DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„Surfing an Impulsive Gravitational Wave“

verfasst von / submitted by

Judith Painsi

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 190 406 412
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Lehramtsstudium UF Mathematik UF Physik
Betreut von / Supervisor:	a.o. Univ.-Prof. Mag. Dr. Roland Steinbauer

Abstract

In this thesis we discuss gravitational waves and, in particular, their subclass of impulsive waves. These are idealised models of short but violent bursts of gravitational radiation. We begin, however, with a discussion of Special and General Theory of Relativity, with focus on their physical motivation. Next, we give a rough guide on the mathematical formalism underlying General Relativity, i.e. semi-Riemannian geometry. Gravitational waves are then introduced as solutions to the linearised Einstein equations and we discuss their definition in the full non linear theory. Finally we turn to a simple class of such solutions, the pp-waves. It is this family of solutions which allows for impulsive profile functions amounting to impulsive waves. Finally, we give some methods of construction.

Zusammenfassung

In dieser Diplomarbeit werden Gravitationswellen und im Speziellen deren Unterkategorie der impulsiven Wellen diskutiert. Es handelt sich dabei um idealisierte Modelle von kurzen, starken Pulsen von Gravitationswellen. Zu Beginn wird auf die Spezielle Relativitätstheorie und die Allgemeine Relativitätstheorie mit Fokus auf deren physikalischer Motivation eingegangen. Weiters werden grundlegende Konzepte der semi-Riemannschen Geometrie diskutiert, um ein mathematisches Gerüst aufzubauen. Gravitationswellen werden als Lösungen der linearisierten Einsteingleichungen eingeführt, gefolgt von einer Defintion in der vollen nicht linearen Theorie. Abschließend wird eine einfache Klasse dieser Lösungen, namentlich pp-Wellen, diskutiert. Diese Familie von Lösungen kann impulsive Komponenten aufweisen und wird dann als impulsive Welle bezeichnet. Für diese werden schlussendlich Konstruktionsmethoden angegeben.

Acknowledgements

This was a hard one.

Thank you all for your patience. Special thanks go out to Roland, Barbara and Diethelm, who probably needed to show the most of it.

And the moon is swimming naked
And the summer night is fragrant
With a mighty expectation of relief
So we struggle and we stagger
Down the snakes and up the ladder

Leonard Cohen

Table of Contents

1 Introduction

A warm welcome to the esteemed reader. In this diploma thesis I want to take you on a short ride through space and time. We will start with Galileo Galilei at a point, where those two notions could easily be told apart. But after introducing Einstein's theory of Special Relativity, we will see, that a union of both, the so called *spacetime*, is a better fitting tool to describe nature. To see the consequences of this new concept, we will undergo a quick interlude devoted to the discussion of the newly arisen notion of *Minkowski* spacetimes.

Back on track, the next stop will be a heuristic motivation for Einstein's theory of General Relativity, which can be seen as an extension to Special Relativity, in the sense that it uses its framework, but is able to include gravity. Eventually we will find out that the universality of the laws of gravity inevitably leads to Einstein's idea, that gravitation is a property of spacetime and can be traced back to the presence of matter and energy curving said spacetime.

While the discussion of the fundamental ideas heels a little in favour of 'physicists slang' the subsequent chapters will provide us with a proper mathematical formulation, albeit introducing a noticeable change of style. Since whole doorstoppers have been written about the mathematics of General Relativity this short summary is merely a skirmish of Differential Geometry, if you will.

Nonetheless, after we have ensured us a gravitational toolkit, we can start to discuss the notion of *gravitational waves*, travelling through spacetime at the speed of light. First we will trace back the footsteps of Einstein and solve his linearised field equations, to see that they admit for wavelike solutions. Continuing, we give a definition of gravitational waves in the full non linear theory, so called pp-*waves*. It turns out, the function governing the wave profile is not required to be nice in any mathematical sense. So, the theory allows for so called *impulsive gravitational waves*, which we will discuss in the last chapter. Concluding, we give a short answer to the question: What would be the destiny of a ship full of daring spacefarers trying to 'surf' an impulsive gravitational wave?

Welcome aboard and enjoy your ride.

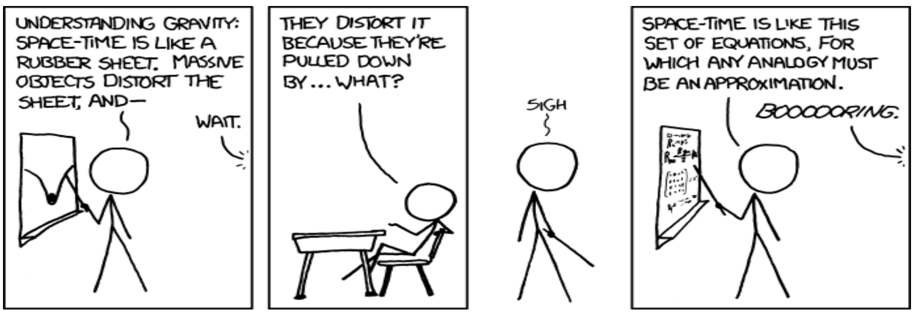

Figure 1: A comic sketching the difficulties of a heuristic approach to General Relativity. XKCD. Taken from https://xkcd.com/895/ (visited on 17/02/2020).

2 Fundamental Ideas

2.1 From Galilean Invariance to Special Relativity

In this section we are undertaking a historical approach to Special Relativity, following in the footsteps of famous mathematician Galileo Galilei. In 1632, he was the first to discuss the principle of relativity in his work *Dialogue Concerning the Two Chief World Systems* [12]. According to him, it is only meaningful to talk about a body being at rest or having a certain velocity if it is related to another body. In other words, there is no absolute notion of velocity. We will follow Galileos *gedankenexperiment* to derive some central aspects of relativity:

Consider a person in a windowless hull of a ship. This person has no way of finding out whether the ship is moving with a constant velocity or whether it is at rest. No experiment, such as tossing a ball in to the air or observing the fall of water drops in a bucket could help distinguishing between the two possibilities. The person needs a reference point outside the system, in this case the ship, to compare her situation to.

Building a mathematical background for this gedankenexperiment we take three-dimensional Euclidean space and add one additional time coordinate to obtain a continuum composed of events. Assigning four numbers (one for time and three for the space components) to each event gives a four dimensional *coordinate system* (t, x, y, z). The set of all these quadruples is called *spacetime* and as we will see later, manifolds are the correct mathematical tool to describe it. If a coordinate system is non-accelerating, i.e. in uniform motion, it is called an *inertial frame of reference* or *inertial system* for short and (t, x, y, z) are called the corresponding *spacetime coordinates* of an event.

Now, how does the trajectory of a falling water drop on the ship look like for a person standing on the shore? In other words, how can we change from a given inertial frame (t, x, y, z) to a different one (t', x', y', z'), which is in uniform motion relative to the first one? Galileo provides us with a set of transformations, consisting of spatial translations, rotations, time translations and velocity transformations. Further, he states, that for physical laws to be meaningful, they need to hold in all inertial frames. That is, if a ball, tossed straight into the air on the ship, has a certain trajectory, it has to have the same, if a person on the shore would repeat the exact same throw, relative to the respective starting point.

The *Galilean transformations* from one inertial frame of reference to another one moving with constant relative velocity v in x-direction (where without loss of generality, we assume the spatial origins to coincide initially, at $t = 0$) are,

$$\begin{aligned} t' &= t, \\ x' &= x + vt, \\ y' &= y, \\ z' &= z. \end{aligned} \tag{1}$$

Recall that all inertial frames are on equal footing. Stressing another picture, it no longer makes a physical or mathematical difference, whether you are sitting on a train, watching the landscape pass by the window or whether you are standing outside along the railway lines and wave at the bypassing passengers inside the train.

We now can formulate the

Gallilean Principle of Relativity

Physical laws need to be *invariant*, i.e. remain unchanged, under the Galilean transformations (1).

Indeed, the laws of classical mechanics and especially Newtons law of gravity obey this principle. Hence, for quite some time transformations (1) were believed to provide a correct description of nature. Only roughly 200 years after Galileo Galilei published his work concerning the principle of relativity, first problems arose as James C. Maxwell developed the infamous *Maxwell equations* [14, pp. 459-512]. They give a fundamental and experimentally well established description of the laws of electro-magnetism. But surprisingly enough, they are not Galilei-invariant. Many physicists and mathematicians of that time, such as Lorentz and Poincaré, started to work on this problem. They found a fitting set of transformations, which will be introduced shortly. But they did not find them without introducing the so-called *aether*, an always distinguishable background coordinate system which is at rest all the time. To prove the existence of said *aether*, in 1887, Albert Abraham Michelson and Edward Morley started a series of experiments [15]. They measured the relative velocity of light, which should change depending on the direction of propagation. Finding no shift was felt as 'serious embarrassment' by Einstein [11, p. 219]. For, the validity of Maxwell's theory was confirmed impressively and therefore presumably correct. Was it possible that the mistake was rather in the general understanding of space and time? It turned out that was very much the case.

Einstein himself was able to solve this rather painful problem. He chose an approach in which he assumes the fundamental principle of relativity to be correct but rejected the Galilean transformations. Instead, he believed in the correctness of the transformations found by Lorentz and Poincaré.

In 1905, he formulated two postulates [8], which are known as the foundation of special relativity. They are partly in accordance with the ideas of Galileo, as the first one is merely a slight modification of the Galilean principle of relativity.

Postulates of Special Relativity

- All inertial systems are on equal footing and physical laws need to have the same form in all of them. Being *at rest* or *in motion* are only meaningful notions if relating two coordinate systems to one another.
- The velocity of light denoted by c is a universal constant.

Using solely these two assumptions, Einstein was able to derive the correct transformations without introducing an aether system. In the situation analogous to (1) they take the form

$$\begin{aligned} t' &= (t + \frac{v}{c^2}x)\gamma, \\ x' &= (x + vt)\gamma, \\ y' &= y, \\ z' &= z, \end{aligned} \tag{2}$$

where $\gamma = (1 - \frac{v^2}{c^2})^{-\frac{1}{2}}$, and c is the speed of light. These equations are called *Lorentz transformations*. For simplicity we set $c = 1$ in the following.

The crucial difference between transformations (1) and (2) is that in the new set of equations time is no longer an absolute quantity and thus the non-trivial first transformation in (2) replaces the trivial equation $t = t'$ in the Galilean case.

Equations (2) describe the change from one inertial frame moving with constant velocity v in x-direction with respect to another and movement in y- and z-direction is described analogously. Casually speaking this kind of transformation is called a *Lorentz boost in x-direction*.
Adding spatial translations, time translations and spatial rotations gives the Poincaré transformations, the set of all possible transformations between inertial coordinate systems.

The derivation of these transformations without introducing the peculiar notion of a background aether drastically changed physic's perception of space and time. From this point on, the concept of time is ultimately connected to the concept of space. These two entities are no longer independent objects, or, as Hermann Minkowski put it in 1906, [16]: '*Von Stund' an sollen Raum für sich und Zeit für sich völlig zu Schatten herabsinken und nur noch eine Art Union der beiden soll Selbständigkeit bewahren.*'

Even though transformations (2) might not be intuitive, their accuracy has been well established in numerous experiments. An impressive confirmation of the theory is the observation of muon flux at the earth's surface. If we wouldn't take Special Relativity into account the number of particles reaching earth would be far too high considering their short lifespan. But when taking a relativistic effect called *time dilatation* (the reason for the famous saying 'moving clocks run slower') into account, the theory is in good accordance with the measurements [26].

2.2 Interlude: Minkowski Spacetime

Since we just followed a historical approach to derive the Lorentz transformations (2), in this chapter we want to discuss the geometrical and so called *causal* properties of Minkowskian spacetime, which constitutes the playground for Special Relativity. For now, we will focus on the heuristics. Rigorous definitions will be given in chapter 3.

In pre-relativistic physics, one thought of the causality structure of spacetime as shown in figure 2. It is divided into three areas. 1) The future of (an event at) point p. 2) it's past and 3) the simultaneity plane. All coordinate systems agree on what it means for two events to happen 'at the same time'. Special Relativity paints a different picture (confer figures 3 and 4 below).

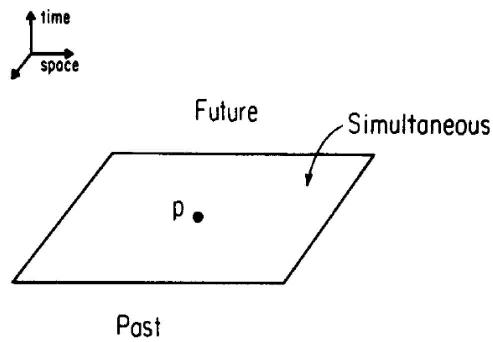

Figure 2: Structure of space and time as pictured in the prerelativistic era. Taken from [29, p. 5].

The causality structure of a spacetime is directly connected to its geometry. To study it, a way to measure lengths and angles is required. In Euclidean geometry, the familiar scalar product determines the squared distant function $ds^2 = dx^2 + dy^2 + dz^2$, which is invariant under the Galilei transformations (1). In Special Relativity though, the scalar product will no longer be positive definite, as we would require of a Euclidean scalar product.

