gamma^{\alpha}_{\sigma\alpha}(h)\Gamma^{\sigma}_{\mu\nu}(g^0) - \Gamma^{\alpha}_{\sigma\nu}(g^0)\delta\Gamma^{\sigma}_{\mu\alpha}(h) - \delta\Gamma^{\alpha}_{\sigma\nu}(h)\Gamma^{\sigma}_{\mu\alpha}(g^0).
\end{aligned} \tag{3.5}$$

Working out the right hand side of eq.(3.3) we find that:

$$T = g^{\mu\nu}T_{\mu\nu} = g^{0\mu\nu}T^0_{\mu\nu} - h^{\mu\nu}T^0_{\mu\nu} - g^{0\mu\nu}\delta T_{\mu\nu} + \mathcal{O}(h^2) \equiv T^0 + \delta T + \mathcal{O}(h^2). \tag{3.6}$$

Consequently:

$$T_{\mu\nu} - \frac{1}{2}g_{\mu\nu}T = T^0_{\mu\nu} - \frac{1}{2}g^0_{\mu\nu}T^0 + \delta T_{\mu\nu} - \frac{1}{2}(g^0_{\mu\nu}\delta T + h_{\mu\nu}T^0) + \mathcal{O}(h^2). \tag{3.7}$$

Now combining eq.(3.5) and eq.(3.7) and remembering that the exact solution satisfies $R_{\mu\nu}(g^0) = \frac{8\pi G}{c^4}(T^0_{\mu\nu} - \frac{1}{2}g^0_{\mu\nu}T^0)$, the Einstein equations for the perturbation reduce to:

$$\delta R_{\mu\nu}(h) = \frac{8\pi G}{c^4}\left[\delta T_{\mu\nu} - \frac{1}{2}(g^0_{\mu\nu}\delta T + h_{\mu\nu}T^0)\right] + \mathcal{O}(h^2). \tag{3.8}$$

3.1.1 Flat spacetime

Because we are working under the assumption that gravitational waves are very weak[2] compared to the gravitational fields generated by the Earth and the Sun, we can approximate the regions where we observe gravitational waves to be flat with a small perturbation. Therefore consider $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ where h is an arbitrary 4×4 matrix with components $|h_{\mu\nu}| \ll 1$ and $\eta_{\mu\nu}$ is the Minkowski metric. Because in flat spacetime the connection coefficients vanish, we find that:

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2}\eta^{\lambda\rho}\left[\partial_\mu h_{\rho\nu} + \partial_\nu h_{\rho\mu} - \partial_\rho h_{\mu\nu}\right] + \mathcal{O}(h^2), \tag{3.9}$$

and the equations in eq.(3.8) reduce to:

$$\begin{aligned}
&\partial_\alpha\Gamma^{\alpha}_{\mu\nu}(h) - \partial_\nu\Gamma^{\alpha}_{\mu\alpha}(h) + \mathcal{O}(h^2) \\
&\quad = \frac{1}{2}\left\{-\Box_F h_{\mu\nu} + \left[\partial_\lambda\partial_\mu h^\lambda_\nu + \partial_\lambda\partial_\nu h^\lambda_\mu - \partial_\mu\partial_\nu h^\lambda_\lambda\right]\right\} + \mathcal{O}(h^2).
\end{aligned} \tag{3.10}$$

[2]It will be seen later on that the signal of GWs is proportional to $1/r$, and the nearest sources of GWs are at such distances that we can consider the amplitude to be small w.r.t. the background metric.

The operator $\Box_F$ is the **D'Alembertian for flat spacetime**, defined as:

$$\Box_F = \eta^{\alpha\beta}\partial_\alpha\partial_\beta = -\frac{\partial^2}{\partial t^2 c^2} + \nabla^2. \tag{3.11}$$

Because for flat spacetime also $T^0 = 0$, we find that the Einstein equations for the perturbation $h_{\mu\nu}$ become:

$$\Box_F h_{\mu\nu} - \left[\partial_\lambda\partial_\mu h^\lambda_\nu + \partial_\lambda\partial_\nu h^\lambda_\mu - \partial_\mu\partial_\nu h^\lambda_\lambda\right] = -\frac{16\pi G}{c^4}\left(\delta T_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\delta T\right). \tag{3.12}$$

As discussed before, there is a gauge freedom in the Einstein equations dependent on invariance of the solution under coordinate transformations. If we make an arbitrary coordinate transformation, the transformed metric tensor is still a solution. But since we assume that $|h_{\mu\nu}| \ll 1$, we have to assume that $|h'_{\mu\nu}| \ll 1$ as well, if $h'_{\mu\nu}$ is the transformed perturbation. This means that we need to perform a coordinate transformation such that the new coordinates are $x'^\mu = x^\mu + \epsilon^\mu(x)$, where ϵ^μ is an arbitrary vector such that $\partial_\nu \epsilon^\mu$ is of the same order as $h_{\mu\nu}$. For the new perturbation it then holds that:

$$h'_{\mu\nu} = h_{\mu\nu} - \partial_\nu\epsilon_\mu - \partial_\mu\epsilon_\nu. \tag{3.13}$$

As mentioned before, the harmonic gauge simplifies the equation in eq.(3.12). The following theorem states that it is always possible to transform to this gauge:

Theorem 3.1.1. *If the harmonic gauge condition (eq.(2.30)) is not satisfied in a reference frame, it is always possible to perform an infinitesimal coordinate transformation $x'^\lambda = x^\lambda + \epsilon^\lambda$ such that in the new frame the condition is satisfied. This is only possible provided:*

$$\Box_F \epsilon_\rho = \partial_\beta h^\beta_\rho - \frac{1}{2}\partial_\rho h^\beta_\beta. \tag{3.14}$$

Proof: The object Γ^λ transforms according to eq.(2.32) where, as seen from the form of the transformation:

$$\partial_\rho x'^\lambda = \delta^\lambda_\rho + \partial_\rho \epsilon^\lambda. \tag{3.15}$$

We see that:

$$\begin{aligned}\Gamma^\lambda = g^{\mu\nu}\Gamma^\lambda_{\mu\nu} &= \frac{1}{2}\eta^{\mu\nu}\eta^{\lambda\kappa}\left[\partial_\nu h_{\kappa\mu} + \partial_\mu h_{\kappa\nu} - \partial_\kappa h_{\mu\nu}\right] \\ &= \frac{1}{2}\eta^{\lambda\kappa}[\partial_\nu h^\nu_\kappa + \partial_\mu h^\mu_\kappa - \partial_\kappa h^\nu_\nu] = \eta^{\rho\kappa}\left[\partial_\mu h^\mu_\kappa - \frac{1}{2}\partial_\kappa h^\nu_\nu\right].\end{aligned} \tag{3.16}$$

Additionally, we see that:

$$\begin{aligned}g^{\rho\sigma}\partial_\rho\partial_\sigma x'^\lambda &= g^{\rho\sigma}\partial_\rho\left(\partial_\sigma x^\lambda + \partial_\sigma\epsilon^\lambda\right) \\ &\simeq \eta^{\rho\sigma}\partial_\rho\partial_\sigma\epsilon^\lambda = \Box_F\epsilon^\lambda.\end{aligned} \tag{3.17}$$

The $\simeq$ arises from the fact that $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ and we leave out terms of order h^2 and higher. Consequently the gauge condition in the new coordinate system becomes:

$$\begin{aligned}0 = \Gamma'^\lambda &= \left[\delta^\lambda_\rho + \partial_\rho\epsilon^\lambda\right]\eta^{\rho\kappa}\left[\partial_\mu h^\mu_\kappa - \frac{1}{2}\partial_\kappa h^\nu_\nu\right] - \Box_F\epsilon^\lambda \\ &\simeq \eta^{\lambda\kappa}\left[\partial_\mu h^\mu_\kappa - \frac{1}{2}\partial_\kappa h^\nu_\nu\right] - \Box_F\epsilon^\lambda \\ &\Rightarrow \Box_F\epsilon_\alpha = \partial_\mu h^\mu_\alpha - \frac{1}{2}\partial_\alpha h^\nu_\nu.\end{aligned} \tag{3.18}$$

The $\simeq$ again is where we leave out terms of order h^2 or higher. The implication follows after contracting with $\eta_{\lambda\alpha}$. $\square$

Up to first order terms in $h_{\mu\nu}$, the harmonic gauge condition implies that in eq.(3.12):

$$\partial_\mu h^\mu_\nu = \frac{1}{2}\partial_\nu h^\mu_\mu. \Rightarrow \square_F h_{\mu\nu} = -\frac{16\pi G}{c^4}\left(\delta T_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\delta T\right). \tag{3.19}$$

Introducing the tensor $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$, where $h = \eta^{\mu\nu}h_{\mu\nu}$, this becomes:

$$\begin{cases} \square_F \bar{h}_{\mu\nu} & = -\frac{16\pi G}{c^4}\delta T_{\mu\nu} \\ \partial_\mu \bar{h}^\mu_\nu & = 0 \end{cases} \tag{3.20}$$

This is a typical form of a wave-equation, meaning a perturbation in a flat metric indeed propagates as a wave.

Example 3.1.1. The simplest solution of the wave equation in vacuum ($\delta T_{\mu\nu} = 0$ in eq.(3.20)) is a so-called **monochromatic plane wave** of the form:

$$\bar{h}_{\mu\nu} = \Re\left\{A_{\mu\nu}e^{ik_\alpha x^\alpha}\right\}, \tag{3.21}$$

where $A_{\mu\nu}$ is the **polarization tensor** (comparable with the wave amplitude) and $\vec{k}$ the **wave vector**, indicating the direction the wave is travelling. Direct substitution of eq.(3.21) in the first equation of eq.(3.20) yields:

$$\square_F \bar{h}_{\mu\nu} = \eta^{\alpha\beta}\partial_\alpha\partial_\beta e^{ik_\gamma x^\gamma} = -\eta^{\alpha\beta}k_\alpha k_\beta e^{ik_\gamma x^\gamma} \Rightarrow \eta^{\alpha\beta}k_\alpha k_\beta = 0. \tag{3.22}$$

Thus $\vec{k}$ should be a null-vector, meaning the waves travel at light-speed. The harmonic gauge condition implies furthermore that:

$$\partial_\mu \bar{h}^\mu_\nu = 0 \Rightarrow \eta^{\mu\alpha}\partial_\mu \bar{h}_{\alpha\nu} = 0 \Rightarrow \eta^{\mu\alpha}A_{\alpha\nu}k_\mu = 0 \Rightarrow k_\mu A^\mu_\nu = 0. \tag{3.23}$$

Consequently, the polarization tensor and the wave vector should be orthogonal, which is analogous to the way electromagnetic waves propagate in ordinary space. ■

3.2 TT-gauge

To understand how many of the 10 components of $h_{\mu\nu}$ have a physical meaning, we need to find the degrees of freedom for a gravitational plane wave. Thus in the following we consider a wave propagating in flat spacetime along the $x^1 = x$-direction.
Since $h_{\mu\nu}$ is independent of y and z in this case, eq.(3.20) becomes:

$$\begin{cases} \left(-\frac{\partial^2}{\partial t^2 c^2} + \frac{\partial^2}{\partial x^2}\right)\bar{h}^\mu_\nu & = 0 \\ \partial_\mu \bar{h}^\mu_\nu & = 0 \end{cases}, \tag{3.24}$$

where we raised and lowered indices with $\eta_{\mu\nu}$. It is clear now that $\bar{h}^\mu_\nu$ is an arbitrary function of $\chi = t \pm x/c$. This implies that:

$$\partial_\mu \bar{h}^\mu_\nu = \frac{1}{c}\frac{\partial \bar{h}^t_\nu}{\partial t} + \frac{\partial \bar{h}^x_\nu}{\partial x} = \frac{1}{c}\frac{\partial}{\partial\chi}(\bar{h}^t_\nu - \bar{h}^x_\nu) = 0. \tag{3.25}$$

By integrating eq.(3.25) and by choosing the integration constants to be equal to zero (because we only have interest in the time-dependent part of the solution), we find that $\bar{h}^t_\mu = \bar{h}^x_\mu$ for $\mu = t, x, y, z$, meaning that we are left with six independent components.

We will now show that there exist another four degrees of freedom by making an infinitesimal coordinate transformation $x'^\mu = x^\mu + \epsilon^\mu$ such that $\Box_F \epsilon^\mu = 0$. From this we find that, for a solution $\bar{h}_{\mu\nu}$ of the wave equation in eq.(3.20) for $\delta T_{\mu\nu} = 0$, the perturbations in the new gauge ($h'_{\mu\nu} = h_{\mu\nu} - \partial_\mu \epsilon_\nu - \partial_\nu \epsilon_\mu$) satisfy:

$$\bar{h}'_{\mu\nu} = \bar{h}_{\mu\nu} - \partial_\mu \epsilon_\nu - \partial_\nu \epsilon_\mu + \eta_{\mu\nu} \partial^\alpha \epsilon_\alpha. \tag{3.26}$$

This implies that $\Box_F \bar{h}'_{\mu\nu}$, thus that $\bar{h}'_{\mu\nu}$ also satisfies the wave equation.
Now, using these four degrees of freedom to set the four quantities $\bar{h}^t_i = 0$ ($i = x, y, z$) and $\bar{h}^y_y + \bar{h}^z_z = 0$, it follows that the remaining non-vanishing components are $\bar{h}^z_y$ and $\bar{h}^y_y - \bar{h}^z_z$. We have used all of our gauge freedom, and reduced the independent components of $\bar{h}_{\mu\nu}$ to the two functions $\bar{h}_{zy} = \bar{h}_{yz}$ and $\bar{h}_{yy} = -\bar{h}_{zz}$.
From the fact that we chose the four components above to be zero we see that:

$$\bar{h}^x_i = \bar{h}^t_t = 0; \quad i = x, y, z, \tag{3.27}$$

implying that:

$$\bar{h}^\mu_\mu = 0. \tag{3.28}$$

Since $\bar{h}^\mu_\mu = h^\mu_\mu - 2h^\mu_\mu = -h^\mu_\mu$ it follows that $h^\mu_\mu = 0$, thus in the gauge we transformed to in eq.(3.26) $h_{\mu\nu}$ and $\bar{h}_{\mu\nu}$ coincide and are traceless.
In conclusion, a gravitational wave in vacuum only has two physical degrees of freedom which correspond to two possible polarization states. The gauge in which this is clearly manifested in the way derived above is called the "transverse and traceless" gauge, or TT-gauge. In the TT-gauge the components of $h_{\mu\nu}$ are only different from zero on the plane orthogonal to the direction of propagation and that $h_{\mu\nu}$ is traceless. We can always transform to the TT-gauge by the manner of using the projection operators (see below). This operation is equivalent to finding an infinitesimal coordinate transformation ϵ^α such that $x'^\alpha = x^\alpha + \epsilon^\alpha$ and $\Box_F \epsilon^\alpha = 0$ and to imposing that in the new frame, $\bar{h}'_{\alpha\beta} n^\beta = 0$ and $\bar{h}'_{\alpha\beta} \delta^{\alpha\beta} = 0$, to ensure transverseness and tracelessness, when n^β is the vector along which the wave propagates.

We now define the **orthogonal projector** $P_{jk} = \delta_{jk} - n_j n_k$ that projects a vector onto the plane orthogonal to the direction of $\mathbf{n} = (n_x, n_y, n_z)$. Properties of this operator include symmetry and the facts that $P_{jk} V^k n^j = 0$, $n^j P_{jk} = 0$ and $P^j_k P^k_l V^l = P^j_l V^l$.
Now define the **transverse-traceless projector** $\mathcal{P}_{jkmn} = P_{jm} P_{kn} - \frac{1}{2} P_{jk} P_{mn}$. It satisfies the following identities:

1. $\mathcal{P}_{jklm} = \mathcal{P}_{lmjk}$.
2. $\mathcal{P}_{jklm} = \mathcal{P}_{kjml}$.
3. $\mathcal{P}_{jkmn} \mathcal{P}_{mnrs} = \mathcal{P}_{jkrs}$.
4. Transverseness: $n^j \mathcal{P}_{jkmn} = n^k \mathcal{P}_{jkmn} = n^m \mathcal{P}_{jkmn} = n^n \mathcal{P}_{jkmn} = 0$.
5. Tracelessness: $\delta^{jk} \mathcal{P}_{jkmn} = \delta^{mn} \mathcal{P}_{jkmn} = 0$.

Since h_{mn} and $\bar{h}_{mn}$ are different only in the sense that their traces are different, the metric perturbation in the TT-gauge is found by applying the transverse-traceless operator to either one of them:

$$h_{jk}^{TT} = \mathcal{P}_{jkmn} h_{mn} = \mathcal{P}_{jkmn} \bar{h}_{mn}. \tag{3.29}$$

3.3 Effect on test masses

To detect GWs, we need to know what the effect of GWs is on test particles. First, we will consider one particle, with a locally inertial frame attached to it and with the x-axis coincident with the direction of propagation of an incoming GW, observed from the TT-gauge[3]. The geodesic equation tells us that:

$$\frac{d^2 x^\alpha}{d\tau^2} + \Gamma^\alpha_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 0. \tag{3.30}$$

Assuming that the particle is at rest at $t = 0$, its acceleration due to the GW is:

$$\left.\frac{d^2 x^\alpha}{d\tau^2}\right|_{t=0} = -\Gamma^\alpha_{00} = -\frac{1}{2}\eta^{\alpha\beta}[\partial_0 h_{\beta 0} + \partial_0 h_{0\beta} - \partial_\beta h_{00}] = 0. \tag{3.31}$$

This follows from the form of $h_{\mu\nu}$ in the TT-gauge.

Thus we conclude that GWs can not be studied by observing the motion of a single particle. Therefore, we study the relative motion of two test particles induced by a GW. To do this, consider two infinitesimally close particles A and B at coordinates x^μ_A and x^μ_B, initially at rest with **coordinate separation** $\delta x^\mu = x^\mu_B - x^\mu_A$. The wave reaches the particles at $t = 0$, propagating along the x-axis in the TT-gauge. Since $g_{00} = \eta_{00} = -1$ both particles have the same proper time $\tau = ct$. Note that the coordinate distance remains constant, but that the proper distance changes. For example, assuming the particles both on the z-axis, we see:

$$\Delta l = \int ds = \int_{z_A}^{z_B} |g_{zz}|^{1/2} dz = \int_{z_A}^{z_B} |1 + h_{zz}(t - x/c)|^{1/2} dz. \tag{3.32}$$

Changing coordinates to a LIF centred on the geodesic of particle A, we obtain a metric different from the Minkowski metric only in terms of order $|\delta x|^2$ (note that we can always choose such a frame as mentioned in previous chapters). In this frame $t_A = \tau/c$, $(dx^\mu/d\tau)_A = (1, 0, 0, 0)$, $g_{\mu\nu}\Big|_A = \eta_{\mu\nu}$ and $\Gamma^\alpha_{\mu\nu}\Big|_A$. $x^i_B = \delta x^i$ for $i = 1, 2, 3$ (space components only) are the coordinates of B.
The separation vector satisfies the geodesic deviation from eq.(1.67). Evaluating those equations along the geodesic of particle A we find:

$$\frac{1}{c^2}\frac{d^2 \delta x^i}{dt^2} = R^i_{00j}\delta x^j. \tag{3.33}$$

[3] In the following, we will do all calculations in the TT-gauge, thus we will not use the notation $h^{TT}_{\mu\nu}$ and just write $h_{\mu\nu}$.

The Riemann tensor for $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ satisfies:

$$\begin{aligned}
R_{\alpha\kappa\lambda\mu} &= \frac{1}{2}\left(\partial_\kappa\partial_\lambda g_{\alpha\mu} + \partial_\alpha\partial_\mu g_{\kappa\lambda} - \partial_\kappa\partial_\mu g_{\alpha\lambda} - \partial_\alpha\partial_\lambda g_{\kappa\mu}\right) \\
&\quad + g_{\nu\sigma}(\Gamma^\nu_{\kappa\lambda}\Gamma^\sigma_{\alpha\mu} - \Gamma^\nu_{\kappa\mu}\Gamma^\sigma_{\alpha\lambda}) \\
\Rightarrow R_{\alpha\kappa\lambda\mu} &= \frac{1}{2}\left(\frac{\partial^2 h_{\alpha\mu}}{\partial x^\kappa \partial x^\lambda} + \frac{\partial^2 h_{\kappa\kappa}}{\partial x^\alpha \partial x^\mu} - \frac{\partial^2 h_{\alpha\lambda}}{\partial x^\kappa \partial x^\mu} - \frac{\partial^2 h_{\kappa\mu}}{\partial x^\alpha \partial x^\lambda}\right) + \mathcal{O}(h^2) \\
\Rightarrow R_{i00m} &= \frac{1}{2}\partial_0\partial_0 h_{im}.
\end{aligned} \tag{3.34}$$

The first implication arises after neglecting terms of order h^2 and higher. The second implication follows from the fact that in the TT-gauge, $h_{i0} = h_{00} = 0$. Note that i and m can only assume the values 2 and 3, yielding the geodesic deviation as:

$$\frac{d^2}{dt^2}\delta x^\lambda = \frac{1}{2}\eta^{\lambda i}\frac{\partial^2 h_{im}}{\partial t^2}\delta x^m. \tag{3.35}$$

For $t \leq 0$ the particles are still at rest with respect to each other, but for $t > 0$ we describe the change in relative position with $\delta x_1^j(t)$, a small perturbation with respect to the initial position δx_0^j. This assumption is valid because $|h_{\mu\nu}| \ll 1$.
Substituting this in the equation for geodesic deviation and neglecting terms of order h^2 and higher, we find:

$$\frac{d^2}{dt^2}\delta x_1^j = \frac{1}{2}\eta^{ji}\frac{\partial^2 h_{ik}}{\partial t^2}\delta x_0^k \Rightarrow \delta x^j = \delta x_0^j + \frac{1}{2}\eta^{ji}h_{ik}\delta x_0^k. \tag{3.36}$$

In order to study the effect of the GW on test particles we start by assuming the non-vanishing components equal to:

$$\begin{aligned}
h_{yy} = -h_{zz} &= 2\Re\left\{A_+ e^{i\omega(t-x/c)}\right\} \\
h_{yz} = h_{zy} &= 2\Re\left\{A_\times e^{i\omega(t-x/c)}\right\}
\end{aligned} \tag{3.37}$$

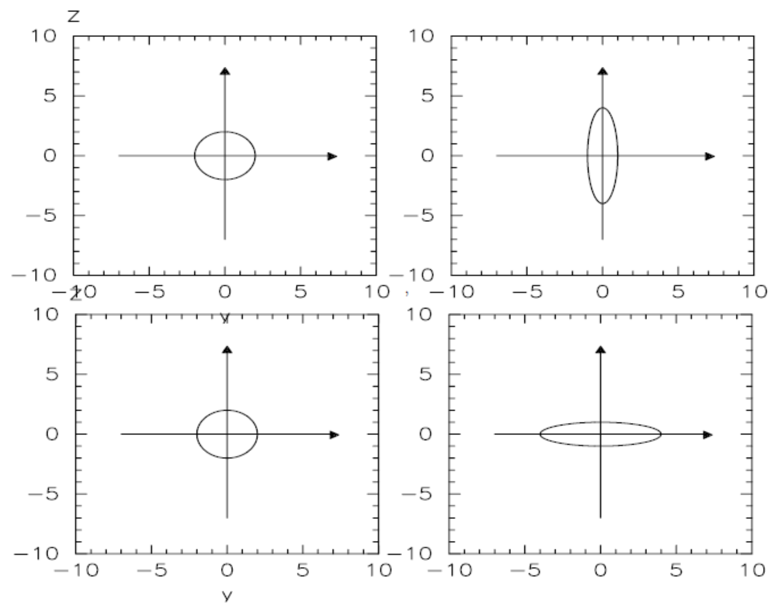
We consider the effects of the A_+ and $A_\times$ separately. Assume only $A_+ \neq 0$ with only real components and consider two particles, 1) at $(0, y_0, 0)$ and 2) at $(0, 0, z_0)$. The form of the displacements then equals:

$$\begin{aligned}
&1) z = 0; \quad y = y_0 + \frac{1}{2}h_{yy}y_0 = y_0[1 + A_+\cos(\omega(t - x/c))]; \\
&2) y = 0; \quad z = z_0 + \frac{1}{2}h_{zz}z_0 = z_0[1 - A_+\cos(\omega(t - x/c))].
\end{aligned} \tag{3.38}$$

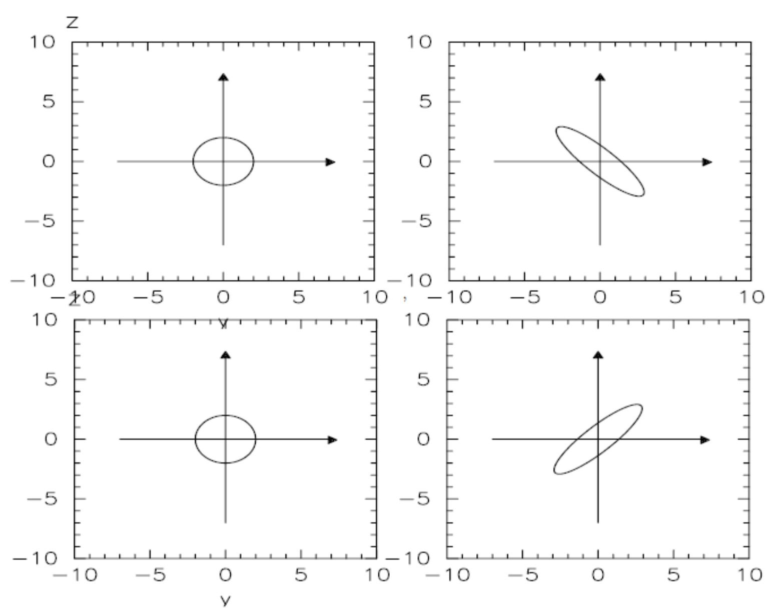
Now for $A_+ = 0$ and $A_\times \neq 0$ and real, we see that the equations for an arbitrary particle at $(0, y_0, z_0)$ become:

$$\begin{aligned}
y &= y_0 + \frac{1}{2}h_{yz}z_0 = y_0 + z_0 A_\times \cos(\omega(t - x/c)); \\
z &= z_0 + \frac{1}{2}h_{zy}y_0 = z_0 + y_0 A_\times \cos(\omega(t - x/c)).
\end{aligned} \tag{3.39}$$

These expressions for the effect of the "+"- and the "×"-polarization are displayed in figure 3.1 for the instances where $t = 0$, $t = P/4$, $t = P/2$ and $t = 3P/4$.



(a) Effect of a plus-polarized ($h_\times = 0, h_+ \neq 0$) GW on a ring of test masses.



(b) Effect of a cross-polarized ($h_+ = 0, h_\times \neq 0$) GW on a ring of test masses.

Figure 3.1: The effect of both polarizations of GWs on a ring of test masses. The propagation direction is taken to be x, such that the effect is limited to the y-z-plane. The pictures shown are for $t = 0$, $t = P/4$, $t = P/2$ and $t = 3P/4$. These effects have been calculated in chapter 3. The exact expressions for the positions of the particles can be found in section 13.6 of [8].

3.4 Quadrupole Formalism

To study the form of GWs generated by binary systems we need to explore the so-called *Quadrupole Formalism.* The gravitational energy and waveforms emitted by an evolving binary system will depend on the stress-energy tensor of the system.
We shall solve eq.(3.20) under the assumption that the source is confined in a region much smaller than the wavelength of the emitted radiation. This is also called the *slow-motion approximation* because it implies that the typical velocities in the physical system are much smaller than the speed of light. These assumptions significantly simplify the solution of the Einstein equations. More importantly, these assumptions are physically viable because in most binary systems not close to coalescing the orbital speed is much slower than the speed of light. Furthermore, this approximation is valid for most observed binary systems of white dwarfs. For example, the two white dwarfs in the system $J065133.34 + 284423.4$ $(J06)$[4] have Keplerian velocities of $a\omega_K/c = 3.18 \cdot 10^{-3} \ll 1$, where a is the orbital separation of the stars, and ω_K the Kepler angular velocity, equal to $2\pi/P$, where P is the orbital period.

One important result to be used later on is the Tensor Virial Theorem:

Theorem 3.4.1. *(Tensor Virial Theorem): For $k, n = 1, 2, 3$ it holds that:*

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\int_V T^{00}x^k x^n d^3x = 2\int_V T^{kn} d^3x, \tag{3.40}$$

when assuming $T_{\mu\nu}$ vanishes at the boundary of the source, which is contained in the volume V.

Proof: A full proof can be found in paragraph 14.1 in [8]. In short it reads:
Using the conservation law for T and only considering the spatial components, we see that:

$$\frac{\partial T^{n0}}{\partial x^0} = -\frac{\partial T^{ni}}{\partial x^i} \Rightarrow \frac{1}{c}\frac{\partial}{\partial t}\int_V T^{n0}x^k d^3x = \int_V T^{nk} d^3x. \tag{3.41}$$

The implication holds after multiplying by x^k and integrating over the source volume. The equality holds after using the product rule, using Gauss' Theorem and using that $T = 0$ at the boundary of the source. Because $T^{\mu\nu}$ is symmetric this can be written as:

$$\frac{1}{2c}\frac{\partial}{\partial t}\int_V (T^{n0}x^k + T^{k0}x^n) d^3x = \int_V T^{nk} d^3x. \tag{3.42}$$

Now multiplying the 0-component of the conservation law by $x^k x^n$ and integrating over the volume as we did with the spatial part before:

$$\frac{1}{c}\frac{\partial T^{00}}{\partial t} = -\frac{\partial T^{0i}}{\partial x^i} \Rightarrow \frac{1}{c}\frac{\partial}{\partial t}\int_V T^{00}x^k x^n d^3x = \int_V (T^{0k}x^n + T^{0n}x^k) d^3x. \tag{3.43}$$

Now using the result from above after differentiating w.r.t. x^0 we find (for $n, k = 1, ..., 3$):

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\int_V T^{00}x^k x^n d^3x = 2\int_V T^{kn} d^3x. \tag{3.44}$$

This proves the theorem. □

[4] We consider binary parameters as listed in [14], table B1.

We will use this result to simplify the equations for GWs, which will be derived later. First we will transform the expressions for $\bar{h}_{\mu\nu}$ and $T_{\mu\nu} \equiv \delta T_{\mu\nu}$ (we will omit the δ in this analysis) into the Fourier domain, which results in the following form of eq.(3.20):

$$[\nabla^2 + \omega^2/c^2]\bar{h}_{\mu\nu}(\omega, x^i) = -KT_{\mu\nu}(\omega, x^i); \quad i = 1, 2, 3, \tag{3.45}$$

where $K = 16\pi G/c^4$ and ω is such that $\lambda_{GW} = 2\pi c/\omega$.