The quantity invariant under Lorentzian transformations (2)

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \tag{3}$$

is called the *Minkowski line element* and determines the *Minkowski metric.* A precise definition will be given in chapter 3.1.3. The matrix representation of the metric corresponding to the line element (3) is

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4}$$

Contrary to the Euclidean case, the minus sign takes $\eta_{\mu\nu}$ to be indefinite, i.e. $\exists v, w$ such that $\eta_{\mu\nu}v^\mu v^\nu < 0$[1]and $\eta_{\mu\nu}w^\mu w^\nu > 0$. As a symmetric bilinear form the matrix is diagonalizable and the number of negative eigenvalues determines the so called signature of the metric. The signature of the Minkowski metric (4) is $(-,+,+,+)$ and its indefiniteness allows to divide the set of vectors into three categories:

- A vector v is said to be *timelike* if $\eta_{\mu\nu}v^\mu v^\nu = -1$.
- A vector v is said to be *spacelike* if $\eta_{\mu\nu}v^\mu v^\nu = 1$.
- A vector v is said to be *lightlike* or *null* if $\eta_{\mu\nu}v^\mu v^\nu = 0$.

Now, let A and B be two events corresponding to distinct points in spacetime. If we fix the event A at point p the event B can be related to it again in three ways, depending on which category the connection vector is in. If it is timelike, then one event can, in principle, influence the other, i.e. there exists a coordinate system where these two events occur at the same place. If that is the case, their relation is called *chronological*. If the connection vector is lightlike, one event can only influence the other with a signal travelling at the speed of light. Both cases are subsumed under the term *causally related*. The final case, the connection vector being spacelike, means that the two events can not influence each other, i.e. there exists a coordinate system in which these two events take place at the same time.

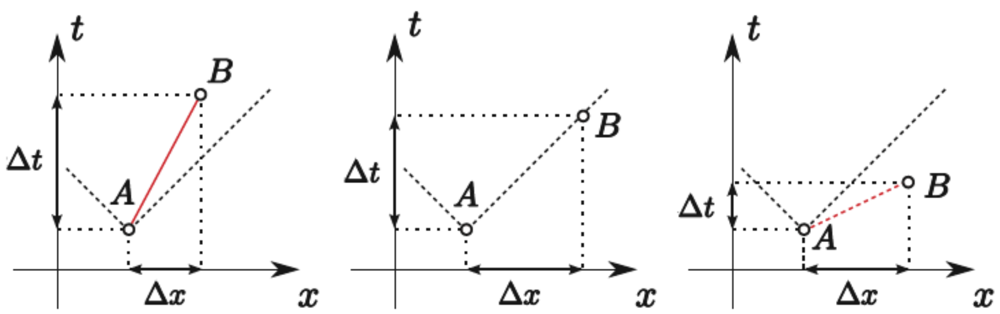

Figure 3: Different causality situations in Minkowski spacetime. Adapted from [9, p. 183].

Figure 3 is a visualisation of Minkowski spacetime omitting all spatial coordinates besides one. The first graph shows two events A and B related in a timelike way, the second two lightlike related events and the third one shows the spacelike case. The dotted lines emerging from A with an angle of $45°$ depict the so-called *future null cone* of A. If we extend them below A, the *past null cone* of A is obtained. Together they are called the *light cone* of A. We can define the light cone C as $C := \{\boldsymbol{v} \in \mathbb{R}^4 | \eta_{\mu\nu}v^\mu v^\nu = 0\}$.

[1]Here, the *Einstein summation convention* is used, i.e. summation over every index that appears as a sub- and as a superscript is implied ($a^i b_i = \sum_{i=0}^{3} a^i b_i = a^0 b_0 + a^1 b_1 + a^2 b_2 + a^3 b_3$ for $i = 0, 1, 2, 3$). This convention will be used throughout this thesis.

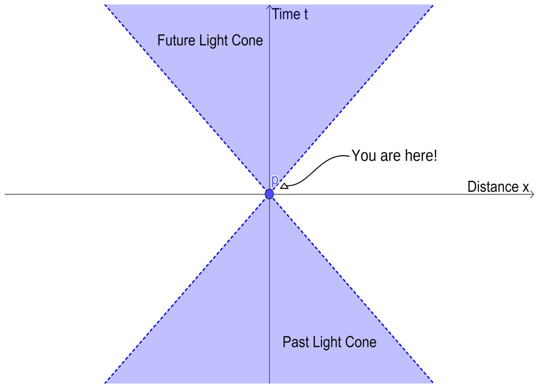

Figure 4: The blue area depicts the causal past and future of event A at point p.

Minkowski spacetime has a maximum number of symmetry transformation for vectors, which are represented by four translations, three rotations and three Lorentz boosts. Together, they form the Lorentz group.

Employing coordinate transformations the Minkowski line element ds^2 can take on different forms. For example, if picking the spatial part of spherical polar coordinates, i.e. $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$ and keeping $t = t$, the line element (3) transforms as:

$$ds^2 = -dt^2 + dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{5}$$

To see this, we calculate the coordinate differentials

$$\begin{aligned} dt &= dt \\ dx &= \sin\theta\cos\phi dr + r\cos\theta\cos\phi d\theta - r\sin\theta\sin\phi d\phi \\ dy &= \sin\theta\sin\phi dr + r\cos\theta\sin\phi d\theta + r\sin\theta\cos\phi d\phi \\ dz &= \cos\theta dr - r\sin\theta d\theta \end{aligned}$$

and substitute their squares into (3). After reorganizing and factoring out we obtain

$$\begin{gathered} ds^2 = -dt^2 + (\sin^2\theta\cos^2\phi + \sin^2\theta\sin^2\phi + \cos^2\theta)dr^2 \\ +(\cos^2\theta\cos^2\phi + \cos^2\theta\sin^2\phi + \sin^2\theta)r^2 d\theta^2 \\ +(\sin^2\phi + \cos^2\phi)r^2\sin^2\theta d\phi^2 + (\cos^2\phi + \sin^2\phi - 1)2r\sin\theta\cos\theta d\theta dr. \end{gathered}$$

Using the identity $\cos^2 x + \sin^2 x = 1$ gives the above line element (5).

Because the scalar product is preserved when changing from Cartesian coordinates to spherical coordinates, the Minkowski metric in the former and the Minkowski metric in the latter coordinates are called *isometric* to each other. This is a good opportunity to introduce another set of isometric coordinates for Minkowski spacetimes, albeit we will only use them a little later on.

Let u and r be so-called *double null* or *light cone coordinates*, i.e.

$$u = \frac{1}{\sqrt{2}}(t - z) \quad \text{and} \quad r = \frac{1}{\sqrt{2}}(t + z).$$

The set of all vectors $u = 0$ and $r = 0$ constitute families of *null planes*, which give the boundary of the nullcone specified above (confer figure 4).

In these coordinates the line element amounts to

$$ds^2 = -2dudr + dx^2 + dy^2. \tag{6}$$

This can easily be shown via plugging in the differentials du and dr in (3):

$$du = \frac{1}{\sqrt{2}}(dt - dz) \qquad\qquad dr = \frac{1}{\sqrt{2}}(dt + dz)$$

$$ds^2 = -2\frac{1}{\sqrt{2}}\frac{1}{\sqrt{2}}(dt - dz)(dt + dz) + dx^2 + dy^2 = -dt^2 + dx^2 + dy^2 + dz^2.$$

The Minkowski metric in double null coordinates amounts to

$$\eta_{\mu\nu} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{7}$$

2.3 Heuristic Motivation for General Relativity

After Albert Einstein introduced Special Relativity, he was far from finished, for the theory explicitly does not include gravity. It took him approximately 15 years of work to achieve this, thereby developing General Relativity. In this section we want to discuss some heuristic ideas motivating the connection between gravity and the geometry of spacetime, since this is the central aspect of the new theory. The following discussion uses some ideas of [30].

First we want to introduce the *weak equivalence principle*, which goes back as far as to the first page of Newtons Principia Mathematica [18]. Recall the equations of motion for a particle in a gravitational field of a mass M_a of classical mechanics:

$$m_i\ddot{r} = \frac{Gm_aM_a}{r^2} = m_a g(r). \tag{8}$$

Here, m_i denotes the initial mass of the particle, m_a is the gravitational mass, G the gravitational constant, $g(r) = \frac{GM_a}{r^2}$ the gravitational acceleration and $\ddot{r}$ is the second derivative of the position vector of the particle $r = (x, y, z)$ with respect to time.

The *weak equivalence principle* states that in equation (8) we have $m_i = m_a$. This equality was famously tested in the Eötvös experiment in 1885 [10, pp. 65-68]. The results showed that the difference between initial and gravitational mass is no bigger than 10^{-6}, i.e. $|m_i - m_a| \leq 10^{-6}$. Modern experiments decreased the difference even more, namely to 10^{-11}.

It is possible to formulate the weak equivalence principle in many slightly different ways, in fact $m_i = m_a$ is already one of them. We choose:

Weak Equivalence Principle

All bodies in free fall in a gravitational field experience the same acceleration, namely $g(r)$, independent of their corresponding masses.

Because the respective masses in (8) cancel, the equation then amounts to $\ddot{r} = g(r)$. This gives rise to questions about the physical nature of gravity itself:

Consider a transformation $t = t'$, $x = x'$, $y = y'$, $z = z' + \frac{1}{2}at^2$, i.e. a system being accelerated uniformly relative to the original system in z-direction.

Then we have $\ddot{r}' = \ddot{r} - (0, 0, a)^t$. Without loss of generality we may assume the gravitational acceleration has only one non-vanishing constant component in the third slot, i.e. $g(r) = (0, 0, g)$. We can make the effects of any apparent gravitational field vanish at a point in an inertial frame, by setting $a = -g$.

In other words, we can always pick a coordinate system 'whithout gravity'. This results in the

Strong Equivalence Principle

Locally, one can not distinguish between the effects of motion in a gravitational field and mere acceleration due to other reasons.

The corresponding *gedankenexperiment* is, to imagine a person in a rocket being in free fall in a gravitational field. The person in the rocket can in no way determine whether it is under the influence of a gravitational field or the rocket is merely accelerating with $a = -g$ in the other direction. So, gravity is unobservable in systems in free fall.

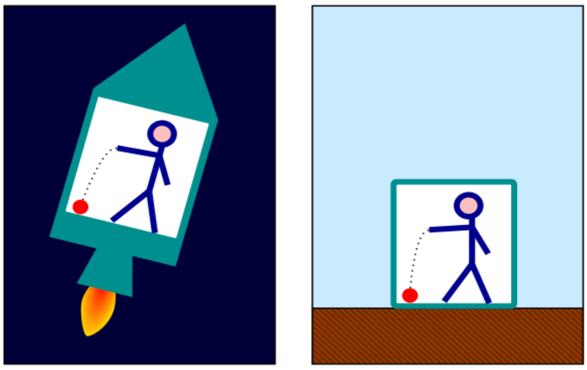

Figure 5: A sketch of Einsteins famous *gedankenexperiment*. Pbroks13, 2008. Taken from https://de.wikipedia.org/wiki/Datei:Elevator_gravity.svg (visited on 14/02/2020).

These findings can be practically experienced by taking a ride on so-called 'vomit comets', reduced gravity-aircrafts. The catchy name originating from the revolting effects of weightlessness on the human stomach, as eyes and vestibular systems send contradicting information to the human brain. Airplanes fly on parabolic trajectories, to create the sensation of zero gravity for the passengers. Around the vertex of the parabola, the plane moves on a ballistic path and is thus in free fall around the earth for a few seconds [28].

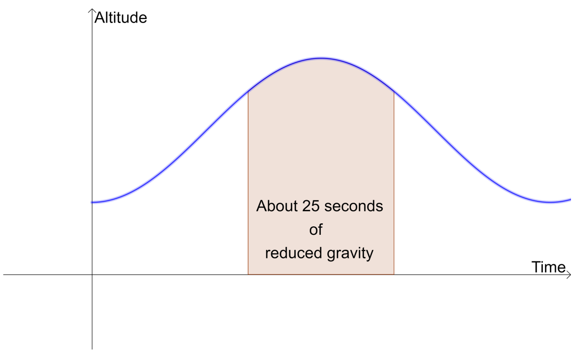

Figure 6: The trajectory of a reduced gravity-aircraft manoeuvre.

So, how can we include gravity in the Special Theory of Relativity? How is an inertial frame of reference to be chosen, if there is no way to distinguish between acceleration and the affects of gravity? For example, in the theory of electro-magnetism this is done using electrically neutral particles, as their movement is independent of electro-magnetic forces. Such an approach is not possible in case of gravity, because there are no objects not affected by it. Fortunately, we already have found a way out: We will assign the role of inertial systems to bodies in free fall. They are natural candidates, as they do not experience any effects of acceleration or gravity. So, Special Relativity holds in systems in free fall and locally we can always find such a system.

As all bodies experience the same effects of gravity, we can understand it to be universal. The infamous idea of Einstein was, that it therefore is a property of spacetime itself, called curvature, and that it is induced by the presence of matter and energy. The relationship between curvature and the matter-energy distribution is given by the Einstein equations (see chapter 3.2).

In Newtonian physics, particles not under the influence of any force move on straight lines. Generalizing this to particles subjected only to gravity, i.e. in free fall, they move on so-called geodesics, which are 'the straightest possible lines on a curved surface', heuristically speaking. Thus, gravity can be included in the theory of special relativity as the effects that curvature of spacetime has on the trajectory of particles. Namely, in the relative acceleration between two inertial observers moving on geodesics.

The next sections will specify the developed concepts and endow them with a precise mathematical meaning.

3 Mathematical Formulation of General Relativity – Differential Geometry

In this chapter we will introduce some general mathematical concepts needed to formulate General Relativity and to study gravitational waves in its framework. Technically speaking, gravitational waves are solutions of the Einstein Equations. Recall, that these equations give a proportionality between curvature and matter-energy distribution. First, we will discuss some concepts in Differential Geometry, to be able to get a firm grip on the notion of curvature later on. Interestingly, this mathematical field was developed much earlier than Einstein actually published his work about gravity. An urban legend says, he needed the time period between the publications of Special Relativity and of General Relativity to study the mathematics needed.