Outside the source, where $T_{\mu\nu} = 0$ we see that the simplest solution of this equation is a spherical wave. We are only interested in emitted waves, so only consider the outgoing part. This is all represented in a form such as:

$$\bar{h}_{\mu\nu}(\omega, r) = \frac{A_{\mu\nu}(\omega)}{r} e^{ir\omega/c}, \tag{3.46}$$

independent of the spherical coordinates θ and ϕ. The exact expression for $A_{\mu\nu}$ depends on the solution of the Einstein equations inside the source, which we will now explore.

3.4.1 Interior solution

Integrating over the source volume (which is assumed to be contained within a sphere of radius $\epsilon \ll 2\pi c/\omega$ for the slow-motion approximation) gives:

$$\int_V \left[\nabla^2 + \frac{\omega^2}{c^2}\right] \bar{h}_{\mu\nu}(\omega, x^i) d^3x = -K \int_V T_{\mu\nu}(\omega, x^i) d^3x. \tag{3.47}$$

The first term in eq.(3.47) is equal to:

$$\int_V \text{div}[\nabla \bar{h}_{\mu\nu}] d^3x = \int_S \left(\nabla \bar{h}_{\mu\nu}\right)^k dS_k, \tag{3.48}$$

due to Gauss' theorem. This is approximately equal to (when only retaining terms of order ϵ and noting that $e^{ir\omega/c} \sim 1$ because $\lambda_{GW} >> \epsilon$):

$$\begin{aligned} 4\pi\epsilon^2 \left(\frac{d}{dr}\frac{A_{\mu\nu}}{r} e^{i\omega r/c}\right)_{r=\epsilon} &= 4\pi\epsilon^2 \left[-\frac{A_{\mu\nu}}{r^2} e^{i\omega r/c} + \frac{A_{\mu\nu}}{r}(i\omega/c) e^{i\omega r/c}\right]_{r=\epsilon} \\ &\approx -4\pi A_{\mu\nu}(\omega). \end{aligned} \tag{3.49}$$

The second term in eq.(3.47) satisfies:

$$\int_V \frac{\omega^2}{c^2} \bar{h}_{\mu\nu} d^3x < |\bar{h}_{\mu\nu}|_{max} \frac{\omega^2}{c^2} \frac{4}{3}\pi\epsilon^3 \approx 0. \tag{3.50}$$

Here, $|\bar{h}_{\mu\nu}|_{max}$ is the maximum reached bij $\bar{h}_{\mu\nu}$ on the volume V[5].This gives the final expression for $A_{\mu\nu}$:

$$A_{\mu\nu}(\omega) = \frac{4G}{c^4} \int_V T_{\mu\nu}(\omega, x^i) d^3x. \tag{3.51}$$

This implies that after performing the inverse Fourier transform we obtain:

$$\bar{h}_{\mu\nu}(t, r) = \frac{4G}{c^4} \frac{1}{r} \int_V T_{\mu\nu}(t - r/c, x^i) d^3x. \tag{3.52}$$

[5]This volume is closed and bounded, the functions $h_{\mu\nu}$ are continuous, so indeed, this maximum exists due to the Extreme Value Theorem.

We will simplify this by introducing the quadrupole moment tensor. Firstly, note that, to extract the physical components we still need to project $\bar{h}_{\mu\nu}$ onto the TT-gauge. Secondly, we should keep in mind that this solution is derived on two assumptions. These are the weak field approximation and the slow-motion approximation. Lastly note that the solution above satisfies both conditions from eq.(3.20), while we only used the first one to derive it. This can be checked as done at the beginning of chapter 14 of [8]. Essentially, this is due to the fact that $T_{\mu\nu}$ satisfies the conservation law and vanishes at the boundary, and to the fact that the source is contained in a very small volume compared to the typical wavelength of the GW.

3.4.2 The quadrupole moment

To simplify the expression in eq.(3.52) we use the conservation law that $T_{\mu\nu}$ satisfies:

$$\frac{1}{c}\partial_t T^{\mu 0} = -\partial_k T^{\mu k}; \quad \mu = 0, 1, 2, 3; \quad k = 1, 2, 3. \tag{3.53}$$

Integrating these expressions over the source volume V and using Gauss' Theorem on the right hand side we see that:

$$\int_V T^{\mu 0} d^3x = \text{constant} \Rightarrow \bar{h}^{\mu 0} = \text{constant} = 0. \tag{3.54}$$

The implication here follows from eq.(3.52), and we can choose the constant to be zero because we are only interested in the time-dependent part of the field.

Now we define the **quadrupole moment tensor** of the system:

$$q^{kn}(t) = \frac{1}{c^2}\int_V T^{00}(t, x^i) x^k x^n d^3x; \quad k, n = 1, 2, 3. \tag{3.55}$$

This means that:

$$\begin{cases} \bar{h}^{\mu 0} & = 0, \quad \mu = 0, 1, 2, 3 \\ \bar{h}^{ik}(t, r) & = \dfrac{2G}{c^4 r}\left[\dfrac{d^2}{dt^2} q^{ik}(t - r/c)\right] \end{cases}. \tag{3.56}$$

This we can project on the TT-gauge by applying the operator $\mathcal{P}_{jkmn}$ on q_{mn}, defining $Q^{TT}_{jk} = \mathcal{P}_{jkmn} q_{mn}$, the **transverse-traceless part of the quadrupole moment**.
Sometimes it is useful to use the **reduced quadrupole moment** $Q_{jk} = q_{jk} = \frac{1}{3}\delta_{jk} q^m_m$, whose trace is zero by definition.

3.4.3 Energy and Flux of GWs

Gravitational waves contain energy, meaning that systems emitting them will evolve over time due to losing energy. To describe this effect, we need to know how to express the energy and flux of GWs.

Energy

To describe the energy radiated in GWs by evolving systems we need a tensor describing the energy content of the gravitational field. In a LIF we want to find a tensor satisfying the law $\partial_\nu T^{\mu\nu} = 0$, that means finding a quantity antisymmetric in μ and ν, $\eta^{\mu\nu\alpha}$ such that $T^{\mu\nu} = \partial_\alpha \eta^{\mu\nu\alpha}$, this expression can be found by using the Einstein equations in the form:

$$T_{\mu\nu} = \frac{c^4}{8\pi G}\left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R\right). \tag{3.57}$$

By using the fact that we evaluate these in a LIF, we find:

$$T^{\mu\nu} = \partial_\alpha \left\{ \frac{c^4}{16\pi G} \frac{1}{(-g)} \frac{\partial}{\partial x^\beta} [(-g)(g^{\mu\nu} g^{\alpha\beta} - g^{\mu\alpha} g^{\nu\beta})] \right\} =: \partial_\alpha \eta^{\mu\nu\alpha}. \tag{3.58}$$

Introducing the quantity $\zeta^{\mu\nu\alpha} = (-g)\eta^{\mu\nu\alpha}$, we see that, in a LIF: $\partial_\alpha \zeta^{\mu\nu\alpha} = (-g)T^{\mu\nu}$. In any other frame this is not generally true, so we define the quantity $t^{\mu\nu}$ as:

$$(-g)t^{\mu\nu} = \partial_\alpha \zeta^{\mu\nu\alpha} - (-g)T^{\mu\nu}. \tag{3.59}$$

Because $\zeta^{\mu\nu\alpha}$ is antisymmetric in μ and α, the following law holds: $\dfrac{\partial}{\partial x^\mu}[(-g)(t^{\mu\nu} + T^{\mu\nu})] = 0$. Since $t^{\mu\nu}$ when added to $T^{\mu\nu}$ satisfies a conservation law (the law above has the form of a vanishing ordinary divergence) and since it vanishes in a LIF, we interpret $t^{\mu\nu}$ as the quantity containing the information about energy and momentum of the gravitational field. It should be noted that $t^{\mu\nu}$ is not a tensor, but it behaves as a tensor under linear coordinate transformations.

Flux

Consider an emitting system and the associated 3-dimensional coordinate frame O. Let an observer be situated at an arbitrary point, in the direction of a vector $\mathbf{n}$. Consider the frame O' with origin coincident with that of O and x'-axis aligned with $\mathbf{n}$. If we want to measure the flux across the surface orthogonal to x', $t^{0x'}$, we need to compute the Christoffel symbols, which we find by finding the derivatives of $h^{TT}_{\mu'\nu'}$. As derived in eq.(3.56) the metric perturbation, even after applying $\mathcal{P}_{jkmn}$, has the form of $h^{TT}(t, x') = \dfrac{\text{constant}}{x'} f(t - x'/c)$. The only derivatives which matter are those with respect to t and x', and we see that those, after neglecting terms of order $1/x'^2$ and higher, are equal to:

$$\partial_t h^{TT} = \frac{\text{constant}}{x'} \partial_t f; \quad \partial_{x'} h^{TT} \sim -\frac{1}{c} \frac{\text{constant}}{x'} \partial_t f = -\frac{1}{c} \partial_t h^{TT}. \tag{3.60}$$

This means that, in the case of linear polarization with only one polarization unequal to zero, where:

$$g_{\mu'\nu'} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 + h^{TT} & 0 \\ 0 & 0 & 0 & 1 - h^{TT} \end{pmatrix}, \tag{3.61}$$

the Christoffel symbols that do not vanish are:

$$\begin{aligned} \Gamma^0_{y'y'} &= -\Gamma^0_{z'z'} = \frac{1}{2} \partial_t h^{TT}; \qquad \Gamma^{y'}_{0y'} = -\Gamma^{z'}_{0z'} = \frac{1}{2} \partial_t h^{TT}; \\ \Gamma^{x'}_{y'y'} &= -\Gamma^{x'}_{z'z'} = \frac{1}{2c} \partial_t h^{TT}; \quad \Gamma^{y'}_{y'x'} = -\Gamma^{z'}_{z'x'} = -\frac{1}{2c} \partial_t h^{TT}. \end{aligned} \tag{3.62}$$

By substituting these in $t^{\mu\nu}$ we find:

$$ct^{0x'} = \frac{dE_{GW}}{dtdS} = \frac{c^3}{16\pi G} \left(\frac{dh^{TT}(t, x')}{dt} \right)^2. \tag{3.63}$$

Where E_{GW} is the **gravitational wave energy**. In the case of both polarizations non-zero this result generalizes to:

$$\begin{aligned} ct^{0x'} &= \frac{c^3}{16\pi G}\left[\left(\frac{dh_{+}^{TT}(t,x')}{dt}\right)^2 + \left(\frac{dh_{\times}^{TT}(t,x')}{dt}\right)^2\right] \\ &= \frac{c^3}{32\pi G}\sum_{j,k=1}^{3}\left(\frac{dh_{jk}^{TT}(t,x')}{dt}\right)^2. \end{aligned} \tag{3.64}$$

In GR the energy of the field can not be defined locally, so we need to average over several periods to find the GW-flux, meaning that for an arbitrary direction $\mathbf{r} = r\mathbf{n}$ (an observer located at another position finds the same expression but with a different projection operator, therefore we can just consider an arbitrary direction):

$$\frac{dE_{GW}}{dtdS} = \langle ct^{0r}\rangle = \frac{G}{8\pi c^5 r^2}\left\langle \sum_{jk}\left(\dddot{Q}_{jk}^{TT}(t-r/c)\right)^2\right\rangle. \tag{3.65}$$

To now find the **gravitational wave luminosity** L_{GW} we need to integrate over the complete solid angle as seen from the source. Using the properties of $\mathcal{P}_{jkmn}$ and the simple integrals $\frac{1}{4\pi}\int n_m n_r d\Omega = \frac{1}{3}\delta_{mr}$ and $\frac{1}{4\pi}\int n_m n_n n_r n_s d\Omega = \frac{1}{15}(\delta_{mn}\delta_{rs} + \delta_{mr}\delta_{ns} + \delta_{ms}\delta_{nr})$ this integral is equal to:

$$\begin{aligned} L_{GW} &= \frac{G}{8\pi c^5}\int\left\langle \sum_{j,k=1}^{3}\left(\mathcal{P}_{jkmn}\dddot{Q}_{mn}\right)^2\right\rangle d\Omega \\ &= \frac{G}{5c^5}\left\langle \sum_{j,k=1}^{3}\dddot{Q}_{kn}(t-r/c)\dddot{Q}_{kn}(t-r/c)\right\rangle. \end{aligned} \tag{3.66}$$

The exact details of this calculation can be found in [8]. Next we will apply the results above to binary systems.

3.5 Binary Systems (Part 1)

In this section we will use the results stated in the last section to estimate the gravitational signal emitted by binary systems of two stars (masses m_1 and m_2) moving on a circular orbit around their common center of mass. We consider both masses as point masses for simplicity.
Let a be the orbital separation, the total mass $M = m_1 + m_2$, the reduced mass $\mu = m_1 m_2/M$. With Kepler's law we determine that the binary orbital frequency is equal to $\omega_K = \sqrt{GM/a^3}$. Let the center of mass be the origin of a coordinate system in the orbital plane. We find that the coordinates (x_1, y_1) and (x_2, y_2) follow:

$$\begin{aligned} x_1 &= \frac{m_2}{M}a\cos(\omega_K t); \quad x_2 = -\frac{m_1}{M}a\cos(\omega_K t); \\ y_1 &= \frac{m_2}{M}a\sin(\omega_K t); \quad y_2 = -\frac{m_1}{M}a\sin(\omega_K t). \end{aligned} \tag{3.67}$$

The 00-component of the stress-energy tensor is given by:

$$T^{00} = c^2\sum_{n=1}^{2} m_n\delta(x-x_n)\delta(y-y_n)\delta(z). \tag{3.68}$$

This implies that the non-vanishing components of the quadrupole moment are:

$$\begin{aligned}
q_{xx} &= m_1 \int_V x^2 \delta(x-x_1)\delta(y-y_1)\delta(z) dxdydz \\
&\quad + m_2 \int_V x^2 \delta(x-x_2)\delta(y-y_2)\delta(z) dxdydz \\
&= m_1 x_1^2 + m_2 x_2^2 = \mu a^2 \cos^2(\omega_K t) = \frac{\mu}{2} a^2 \cos(2\omega_K t) + C \\
q_{yy} &= -\frac{\mu}{2} a^2 \cos(2\omega_K t) + D \\
q_{xy} &= \frac{\mu}{2} a^2 \sin(2\omega_K t) \\
q_k^k &= q_{xx} + q_{yy} = \text{constant},
\end{aligned} \tag{3.69}$$

Note that the time-varying part of q_{ij} and $Q_{ij} = \mathcal{P}_{ijkl} q^{kl}$ are equal here. The constants C and D arise from the rule of goniometry that $\cos^2 x = \frac{1}{2} + \frac{1}{2}\cos(2x)$.
Defining the matrix A_{ij} as:

$$A_{ij}(t) = \begin{pmatrix} \cos 2\omega_K t & \sin 2\omega_K t & 0 \\ \sin 2\omega_K t & -\cos 2\omega_K t & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{3.70}$$

we may write $q_{ij} = \frac{\mu}{2} a^2 A_{ij} + \text{constant}$. Projecting the wave along a generic direction $\mathbf{n}$ onto the TT-gauge we find:

$$h_{ij}^{TT} = -\frac{4\mu M G^2}{rac^4} \mathcal{P}_{ijkl} A_{kl}(t - r/c) \equiv -h_0 A_{ij}^{TT}(t - r/c). \tag{3.71}$$

Here, $h_0 = \dfrac{4\mu M G^2}{rac^4}$.

Example 3.5.1.

- If $\mathbf{n} = \mathbf{z}$, we see that P_{ij} =diag$(1,1,0)$, so $A_{ij}^{TT} = A_{ij}$ and:

$$\begin{aligned} h_{xx}^{TT} = -h_{yy}^{TT} &= -h_0 \cos\left(2\omega_K [t - z/c]\right); \\ h_{xy}^{TT} = h_{yx}^{TT} &= -h_0 \sin\left(2\omega_K [t - z/c]\right), \end{aligned} \tag{3.72}$$

  meaning that the wave is circularly polarized.

- If $\mathbf{n} = \mathbf{x}$, P_{ij} =diag$(0,1,1)$, so:

$$A_{ij}^{TT} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\frac{1}{2}\cos 2\omega_K t & 0 \\ 0 & 0 & \frac{1}{2}\cos 2\omega_K t \end{pmatrix} \Rightarrow h_{yy}^{TT} = \frac{1}{2} h_0 \cos\left(2\omega_K [t - z/c]\right), \tag{3.73}$$

  implying a linear polarization. ■

We now find that:

$$\begin{aligned} \sum_{k,n=1}^{3} \dddot{Q}_{kn} \dddot{Q}_{kn} &= 32 \mu^2 a^4 \omega_K^6 = 32 \mu^2 G^3 M^3 a^{-5} \\ \Rightarrow L_{GW} &= \frac{32 G^4 \mu^2 M^3}{5 c^5 a^5}. \end{aligned} \tag{3.74}$$

Making the *adiabatic approximation*, which assumes that orbital parameters do not change significantly over the time interval the average is taken over, the energy lost by GW emission is compensated by a change in orbital energy, $dE_{orb}/dt = L_{GW}$, where $E_{orb} = E_K + U$, where E_K is the kinetic energy, and U the potential energy. These can be found to be equal to:

$$E_K = \frac{1}{2}\frac{G\mu M}{a}; \qquad U = -\frac{G\mu M}{a}; \qquad E_{orb} = -\frac{1}{2}\frac{G\mu M}{a}. \tag{3.75}$$

Expressing the time derivative of E_{orb} in terms of the orbital period by expressing da/dt in terms of $d\omega_K/dt$ and using $\omega_K = 2\pi P^{-1}$ and by using $dE_{orb}/dt = -L_{GW}$, we find:

$$\frac{dP}{dt} = \frac{3}{2}\frac{P}{E_{orb}}L_{GW}. \tag{3.76}$$

Using this expression and assuming $a(t=0) = a^{in}$, we find that:

$$a(t) = a^{in}\left[1 - \frac{t}{t_{coal}}\right]^{1/4}; \qquad t_{coal} = \frac{5}{256}\frac{c^5}{G^3}\frac{(a^{in})^4}{\mu M^2}. \tag{3.77}$$

Here we see that when $t = t_{coal}$, $a(t_{coal}) = 0$. Note that the bodies are actually not point-like, so they start merging before $t = t_{coal}$ is reached in reality.

Example 3.5.2. The binary system *J06* introduced in the beginning of the last section has two stars, of masses $0.55M_\odot$ and $0.25M_\odot$ ($M_\odot$ is the solar mass), and an orbital period of $P = 765.4s$. This leads to $a = 1.16 \cdot 10^8 m$, $M = 0.8M_\odot$ and $\mu = 0.172M_\odot$. It is at a distance of $1000pc$, meaning that $h_0 = 3.34 \cdot 10^{-22} \ll 1$, so we can indeed use the linearised theory for this system and its analogues. It holds that $L_{GW} = 1.174 \cdot 10^{34}$erg/s, and as such $t_{coal} = 3.34 \cdot 10^{13}s$ or 1.1 Megayears. ■

In the adiabatic regime we find the change of ω_K in time, as:

$$\omega_K(t) = \sqrt{\frac{GM}{a(t)^3}} = \omega_K^{in}\left[1 - \frac{t}{t_{coal}}\right]^{-3/8}, \tag{3.78}$$

where $\omega_K^{in} = \omega_K(a^{in}) = \omega_K(0)$. The amplitude h_0 now also changes over time, by introducing $\nu_{GW}(t) = \omega_K(t)/\pi$ and $\mathcal{M} = \mu^{3/5}M^{2/5}$ (the **chirp mass**) we find that:

$$h_0(t) = \frac{2\pi^{2/3}G^{5/3}\mathcal{M}^{5/3}}{c^4 r}\nu_{GW}^{2/3}(t). \tag{3.79}$$

Since ω_K now changes in time, the phase appearing in A_{kl} in eq.(3.70) has to be substituted by an integrated phase:

$$\Phi(t) = \int_0^t 2\omega_K(t)dt = \int_0^t 2\pi\nu_{GW}(t)dt + \Phi(t=0). \tag{3.80}$$

We have that:

$$\nu_{in}t_{coal}^{3/8} = 5^{3/8}\frac{1}{8\pi}\left(\frac{c^3}{G\mathcal{M}}\right)^{5/8} \Rightarrow \nu_{GW}(t) = \frac{1}{8\pi}\left(\frac{c^3}{G\mathcal{M}}\right)^{5/8}\left[\frac{5}{t_{coal} - t}\right]^{3/8}, \tag{3.81}$$

which, when substituted in the previous expression, gives:

$$\Phi(t) = -2\left[\frac{c^3(t_{coal} - t)}{5G\mathcal{M}}\right]^{5/8} + \Phi(t=0). \tag{3.82}$$

Hence,

$$h_{ij}^{TT} = -\frac{4\pi^{2/3}G^{5/3}\mathcal{M}^{5/3}}{c^4 r}\nu_{GW}^{2/3}(t)\left[\mathcal{P}_{ijkl}A_{kl}(t-r/c)\right], \tag{3.83}$$

where

$$A_{ij}(t) = \begin{pmatrix} \cos\Phi(t) & \sin\Phi(t) & 0 \\ \sin\Phi(t) & -\cos\Phi(t) & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.84}$$

This is the polarization tensor we expect to find with LISA for binary white dwarfs.

3.6 Signal Processing

In the previous sections, we explored the details of the mathematics and physics of Gravitational Waves. The most important result is the waveform we derived in equation eq.(3.71). This equation can now be used to find the theoretical signal we would expect to see from a GW source. The aim of this chapter is to find the optimal signal-to-noise ratio (SNR) of such observations. This can be done with so-called "matched filtering".

To optimise the SNR, which determines the significance of our measurement, we need to apply the theory of signals and systems. To do this, we assume that the output signal, with noise $n(t)$, is given as $s(t) = h(t) + n(t)$, where $h(t) = D^{ij}h_{ij}$, for $i, j = 1, 2, 3$, where h_{ij} is the tensor describing the GW and D^{ij}, the **detector tensor**, is a quantity that determines the response of the detector. If, for example, it only reacts on the (x, x)-component of h_{ij}, we have $D^{11} = 1$ and the other components are zero.

When $h(t)$ would be of the order of $n(t)$ or bigger, we would not have any problems determining the signal. Typically however, one expects $|h(t)| \ll |n(t)|$, meaning we need more advanced methods to filter through the noise.

If we multiply $s(t) = h(t)+n(t)$ with the theoretical form for $h(t)$ and integrate over the observation time T and divide by it, we find that:

$$\frac{1}{T}\int_0^T s(t)h(t)dt = \frac{1}{T}\int_0^T h^2(t)dt + \frac{1}{T}\int_0^T n(t)h(t)dt. \tag{3.85}$$

If $h(t)$ and $n(t)$ are uncorrelated oscillating functions, we find that the first integral on the right hand side grows as T for large T and its value averaged over time is of order one in T, meaning we can approximate it as h_0^2, where h_0 is the characteristic amplitude of $h(t)$. Due to the product of $n(t)$ and $h(t)$ in the second integral, it will grow as $T^{1/2}$ for large T, so that we approximate it by $(\tau_0/T)^{1/2}n_0h_0$, where τ_0 is a typical characteristic time (i.e. the period of h_0) and n_0 the characteristic amplitude of $n(t)$. This means that we could detect a signal when $h_0 > (\tau_0/T)^{1/2}n_0$.

Example 3.6.1. For our example in the last chapter, $J06$, with a period of about 12 minutes, it holds that, if we observe for $T = 1yr$, $(\tau_0/T)^{1/2} \sim 0.0048$, so we can detect a signal that is about 3-4 orders of magnitude lower than the noise. ■

This idea can be made more precise by using a so-called **filter-function** $K(t)$, which we will choose to maximize the SNR when we know the form of $h(t)$. As this procedure is called *matched filtering*, we choose a filter that matches the theoretically expected signal. To do this we define:

$$\hat{s}(t) = \int_{-\infty}^{\infty} s(t)K(t)dt. \tag{3.86}$$

The SNR is now defined as $SNR := S/N$ where S is the expected value of $\hat{s}$ when a signal is present, and N the root mean squared value of $\hat{s}$ without a signal present. We recall the definition of the ensemble average of a random process:

Definition 3.6.1. Given an ensemble of N realizations $x^{(i)} : [0,T] \to [0,1]$ for $1 \leq i \leq N$ of a random process over a time interval of length T, the **ensemble average** at time t_1, $\langle x(t_1)\rangle : [0,T] \to [0,1]$ is defined as:

$$\langle x(t_1)\rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} x^{(i)}(t_1) \tag{3.87}$$

Remark.

- Note that the given limit converges because $x^{(i)}(t_1) \in [0,1]$ for every $t_1 \in [0,T]$, which means that taking a sum of N ones and dividing by N again gives an element of $[0,1]$. It can be shown that this is a Cauchy-row

- When defining $x_1 = x(t_1)$, we see that the probability density function $p_1(x_1,t_1)$ gives that $p_1(x_1,t_1)dx_1$ is the probability of finding x in the range $x_1, x_1 + dx_1$ at a time t_1. This means that we can write $\langle x(t_1)\rangle = \int\limits_{-\infty}^{\infty} x_1 p_1(x_1,t_1)dx_1$.

- We can now also define the ensemble average of a product as:

$$\langle x(t_1)x(t_2)\rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} x^{(i)}(t_1)x^{(i)}(t_2) = \int\limits_{\infty}^{\infty}\int\limits_{\infty}^{\infty} x_1x_2p_2(x_1,x_2,t_1,t_2)dx_1dx_2 \tag{3.88}$$

  Here p_2 is the second-order joint probability density function, such that $p_2(x_1,x_2,t_1,t_2)dx_1dx_2$ is the probability of finding x between x_1 and $x_1 + dx_1$ at t_1 as well as finding x between x_2 and $x_2 + dx_2$ at t_2.

- Physically, we can not take an ensemble average, so we define the ensembles by measuring the signal over a given time interval of length T and considering this one "realization" (using this procedure assumes the system to be ergodic[6]). Repeating this and separating the measured intervals such that the intervals can be assumed to be uncorrelated (also assuming the noise to be completely random and Gaussian, giving that $\langle n(t)\rangle = 0$ without loss of generality for the results about the signal), we can average over all these "independent" realisations. Mathematically, you could define this ensemble average taken over one signal by defining sample $x^{(i)}(t)$ as the signal limited to the interval $[iT + (i-1)S, (i+1)T + (i-1)S]$, where T is the length of a sample, and S is the length we take between samples, to allow for uncorrelated samples.
  We should now still note that we can not define infinitely many samples, for we observe for a finite amount of time, but for theoretic purposes, we assume we can select many samples, such that our approximations hold. ▲

[6] A system is ergodic when:

$$\langle x(t)\rangle = \lim_{T\to\infty} \frac{1}{T} \int\limits_{-T/2}^{T/2} x(t)dt.$$

For our periodic signal and the completely random noise we can assume ergodicity.

As stated before, we assume $\langle n(t)\rangle = 0$. Now we find that the ensemble average of our GW-signal $\langle h(t)\rangle$ equals $h(t)$ again, for it can be assumed to be of the same form during our observation.

We denote the Fourier transform of a function $F(t)$ with $\tilde{F}(f)$ such that:

$$\begin{aligned} F(t) &= \int_{-\infty}^{\infty} \tilde{F}(f)e^{-2\pi i f t}df; \\ \tilde{F}(f) &= \int_{-\infty}^{\infty} F(t)e^{2\pi i f t}dt. \end{aligned} \tag{3.89}$$

This way, we see that the convolution of two functions h and K is given as:

$$\begin{aligned} (h * K)(\tau) &= \int_{-\infty}^{\infty} h(t)K(t-\tau)dt = \int_{-\infty}^{\infty} h(t) \int_{-\infty}^{\infty} \tilde{K}(f)e^{-2\pi i f(t-\tau)}df\,dt \\ &= \int_{-\infty}^{\infty} \tilde{K}^*(f) \int_{-\infty}^{\infty} h(t)e^{2\pi i f t}dt e^{-2\pi i f \tau}df = \int_{-\infty}^{\infty} \tilde{K}^*(f)\tilde{h}(f)e^{-2\pi i f \tau}df. \end{aligned} \tag{3.90}$$

Here, the second line is found after changing the integral from $f \to -f$, such that $df \to -df$ and we still integrate from $-\infty$ to ∞. Also, it is implicitly assumed $K(t)$ is real, such that $\tilde{K}(-f) = \tilde{K}^*(f)$[7]. Using this result, we finally find that:

$$\begin{aligned} S = \langle \hat{s}\rangle &= \int_{-\infty}^{\infty} \langle s(t)\rangle K(t)dt = \int_{-\infty}^{\infty} h(t)K(t)dt \\ &= (h * K)(\tau = 0) = \int_{-\infty}^{\infty} \tilde{h}(f)\tilde{K}^*(f)df. \end{aligned} \tag{3.91}$$

Using the same method while assuming $n(t)$ to be real as well and defining the **noise spectral density**[8] $S_n(f)$ by $\langle \tilde{n}^*(f)\tilde{n}(f)\rangle = \frac{1}{2}\delta(f-f')S_n(f)$, we find that[9]:

$$N^2 = \int_{-\infty}^{\infty} \frac{1}{2}S_n(f)|\tilde{K}(f)|^2 df. \tag{3.92}$$

A proof of this fact is given in the appendix.
This has as a consequence that:

$$SNR := \frac{S}{N} = \frac{\int_{-\infty}^{\infty} \tilde{h}(f)\tilde{K}^*(f)df}{\left[\int_{-\infty}^{\infty} \frac{1}{2}S_n(f)|\tilde{K}(f)|^2 df\right]^{1/2}}. \tag{3.93}$$

[7]The asterisk as superscript denotes the complex conjugate of a function.

[8]The noise spectral density can be interpreted as a function that gives the amplitude of each component of the signal, as function of the frequency of that component.

[9]It should be noted that this definition is only rigorous when $n(t)$ has a well defined Fourier transform. We assume $\tilde{n}(f)$ exists in the following.

To now find the filter $K(t)$ maximizing S/N, we define the **scalar product between two real functions** $A(t)$ and $B(t)$ as:

$$(A|B) = \Re\left\{\int_{-\infty}^{\infty} \frac{\tilde{A}^*(f)\tilde{B}(f)}{\frac{1}{2}w(f)}df\right\} = 4\Re\left\{\int_{0}^{\infty} \frac{\tilde{A}^*(f)\tilde{B}(f)}{w(f)}\right\}. \tag{3.94}$$

The second equality holds because $A(t)$ and $B(t)$ are real. Also, $w(f)$ is a real-valued weighting function. In our case, this will be $S_n(f)$. This means that we can write for eq.(3.93):

$$\frac{S}{N} = \frac{(u|h)}{(u|u)^{1/2}}, \tag{3.95}$$

Where u is a function whose Fourier transform is:

$$\tilde{u}(f) = \frac{1}{2}S_n(f)\tilde{K}(f). \tag{3.96}$$

Since $S_n(f) > 0$, the scalar product is positive definite, so it can be seen as "inner product" on a "vector space" and we are thus searching for the "vector" $u/(u|u)^{1/2}$ such that its scalar product with h is maximum, meaning we want h and $u/(u|u)^{1/2}$ to be parallel, so:

$$\tilde{K}(f) = C\frac{\tilde{h}(f)}{S_n(f)}, \tag{3.97}$$

where C is a constant, which we can leave out of our calculations because rescaling $\hat{s}$ does not change the SNR. Using this filter, we get that:

$$\frac{S}{N} = (h|h)^{1/2} \Rightarrow \left(\frac{S}{N}\right)^2 = 4\int_0^\infty \frac{|\tilde{h}(f)|^2}{S_n(f)}df. \tag{3.98}$$

This is an arbitrary result, and will be applied to binary systems in the next section.