After discussing the basic tools of Differential Geometry, which is going to be quite technical, we will put those new instruments to use. We are looking for a way to relate the Riemannian tensor, in which the curvature of a spacetime is encoded, to the stress-energy tensor, which governs the matter distribution, to be able to write down the Einstein Equations.

All of the definitions and results are retrieved from [19]. Due to the depth of the topic all proofs are omitted, but can be found in said book.

3.1 Foundations of Differential Geometry

3.1.1 Manifold Theory

As it turns out, the most fitting way to describe spacetime is in terms of manifolds. Thus, we will give a precise definition in this section. Casually speaking, manifolds are (not necessarily) higher dimensional generalizations of possibly curved surfaces and could be seen as the playground of differential geometry. More technically, a manifold is a topological space that locally resembles Euclidean space. This fits perfectly our physical demands, that Special Relativity holds in systems in free fall. Therefore, we will first give a definition of smooth manifolds and briefly discuss functions 'living' on such manifolds.

However, to do analysis, we will have to introduce a linear structure locally at any point on the manifold. Generalizing the notion of tangent vectors will provide us with such a structure, namely tangent spaces at points $p \in M$.
Further, we do not want to impose any eventual embedding space of the manifold. In order to do that, we are looking to construct these tangent spaces solely based on intrinsic structures of the manifold.

Given a function $\phi : M \to N$ between manifolds, we will discuss its differential $d\phi$ and concluding we will look at some important maps on manifolds, as curves and vector fields.

3.1.1.1 Smooth Manifolds

First we will clarify some general notations:
For $1 \leq i \leq n$, let $u^i : \mathbb{R}^n \to \mathbb{R}$ be the function that sends each point $p = (p_1, ..., p_n)$ to its ith coordinate p_i. Then $u^1, ...u^n$ are the *natural coordinate functions* of $\mathbb{R}^n$.
A function ϕ from an open set $\mathcal{U}$ of $\mathbb{R}^m$ to $\mathbb{R}^n$ is *smooth* provided each real valued function $u^i \circ \phi$ is smooth $(1 \leq i \leq n)$.

Starting with the general notion of a topological space S, we need a way to 'transport' points from S to $\mathbb{R}^n$, where we can utilize its vector space structure, i.e. where we understand how to do calculus. This is exactly what so-called charts are able to provide:

A *coordinate system* or *chart* in a topological space S is a homeomorphism ξ (a continuous bijection with continuous inverse) of an open set $\mathcal{U} \subseteq S$ onto an open set $\xi(\mathcal{U})$ of $\mathbb{R}^n$.
If we write $\xi(p) = (x^1(p), ..., x^n(p))\ \forall p \in \mathcal{U}$, the resulting functions $x^1, ..., x^n$ are called *coordinate functions* of ξ. Thus $\xi = (x^1, ..., x^n) : \mathcal{U} \to \mathbb{R}^n$.

Two n-dimensional charts ξ and η in S overlap smoothly provided the functions $\xi \circ \eta^{-1}$ and $\eta \circ \xi^{-1}$ are both smooth. That is, if $\xi : \mathcal{U} \to \mathbb{R}^n$ and $\eta : \mathcal{V} \to \mathbb{R}^n$, then $\eta \circ \xi^{-1} : \xi(\mathcal{U} \bigcap \mathcal{V}) \to \eta(\mathcal{U} \bigcap \mathcal{V})$, while $\xi \circ \eta^{-1}$ runs in the opposite direction (see figure 7 below), and these are smooth functions as functions on subsets of $\mathbb{R}^4$.

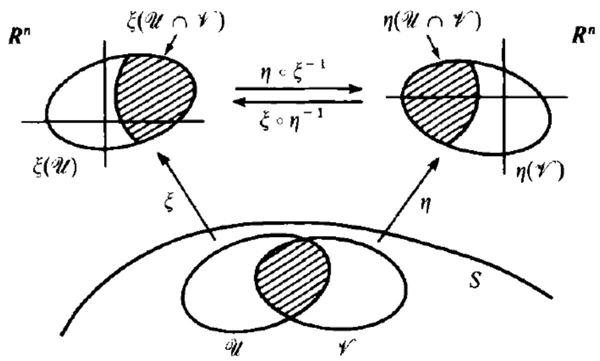

Figure 7: How charts map points from S to $\mathbb{R}^n$. Taken from [19, p. 2].

Definition 3.1.1

An *atlas* $\mathcal{A}$ of dimension n on a topological space S is a collection of n-dimensional charts in S such that
(1) each point of S is contained in the domain of some chart in $\mathcal{V}$
(2) any two charts in $\mathcal{A}$ overlap smoothly.

Now, we can give a rigorous definition of a smooth manifold:

Definition 3.1.2

A *smooth manifold* M is a topological Hausdorff[2]space furnished with a complete atlas (it contains all charts in S that overlap smoothly with every chart in said atlas).

Now, that we have established the notion of manifolds, two natural questions arise. How can we describe functions living on M and what does it mean for them to be smooth? Further, what about maps between two manifolds and what exactly does it mean for such a map to be smooth?

Consider real-valued functions f on M, i.e. $f : M \to \mathbb{R}$. If $\xi : \mathcal{U} \to \mathbb{R}^n$ is a chart in M, then the composite function $f \circ \xi^{-1} : \xi(\mathcal{U}) \to \mathbb{R}$ is called *coordinate expression* for f in terms of ξ and $f = (f \circ \xi^{-1})(x^1, ..., x^n)$ on $\mathcal{U}$.

Definition 3.1.3

Let M^m and N^n be manifolds. A mapping $\phi : M \to N$ is smooth provided that for every chart ξ in M and η in N the coordinate expression $\eta \circ \phi \circ \xi^{-1}$ is Euclidean smooth.

Let $\mathcal{F}(M)$ denote the set of all smooth real-valued functions on M.

An important property of manifolds is their so called orientability:

Definition 3.1.4

A manifold M is *orientable* provided there exists a collection $\mathcal{O}$ of charts in M whose domains cover M and such that for each $\xi, \eta \in \mathcal{O}$ the Jacobian determinant of the chart change $J(\xi, \eta) = \det\left(\frac{\partial y^i}{\partial x^j}\right)$ is positive.

3.1.1.2 Tangent Vectors

Because manifolds do not have a vector space structure, we need a way to assure that calculus is retained. This is possible, equipping points with their corresponding tangent spaces. For this, we use the concept of tangent vectors as directional derivatives.

[2]The Hausdorff property, also called second separation axiom, secures that all distinct points $p \neq q \in S$ can be separated by neighbourhoods.

Definition 3.1.5: Axiomatization of Directional Derivatives

Let $p \in M$ be a point. A *tangent vector* to M at p is a real-valued function $v : \mathcal{F}(M) \to \mathbb{R}$, that satisfies
(1) $\mathbb{R}$-Linearity: $v(af + bg) = av(f) + bv(g)$
(2) Leibniz-rule: $v(fg) = v(f)g(p) + f(p)v(g)$
$\forall a, b \in \mathbb{R}$ and $\forall f, g \in \mathcal{F}(M)$.

At each $p \in M$ let T_pM be the set of all tangent vectors to M at p. With functional addition and scalar multiplication defined in terms of charts T_pM is a vector space over $\mathbb{R}$.

To define partial differentiation we move $f \in \mathcal{F}(M)$ back to Euclidean space with a chart and take the usual derivative.

Definition 3.1.6

Let $\xi = (x^1, ..., x^n)$ be a chart in M at p.
If $f \in \mathcal{F}(M)$, let $\frac{\partial f}{\partial x^i}(p) = \frac{\partial(f \circ \xi^{-1})}{\partial u^i}(\xi(p)) \quad (1 \leq i \leq n)$.

Computation shows that the function $\partial_i|_p = \frac{\partial}{\partial x^i}\Big|_p : \mathcal{F}(M) \to \mathbb{R}$, $f \mapsto \frac{\partial f}{\partial x^i}(p)$ is a tangent vector to M at p.

The next theorem gives the link between charts and tangent spaces. Partial derivatives of the chart form a basis for the tangent space.

Theorem 3.1.1: Basis Theorem

If $\xi = (x^1, ..., x^n)$ is a chart in M at p, then we call $\partial_1|_p, ..., \partial_n|_p$ the*coordinate vectors* of ξ and they form a basis of T_pM and we may write $v = \sum_{i=1}^n v(x^i)\ \partial_i|_p\ \forall v \in T_pM$.

We stated above that T_pM is a vector space and so have reached our goal to obtain a meaningful notion of calculus on manifolds.

3.1.1.3 Differential Maps

A so called differential map can be understood to be a map between tangents spaces of smooth manifolds. It can be seen as a generalization of total derivatives.

Let $\phi : M \to N$ be a function. If $v \in T_pM$, then $v_\phi : \mathcal{F}(N) \to \mathbb{R}$ defined via $g \mapsto v(g \circ \phi)$ is a tangent vector to N at $\phi(p)$.

Definition 3.1.7

Let $\phi : M \to N$ be smooth.
Then $\forall p \in M$ $d\phi_p : T_pM \to T_{\phi(p)}N$, $v \mapsto v_\phi$ is called *differential map* of ϕ at p.

Thus, $d\phi_p(v)(g) = v(g \circ \phi)$ $\forall v \in T_pM$ and $\forall g \in \mathcal{F}(N)$.

The next lemma shows what the differential map of ϕ looks like, using a coordinate system.

Lemma 3.1.1

Let $\phi : M^m \to N^n$ be smooth.
If ξ is a chart at p in M and η a chart at $\phi(p)$ in N, then
$d\phi_p \left(\frac{\partial}{\partial x^j}\Big|_p \right) = \sum_{i=1}^n \frac{\partial(y^i \circ \phi)}{\partial x^j}(p) \, \frac{\partial}{\partial y^i}\Big|_{\phi(p)} \quad (1 \leq j \leq m).$

Thus, the matrix of $d\phi_p$ with respect to this coordinate basis is

$$\left(\frac{\partial(y^i \circ \phi)}{\partial x^j}(p) \right)_{1 \leq i \leq n, 1 \leq j \leq m}.$$

3.1.1.4 Curves, Vector Fields, One-Forms and Integral Curves

Next we will discuss important special cases of mappings on manifolds.

Curves:
A *curve* in M is a smooth mapping $\alpha : I \to M$. As an open submanifold of $\mathbb{R}$, the interval I has a chart consisting of the identity map u of I. At eacht $t \in \mathbb{R}$ we can picture the coordinate vector $\left(\frac{d}{du}\right)(t) \in T_t\mathbb{R}$ as the unit vector at t in positive u direction.

Definition 3.1.8

Let $\alpha : I \to M$ be a (parametrized) curve.
The *velocity vector* of α at $t \in I$ is $\alpha'(t) = d\alpha \left(\frac{d}{du}\Big|_t \right) \in T_{\alpha(t)}M$.

1. Directional derivative: By definition of $d\alpha$ the tangent vector $\alpha'(t)$ applied to $f \in \mathcal{F}(M)$ gives $\alpha' f = \frac{d(f \circ \alpha)}{du}(t)$.

   Thus, if α is any curve with $\alpha'(0) = v$, then $v(f) = \frac{d(f \circ \alpha)}{dt}(0)$.

2. Coordinate expression: Let $x^1, ..., x^n$ be a chart in M at $\alpha(t)$ of α.

   Then $\alpha'(t) = \sum_i \frac{d(x^i \circ \alpha)}{du}(t) \; \partial_i|_{\alpha(t)}$.

3. Reparametrization: Let $\alpha : I \to M$ be a curve and $h : J \to I$ a smooth function.

Then $\beta = \alpha(h) : J \to M$ is a curve called *reparametrization*.

Further, $\beta'(s) = \frac{dh}{du}(s)\alpha'(h(s)) \; \forall s \in J$.

4. Effect of mapping: Let $\phi : M \to N$ carry α to the curve $\phi \circ \alpha : I \to N$ in N.

   The differential map of ϕ transports velocities accordingly, i.e. $d\phi(\alpha'(t)) = (\phi \circ \alpha)'(t)$ $\forall t \in I$.

Vector Fields:
A *vector field* V on a manifold M is a function that assigns to each $p \in M$ a tangent vector V_p to M at p. With $f \in \mathcal{F}(M)$, Vf denotes $(Vf)(p) = V_p(f) \; \forall p \in M$. Then V is smooth provided Vf is smooth $\forall f \in \mathcal{F}(M)$.

Let $\mathfrak{X}(M)$ denote the *set of all smooth vector fields* on M.

If $\xi = (x^1, ..., x^n)$ is a chart on $\mathcal{U} \subset M$, then $\forall 1 \leq i \leq n$ the vector field ∂_i on $\mathcal{U}$, sending each p to $\partial_i|_p$, is called the *ith coordinate vector field* of ξ. From the basis theorem it follows, that for any vector field V: $V = \sum Vx^i \partial_i$ on $\mathcal{U}$.

A *derivative* on $\mathcal{F}(M)$ is a function $\mathcal{D} : \mathcal{F}(M) \to \mathcal{F}(M)$ with the properties

1. $\mathbb{R}$-Linearity: $\mathcal{D}(af + bg) = a\mathcal{D}(f) + b\mathcal{D}(g)$ $(a, b \in \mathbb{R})$
2. Leibniz rule: $\mathcal{D}(fg) = \mathcal{D}(f)g + f\mathcal{D}(g)$.

The definition of tangent vectors shows, that for a vector field $V \in \mathfrak{X}(M)$ the function $f \to Vf$ is a derivative on $\mathcal{F}(M)$. Conversely, every derivative $\mathcal{D}$ on $\mathcal{F}(M)$ comes from a vector field.

One-Forms:
One-forms on a smooth manifold are objects dual to vector fields. At $p \in M$ the dual space $T_pM^\star$ of T_pM is called *cotangent space*. Elements of $T_pM^\star$ called *covectors* are linear maps of T_pM into $\mathbb{R}$.

Definition 3.1.9

A *one-form* Θ on a manifold M is a function that assigns to each point p an element Θ_p of the cotangent space $T_pM^\star$.

A one-form Θ is smooth provided ΘX is smooth $\forall X \in \mathfrak{X}(M)$.

It is possible to transform functions into one forms:

Definition 3.1.10

The *differential* of $f \in \mathcal{F}(M)$ is the one-form df such that $(df)(v) = v(f)$ $\forall$ tangent vectors v in T_pM for all $p \in M$.

If $x^1, ..., x^n$ is a chart on $\mathcal{U} \subset M$ we thus have the coordinate one-forms $dx^1, ..., dx^n$ on $\mathcal{U}$. At each point of $\mathcal{U}$ these provide a dual basis to the coordinate vector fields $\partial_1, ..., \partial_n$, since $dx^j(\partial_j) = \frac{\partial x^i}{\partial x^j} = \delta_{ij}$. It follows that $\Theta = \sum \Theta(\partial_i)dx^i$ on $\mathcal{U}$. In particular, if $f \in \mathcal{F}(M)$, then $df = \sum \frac{\partial f}{\partial x^i} dx^i$ on $\mathcal{U}$ (since $df(\partial_i) = \frac{\partial f}{\partial x^i}$).

Integral Curves:
A vector field on a manifold gives rise to a differential equation.

Definition 3.1.11

A curve $\alpha : I \to M$ is an *integral curve* of $V \in \mathfrak{X}(M)$ provided $\alpha' = V_\alpha$, i.e. $\alpha' = V_{\alpha(t)}$ $\forall t \in I$.

Thus, at each point, the curve α has the velocity prescribed by V. Using a chart ξ, we obtain a system of first-order differential equations:

$$\frac{d(x^i \circ \alpha)}{dt} = F^i(x^1 \circ \alpha, ..., x^n \circ \alpha) \quad (1 \leq i \leq n),$$

where F^i is the coordinate expression for V with respect ot the chart x^i.

The uniqueness and existence theorem for such systems of differential equations give in our case:

Proposition 3.1.1

If $V \in \mathfrak{X}(M)$, then for each $p \in M$ there exists an interval I around 0 and a unique integral curve $\alpha : I \to M$ of V such that $\alpha(0) = p$.

3.1.2 Tensors

In this section we will define the notions of tensors and tensor fields. It is merely a generalization of real-valued functions, vector fields and one-forms on manifolds. This will enable us to introduce metrics as a central tool for studying the geometry of manifolds. Metrics are symmetric nondegenerate bilinear forms and these themselves are tensor fields. A metric determines a scalar product on T_pM for all $p \in M$. Casually put, with a scalar product at hand we can start practising geometry, i.e. measuring lengths and angles.

Additionally, tensors allow us to formulate the principle of general covariance of the laws of physics. Already discussed in the heuristic motivations in the beginning, we want for physical laws to hold in all inertial frames of reference. So they should be coordinate independent. Equations that can be written in tensorial form give exactly that. They are independent of a chart and therefore are correct in all frames of reference.

3.1.2.1 Basic Algebra

We are going to focus on two main cases, i.e. the module $\mathfrak{X}(M)$ over $\mathcal{F}(M)$ and the vector space T_pM over $\mathbb{R}$.

In the simplest case we can picture a tensor as a multilinear map from a product of a certain number of vector spaces and dual spaces into the field of real numbers. For the sake of generality we will choose a more abstract approach and replace 'vector space' with 'module' and 'field' with 'ring'. Nevertheless we have to clarify the notion of *multilinearity* first:

Let $V_1, ..., V_s$ be modules over a ring K. Then $V_1 \times \cdots \times V_s$ is the set of all s-tupels $(v_1, ..., v_s)$ with $v_i \in V_i$. Usual addition and multiplication make $V_1 \times \cdots \times V_s$ a module over K, called direct product. If W is another module over K, a function $A : V_1 \times \cdots \times V_s \to W$ is K-multilinear provided A is K-linear in each slot, i.e. for $1 \leq i \leq s$ and $v_j \in V_j$ $(i \neq j)$, the function $v \to A(v_1, ..., v_{i-1}, v, v_{i+1}, ..., v_s)$ is K-linear.

For V a K-module, let $V^\star$ be the set of all K-linear functions from V to K. Usual addition and multiplication make $V^\star$ a module over K, called the *dual module* of V. If $V_i = V$ for $1 \leq i \leq s$, the notation $V_1 \times \cdots \times V_s$ is abbreviated to V^s.

Definition 3.1.12

For integers $r \geq 0$, $s \geq 0$ not both zero, a K-multilinear function

$$A : (V^\star)^r \times V^s \to K$$

is called a *tensor* of type (r, s) over V.

The *set of all tensors* of type (r, s) over V is denoted by $\mathcal{T}^r_s(V)$ and again is a module over K.

3.1.2.2 Tensor Fields

A *tensor field* A on a manifold M is a tensor over the $\mathcal{F}(M)$-module $\mathfrak{X}(M)$. Thus, if A is of type (r, s) it is an $\mathcal{F}(M)$-multilinear function $A : \mathfrak{X}^\star(M)^r \times \mathfrak{X}(M)^s \to \mathcal{F}(M)$.
Heuristically speaking A is a 'machine' which when fed r one-forms $\Theta^1, ..., \Theta^r$ and s vector fields $X_1, ..., X_s$ produces a real-valued function $f = A(\Theta^1, ..., \Theta^s, X_1, ..., X_s) \in \mathcal{F}(M)$.

The local forms for a vector field $X = \sum X(x^i)\partial_i$ and a one-form $\Theta = \sum \Theta(\partial_i)dx^i$ extend to tensor fields in the following way:

Definition 3.1.13

Let $\xi = (x^1, ..., x^n)$ be a chart on $\mathcal{U} \subset M$. If $A \in \mathcal{T}_s^r$, the components of A relative to ξ are the real-valued functions, $A_{j_1 \cdots j_s}^{i_1 \cdots i_r} = A(dx^{i_1}, ..., dx^{i_r}, \partial_{j_1}, ..., \partial_{j_s})$ on $\mathcal{U}$.

Giving an example, the tensor field $A \in \mathcal{T}_2^1$ can be expressed in coordinates in the following way:

$$A(\Theta, X, Y) = A(dx^k, \partial_i, \partial_j)\Theta_k X^i Y^j = A_{ij}^k \Theta_k X^i Y^j.$$

The components of a tensor product between two tensors is given by

$$(A \otimes B)_{j_1,...,j_{s+s'}}^{i_1,...,i_{r+r'}} = A_{j_1,...,j_s}^{i_1,...,i_r} \cdot B_{j_{s+1},...,j_{s+s'}}^{i_{r+1},...,i_{r+r'}},$$

where indices run from 1 to $n = dim M$.

In coordinates for $A \in \mathcal{T}_2^1$ and $B \in \mathcal{T}_1^1$ we can write for the product $A \otimes B \in \mathcal{T}_3^2$:

$$(A \otimes B)_{ijp}^{kq} = (A \otimes B)(dx^k, dx^q, \partial_i, \partial_j, \partial_p) = A(dx^k, \partial_i, \partial_j) \cdot B(dx^q, \partial_p) = A_{ij}^k B_p^q.$$

3.1.2.3 Operations on Tensors: Contraction, Pull Back and Tensor Derivations

This operation reduces the degree of (r, s)-tensors to $(r-1, s-1)$. The general definition derives from following special case:

Lemma 3.1.2

There is a unique $\mathcal{F}(M)$-linear function $C : \mathcal{T}_1^1(M) \to \mathcal{F}(M)$, called $(1, 1)$-contraction such that $C(X \otimes \Theta) = \Theta X \ \forall X \in \mathfrak{X}(M), \Theta \in \mathfrak{X}^\star(M)$.

We discuss the general case by giving an example:
If A is a $(2, 3)$-tensor field, then $C_3^1(A)$ is the $(1, 2)$-tensor field given by $(C_3^1 A)(\Theta, X, Y) = C\{A(\cdot, \Theta, X, Y, \cdot)\}$. Relative to a chart the components of $C_3^1 A$ are:

$$\begin{gathered}(C_3^1 A)_{ij}^k = (C_3^1 A)(dx^k, \partial_i, \partial_j) = \\ C\{A(\cdot, dx^k, \partial_i, \partial_j, \cdot)\} = \\ \sum_m A(dx^m, dx^k, \partial_i, \partial_j, \partial_m) = \sum_m A_{ijm}^{mk}.\end{gathered}$$

By means of a map from M to N any covariant tensor on N can be pulled back to M.

Definition 3.1.14

Let $\phi : M \to N$ be smooth. If $A \in \mathcal{T}^0_s(N)$ with $s \geq 1$, let

$$(\phi^\star A)(v_1, ..., v_s) = A(d\phi v_1, ..., d\phi v_s) \quad \forall v_i \in T_pM, p \in M$$

Then $\phi^\star(A)$ is called *pullback* of A by ϕ and $\phi^\star : \mathcal{T}^0_s(N) \to \mathcal{T}^0_s(M)$.

We conclude this section about operations on tensors with discussing their derivations:

Definition 3.1.15

A tensor derivation $\mathcal{D}$ on a smooth manifold M is a set of $\mathbb{R}$-linear functions

$$\mathcal{D} = \mathcal{D}^r_s : \mathcal{T}^r_s(M) \to \mathcal{T}^r_s(M) \quad (r \geq 0, s \geq 0)$$

such that for any tensors A and B:

1. $\mathcal{D}(A \otimes B) = \mathcal{D}A \otimes B + A \otimes \mathcal{D}B$
2. $\mathcal{D}(CA) = C(\mathcal{D}A)$ for any contraction.

As a derivative $\mathcal{D}$ is $\mathbb{R}$-linear, preserves tensor types, obeys the product rule and commutes with all contractions. We now introduce a special tensor derivation:

Proposition 3.1.2

If $\mathcal{D}$ is a tensor derivation on M and $\mathcal{U}$ is an open subset, then there is a unique tensor derivation $\mathcal{D}_\mathcal{U}$ on $\mathcal{U}$ such that

$$\mathcal{D}_\mathcal{U}(A|_\mathcal{U}) = (\mathcal{D}A)|_\mathcal{U},$$

for all tensors A on M.

In the following we give another useful special case:

Definition 3.1.16

If $V \in \mathfrak{X}(M)$ the tensor derivative L_V defined by

- $L_V(f) = Vf \quad f \in \mathcal{F}(M)$
- $L_V(X) = [V, X] \quad \forall X \in \mathfrak{X}(M)$

is called *Lie derivative relative to* V.

Here $[.,.]$ is the usual Lie bracket given by $[X, Y](f) = X(Y(f)) - Y(X(f))$ for smooth vector fields X and Y.

3.1.2.4 The Scalar Product as Bilinear Form

In the broad field of geometry a special form of tensors plays a very important role. A scalar product is needed to measure lengths and angles. Generalizing its well known Euclidean form gives rise to different and probably unintuitive properties.

Let V be the finite dimensional real vector space $\mathbb{R}^n$. A *bilinear form* on V is an $\mathbb{R}$-bilinear function $b : V \times V \to \mathbb{R}$. We consider only the symmetric case, i.e. $b(v, w) = b(w, v)\ \forall v, w$.

We now drop the requirement for a scalar product to be positive definite and define it in terms of bilinear forms, which can have one of the three properties:

Definition 3.1.17

A *symmetric bilinear form* b on V is either

1. positive (negative) definite if $v \neq 0$ implies $b(v, v) > 0\ (< 0)$
2. positive (negative) semidefinite if $b(v, v) \geq 0\ (\leq 0)\ \forall v \in V$
3. nondegenerate if $b(v, w) = 0\ \forall w \in V$ implies $v = 0$.

In the Lorentzian case this has an impact on the causality structure of a spacetime, as we have seen in chapter 2.2.

Definition 3.1.18

A *scalar product* g on a vector space V is a nondegenerate symmetric bilinear form on V.

Euclidean Geometry therefore uses merely a special case of the scalar product, i.e. a positive definite one, which in our present terminology is called *inner product*.

A way of categorising nondegenerate bilinear forms is to indicate its so called *index*.

Definition 3.1.19

The *index* ν of a symmetric bilinear form b on V is the largest integer that is the dimension of a subspace $W \subset V$ on which $b|_W$ is negative definite.

Thus $0 \leq \nu \leq \dim V$, and $\nu = 0$ if and only if b is positive definite. The Minkowski metric has index $\nu = 1$

If $e_1, ..., e_n$ is a basis for V, the $(n \times n)$-matrix $(b_{ij}) = b(e_i, e_j)$ is called the matrix of b relative to $(e_1, ..., e_n)$.

Lemma 3.1.3

A symmetric bilinear form is nondegenerate if and only if its matrix is invertible.

Closely related to the bilinear form b is the function $q : V \to \mathbb{R}$ given by $q(v) = b(v, v)$, called *quadratic form* of b.

Generally, scalar products admit for so called *null vectors*. A vector $v \in V$ is *null* provided $q(v) = 0$. Evidently, these exist if and only if g is not definite.

Vectors $v, w \in V$ are called *orthogonal* provided $g(v, w) = 0$. Note, when g is not definite we can no longer picture orthogonal vectors as being at right angles to each other.

Choosing a basis the scalar product g has a matrix-representation. Relative to an orthonormal basis $e_1, ..., e_n$ for V, i.e. $\langle e_i, e_j \rangle = \delta_{ij} e_j$, it will be diagonal. In fact, $g(e_i, e_j) = \delta_{ij} \epsilon_j$, where $\epsilon_j = g(e_j, e_j) = \pm 1$. We will call $(\epsilon_1, ..., \epsilon_n)$ the signature of g (negative signs first) and it relates to the index of g as follows:

Lemma 3.1.4

For any orthonormal basis $e_1, ..., e_n$ for V, the number of negative signs in the signature $(\epsilon_1, ..., \epsilon_n)$ is the index ν of V.