3.7 Binary Systems (Part 2)

In the last part, we derived the waveform for binary systems. We will now apply the signals and systems theory above to these waveforms we expect to find due to binary systems. To do this, we can define the **detector pattern functions**:

$$F_A(\hat{\mathbf{n}}) = D^{ij}e^A_{ij}(\hat{\mathbf{n}}), \tag{3.99}$$

which depend on the direction of propagation of the wave. If $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ are unit vectors orthogonal to each other and to $\hat{\mathbf{n}}$, we have that:

$$e^+_{ij}(\hat{\mathbf{n}}) = \hat{\mathbf{u}}_i\hat{\mathbf{u}}_j - \hat{\mathbf{v}}_i\hat{\mathbf{v}}_j; \quad e^\times_{ij}(\hat{\mathbf{n}}) = \hat{\mathbf{u}}_i\hat{\mathbf{v}}_j + \hat{\mathbf{v}}_i\hat{\mathbf{u}}_j. \tag{3.100}$$

This means that we can write our wave-signal as:

$$h(t) = h_+(t)F_+(\tau,\phi) + h_\times F_\times(\tau,\phi). \tag{3.101}$$

It should be noted that here we will have to use an orbital average if we observe for a substantial amount of time. This longer time of observation is needed to accurately determine the location of

the object on the sky, so we will assume observation times generally longer than one year (of course divided into smaller runs).
Now assuming a monochromatic wave with frequency f_0 and amplitudes $h_{0,+}$ and $h_{0,\times}$ of the respective polarizations:

$$\begin{aligned} h_+(t) &= h0,+\cos(2\pi f_0 t);\\ h_\times(t) &= h_{0,\times}\cos(2\pi f_0 t+\alpha);\\ h(t) &= F_+ h_{0,+}\cos(2\pi f_0 t) + F_\times h_{0,\times}\cos(2\pi f_0 t+\alpha)\\ &= \Re\{(F_+ h_{0,+} + F_\times h_{0,\times} e^{i\alpha})e^{2\pi i f_0 t}\}\\ &= |F_+ h_{0,+} + F_\times h_{0,\times} e^{i\alpha}|\cos(2\pi f_0 t). \end{aligned} \tag{3.102}$$

where α is a relative phase between the two polarizations. This results in the Fourier transform being equal to:

$$\begin{aligned} \tilde{h}(f) &= \int_{-\infty}^{\infty} h(t)e^{-2\pi i f t}dt = |F_+ h_{0,+} + F_\times h_{0,\times} e^{i\alpha}| \int_{-\infty}^{\infty} \cos(2\pi f_0 t)e^{-2\pi i f t}dt\\ &= |F_+ h_{0,+} + F_\times h_{0,\times} e^{i\alpha}| \int_{-\infty}^{\infty} \left(\frac{e^{2\pi f_0 t} + e^{-2\pi i f_0 t}}{2}\right) e^{-2\pi i f t}dt\\ &= |F_+ h_{0,+} + F_\times h_{0,\times} e^{i\alpha}| \frac{1}{2}\left(\int_{-\infty}^{\infty} e^{2\pi i(f_0 - f)t}dt + \int_{-\infty}^{\infty} e^{2\pi i(-f_0-f)t}dt\right)\\ &= |F_+ h_{0,+} + F_\times h_{0,\times} e^{i\alpha}| \frac{1}{2}(\delta(f-f_0)+\delta(f+f_0))\\ &= |F_+ h_{0,+} + F_\times h_{0,\times} e^{i\alpha}| \frac{1}{2}\delta(f-f_0). \end{aligned} \tag{3.103}$$

The last equality holds because we take $f > 0$, for we observe physical systems, so $\delta(f + f_0) = 0$ always. Note that we used the Fourier-definition of the δ-function:

$$\delta(f) = \int_{-\infty}^{\infty} e^{2\pi i f t}dt. \tag{3.104}$$

If we measure for a limited time T, we have to alter this definition to:

$$\delta(f) = \int_{-T/2}^{T/2} e^{2\pi i f t}dt \Rightarrow \delta(0) = T. \tag{3.105}$$

Using this result, we find the optimal SNR for a binary system to be:

$$\begin{aligned} \left(\frac{S}{N}\right)^2 &= 4\int_0^{\infty} \frac{|\tilde{h}(f)|^2}{S_n(f)}df = 4\int_0^{\infty} \frac{1}{4}|F_+ h_{0,+} + F_\times h_{0,\times} e^{i\alpha}|^2 \frac{\delta(f-f_0)^2}{S_n(f)}df\\ &= |F_+ h_{0,+} + F_\times h_{0,\times} e^{i\alpha}|^2 \int_0^{\infty} \frac{\delta(f-f_0)}{S_n(f)}\delta(f-f_0)df\\ &= |F_+ h_{0,+} + F_\times h_{0,\times} e^{i\alpha}|^2 \frac{\delta(0)}{S_n(f_0)} = |F_+ h_{0,+} + F_\times h_{0,\times} e^{i\alpha}|^2 \frac{T}{S_n(f_0)}. \end{aligned} \tag{3.106}$$

When we thus observe, for example a binary system along the optimal orientation, $F_+ = 1$ and $F_\times = 0$, and:

$$h_+ = \frac{4\mu M G^2}{rac^4} = \frac{4}{r}\left(\frac{G\mathcal{M}}{c^2}\right)^{5/3}\left(\frac{\pi f_0}{c}\right)^{2/3}. \tag{3.107}$$

That way, we see that:

$$\frac{S}{N} = \frac{\sqrt{T}}{\sqrt{S_n(f_0)}}\left(\frac{4}{r}\left(\frac{G\mathcal{M}}{c^2}\right)^{5/3}\left(\frac{\pi f_0}{c}\right)^{2/3}\right). \tag{3.108}$$

Thus, given an orientation and parameters, we can find the optimal SNR for a binary system.

Example 3.7.1. In the case of our example $J06$, the GW amplitude has been given in previous examples. When we perform an orbital average (over the orbit of the detector around the sun), we expect a SNR of 19.67. Without performing the orbital average, the SNR would be 33.48, but not performing the average is not realistic, for we have to observe for a longer time to establish the location of the source completely.
For this calculation, we have used $S_n(f)$ from [2], which is shown in 1. We integrated for an observation time of one year. The SNR increases for longer observation times, for 4 years it is 103.75 and for 10 years it is 176.69. ■

3.8 Summary

In this chapter we defined the theoretical background to describe Gravitational Waves as propagating perturbations of the metric of spacetime. To determine physical properties of these GWs the gauge-invariance of Einstein's equations is used. The quadrupole formalism for finding the wave-form emitted by binary systems is described, to be used in analysis of detected signals by GW observatories like LISA. The results are used for an analysis to determine the optimal SNR for binary systems when detecting them with LISA. In the next chapter, these results will be used for calculations.

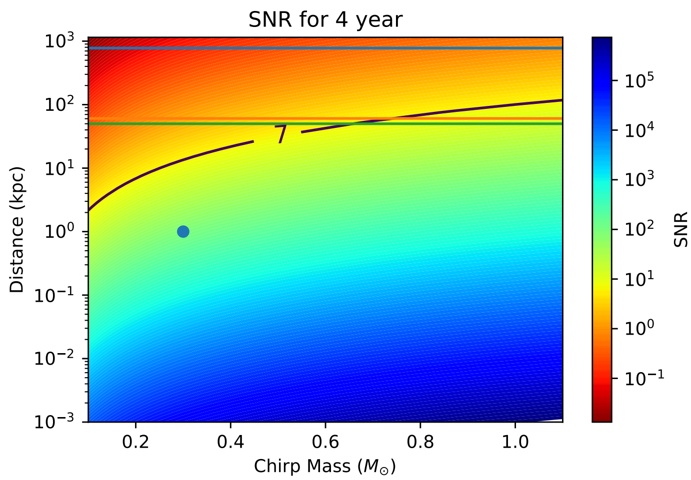
Chapter 4

Results

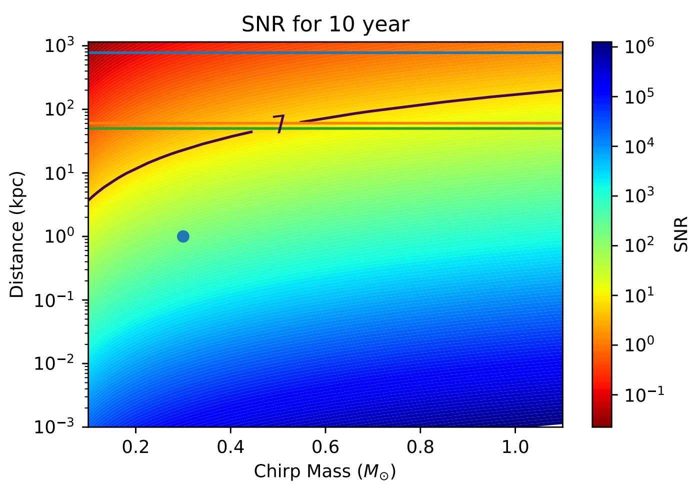
In the previous chapters we set up the theoretical tools to describe and detect GWs. We used the example of GWs emitted by binary systems as red line through the third chapter, because we will now limit our discussion to binary systems of White Dwarfs (DWDs). These systems are very common in the Milky Way, more then $25 \cdot 10^5$ are expected to be detected with LISA[14],[13]. We will first determine the maximal distance at which DWDs can be observed, given a certain parameter space. Using this, we can determine the properties of visible systems in nearby galaxies like the Small and Large Magellanic Clouds (SMC and LMC) and the Andromeda galaxy (M31). These are some of the nearest galaxies to the Milky Way.

4.1 Maximal Distance

To determine the maximal distance at which binary star systems of double white dwarfs (DWD) can be observed with LISA we observe the known object $J06$, as used in our previous examples. In [14] (Table B1) this source is the only DWD for which the SNR is higher than the nominal threshold of 7, which is taken as a limit for LISA. We investigate the ability to observe DWDs in the Small and Large Magellanic Clouds (SMC and LMC) or in M31. After varying the distance and chirp mass of $J06$-like systems, we found the SNR for these parameters for four and ten years of observation in figure 4.1. The chirp mass is constrained from 0.1 to $1.1M_\odot$, because all stars from the simulation have a chirp mass between these limits. We also constrain that both m_1 and m_2 can not be bigger than $1.4M_\odot$, the Chandresekhar Limit, which is the maximal allowed mass for a White Dwarf[4]. This implies that $m_1 + m_2$ lies between 0.01 and 2.8 solar masses.
The mass-ratio $q = m_2/m_1$, where m_2 is the smaller mass, is of impact on the chirp Mass. In the case of $J06$, $m_2 = 0.25M_\odot$ and $m_1 = 0.55M_\odot$, so $q = 0.45$ and $M_{chirp} \approx 0.3M_\odot$. It emits at a GW frequency that is equal to twice the inverse of the period, as seen in chapter 3 (eq.(3.70)). In the case of $J06$, it is 2.61mHz.
It can be seen in figures 4.1a and 4.1b that after four years of observation, DWDs with the same orbital period and galactic coordinates as $J06$ can not be observed at the distance of M31, but systems that are massive enough (chirp mass higher than about $0.65M_\odot$) can be found in the SMC and LMC. Even longer observations lower this limit to about $0.5M_\odot$.
When halving the orbital period, we have to assume the components of the binary become more massive, otherwise we can not use the point-mass approximation anymore. This means that smaller orbital periods induce heavier masses and thus stronger signals, thus we should be able to push the distance of visible binaries with those parameters further again.



(a) 4 years.



(b) 10 years.

Figure 4.1: The SNR for $J06$-like DWD after four and ten years of observation for different distances and chirp masses. The green horizontal line signifies the distance of the Large Magellanic Cloud (LMC), the red line that of the Small Magellanic Cloud (SMC) and the blue line that of M31. The blue dot shows the parameters of $J06$. All combinations of parameters below the curve labelled by "7" are observable with $SNR > 7$.

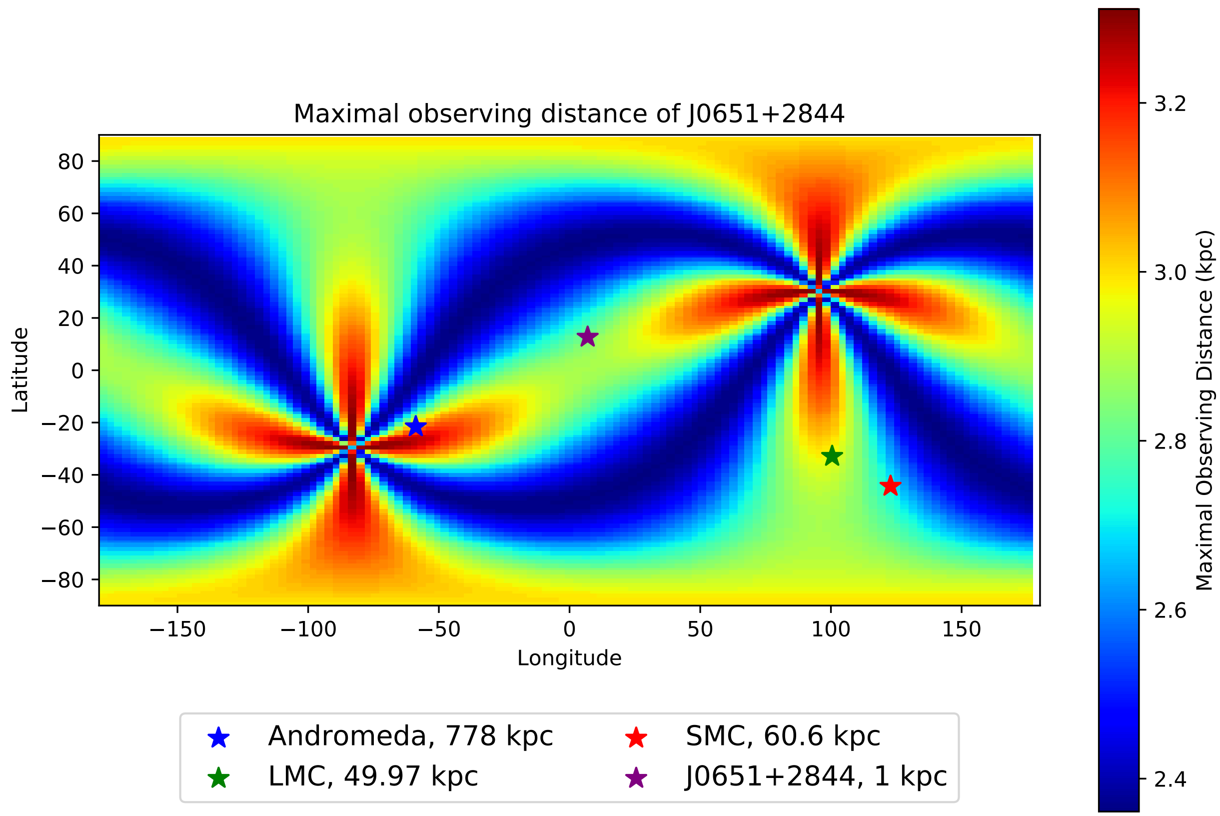


Figure 4.2: The maximal distance at which $J06$ could be observed for different galactic coordinates. The two singularities arise from the orientation of LISA w.r.t. the galactic coordinate system. Their locations are due to the tilt of the configuration w.r.t. the axis of the Earth. The four stars represent the galactic coordinates of SMC, LMC, Andromeda and the actual location of $J06$.