Some final words on how scalar products behave under linear transformations:

Let V and $\bar{V}$ have scalar products g and $\bar{g}$. A linear transformation $T : V \to \bar{V}$ preserves scalar products if $\bar{g}(Tv, Tw) = g(v, w)\ \forall v, w \in V$. In this case T is necessarily one-to-one, because if $Tv = 0$, then $g(v, w) = 0\ \forall w$, hence $v = 0$. T preserves scalar products if and only if it preserves their associated quadratic forms, i.e. $\bar{q}(Tv) = q(v)\ \forall v \in V$. A linear isomorphism $T : V \to W$ that preserves scalar products is called a *linear isometry*.

3.1.3 Semi-Riemannian Manifolds

Now, finally we will introduce the tools of geometry, defining a scalar product on each tangent space of a point $p \in M$ using a metric g. Thus, we construct semi-Riemannian manifolds. Etymologic 'semi' refers to the nondegenerateness of the scalar product and 'Riemannian' is to be understood as distinction from the flat Euclidean case.

As a first approach to curvature, we are going to investigate the success or failure of a vector to retain its initial value after a so called *parallel transport* on a semi-Riemannian manifold M. This requires more than just a manifold structure, namely a connection on M.

With the notion of parallel translation at hand we can define geodesics, 'lines that curve as little as possible'. Recall, that they give the trajectories of particles in free fall.

Concluding this chapter, we will introduce a way to find locally nice coordinate patches for a given manifold M. The so called exponential map, is a very powerful tool, when it comes to computations.

3.1.3.1 Metrics

Definition 3.1.20

A *metric tensor* g on a smooth manifold M is a symmetric nondegenerate $(0, 2)$ tensor field on M of constant index.

So, a metric $g \in \mathcal{T}_2^0(M)$ smoothly assigns to each point $p \in M$ a scalar product g_p on the tangent space T_pM and the index of g_p is the same $\forall p$. We will employ $\langle \cdot, \cdot \rangle$ as an alternative notation for g_p, i.e. $g(v, w) = \langle v, w \rangle \in \mathbb{R}$.

Definition 3.1.21

A *semi-Riemannian manifold* (M, g) is a smooth manifold M furnished with a metric tensor g.

If $x^1, ..., x^n$ is a chart on $\mathcal{U} \subset M$ the components of the metric tensor g on $\mathcal{U}$ are $g_{ij} \langle \partial_i, \partial_j \rangle$ (for $1 \leq i, j \leq n$). Thus, for vector fields $V = V^i \partial_i$ and $W = W^j \partial_j$ we have

$$g(V, W) = \langle V, W \rangle = g_{ij} V^i W^j.$$

Since g is nondegenerate, at each $p \in \mathcal{U}$ the matrix $(g_{ij}(p))$ is invertible and its inverse matrix is denotes by $(g^{ij}(p))$. Since g is symmetric, $g_{ij} = g_{ji}$ and hence $g^{ij} = g^{ji}$ for $1 \leq i, j \leq n$. Finally, on $\mathcal{U}$ the metric tensor can be written as $g = g_{ij} dx^i \otimes dx^j$.

Let $q(v) = \langle v, v \rangle$ for all tangent vectors v to M. This is called the *line element* of M, denoted by ds^2. At each $p \in M$, q gives the associated quadratic form of the scalar product at p this way determining the metric tensor.

If $V \in \mathfrak{X}(M)$ and $f \in \mathcal{F}(M)$, then $q(fV) = f^2 q(V) \in \mathcal{F}(M)$, so q is not a tensor field. In coordinates we have

$$q = ds^2 = g_{ij} dx^i dx^j. \tag{9}$$

Using q on a vector field amounts to

$$q(V) = g_{ij} dx^i(V) dx^j(V) = g_{ij} V^i V^j.$$

The norm of a tangent vector $v \in T_p M$ is $|q(v)|^{\frac{1}{2}} = |\langle v, v\rangle|^{\frac{1}{2}}$.

We already know the Minkowski line element $ds^2 = -dt^2 + dx^2 + dy^2 + dz^2$. It determines the Minkowski metric (4). It is a nondegenerate scalar product of index $\nu = 1$.

Definition 3.1.22

Let M, N be semi-Riemannian manifolds with metric tensors g_M, g_N.
An *isometry* is a diffeomorphism $\phi : M \to N$ that preserves the metric tensors: $\phi^\star(g_N) = g_M$.
Explicitly $\langle d\phi(v), d\phi(w)\rangle = \langle v, w\rangle \ \forall v, w \in T_p M, p \in M$.

An object preserved by all isometries is called an isometric invariant. Semi-Riemannian geometry is traditionally described as the study of such invariants. Roughly speaking isometric manifolds are geometrically the same.

3.1.3.2 The Levi-Civita Connection

Since we are aiming at giving an intrinsic notion of the curvature of a manifold, by investigating parallel transport, we want to compare two possibly different vectors with each other. But given solely a manifold structure this is an impossible task, for given two distinct points p and q both in M, the corresponding tangent spaces V_p and V_q will possibly be different. It is clear that a notion of parallel transport will involve taking derivatives of vector fields on manifolds, as we want to investigate the rate of change of these vector fields. That is what we are going to discuss next.

Let V and W be vector fields on a semi-Riemannian manifold M. We want to define a new vector field $D_V W$ on M whose value at each point p is the vector rate of change of W in the V_p direction. There is a natural way to do this on $\mathbb{R}^n_\nu$:

Definition 3.1.23

Let $u^1, ..., u^n$ be the natural coordinates on $\mathbb{R}^n_\nu$. If V and $W = \sum W^i \partial_i$ are vector fields on $\mathbb{R}^n_\nu$, the vector field $D_V W = \sum V(W^i)\partial_i$ is called the *natural covariant derivative* of W with respect to V.

We are now going to axiomatize the key properties of the natural derivative. That is, we take a step back from vector spaces and fields and onto modules and rings.

Definition 3.1.24

A *connection* D on a smooth manifold M is a function $D : \mathfrak{X}(M) \times \mathfrak{X}(M)$ such that

- (1) $D_V W$ is $\mathcal{F}(M)$-linear in V
- (2) $D_V W$ is $\mathbb{R}$-linear in W
- (3) $D_V(fW) = (Vf)W + fD_V W$ for $f \in \mathcal{F}(M)$.

$D_V W$ is called the *covariant derivative* of W with respect to V for the connection D.

Next proposition will give a useful algebraic tool:

Proposition 3.1.3

Let M be a semi-Riemannian manifold.
If $V \in \mathfrak{X}(M)$, let $V^\star$ be the one-form on M such that $V^\star(X) = \langle V, X\rangle\ \forall X \in \mathfrak{X}(M)$.
Then the function $V \to V^\star$ is an $\mathcal{F}(M)$-linear isomorphism from $\mathfrak{X}(M)$ to $\mathfrak{X}^\star(M)$.

Thus, in semi-Riemannian geometry we can freely transform a vector field in a one-form, and vice versa. Corresponding pairs $V \leftrightarrow \Theta$ contain exactly the same information and are called *metrically equivalent*.

The following theorem is sometimes called 'the miracle of semi-Riemannian geometry' as in [19]. On every semi-Riemannian manifold there exists a unique connection with the additional properties (D4) and (D5).

Theorem 3.1.2

On a semi-Riemannian manifold M there is a unique connection D such that

- $(D4)\ [V, W] = D_V W - D_W V$
- $(D5)\ X\langle V, W\rangle = \langle D_X V, W\rangle + \langle V, D_X W\rangle\ \forall X, V, W \in \mathfrak{X}(M)$.

D is called the *Levi-Civita connection* of M.

From now on, given a semi-Riemannian manifold (M, g) we will always equip it with its Levi-Civita connection.

Naturally, we can give a local expression, choosing a coordinate system:

Definition 3.1.25

Let $x^1, ..., x^n$ be a chart on a neighbourhood $\mathcal{U}$ in a semi-Riemannian manifold M. The *Christoffel symbols* for this chart are the real-valued functions Γ^k_{ij} on $\mathcal{U}$ such that $D_{\partial_i}(\partial_j) = \Gamma^k_{ij}\partial_k \; 1 \leq i, j \leq n.$

Since $[\partial_i, \partial_j] = 0$, it follows from $(D4)$ that $D_{\partial_i}(\partial_j) = D_{\partial_j}(\partial_i)$ and hence $\Gamma^k_{ij} = \Gamma^k_{ji}$.

Since the Christoffel symbols depend on the chosen chart, it is clear, that they are not tensorial, i.e. coordinate independent, quantities. Nonetheless, especially in older physics literature, a great significance is attributed to this very obvious fact. This issue is reflecting the physicist's way of defining (or trying to define) global objects, as tensors for example, by patching together local objects that have the 'correct' transformation behaviour and not defining them globally, as we did in chapter 3.1.2.

Figure 8: A little thrust in the style of René Magritte at the physicist community, where some refuse to define tensor formalism in a modern way.

The local formula of the Christoffel symbols amounts to:

Proposition 3.1.4

For a chart $x^1, ..., x^n$ on $\mathcal{U}$,

1. $D_{\partial_i}(W^j \partial_j) = (\frac{\partial W^k}{\partial x^i} + \Gamma^k_{ij} W^j)\partial_k$ with
2. $\Gamma^k_{ij} = \frac{1}{2} g^{mk}(\frac{\partial g_{jm}}{\partial x^i} + \frac{\partial g_{im}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^m})$.

A vector field V is called *parallel* if its covariant derivatives $D_X V$ are zero $\forall X \in \mathfrak{X}(M)$. In general the Christoffel symbols of a chart measure the failure of is coordinate vector fields to be parallel. Thus, the vanishing of the Christoffel symbols means that the natural coordinate vector fields on $\mathbb{R}^n_\nu$ are parallel.

The covariant derivative D_V can be extended to operate on arbitrary tensor fields:

Definition 3.1.26

Let V be a vector field on a semi-Riemannian manifold M. The *(Levi-Civita) covariant derivative* D_V is the unique derivative on M such that
$D_V f = V f$ for $f \in \mathcal{F}(M)$, and $D_V W$ is the Levi-Civita covariant derivative $\forall W \in \mathfrak{X}(M)$.

We also define a differential of tensors:

Definition 3.1.27

The *covariant differential* of an (r, s) tensor A on M is the $(r, s+1)$ tensor DA such that $(DA)(\Theta^1, ..., \Theta^r, K_1, ..., X_s, V) = (D_V A)(\Theta^1, ..., \Theta^r, K_1, ..., X_s)$ $\forall V, X_i \in \mathfrak{X}(M)$ and $\Theta^j \in \mathfrak{X}^\star(M)$.

In the case $r = s = 0$ the covariant differential of a function f is its usual differential $df \in \mathfrak{X}^\star(M)$, since

$$(Df)(V) = D_V f = V f = df(V) \quad \forall V \in \mathfrak{X}(M).$$

3.1.3.3 Parallel Translation

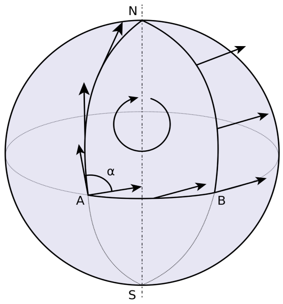

Figure 9: Parallel transport of a vector on the surface of a sphere. Fred the Oyster, 2014. Taken from https://commons.wikimedia.org/wiki/File:Parallel_Transport.svg (visited on 17/02/2020).

Consider a vector field Z on a curve $\alpha : I \to M$. Z smoothly assigns to each $t \in I$ a tangent vector to M at $\alpha(t)$. For example, the velocity α' is a vector field on α. The set $\mathfrak{X}(M)$ of all smooth vector fields on α is a module over $\mathcal{F}(I)$. When M is a semi-Riemannian manifold there is a natural way to define the vector rate of change Z' of a vector field $Z \in \mathfrak{X}(\alpha)$:

Proposition 3.1.5

Let $\alpha : I \to M$ be a curve in a semi-Riemannian manifold M.
Then there exists a unique function $Z \to Z' \equiv \frac{DZ}{dt}$ from $\mathfrak{X}(\alpha)$ to $\mathfrak{X}(\alpha)$, called the *induced covariant derivative* such that

1. $(aZ_1 + bZ_2)' = aZ_1' + bZ_2'$ $(a, b \in \mathbb{R})$
2. $(hZ)' = \left(\frac{dh}{dt}\right) Z + hZ'$ $(h \in \mathcal{F}(I))$
3. $(V_\alpha)'(t) = D_{\alpha'(t)}V$ $(t \in I, V \in \mathfrak{X}(M))$
4. $\left(\frac{d}{dt}\right) \langle Z_1, Z_2\rangle = \langle Z_1', Z_2\rangle + \langle Z_1, Z_2'\rangle$.

Introducing Christoffel symbols in its local form yields

$$Z' = \left(\frac{dZ^k}{dt} + \Gamma^k_{ij}\frac{d(x^i \circ \alpha)}{dt} Z^j\right) \partial_k.$$

If $Z' = 0$, then Z is parallel. This is equivalent to a system of linear first order differential equations. With the existence and uniqueness theorem we obtain:

Proposition 3.1.6

For a curve $\alpha : I \to M$, let $a \in I$ and $z \in T_{\alpha(a)}M$. Then there is a unique parallel vector field Z on α such that $Z(a) = z$.