It should be noted that the exact limiting distance is also dependent on the galactic coordinates of the structure, because of the orientation of LISA during its orbit around the sun. This dependence is shown in figure 4.2.
In this figure we see that M31 is at a favourable position on the sky for observations with LISA, so we will still consider the Andromeda as structure with possibilities of observable DWDs.
There are many different DWDs that could be visible in the LMC and SMC, which is not surprising, for these are a lot less distant (a GW signal decays with $1/r$).
We have done this study for one point in the Galactic coordinate-system, but by looking at the colour scale in figure 4.2, we conclude these results will all be comparable for other coordinates, and the galaxies we studied are at coordinates where the visibility of GW signals by LISA is intermediate after one year of observing. The singularities in this figure are due to the configuration LISA has with respect to the Galactic coordinate system.
In conclusion, structures like SMC, LMC and Andromeda are the most likely places to find DWDs that are still observable with LISA, so we will further analyse the visibility of these DWDs in these structures in the next section. Of course intergalactic DWDs could also exist, so for these we now conclude that the maximum distance at which they are observable depends strongly on their galactic coordinates and properties, but the distance of the Andromeda is a viable estimate of a maximal distance for most DWDs.

4.2 Nearby Structures

Because we saw in the last section that the distance limit of detectable DWD targets does not reach far beyond the Andromeda galaxy, we limit ourselves to the nearby galaxies named before. We can determine the parameters of visible DWDs in these structures using a parameter space and analysing the SNR belonging to all combinations, just as we did when creating the plots in the last section. The parameter space we use is constrained by the chirp mass of the system being between $0.1M_\odot$ and $1.1M_\odot$, and the orbital separation being between $3R_*$ and $30R_*$, where R_* is the mean radius of a white dwarf, (meaning that the orbital period ranges between 81 and 8500 seconds), because when the orbital separation is smaller than three times the mean radius of a white dwarf (about $0.014R_\odot$), the approximation we made of point sources for the masses is not viable anymore. When the orbital separation becomes too big, the system will not be of interest, for the amplitude of GWs emitted is too small, or the system will not be gravitationally bound anymore.
The galactic coordinates of the named structures are used to find the results for an inclination of 45°[1], shown in figure 4.3. When the orbital separation is small, a whole range of chirp masses can produce visible DWD systems, because of the high GW frequencies. There is a decay, following a power law, in possible masses as the orbital separation grows, due to the dependence of h_0 on these properties. We are hopeful for observable binaries in the SMC and LMC because all the combinations of parameters above the curves are visible with $SNR > 7$. Also we here find that there is still hope for M31, because still a substantial part of the parameter space produces visible DWDs at that distance as seen.
These curve tend to heavily depend in the inclination. For example for inclinations nearing 90° these curves shift to the upper left.
Before actually finding results, we can produce an expectation of which types of DWDs we could find in these nearby structures, which will be done in the next section.

[1] We take this inclination because then both polarizations are present in the signal, and they both are at their mean values when changing the inclination i over the range of $\sin(i)$.

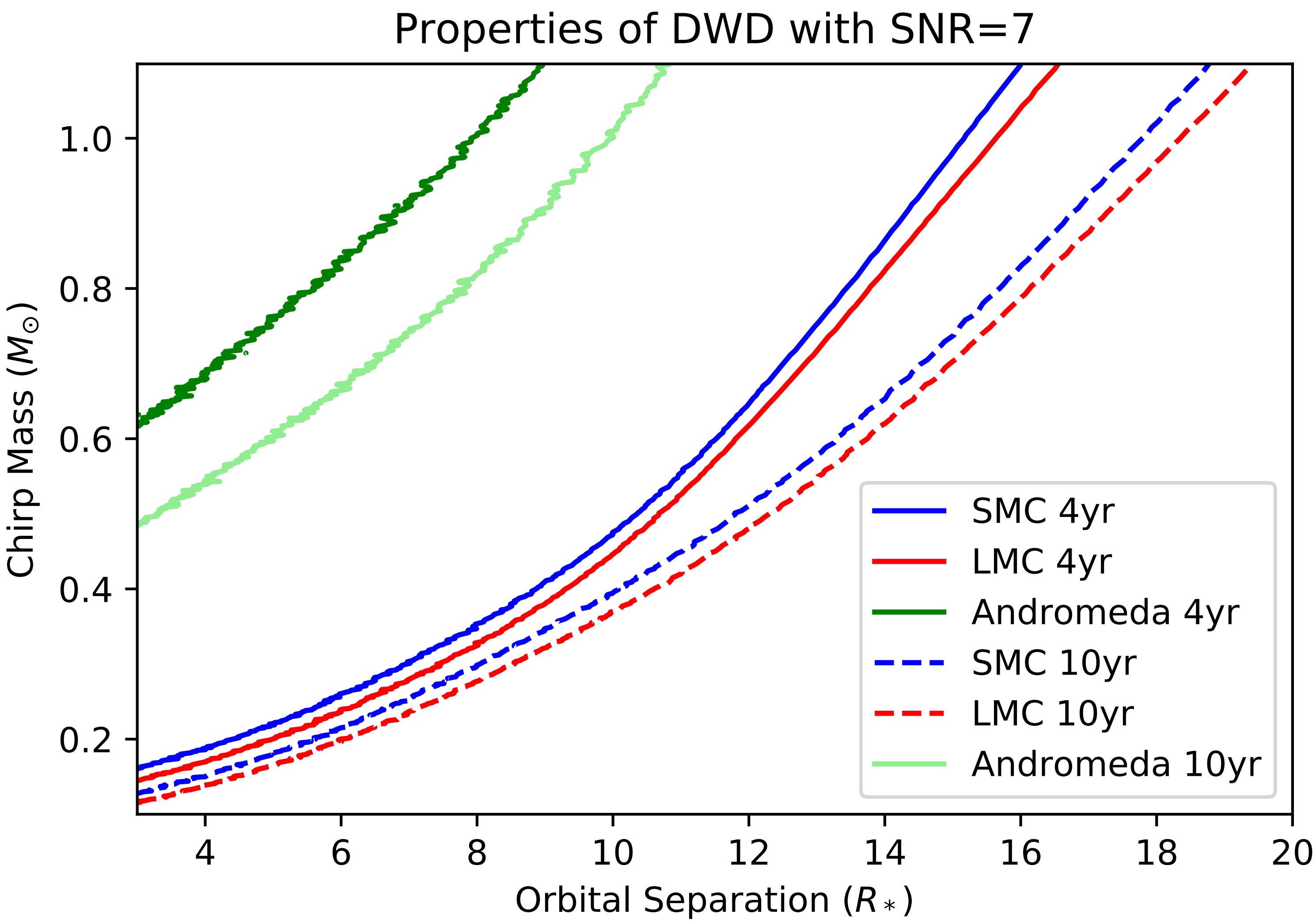


Figure 4.3: The contours where $SNR = 7$ for observing targets in the SMC, LMC and Andromeda for 4 and 10 years observation time. The regions to the upper left are the combinations of parameters for which the DWDs are observable. R_* is the mean radius of a White Dwarf, equal to about $0.014R_\odot$. An inclination of 45° is used.

Populations

By analysing a synthetic population of DWDs for nearby structures, we can produce an expectation of the types of visible DWDs, to limit our search to those types when actually observing. We can of course not assume these to be completely correct, for there are assumptions involved in creating these synthetic populations. They do however create a viable categorization of possibilities of visible DWDs. When actually observing corresponding populations, we could tell more about the actual population of nearby structures with statistical arguments. Knowing where DWDs are situated in these nearby structures will provide better information to be used in simulations to learn more about the formation history of the galaxies, and thus about the history of our Local Group.

Unfortunately, no theoretical populations of DWDs have been constructed for the SMC and LMC, so we will restrict our discussion to M31. We can now produce a contour plot as in fig.4.3, but for different thresholds for the SNR. Additionally, we make use of the properties we know of Andromeda. The mass of this galaxy is estimated to be approximately the same as that of the Milky Way, both around $0.8–1.5 \cdot 10^{12} M_{\odot}$. This leads to the assumption that the stellar formation histories of the two galaxies are roughly equal. This would mean that we expect comparable populations of DWDs in both galaxies. Based on the star formation history and observations of the Milky Way, a population of about $26.4 \cdot 10^6$ DWDs has been constructed in [14] and references therein. We will simulate these DWDs to be at the distance and galactic coordinates of Andromeda and determine the expected SNR for each target. This way we can determine the amount of stars visible above a certain threshold for the SNR. These results are found in figures 4.4 and 4.5. The contours are plotted for an inclination of $45°$.

It can be seen that there are more and more possible combinations of parameters visible while we lower our threshold, as would be expected. The levels are indicated at the top of the figure in the curve.

We can further characterize the visibility of stars in the synthetic population by looking at figures 4.4 and 4.5, where we also plotted the stars from the population in our parameter space constrained by the chirp mass being between $0.1 M_{\odot}$ and $1.1 M_{\odot}$ (in [14] it is found that all DWDs from the population lie within this range) and the orbital period between one and twenty minutes (shorter would mean an orbital separation that is too small and longer would not give a sufficient amplitude). The value of the SNR given to all stars is the most common SNR when assigning a sample from a random distribution of inclinations to each star.

We left out all members of the population for which $SNR < 1$, because these are not of any interest for LISA and only crowd the figure.

The amount of stars from the population visible above a certain threshold is displayed at the left of the figure. We should note here that those numbers do not exactly correspond to the actual data points, for we took away the randomness of the inclination of the DWDs and used the standard $45°$ for the model.

In conclusion, with a threshold of $SNR = 7$ we expect to find about 4 DWDs in M31 if we observe for 4 years, and 30 if we observe for 10 years. Actually observing these systems would imply that the Star Formation History of M31 and the Milky Way are comparable.

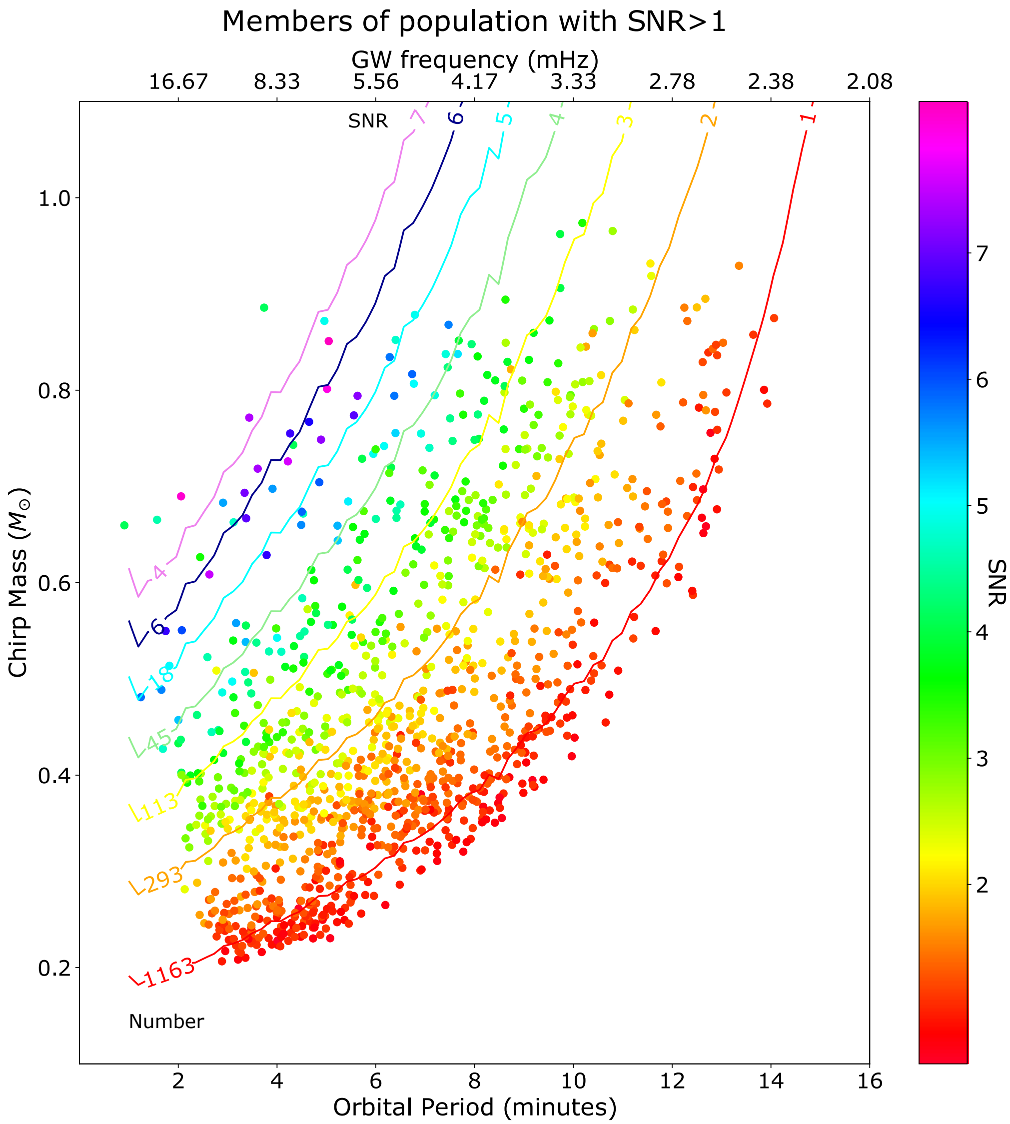


Figure 4.4: The DWDs from the generated population for which $SNR > 1$ when observed in M31 for 4 and 10 years. The ratio's were calculated by using finding the most common SNR per star for various inclinations. The contours are given by the theoretical division of the parameter space into areas where the $SNR = 1, 2, 3, 4, 5, 6, 7$ (as specified by the numbers at the upper side of the graph), for an inclination of 45°. Added to the contours on the left side are are the number of stars from the synthetic population visible with a SNR higher than the given value of that contour.