If $b \in I$ then the function $P = P_a^b(\alpha) : T_pM \to T_qM$ sending each z to $Z(b)$ is called *parallel translation* along α from $p = \alpha(a)$ to $q = \alpha(b)$.

Thus, we found a way to compare two tangent spaces of distinct points on a manifold and hence know how to measure the failure or success of vector to retain its initial value after parallel transport on a closed curve. Note that parallel translation from p to q depends on the particular curve joining the two points.

3.1.3.4 Geodesics

Heuristically speaking, geodesics are the 'straightest possible lines' on a curved surface. They only bend the way they need to, due to the curvature of the manifold. In principle, there are some different approaches to geodesics, which nonetheless give the same notion. The more physically focused approach would be via Lagrange functions, but we are choosing the mathematical way:

A *geodesic* in a semi-Riemannian manifold M is a curve $\gamma : I \to M$ whose vector field γ' is parallel. That is, geodesics are the curves of acceleration zero, i.e. $\gamma'' = 0$:

Corollary 3.1.1

Let $x^1, ..., x^n$ be a chart on $\mathcal{U} \subset M$. A curve γ in $\mathcal{U}$ is a geodesic of M if and only if its coordinate functions $x^k \circ \gamma$ satisfy

$$\frac{d^2(x^k \circ \gamma)}{dt^2} + \Gamma_{ij}^k(y) \frac{d(x^i \circ \gamma)}{dt} \frac{d(x^j \circ \gamma)}{dt} = 0 \ (1 \leq k \leq n).$$

These expressions are the components of γ'' relative to the coordinate vector fields $\partial_1, ..., \partial_n$. Basic ODE-theory gives the following:

Lemma 3.1.5

If $v \in T_pM$ there exists an intervall I about 0 and a unique geodesic $\gamma : I \to M$ such that $\gamma'(0) = v$.

The last equation implies $\gamma(0) = p$. In words, there is always a geodesic γ starting at p with initial velocity v. There is even a maximal one:

Proposition 3.1.7

Given any tangent vector $v \in T_pM$ there is a unique geodesic γ_v in M such that

1. the initial velocity of γ_v is v, i.e. $\gamma_v'(0) = v$
2. the domain I_v of γ_v is the largest possible. Hence, if $\alpha : J \to M$ is a geodesic with initial velocity v, then $J \subset I$ and $\alpha = \gamma_v|_J$.

Because of 2. the geodesic γ_v is said to be maximal or geodesically inextendible. In general, geodesics need not be defined for all values of the parameter, i.e. $I_\gamma \neq \mathbb{R}$. However, if that is the case for a semi-Riemannian manifold M, it is said to be *geodesically complete*.

Lemma 3.1.6

Let $\gamma : I \to M$ be a non constant geodesic. A reparametrization $\gamma \circ h : J \to M$ is a geodesic if and only if h has the form $h(t) = at + b$.

A second order differential equation can be converted into a pair of first order differential equations by taking γ' as new variable . Similarly a geodesic in M can be represented by integral curves in the *tangent bundle* TM, which is the union of all T_pM.

Proposition 3.1.8

There is vector field G on TM such that the projection $\pi : TM \to M$ establishes a one-to-one correspondence between integral curves of G and geodesics of M.

3.1.3.5 The Exponential Map

At each point p of a semi-Riemannian manifold M, we collect the geodesics starting at p into a single mapping:

Definition 3.1.28

If $p \in M$, let $\mathcal{D}_p$ be the set of vectors $v \in T_pM$ such that the inextendible geodesic γ_v is defined at least on $[0, 1]$. The *exponential map* of M at p is the function $\exp_p : \mathcal{D}_p \to M$ such that

$$\exp_p(v) = \gamma_v(1) \quad \forall v \in \mathcal{D}_p.$$

Obviously, $\mathcal{D}_p$ is the largest subset of T_pM on which $\exp_p$ can be defined. If M is complete, then $\mathcal{D}_p = T_pM$ for all $p \in M$.

Fixing $v \in T_pM$ and $t \in \mathbb{R}$, the geodesic $s \mapsto \gamma_v(ts)$ has initial velocity $t\gamma_v'(0) = tv$. Hence, $\gamma_{tv}(s) = \gamma_v(ts)$ for all s, t such that either side is well defined. In particular, if $v \in \mathcal{D}_p$ then

$$\exp_p(tv) = \gamma_{tv}(1) = \gamma_v(t).$$

Thus, the exponential map carries lines through the origin of T_pM to geodesics of M through p. We are now going to use the exponential map on a set of nice regions on a manifold, called starshaped.

Proposition 3.1.9

For each $p \in M$ there exists a neighbourhood $\tilde{\mathcal{U}}$ of 0 in T_pM on which the exponential map $\exp_p$ is a diffeomorphism onto a neighbourhood $\mathcal{U}$ of p in M.

A subset S of a vector space is *starshaped* about 0 if $v \in S$ implies $tv \in S$ forall $0 \leq t \leq 1$. If $\mathcal{U}$ and $\tilde{\mathcal{U}}$ are as in the preceding proposition and $\tilde{\mathcal{U}}$ is starshaped about 0, then $\mathcal{U}$ is called a *normal neighbourhood* of p. As a matter of fact, each point $p \in M$ has a normal neighbourhood:

Proposition 3.1.10

If $\mathcal{U}$ is a normal neighbourhood of $p \in M$, then $\forall q \in \mathcal{U}$ there is a unique geodesic $\sigma : [0,1] \to \mathcal{U}$ from p to q in $\mathcal{U}$. Furthermore, $\sigma'(p) = \exp^{-1}(q) \in \tilde{\mathcal{U}}$.

On any normal neighbourhood $\mathcal{U}$ of $p \in M$ there is a special type of chart that is particularly simple. Let $e_1, ..., e_n$ be an orthonormal basis for T_pM, so $\langle e_i, e_j \rangle = \delta_{ij} e_j$. The normal chart $\xi = (x^1, ..., x^n)$ determined by $e_1, ..., e_n$ assigns to each $q \in \mathcal{U}$ the vector coordinates relative to $e_1, ..., e_n$ of the corresponding point $\exp_p^{-1}(q) \in \tilde{\mathcal{U}} \subset T_pM$.
In short,

$$\exp_p^{-1}(q) = x^i(q) e_i \quad (q \in \mathcal{U}).$$

Hence, if $f^1, ..., f^n$ is the dual basis to $e_1, ..., e_n$ then $x^i \circ \exp_p = f^i$ on $\tilde{\mathcal{U}}$. We also have the obvious consequence:

Proposition 3.1.11

If $x^1, ..., x^n$ is a normal chart at $p \in M$ then $\forall i, j, k$

1. $g_{ij}(p) = \delta_{ij} \epsilon_j$
2. $\Gamma^k_{ij}(p) = 0$.

At a point p (in general not elsewhere) the metric tensor and the corresponding Christoffel symbols of a normal chart are Minkowskian. Since tensors are pointwise creatures this is very powerful in computations. The closer we get to p the more M resembles $T_pM \approx \mathbb{R}^n_\nu$. Note, that this cannot be pushed too far.

3.2 Curvature and Einstein Equations

Now, that we know how to parallel transport a vector, we can ask, whether it retains its initial value after translation along a closed curve. This gives an intrinsic notion (see for example [4, p. 160ff]) of the curvature of a manifold and can be used to calculate the so called *Riemannian curvature tensor*, which measures the commutativity of second order differentiations:

Lemma 3.2.1

Let M be a semi-Riemannian manifold with Levi-Civita connection D.
The function $R : \mathfrak{X}(M)^3 \to \mathfrak{X}(M)$ given by

$$R_{XY}Z = D_{[X,Y]}Z - [D_X, D_Y]Z$$

is a $(1,3)$-tensor field on M, called the *Riemannian curvature tensor* of M.

R can be considered as an $\mathbb{R}$-multilinear function on individual tangent vectors. If $x, y \in T_pM$, the linear operater $R_{xy} : T_pM \to T_pM$ sending each z to $R_{xy}z$ is called a curvature operator. The Riemmannian tensor R has following symmetry properties:

Proposition 3.2.1

If $x, y, z, v, w \in T_pM$, then

1. $R_{xy} = -R_{yx}$
2. $\langle R_{xy}v, w\rangle = -\langle R_{xy}w, v\rangle$
3. $R_{xy}z + R_{yz}x + R_{zx}y = 0$ (First Bianchi identity)
4. $\langle R_{xy}v, w\rangle = \langle R_{vw}x, y\rangle$ (Symmetry by pairs).

These lead to a less obvious symmetry property, namely the second Bianchi identity. It concerns itself with the covariant differential DR, a (1,4) tensor field assigning the vector field value $(D_Z R)_{XY}V = (D_Z R)(X,Y)V$ to four vector fields.

Proposition 3.2.2

If $x, y, z \in T_pM$, then $(D_z R)(x, y) + (D_x R)(y, z) + (D_y R)(z, x) = 0$.

The local form of the Riemannian tensor amounts to:

Lemma 3.2.2

On the coordinate neighbourhood of a chart $x^1, ..., x^n$: $R_{\partial_k \partial_l}(\partial_j) = R^i{}_{jkl}\partial_i$,
where the components of R are given by

$$R^i_{jkl} = \Gamma^i_{kj,l} - \Gamma^i_{lj,k} + \Gamma^i_{lm}\Gamma^m_{kj} - \Gamma^i_{km}\Gamma^m_{lj}.$$

A semi-Riemannian manifold M for which the curvature tensor R is zero at any point is called flat. For example every semi-Euclidean space $\mathbb{R}^n_\nu$ is flat: for natural coordinates the Christoffel symbols vanish, hence $R = 0$ by the above Lemma.
Contracting the Riemannian curvature tensor gives:

Definition 3.2.1

Let R be the Riemannian curvature tensor of M. The Ricci curvature tensor Ric of M is the contraction $C^1_3(R) \in \mathcal{T}^0_2(M)$.

Its components relative to a chart are $R_{ij} = R^m_{ijm}$.

Because of the symmetries of R the only nonzero contractions of R are $\pm Ric$. If its Ricci tensor is zero, M is called Ricci flat. A flat manifold is certainly Ricci flat but the converse is not true. Further contracting the Ricci tensor gives:

Definition 3.2.2

The scalar curvature S of M is the contraction $C(Ric) \in \mathcal{F}(M)$ of its Ricci tensor.

In coordinates we have $S = g^{ij}R_{ij} = g^{ij}R^k_{ijk}$.

The remaining question is, how can we relate the curvature of space-time to the mass-energy distribution of the universe? The latter can be described with the stress-energy tensor T_{ij}. It is a symmetric and covariantly constant, i.e. $\nabla_k T_{ij} = 0$, $(0, 2)$-tensor field, so we are looking for a quantity that can be produced out of the Riemannian curvature tensor with the same properties. The Ricci tensor would be a natural candidate. Indeed, the first equation Einstein originally postulated was $R_{ij} = 4\pi T_{ij}$ [29]. But the Ricci tensor is not covariantly constant and $\operatorname{div} Ric = \frac{1}{2}S$. Subtracting half of the scalar curvature from the divergence of the Ricci tensor therefore produces a promising candidate. And indeed, $G = Ric - \frac{1}{2}Sg$ is covariantly constant as desired. The Einstein equations then amount to:

Definition 3.2.3

If M is a spacetime containing matter with stress-energy tensor T, then

$$G = Ric - \frac{1}{2}Sg = 8\pi T, \tag{10}$$

where G is called the Einstein gravitational tensor.

The Einstein equations (10) form a system of 10 coupled, non-linear partial differential equations and describe how the presence of matter and energy encoded in T_{ij} relates to the curvature of spacetime encoded in $G_{ij} = R_{ij} - \frac{1}{2}Sg_{ij}$. Generally these equations have to be solved for both sides, but in the following we will work with so called vacuum solutions outside of gravitational sources, i.e. where $T_{ij} = 0$.

Next, we will discuss different forms of gravitational wave metrics that are solutions for (10) and show some methods for construction.

4 Gravitational Waves

4.1 Introduction – Solving the Linearised Einstein Equations

Intuitively speaking, gravitational waves can be seen as small scale ripples in the curvature of spacetime propagating with the speed of light. To obtain a vivid picture just imagine water waves propagating on the surface of a pond [17, p. 924].

Figure 10: Water waves on the surface of a lake. Caltech/MIT/LIGO Laboratory. Taken from https://www.ligo.caltech.edu/page/gravitational-waves (visited on 17/02/2020).

Theory predicts that gravitational waves originate from processes involving strong and rapidly changing gravitational fields, i.e. sources such as binary neutron stars or black hole mergers, supernovae or the gravitational collapse of a massive star.

In fact, in 1974, Russel A. Hulse, at that time Professor at the University of Massachusetts, and his research student Joseph H. Taylor were the first to discover a binary pulsar, i.e. a rotating neutron star emitting electromagnetic radiation, in this case with another neutron star as companion. Due to the big mass of the system, its movement differs strongly from Newton's law of gravity and so is a perfect test object for Einstein's theory. After a long-term observation, Hulse and Taylor found that the orbit of the bodies is gradually decreasing. The system loses energy due to emission of gravitational waves just as predicted. Measurements agreed well with calculations for the orbital decay. For this first indirect observation of gravitational waves Hulse and Taylor won the 1993 Nobel Prize in Physics [24].

More spectacularly, in September 2015, almost exactly 100 years after the introduction of General Relativity, gravitational waves were directly observed by the large scale gravitational wave detectors of LIGO, short for *Laser Interferometer Gravitational-Wave Observatory* [2]. Observing the wave pattern, scientists of LIGO were able to make theoretical predictions about the cause and origin of the wave (in case of the first observation a black hole merger), masses of the two black holes before and of the resulting single black hole after the collision, distance and roughly the direction of the event.