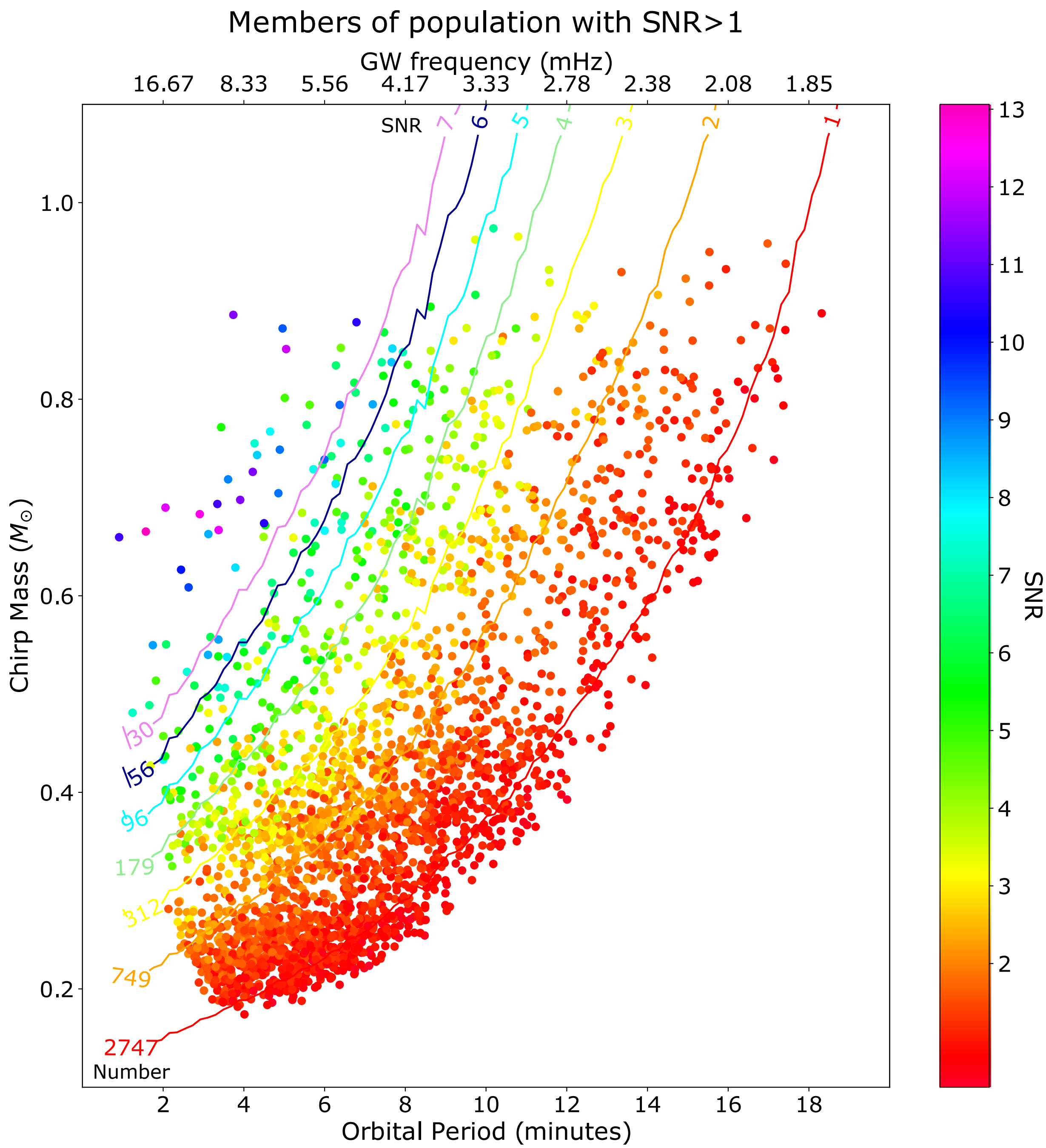


Figure 4.5: The same as in figure 4.4 but then for an observation time of 10 years.

4.3 Prospects

As illustrated in the previous sections, there is a substantial probability to observe extragalactic DWDs with LISA. This would be the first time these types of objects would be resolved as a single source, because they are not visible by only using optical instruments, for their resolution would not be high enough. Even DWDs in globular clusters in the Milky Way have never been observed before due to crowding.
Using LISA to also search for extragalactic DWDs (within a radius of about 10^3kpc) thus will increase our knowledge of the specific population of DWDs in our Local Group and that in nearby galaxies. When we gain information about the distribution of these systems, we can gain information about the history of our Local Group and the structures within, because it would give specific constraints when we compare with cosmological simulations to validate cosmological models. Of course, this would all involve a lot of statistical analysis because of the low number of predicted detections, but knowing part of the outcome simplifies this computationally. Using DWDs for this seems less than ideal because of their weak GW-signal, but they are really numerous as described before, and we should be able to observe many Extragalactic DWDs within the range of satellites of the Milky Way. Mapping the DWDs between galaxies would give a better insight in for example the past collisions between structures in our neighbourhood.

If the SNR for certain measurements is high enough, observing the GW-signals will also increase the accuracy of distance measurements within the Local Group, because of the simple dependence of the amplitude on distance.

By constructing a synthetic population of DWDs expected for the SMC and LMC, a comparable analysis and categorization of visible objects can be produced, enabling us to actively search for the specific associated parameters. Because we expect more types of DWDs to be visible due to the distance, we can even better analyse the history of those galaxies using the distribution of DWDs than we could for M31.
Another possible angle is to reproduce the analysis done in [13] for Andromeda, the SMC and the LMC.

In the near future after the launch of LISA, not many future space-based GW-observatories are planned, and ground-based observatories can presently not reach the range of frequencies needed to observe GWs from DWDs, so LISA is the only option available in short time to perform measurements of Extragalactic DWDs in this way.

4.4 Summary

We presented the maximal distance at which DWDs like $J06$ can be observed, dependent on the calculations derived in the previous chapters. We apply the same theory to a synthetic population of DWDs of the Milky Way, transposed in simulation to be at the location of the Andromeda galaxy. We characterized the properties of the systems we expect to see with a certain threshold in SNR.

Conclusion

By setting up the mathematical environment needed to describe Gravitational Waves from scratch, we determined the optimal signal-to-noise ratio for a theoretical waveform signal received by LISA from a Double White Dwarf system. Using this, the maximal distance for observing DWDs with LISA seemed to heavily depend on the characteristics of the system, including the masses, inclination, galactic coordinates and orbital period.
When trying to observe DWDs in the Andromeda galaxy, we expect to find (with a signal-to-noise ratio higher than 7) 4 with 4 years of observation time, and 30 if we can observe for 10 years.
If we indeed observe these, we can use them to better analyse the history and structure of Andromeda.
Future research may involve characterizing and predicting populations of DWDs for other galaxies or intergalactic DWDs in the Local Group.

Acknowledgements

I would like to thank Dr. Elena Rossi and Valeriya Korol for giving me the opportunity to work on this topic with them. They also supplied the data I used for the analysis. I would also like to thank Dr. Roland van der Veen for helping me find this topic and helping me get through the abstract theory. I also want to thank Prof. Dr. Harold Linnartz for initializing my contact with Elena and having the patience to give me the opportunity to combine both my studies into one thesis. I want to thank my father for correcting and proof-reading my thesis multiple times. Also thanks to all people who listened to me while practising for the multitude of presentations I gave during these months.

Bibliography

[1] Abbott, B. P. e. a. Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters 116*, 6 (Feb. 2016), 061102.

[2] Amaro-Seoane, P. e. a. Laser Interferometer Space Antenna. *ArXiv e-prints* (Feb. 2017).

[3] Braun, M. *Differential Equations and Their Applications: An Introduction to Applied Mathematics (Texts in Applied Mathematics) (v. 11).* Springer, 1992.

[4] Chandrasekhar, S. The Maximum Mass of Ideal White Dwarfs. *74* (July 1931), 81.

[5] Chappell, J. M., Hartnett, J. G., Iannella, N., Iqbal, A., and Abbott, D. Exploring the origin of Minkowski spacetime. *ArXiv e-prints* (Jan. 2015).

[6] Collaboration, L. S. Detections. `https://www.ligo.org/detections.php`, 20-6-2018.

[7] Farmer, A. J., and Phinney, E. S. The gravitational wave background from cosmological compact binaries. *Monthly Notices of the Royal Astronomical Society 346*, 4 (dec 2003), 1197–1214.

[8] Ferrari, V., and Gualtieri, L. Lecture notes on general relativity, black holes and gravitational waves, 2017-2018 (H1-13), 2014-2015 (H14-18).

[9] Frankel, T. *The Geometry of Physics: An Introduction.* Cambridge University Press, 1999.

[10] Franklin, J. Stress tensors, particles and fields. `http://www.reed.edu/physics/courses/Physics411/html/page2/files/Lecture.19.pdf.`, 21-6-2018.

[11] Gourgoulhon, E. Sagemanifolds. `http://sagemanifolds.obspm.fr/index.html`, 2018-04-2018.

[12] Gourgoulhon, E. *3+1 formalism in general relativity : bases of numerical relativity.* Springer, Berlin New York, 2012.

[13] Korol, V., Rossi, E. M., and Barausse, E. A multi-messenger study of the Milky Way's stellar disc and bulge with LISA, Gaia and LSST. *ArXiv e-prints* (June 2018).

[14] Korol, V., Rossi, E. M., Groot, P. J., Nelemans, G., Toonen, S., and Brown, A. G. A. Prospects for detection of detached double white dwarf binaries with Gaia, LSST and LISA. *Monthly Notices of the Royal Astronomical Society 470*, 2 (may 2017), 1894–1910.

[15] Kosenko, D. I., and Postnov, K. A. On the gravitational wave noise from unresolved extragalactic binaries. *336* (Aug. 1998), 786–790.

[16] Lee, J. M. *Riemannian Manifolds: An Introduction to Curvature (Graduate Texts in Mathematics)*. Springer, 1997.

[17] Lübke, M. Introduction to manifolds. Downloaded from the Blackboard page and uploaded to `home.strw.leidenuniv.nl/~koop/Introtomanifolds(2).pdf`, 2017.

[18] Michael Stoll, R. v. L. Linear algebra ii. `http://www.math.leidenuniv.nl/~desmit/edu/la2_2012/`, 2007, 2016.

[19] Ruiter, A. J., Belczynski, K., Benacquista, M., Larson, S. L., and Williams, G. The LISA Gravitational Wave Foreground: A Study of Double White Dwarfs. *The Astrophysical Journal 717*, 2 (jun 2010), 1006–1021.

[20] Ryder, L. *Introduction to General Relativity.* Cambridge University Press, 2009.

[21] Tao, T. Ricci Flow. `https://terrytao.files.wordpress.com/2008/03/ricci.pdf`, 2004.

[22] The Sage Developers. Sagemath, the Sage Mathematics Software System (Version 8.2). `http://www.sagemath.org/`, 2018-04-18, 2018.

[23] Thorpe, J. *Elementary topics in differential geometry.* Springer-Verlag, New York, 1979.

[24] van Holten, J. Relativistic fluid dynamics. `https://www.nikhef.nl/~t32/relhyd.pdf`, 2006.

[25] Weinberg, S. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity.* John Wiley & Sons, Inc., 1972.

Appendix A

Proofs and Additional Derivations

In this appendix, the extensive proofs or derivations left out of the main text are given.

A.1 Chapter 1

A.1.1 Proof of Theorem 1.3.2

We need to prove uniqueness and existence:

- Uniqueness: Due to the assumption of metric compatibility, we have that the following equations hold:

$$\begin{aligned} 0 = \nabla_\rho g_{\mu\nu} = \partial_\rho g_{\mu\nu} - \Gamma^\lambda_{\rho\mu} g_{\lambda\nu} - \Gamma^\lambda_{\rho\nu} g_{\mu\lambda}; \\ 0 = \nabla_\mu g_{\nu\rho} = \partial_\mu g_{\nu\rho} - \Gamma^\lambda_{\mu\nu} g_{\lambda\rho} - \Gamma^\lambda_{\mu\rho} g_{\nu\lambda}; \\ 0 = \nabla_\nu g_{\rho\mu} = \partial_\nu g_{\rho\mu} - \Gamma^\lambda_{\nu\rho} g_{\lambda\mu} - \Gamma^\lambda_{\nu\mu} g_{\rho\lambda}. \end{aligned} \tag{A.1}$$

  Subtracting the second and third from the first and using the symmetry of the connection we see:

$$\begin{aligned} \partial_\rho g_{\mu\nu} - \partial_\mu g_{\nu\rho} - \partial_\nu g_{\rho\mu} + 2\Gamma^\lambda_{\mu\nu} g_{\lambda\rho} = 0 \\ \Rightarrow \Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}(g_{\nu\rho,\mu} + g_{\rho\mu,\nu} - g_{\mu\nu,\rho}). \end{aligned} \tag{A.2}$$

  The last implication follows after multiplication by $g^{\lambda\rho}$. This means that a symmetric and metric-compatible connection is always expressed as given in terms of the metric, so it is unique.

- Existence: The connection coefficients given above in eq.(1.40) are indeed coefficients of a well-defined connection on M, which we show by looking at how the right hand side transforms under a coordinate change. This is of importance, because connection coefficients have to transform in a specific way for the covariant derivative to be a tensor again, as stated above.
  First we should know the way the metric tensor transforms and how then $g_{ij,k}$ transforms. If the new coordinates are denoted by Latin indices and the old by Greek:

$$\begin{aligned} g^{ab} &= \frac{\partial x^a}{\partial x^\alpha}\frac{\partial x^b}{\partial x^\beta} g^{\alpha\beta}; \\ \partial_a g_{bc} &= \partial_a \left(\frac{\partial x^\beta}{\partial x^b}\frac{\partial x^\gamma}{\partial x^c} g_{\beta\gamma} \right) \\ &= \frac{\partial x^\beta}{\partial x^b}\frac{\partial x^\gamma}{\partial x^c}\frac{\partial x^\alpha}{\partial x^a}\partial_\alpha g_{\beta\gamma} + g_{\beta\gamma}\left(\frac{\partial^2 x^\beta}{\partial x^b \partial x^a}\frac{\partial x^\gamma}{\partial x^c} + \frac{\partial^2 x^\gamma}{\partial x^c \partial x^a}\frac{\partial x^\beta}{\partial x^b} \right) \end{aligned} \tag{A.3}$$

This means we can write:

$$
\begin{aligned}
\Gamma^n_{ml} &= \frac{1}{2} g^{an}(g_{am,l} + g_{al,m} - g_{ml,a}) \\
&= \frac{1}{2}\frac{\partial x^a}{\partial x^\alpha}\frac{\partial x^n}{\partial x^\nu} g^{\alpha\nu}\left[\frac{\partial x^\alpha}{\partial x^a}\frac{\partial x^\mu}{\partial x^m}\frac{\partial x^\lambda}{\partial x^l}(g_{\alpha\mu,\lambda} + g_{\alpha\lambda,\mu} - g_{\mu\lambda,\alpha}) + 2g_{\alpha\mu}\frac{\partial x^\alpha}{\partial x^a}\frac{\partial^2 x^\mu}{\partial x^l \partial x^m}\right] \\
&= \frac{\partial x^n}{\partial x^\nu}\frac{\partial x^\mu}{\partial x^m}\frac{\partial x^\lambda}{\partial x^l}\Gamma^\nu_{\mu\lambda} + \frac{\partial x^a}{\partial x^\alpha}\frac{\partial x^n}{\partial x^\nu} g^{\alpha\nu} g_{\alpha\mu}\frac{\partial x^\alpha}{\partial x^a}\frac{\partial^2 x^\mu}{\partial x^l \partial x^m} \\
&= \frac{\partial x^n}{\partial x^\nu}\frac{\partial x^\mu}{\partial x^m}\frac{\partial x^\lambda}{\partial x^l}\Gamma^\nu_{\mu\lambda} + \frac{\partial x^n}{\partial x^\mu}\frac{\partial^2 x^\mu}{\partial x^l x^m} = \frac{\partial x^n}{\partial x^\nu}\frac{\partial x^\mu}{\partial x^m}\frac{\partial x^\lambda}{\partial x^l}\Gamma^\nu_{\mu\lambda} - \frac{\partial x^\mu}{\partial x^m}\frac{\partial x^\lambda}{\partial x^l}\frac{\partial^2 x^n}{\partial x^\mu \partial x^\lambda}.
\end{aligned}
\tag{A.4}
$$

The second equality follows from the fact that the metric tensor is symmetric and by relabelling indices:

$$
\begin{aligned}
& g_{\alpha\mu}\left(\frac{\partial^2 x^\alpha}{\partial x^n \partial x^a}\frac{\partial x^\mu}{\partial x^m} + \frac{\partial^2 x^\mu}{\partial x^m \partial x^n}\frac{\partial x^\alpha}{\partial x^a}\right) + g_{\alpha\nu}\left(\frac{\partial^2 x^\alpha}{\partial x^m \partial x^a}\frac{\partial x^\nu}{\partial x^n} + \frac{\partial^2 x^\nu}{\partial x^n \partial x^m}\frac{\partial x^\alpha}{\partial x^a}\right) \\
&\qquad - g_{\nu\mu}\left(\frac{\partial^2 x^\mu}{\partial x^n \partial x^a}\frac{\partial x^\nu}{\partial x^n} + \frac{\partial^2 x^\nu}{\partial x^n \partial x^a}\frac{\partial x^\mu}{\partial x^m}\right) \\
&= g_{\alpha\nu}\left(\frac{\partial^2 x^\alpha}{\partial x^n \partial x^a}\frac{\partial x^\nu}{\partial x^m} + \frac{\partial^2 x^\nu}{\partial x^m \partial x^n}\frac{\partial x^\alpha}{\partial x^a}\right) + g_{\alpha\nu}\left(\frac{\partial^2 x^\alpha}{\partial x^m \partial x^a}\frac{\partial x^\nu}{\partial x^n} + \frac{\partial^2 x^\nu}{\partial x^n \partial x^m}\frac{\partial x^\alpha}{\partial x^a}\right) \\
&\qquad - g_{\nu\alpha}\left(\frac{\partial^2 x^\alpha}{\partial x^n \partial x^a}\frac{\partial x^\nu}{\partial x^n} + \frac{\partial^2 x^\nu}{\partial x^n \partial x^a}\frac{\partial x^\alpha}{\partial x^m}\right) \\
&= 2g_{\alpha\nu}\frac{\partial^2 x^\nu}{\partial x^m \partial x^n}\frac{\partial x^\alpha}{\partial x^a}.
\end{aligned}
\tag{A.5}
$$

The last equality follows from the fact that:

$$
\begin{aligned}
\frac{\partial x^\mu}{\partial x^m}\frac{\partial x^\lambda}{\partial x^l}\frac{\partial^2 x^n}{\partial x^\mu \partial x^\lambda} &= \frac{\partial x^\mu}{\partial x^m}\left(\frac{\partial x^\lambda}{\partial x^l}\frac{\partial}{\partial x^\lambda}\right)\frac{\partial x^n}{\partial x^\mu} = \frac{\partial x^\mu}{\partial x^m}\left(\frac{\partial}{\partial x^l}\frac{\partial x^n}{\partial x^\mu}\right) \\
&= \frac{\partial}{\partial x^l}\left(\frac{\partial x^\mu}{\partial x^m}\frac{\partial x^n}{\partial x^\mu}\right) - \frac{\partial x^n}{\partial x^\mu}\left(\frac{\partial}{\partial x^l}\frac{\partial x^\mu}{\partial x^m}\right) \\
&= \frac{\partial}{\partial x^l}\left(\frac{\partial x^n}{\partial x^m}\right) - \frac{\partial x^n}{\partial x^\mu}\left(\frac{\partial}{\partial x^l}\frac{\partial x^\mu}{\partial x^m}\right) \\
&= \frac{\partial}{\partial x^l}\left(\delta^n_m\right) - \frac{\partial x^n}{\partial x^\mu}\left(\frac{\partial}{\partial x^l}\frac{\partial x^\mu}{\partial x^m}\right) = -\frac{\partial x^n}{\partial x^\mu}\frac{\partial^2 x^\mu}{\partial x^l \partial x^m}.
\end{aligned}
\tag{A.6}
$$

This proves Theorem 1.3.2. The other properties a connection should have follow from the fact that we defined the coefficients of the connection independently. □

A.2 Chapter 2

A.2.1 Deriving Einsteins Equations using the Variational Principle

Firstly, we introduce the **arclength** between points x and $x+dx$ as integral of the line element defined in chapter 1: $s = \int\limits_{x}^{x+dx} \sqrt{g_{\mu\nu}(x)dx^\mu dx^\nu} =: \int\limits_{t_1}^{t_2} \sqrt{L}dt$ where we integrate along a curve γ where $\gamma(t_1) = x$

and $\gamma(t_2) = x + dx$. This L is also called the **Lagrangian**.
Supposing the curve is varied[1], but with fixed end-points, a **geodesic** is now the path of extremal length, so the variation would be zero. This leads to the Euler-Lagrange equation for a geodesic (if $dx^\mu/dt =: \dot{x}^\mu$):

$$\frac{\partial L}{\partial x^\mu} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^\mu}. \tag{A.7}$$

In [20] chapter 4.1.2, it is derived that, because $L = g_{\kappa\lambda}(x)\dot{x}^\kappa\dot{x}^\lambda$, we can express the Euler-Lagrange equation as $\ddot{x}^\rho + \Gamma^\rho_{\kappa\lambda}\dot{x}^\kappa\dot{x}^\lambda = 0$.
This is the same equation as that of an autoparallel curve from chapter 1. Note that this only holds in manifolds with a metric.

We will now derive Einstein equations by the variational principle, which depend on the Lagrangian principle as well. Hilbert figured that the now called **Einstein-Hilbert action** $S^{EH} = \int \mathcal{L} d^4x$ of the gravitational field defined by the metric tensor should be the integral over spacetime of a Lagrange density, which is a tensor density, written as $\sqrt{-g}$ times a scalar. This g is the determinant of the metric. The scalar we know that is connected to the metric and most easily calculated is the Ricci scalar, so Hilbert then proposed the Lagrange density equal to:

$$\mathcal{L}_H = \frac{c^3}{16\pi G}\sqrt{-g}R, \tag{A.8}$$

such that the Lagrangian $L = \int \mathcal{L}_H d^3x$ and the action is $S^{EH} = \int L dt$. The constants were a result of correcting for dimensionality at the end of the run described below.
In vacuum, the gravitational field is the only object with an energy, so using the fact that $\delta S^{EH} = 0$ should hold for an extremal configuration, we find that:

$$\delta S^{EH} = \frac{c^3}{16\pi G}\int d^4x[\delta(\sqrt{-g})R + \sqrt{-g}\delta R] = 0. \tag{A.9}$$

We now need to prove some Lemma's before we can further investigate this expression:

Lemma A.2.1. *For the variation of the square root of $-g$ where g is the determinant of the metric tensor, it holds that:*

$$\delta(\sqrt{-g}) = -\frac{1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu} \tag{A.10}$$

Proof: Note that $\delta g = \dfrac{\partial g}{\partial g_{\mu\nu}}\delta g_{\mu\nu}$ (when we use the chain rule, which also holds for partial derivatives of tensors), and that the definition of the determinant is $g = \sum_\nu g_{\mu\nu}M_{\mu\nu}(-1)^{\mu+\nu}$, where $M_{\mu\nu}$ is the suitable determinant of a minor matrix, so:

$$\frac{\partial g}{\partial g_{\mu\nu}} = (-1)^{\mu+\nu}M_{\mu\nu} = gg^{\mu\nu}. \tag{A.11}$$

Since

$$\begin{aligned}\delta(g_{\mu\nu}g^{\nu\sigma}) &= \delta\delta^\sigma_\mu = 0\\ &= \delta g_{\mu\nu}g^{\nu\sigma} + g_{\mu\nu}\delta g^{\nu\sigma}\end{aligned} \tag{A.12}$$

[1]A **variation** of a curve γ is a smooth map $\psi : [-\epsilon, \epsilon] \times (-\delta, \delta) \to M$ $(\delta > 0)$ such that $\psi(t, 0) = \gamma(t)$ for all $t \in [a, b]$. It has **fixed endpoints** if $\psi(-\epsilon, s) = \gamma(-\epsilon)$ and $\psi(\epsilon, s) = \gamma(\epsilon)$.

it holds that, when multiplying by $g^{\mu\rho}$:

$$\delta g^{\rho\sigma} = -g^{\mu\rho} g^{\nu\sigma} \delta g_{\mu\nu}. \tag{A.13}$$

This then implies $\delta g = -g g_{\mu\nu} \delta g^{\mu\nu}$, so indeed:

$$\delta(\sqrt{-g}) = -\frac{1}{2}\sqrt{-g} g_{\mu\nu} \delta g^{\mu\nu}. \tag{A.14}$$

□

Lemma A.2.2. *(The Palatini Identity): It holds that:*

$$\delta R_{\mu\nu} = \nabla_\lambda(\delta\Gamma^\lambda_{\mu\nu}) - \nabla_\nu(\Gamma^\lambda_{\mu\lambda}). \tag{A.15}$$

Proof: Note that:

$$R_{\mu\nu} = \partial_\lambda \Gamma^\lambda_{\mu\nu} - \partial_\nu \Gamma^\lambda_{\mu\lambda} + \Gamma^\alpha_{\mu\nu}\Gamma^\lambda_{\alpha\lambda} - \Gamma^\lambda_{\alpha\nu}\Gamma^\alpha_{\mu\lambda}. \tag{A.16}$$

This implies that:

$$\begin{aligned} \delta R_{\mu\nu} = {} & \delta\partial_\lambda \Gamma^\lambda_{\mu\nu} - \delta\partial_\nu \Gamma^\lambda_{\mu\lambda} + \delta\Gamma^\alpha_{\mu\nu}\Gamma^\lambda_{\alpha\lambda} + \Gamma^\alpha_{\mu\nu}\delta\Gamma^\lambda_{\alpha\lambda} \\ & - \delta\Gamma^\lambda_{\alpha\nu}\Gamma^\alpha_{\mu\lambda} - \Gamma^\lambda_{\alpha\nu}\delta\Gamma^\alpha_{\mu\lambda}. \end{aligned} \tag{A.17}$$

We only need to prove that the right side of the identity is a tensor. This is seen by the fact that:

$$\delta\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}[\delta g_{\mu\rho;\nu} + \delta g_{\nu\rho;\mu} - \delta g_{\mu\nu;\rho}]. \tag{A.18}$$

This is a tensor, so its covariant derivative is as well.
The fact above is derived from:

$$\begin{aligned} \delta\left[g^{\lambda\delta} G_{\mu\nu\delta}\right] &= -g^{\rho\lambda} g^{\sigma\delta} \delta g_{\rho\sigma} \Gamma_{\mu\nu\delta} + g^{\lambda\rho} \delta\Gamma_{\mu\nu\rho} \\ &= \frac{1}{2} g^{\lambda\rho}\left[\delta g_{\mu\rho,\nu} + \delta g_{\nu\rho,\mu} - \delta g_{\mu\nu,\rho} - 2\Gamma^\sigma_{\mu\nu}\delta g_{\rho\sigma}\right]. \end{aligned} \tag{A.19}$$

Now we can just expand the right side to show:

$$\begin{aligned} \nabla_\lambda(\delta\Gamma^\lambda_{\mu\nu}) - \nabla_\nu(\Gamma^\lambda_{\mu\lambda}) = {} & \delta\partial_\lambda \Gamma^\lambda_{\mu\nu} - \delta\partial_\nu \Gamma^\lambda_{\mu\lambda} + \delta\Gamma^\alpha_{\mu\nu}\Gamma^\lambda_{\alpha\lambda} + \Gamma^\alpha_{\mu\nu}\delta\Gamma^\lambda_{\alpha\lambda} \\ & - \delta\Gamma^\lambda_{\alpha\nu}\Gamma^\alpha_{\mu\lambda} - \Gamma^\lambda_{\alpha\nu}\delta\Gamma^\alpha_{\mu\lambda} = \delta R_{\mu\nu}. \end{aligned} \tag{A.20}$$

□

Corollary. *Because of the Palatini Identity it holds that:*

$$\delta R = \delta g^{\mu\nu} R_{\mu\nu} + g^{\mu\nu}\delta R_{\mu\nu} = \delta g^{\mu\nu} R_{\mu\nu} + \nabla_\mu\left(g^{\mu\nu}\delta\Gamma^\alpha_{\mu\nu} - g^{\mu\alpha}\delta\Gamma^\lambda_{\mu\lambda}\right). \tag{A.21}$$

Because the second term in this Corollary vanishes if you integrate it over the volume due to Gauss' Theorem, we find that:

$$\delta S^{EH} = \frac{c^3}{16\pi G}\int d^4x\sqrt{-g}[R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R]\delta g^{\mu\nu} = 0. \tag{A.22}$$

This then implies that:

$$G_{\mu\nu} = 0. \tag{A.23}$$

This is Einstein's equation in vacuum.

If there is matter present, we need also to study δS^M, the variation in the action of the matter. This action is characterized by the well-known Euler-Lagrange equations, that generalize to curved spacetime as mentioned before. We also use the result from Lemma A.2.1. We find:

$$\delta S^M = \int d^4x \sqrt{-g} \left[\frac{\partial \mathcal{L}^M}{\partial g^{\mu\nu}} - \frac{1}{2} \mathcal{L}^M g_{\mu\nu} \right] \delta g^{\mu\nu}. \tag{A.24}$$

Here $\mathcal{L}^M$ is the Lagrangian density of the matter fields. When defining the stress-energy tensor as above:

$$T_{\mu\nu} = -2c \left[\frac{\partial \mathcal{L}^M}{\partial g^{\mu\nu}} - \frac{1}{2} \mathcal{L}^M g_{\mu\nu} \right]. \tag{A.25}$$

We find by stating $\delta S = \delta S^{EH} + \delta S^M = 0$ that:

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}. \tag{A.26}$$

These are Einstein equations again. The dimensionality of the equations is correct due to the way Hilbert defined his Lagrangian density and how the stress-energy tensor is defined.

A.3 Chapter 3

A.3.1 Proof of the expression for N^2

We begin with our definition of N as the root mean squared value of $\hat{s}$ when $h = 0$:

$$
\begin{aligned}
N^2 &= [\langle \hat{s}^2 \rangle - \langle \hat{s} \rangle^2]_{h(t)=0} = [\langle \hat{s}^2 \rangle - S^2]_{h=0} = [\langle \hat{s}^2 \rangle]_{h=0} \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(t)K(t') \langle n(t)n(t') \rangle dt dt' \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(t)K(t') \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle \tilde{n}(f)\tilde{n}(f') \rangle e^{-2\pi i (tf + t'f')} df df' dt dt' \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(t)K(t') \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle \tilde{n}(f)\tilde{n}^*(f') \rangle e^{-2\pi i (tf - t'f')} df df' dt dt' \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(t)K(t') \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{2} S_n(f) \delta(f - f') e^{-2\pi i (tf - t'f')} df df' dt dt' \\
&= \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(t)K(t') \int_{-\infty}^{\infty} S_n(f) e^{-2\pi i f (t - t')} df dt dt' \\
&= \frac{1}{2} \int_{-\infty}^{\infty} S_n(f) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(-t)K(t') e^{2\pi f (t + t')} dt dt' df \\
&= \frac{1}{2} \int_{-\infty}^{\infty} S_n(f) \tilde{K}(-f) \tilde{K}(f) df = \frac{1}{2} \int_{-\infty}^{\infty} S_n(f) \tilde{K}^*(f) \tilde{K}(f) df \\
&= \int_{-\infty}^{\infty} \frac{1}{2} S_n(f) |\tilde{K}(f)|^2 df.
\end{aligned}
\tag{A.27}
$$

This indeed proves the expression given in chapter 4 for N^2.