After that, in December 2015, a second observation was made [1], followed by a third in January 2017. With several more observatories around the world taking up their work in the near future, the number and accuracy of the measurements will possibly rise. For this discovery the leading scientists of LIGO were awarded with the Nobel Prize in 2017 [25].

Figure 11 shows the data from one event observed by three observatories. The top graph shows the Signal to Noise Ratio - or roughly how significant the detection was. The middle graph shows the ”chirp” or frequency shift across time. The bottom graph shows the wave form.

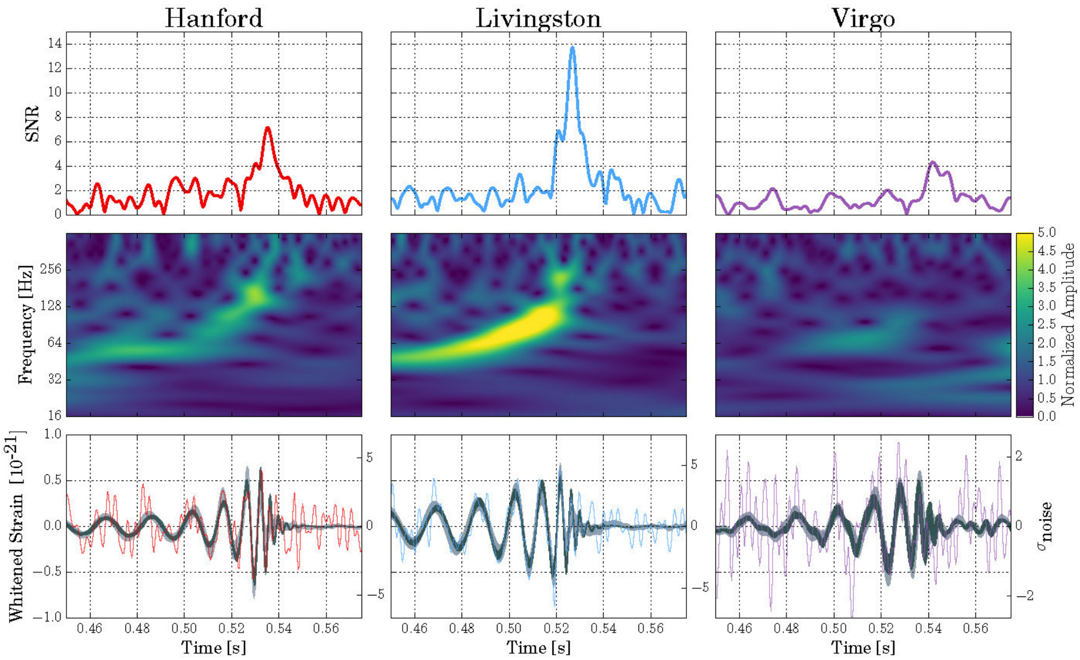

Figure 11: Observation of a gravitational wave event. Caltech/MIT/LIGO Laboratory. Taken from https://www.ligo.caltech.edu/WA/image/ligo20170927a (visited on 17/02/2020).

Though nowadays the existence of gravitational waves seems well established, when they were first considered in theory, the situation was a different one:

Immediately after formulating the gravitational field equations

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} S g_{\mu\nu} = \kappa T_{\mu\nu}, \tag{11}$$

Einstein himself went on to linearise them by assuming, the metric $g_{\mu\nu}$ has the form of a slight perturbation of flat Minkowski spacetime $\eta_{\mu\nu}$, i.e.

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}. \tag{12}$$

Here $h_{\mu\nu}$ is supposed to be 'very small' ($|h_{\mu\nu}| \ll 1$).

Using $\partial_\rho g_{\mu\nu} = \partial_\rho \eta_{\mu\nu} + \partial_\rho h_{\mu\nu} = \partial_\rho h_{\mu\nu}$, we obtain the Christoffel symbols, Riemannian and Ricci tensor and Ricci scalar to be

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2} g^{\rho\sigma}(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}) = \frac{1}{2}\eta^{\rho\sigma}(\partial_\mu h_{\nu\sigma} + \partial_\nu h_{\mu\sigma} - \partial_\sigma h_{\mu\nu}) + \mathcal{O}(h^2) \tag{13}$$

$$R_{\mu\nu\rho\sigma} = \eta_{\mu\tau} R^\tau{}_{\nu\rho\sigma} = \frac{1}{2}(\partial_\rho\partial_\nu h_{\sigma\mu} + \partial_\sigma\partial_\mu h_{\nu\rho} - \partial_\sigma\partial_\nu h_{\mu\rho} - \partial_\rho\partial_\mu h_{\nu\sigma}) \tag{14}$$

$$R_{\nu\sigma} = R^\mu{}_{\nu\mu\sigma} = \frac{1}{2}(\partial^\mu\partial_\nu h_{\sigma\mu} + \partial^\mu\partial_\sigma h_{\nu\mu} - \partial_\sigma\partial_\nu h - \Box h_{\nu\sigma}) \tag{15}$$

$$S = R^\nu{}_\nu = \partial^\mu\partial^\nu h_{\mu\nu} - \Box h, \tag{16}$$

where $h = \eta_{\mu\nu} h^\mu h^\nu$ and $\Box = \eta_{\mu\nu}\partial^\mu\partial^\nu$ is the d'Alembert operator.

The equations can be simplified using harmonic coordinates which satisfy the conditions $\partial^\beta h_{\alpha\beta} = \frac{1}{2}\partial_\alpha h$:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} S = \frac{1}{2}\Box(\frac{1}{2}\eta_{\mu\nu} h - h_{\mu\nu}) = \kappa T_{\mu\nu}. \tag{17}$$

Equations (17) are known as *linearised Einstein equations* and are most transparently written in the variable $\bar{h}_{\mu\nu} = \frac{1}{2}\eta_{\mu\nu} h - h_{\mu\nu}$:

$$\boxed{\Box\bar{h}_{\mu\nu} = 2\kappa T_{\mu\nu}.}$$

Outside the source of the gravitational field, i.e. where $T_{\mu\nu} = 0$, they form a system of decoupled relativistic wave equations. This tells us (e.g. by comparison with electrodynamics) that *linearised* general relativity admits solutions in which the perturbations $h_{\mu\nu}$ of Minkowski spacetime are *plane waves* travelling with the speed of light. This supports the initial heuristic comparison with water waves travelling on a pond.

But Einstein was concerned, whether those wavelike solutions had any real physical meaning or merely were results of the linearisation. Only if the full non linear theory admits for such solutions as well, gravitational waves would be 'real'. With help from Rosen he soon found some candidates but they were discarded as unphysical because the contained singularities. Only shortly after in 1924, Brinkmann was able to show these unwanted singularities were due to an inconvenient coordinate choice. However, it took several decades, simply because the physicist community at that time did not know the work of Brinkmann, until the intricate question of how to properly define gravitational waves in the full theory could be resolved. Main contributors to that solution were Herman Bondi, Felix Pirani, Ivor Robinson, Andrzej Trautman, and others in the mid 1950s. For more details see [7].

The spacetime metric Brinkmann found in 1924 was named pp-*wave* metric by Ehlers and Kundt in 1962 because of its geometrical properties, specifically such gravitational waves have plane wave fronts and parallel rays. Also, these spacetimes already contain plane waves as a subclass, as we will see in the next chapter.

4.2 The Class of pp-Waves

In this section we will formally introduce a set of spacetime metrics, i.e. exact solutions of the full non linear Einstein equations (11), modelling gravitational waves. These are called *plane fronted* waves with *parallel rays*, pp-waves for short. The same class that was already discovered by Brinkmann. The material is quite standard and we roughly follow the approach of [5] and [13, p. 323ff].

pp-Waves may represent gravitational waves, electromagnetic waves or any combinations. It is common practise to *define* pp-wave metrics geometrically by the property that they allow for a (non vanishing) covariantly constant null vector field. In the end, this is simply a geometric transcription of plane fronted and parallel rays.

Definition 4.2.1

A Riemannian manifold (M, g) is called a pp-wave spacetime if it admits a covariantly constant null vector field, i.e. a vector field $\boldsymbol{k} \in \mathfrak{X}(M)$ with

$$k_\mu k^\mu = 0 \quad \text{and} \quad \nabla \boldsymbol{k} = k^\mu{}_{;\nu} = 0.$$

Let (N, h) be a smooth connected two dimensional Riemannian manifold. We consider the spacetime $(M = N \times \mathbb{R}^2, g)$ where $\boldsymbol{x} = x^i = (x^1, x^2)$ are coordinates on N, u and r are light cone coordinates on $\mathbb{R}^2$ as in (6) and the functions $H : N \times \mathbb{R} \to \mathbb{R}$ and $A_i : N \times \mathbb{R} \to \mathbb{R}$ are smooth.

The spacetime given by the line element

$$ds^2 = -2dudr + H(u, \boldsymbol{x})du^2 + 2A_i(u, \boldsymbol{x})dx^i du + h_{ij}(u, \boldsymbol{x})dx^i dx^j \tag{18}$$

possesses such a covariantly constant vector, namely $\partial_r = (1, 0, 0, 0)$.

There are no dr^2-terms in (18) and the components of the u and x^i coordinates in ∂_r are equal to zero, so

$$k_\mu k^\mu = g_{\mu\nu} k^\mu k^\nu = g_{rr} k^r k^r = 0, \tag{19}$$

and ∂_r is null as required.

To calculate the covariant derivative of $\boldsymbol{k}$, recall the formula in chapter 3.1.3.2

$$k^{\mu}{}_{;\nu} = \partial_\nu k^\mu + \Gamma^\mu_{\nu\alpha} k^\alpha. \tag{20}$$

Clearly, the only Christoffel symbols of interest are the ones with lower index r

$$\Gamma^\mu_{\nu r} = \frac{1}{2} g^{\mu\sigma}(g_{r\sigma,\nu} + g_{\nu\sigma,r} - g_{\nu r,\sigma}), \tag{21}$$

because the others vanish, when multiplied with the components of ∂_r that are equal to zero.

Looking at the metric coefficients of (18)

$$g_{\mu\nu} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & H & A_i \\ 0 & A_i & h_{ij} \end{pmatrix}, \tag{22}$$

we see that all of them are independent of r, so $g_{\nu\sigma,r} = 0$. Additionally, the coefficients in the first row and in the first column are constant, so $g_{r\sigma,\nu} = 0$ and $g_{\nu r,\sigma} = 0$. Hence, the corresponding Christoffel symbols vanish and we have

$$k^{\mu}{}_{;\nu} = 0 + \Gamma^\mu_{\nu r} k^r = 0. \tag{23}$$

So, k is covariantly constant and (18) is indeed a pp-wave.

Conversely, any pp-wave spacetime is of this form. More precisely, we have:

Proposition 4.2.1

A Lorentzian manifold is a pp-wave spacetime, if and only if it allows for coordinates in which the metric takes the form (18).

Proof:
It only remains to show the only if part.

Let (N, h) be a smooth connected two dimensional Riemannian manifold. We look for a spacetime $(M = N \times \mathbb{R}^2, g)$ admitting a covariantly constant nowhere vanishing null vector as in definition 4.2.1.

Let $\boldsymbol{k} \neq 0$ everywhere be such a vector. The condition $k^\mu;_\nu = 0$ is then equivalent to the two conditions

$$\begin{aligned} k_{\mu;\nu} + k_{\nu;\mu} &= 0 \\ k_{\mu;\nu} - k_{\nu;\mu} &= 0. \end{aligned} \tag{24}$$

To verify this, we go to a normal coordinate system around a point $p \in M$, where $\Gamma^\mu_{\nu\sigma}(p) = 0$ (see chapter 3.1.3.5). This simplifies equations (24) to

$$
\begin{aligned}
k_{\mu,\nu} + k_{\nu,\mu} &= 0 \\
k_{\mu,\nu} - k_{\nu,\mu} &= 0.
\end{aligned}
\tag{25}
$$

In normal coordinates, this is equivalent to $k_{\mu;\nu} = 0$. Shoving up the first index with the inverse metric we again have $k^{\mu}{}_{;\nu} = 0$.

The first equation in (24) is a so-called *Killing equation*[3], which implies that the metric components are independent of the coordinate r, i.e. $g_{\mu\nu,r} = 0$.

The second tells us that $\boldsymbol{k}$ is locally a gradient vector field, therefore we can take it locally to be $\boldsymbol{k} = \partial_r$ for some coordinate r.

This amounts to choosing a parameter along the integral curves of $\boldsymbol{k}$ as our coordinate. In components we have $\boldsymbol{k} = \delta^{\mu}_{r}$. This gives $k_{\mu} = g_{\mu\alpha}\delta^{\alpha}_{r} = g_{\mu r}$. Since $\boldsymbol{k}$ is null, we have

$$\boxed{k_r = g_{rr} = 0.}$$

The second equation of (25) $k_{\mu,\nu} - k_{\nu,\mu} = 0$ further implies that locally we can write $k_{\mu} = g_{r\mu} = u_{,\mu}$ with u being a function $u = u(x^{\mu})$. This gives

$$\boxed{g_{ur} = g_{ru} = u_{,u} = 1 \text{ and } g_{ri} = g_{ir} = u_{,i} = 0.}$$

Summing up and changing from the x^{μ}-coordinates to (u, r, x^i) as in (18) we obtain the line element of a spacetime metric admitting a covariant constant null vector to be

$$
\begin{aligned}
ds^2 &= g_{\mu\nu} dx^{\mu} dx^{\nu} \\
&= -2dudr + g_{uu}(u, \boldsymbol{x})du^2 + 2g_{iu}(u, \boldsymbol{x})dx^i du + g_{ij}(u, \boldsymbol{x})dx^i dx^j \\
&= -2dudr + H(u, \boldsymbol{x})du^2 + 2A_i(u, \boldsymbol{x})dx^i du + g_{ij}(u, \boldsymbol{x})dx^i dx^j,
\end{aligned}
\tag{26}
$$

where $H : N \times \mathbb{R} \to \mathbb{R}$ and $A_i : N \times \mathbb{R} \to \mathbb{R}$ are smooth functions.

$\square$

In the most prominent subclass of this family, we take the wave surface N to be flat $\mathbb{R}^2$, i.e. $g_{ij} = \delta_{ij}$.

Another transformation lets the off-diagonal terms A_i in (26) vanish. This simplification is possible only locally though and ignores the global properties of such spacetimes. The possible internal rotation of the source of gravitational waves, nammely its spin, is lost. These localized spinning sources are called *gyratons*, as physicists rediscovered such spacetimes as a possible model of a particle. Read more about pp-wave-type qyratons in [23], as we will continue without the

[3] A vector field for which the Lie-derivative of the metric tensor vanishes.

off-diagonal terms.

Keeping H arbitrary, we obtain the pp-wave metric in so-called Brinkmann coordinates:

$$ds^2 = -2dudr + H(u, \boldsymbol{x})du^2 + d\boldsymbol{x}^2. \tag{27}$$

Primary curvature quantities of such spactimes with no off-diagonal terms and general H are

$$g_{\mu\nu} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & H & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad g^{\mu\nu} = \begin{pmatrix} -H & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{28}$$

$$\begin{gathered} \Gamma^r_{uu} = -\frac{1}{2}H_{,u} \quad \Gamma^\alpha_{uu} = -\frac{1}{2}H_{,\alpha} \\ \Gamma^r_{\alpha u} = \Gamma^r_{u\alpha} = -\frac{1}{2}H_{,\alpha} \end{gathered} \tag{29}$$

$$R^\alpha{}_{u\alpha u} = \frac{1}{2}H_{,\alpha\alpha} \quad R^i{}_{uju} = \frac{1}{2}H_{,ij} \quad R^r{}_{iju} = \frac{1}{2}H_{,ij} \tag{30}$$

$$R_{uu} = -\frac{1}{2}H_{,ii} - \frac{1}{2}H_{,jj} = -\frac{1}{2}\Delta H, \tag{31}$$

for $\alpha = i, j$. Note, if $H(\boldsymbol{x}, u)$ vanishes the line element (27) reduces to that of Minkowski space. Just as the linearised theory predicts (confer section 4.1), for small H, these functions can be seen as small perturbations of flat space. For general H the corresponding spacetimes may represent strong gravitational waves with arbitrary profiles.

First, we will discuss a subclass of the line element (27), which are called *plane waves*. Taking the function $H(\boldsymbol{x}, u)$ to be quadratic in $\boldsymbol{x}$, expressed by a symmetric (2×2)-matrix-valued function on $\mathbb{R}$, $H_{ij}(u)$, we obtain:

$$ds^2 = -2dudr + H_{ij}(u)x^i x^j du^2 + d\boldsymbol{x}^2. \tag{32}$$

Equation (32) is the line element of a *plane wave metric*. Note, that $H_{ij}(u)$ is now solely dependent on u. The curvature tensor components of such metrics are constant over each wave surface $u = const.$

Again, we calculate the curvature invariants for the metric, using the symmetry of H_{ij}:

$$g_{\mu\nu} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & H_{ij}x^i x^j & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad g^{\mu\nu} = \begin{pmatrix} -H_{ij}x^i x^j & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{33}$$

$$\begin{gathered} \Gamma^r_{uu} = -\frac{1}{2}H_{ij,u}x^i x^j \quad \Gamma^\alpha_{uu} = -H_{\alpha j}x^j \\ \Gamma^r_{\alpha u} = \Gamma^r_{u\alpha} = -H_{\alpha j}x^j \end{gathered} \tag{34}$$

$$R^\alpha{}_{u\alpha u} = H_{\alpha\alpha} \quad R^i{}_{uju} = H_{ij} \quad R^r{}_{iju} = H_{ij} \tag{35}$$

$$R_{uu} = -H_{ii} - H_{jj} = -\operatorname{Tr} H, \tag{36}$$

where $\alpha = i, j$ and we used $\partial_\alpha H_{ij}x^i x^j = 2H_{\alpha j}x^j$.

4.3 Three Approaches to Impulsive pp-Waves

Next, we will investigate the so-called *impulsive* case of gravitational waves. Recall, that the curvature tensor components (30) of pp-waves in the Brinkmann form (27) do not depend on derivatives of H in u. This allows for discontinuities in said function and even admits distributional dependencies. Hence, we can take $H(\boldsymbol{x}, u) = f(\boldsymbol{x})\delta(u)$. In other words, we let H be some multiple of $\delta(u)$ the Dirac distribution. The resulting line element in this rather naive approach is

$$ds^2 = -2dudr + f(\boldsymbol{x})\delta(u)du^2 + d\boldsymbol{x}^2 \tag{37}$$

and gives the metric of so called *impulsive gravitational waves*.

Intuitively these can be seen as gravitational waves carrying pulses of curvature across the hypersurface, $u = const$. The key curvature quantities of such metrics are ($\alpha = i, j$):

$$\Gamma^r_{uu} = -\frac{1}{2}f\delta_{,u} \quad \Gamma^\alpha_{uu} = -\frac{1}{2}f_{,\alpha}\delta \quad \Gamma^r_{\alpha u} = -\frac{1}{2}f_{,\alpha}\delta \tag{38}$$

$$R^\alpha{}_{u\alpha u} = \frac{1}{2}f_{,\alpha\alpha}\delta \quad R^r{}_{iju} = \frac{1}{2}f_{,ij}\delta \tag{39}$$

$$R_{uu} = -\frac{1}{2}\Delta f\delta. \tag{40}$$

Over all, there are plenty of methods to construct such metrics, other then just exchanging the profile function H with some multiple of the Dirac distribution. We will take a closer look at three of them (for more details, see [22]). Note, that the ordering doesn't imply any hierarchy, whatsoever.

4.3.1 Distributional Limit of Sandwich Wanes

We will begin, discussing impulsive gravitational waves as distributional limits of so called *sandwich waves*. Recall, that we do not impose any continuity requirements on the function $H(u, \boldsymbol{x})$ in the coordinate u in (27) and that therefore it is possible to replace it with some multiple of the Heaviside step function $\Theta(u)$. Such, we obtain so called *shock* or *step waves*.

We now take the support of H to be compact in u, i.e. the function is non-trivial only for a finite period $u_0 \leq u \leq u_1$ of time. The regions behind and in front of the wavefront are represented by flat Minkowski half spaces $\mathcal{M}^-$ and $\mathcal{M}^+$ and the wave propagates along the null hypersurfaces $u = 0$ for $u_0 \leq u \leq u_1$. This gives so called *sandwich waves*, which we could imagine as shock waves squeezed into flat space (see figure 12 below).

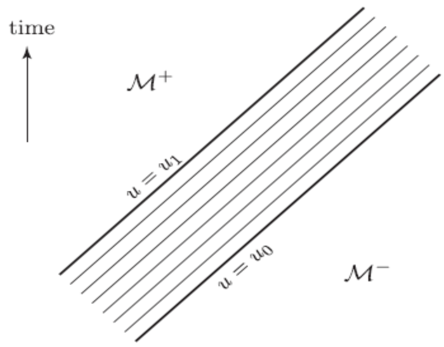

Figure 12: Sandwich wave. Taken from [13, p. 329].

We now take H to be proportional to a so called δ-sequence, i.e. $H_\epsilon = \rho_\epsilon(u)f(x)$, with $\rho_\epsilon(x) = \frac{1}{\epsilon}\rho(\frac{x}{\epsilon})$ and $\int \rho = 1$ and $\mathrm{supp}(\rho) \subseteq [-1, 1]$. Then $H_\epsilon \to f\delta$ in the distributional sense. Casually put, we 'squeeze the wave as hard as possible' and thereby obtain the line element for impuslive pp-waves in Brinkmann coordinates:

$$ds^2 = -2dudr + H(\boldsymbol{x})\delta(u)du^2 + d\boldsymbol{x}^2. \tag{41}$$

Interestingly, this method is intertwined with the one we are going to treat next, in the sense that using a convenient transformation we can construct the same line element (41).

4.3.2 Geometrical Construction – 'Cut and Paste'

The upcoming elegant geometrical approach to impulsive gravitational wave metrics was first discussed by Roger Penrose, for a reference see [21, p. 101-115]. He considered Minkowski spacetime, which is cut in two halves $\mathcal{M}^-(u < 0)$ and $\mathcal{M}^+(u > 0)$ by removing the null hyperplane $\mathcal{N}$ at $u = 0$. The two regions are then identified which each other by deploying a suitable *warp function*.

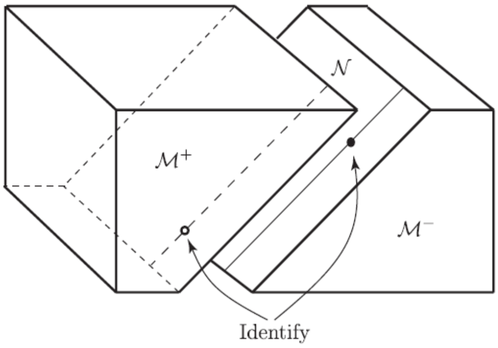

Figure 13: Cutting Minkowski spacetime in two halves and 'glueing' them back together shunted along the null hypersurface. Taken from [13, p. 394].

Penrose gives the following condition (employing null coordinates as in (6)) to re-attach the two parts

$$(u = 0, r, x, y)_{\mathcal{M}^-} = (u = 0, r + H(x, y), x, y)_{\mathcal{M}^+}, \tag{42}$$

where the warp function $H(x, y)$ is an arbitrary real function.

This jump in the coordinates results precisely in (41), as we will see below. In [21] Penrose gives an identification of the two halves of Minkowski spacetime but no explicit form of a transformation. Moreover he found 'remarkable' focussing properties (see [20]).

In the next step we want to obtain an explicit form of the metric (confer [13, p. 394-396]). It is convenient to work with complex coordinates ($\zeta = \frac{1}{\sqrt{2}}(x + iy)$), which brings the Minkowski line element in the form

$$ds^2 = -2dudr + 2d\zeta d\bar{\zeta}. \tag{43}$$

We now apply following coordinate transformation

$$u = U, \quad r = R - H + UH_{,Z}H_{,\bar{Z}}, \quad \zeta = Z - UH_{,\bar{Z}}, \tag{44}$$

where $H(Z, \bar{Z})$ is an arbitrary real function of Z and $\bar{Z}$.

We start with calculating the corresponding differentials to

$$\begin{aligned} du &= dU, \\ dr &= dR + H_{,Z} H_{,\bar{Z}} dU - H_{,Z} dZ - H_{,\bar{Z}} d\bar{Z} \\ &+ U H_{ZZ} H_{,\bar{Z}} dZ + U H_{,Z} H_{,\bar{Z}Z} dZ + U H_{,Z\bar{Z}} H_{,\bar{Z}} d\bar{Z} + U H_{,Z} H_{,\bar{Z}\bar{Z}} d\bar{Z} \\ d\zeta &= dZ - H_{,\bar{Z}} dU - U H_{,\bar{Z}Z} dZ - U H_{,\bar{Z}\bar{Z}} d\bar{Z}, \end{aligned} \tag{45}$$

and compute the products $2dudr$ and $2d\zeta d\bar{\zeta}$. Next, we substract the first from the latter:

$$\begin{aligned} 2d\zeta d\bar{\zeta} = 2[&dZ d\bar{Z} - U H_{,Z\bar{Z}} d\bar{Z} dZ - U H_{,ZZ} dZ dZ - U H_{\bar{Z}Z} dZ d\bar{Z} - U H_{,\bar{Z}\bar{Z}} d\bar{Z} d\bar{Z} \\ &+ U^2 H_{,\bar{Z}Z} H_{,Z\bar{Z}} d\bar{Z} dZ + U^2 H_{,\bar{Z}Z} H_{,ZZ} dZ dZ + U^2 H_{,\bar{Z}\bar{Z}} H_{,Z\bar{Z}} d\bar{Z} d\bar{Z} \\ &+ U^2 H_{\bar{Z}\bar{Z}} H_{,ZZ} dZ d\bar{Z} \\ &+ H_{,Z} H_{,\bar{Z}} dU^2 - H_{,Z} dZ dU - H_{,\bar{Z}} d\bar{Z} dU \\ &+ U H_{,\bar{Z}} H_{,ZZ} dZ dU + U H_{,\bar{Z}} H_{,Z\bar{Z}} d\bar{Z} dU + U H_{,Z} H_{,\bar{Z}Z} dZ dU + U H_{,Z} H_{,\bar{Z}\bar{Z}} d\bar{Z} dU]. \end{aligned} \tag{46}$$

Terms containing dUs all cancel with the corresponding terms generated by multiplying dr by dU. What is left is

$$\begin{aligned} ds^2 &= -2dUdR + 2dZd\bar{Z} \\ &-2U[d\bar{Z}(H_{,Z\bar{Z}} dZ + H_{,\bar{Z}\bar{Z}} d\bar{Z}) + dZ(H_{,ZZ} dZ + H_{,\bar{Z}Z} d\bar{Z})] \\ &+2U^2[(H_{,ZZ} dZ + H_{,Z\bar{Z}} d\bar{Z})(H_{,\bar{Z}\bar{Z}} d\bar{Z} + H_{,\bar{Z}Z} dZ)]. \end{aligned} \tag{47}$$

which, using the complex absolute value $|Z|^2 = Z\bar{Z}$ can conveniently be written as

$$\boxed{ds^2 = -2dUdR + 2|dZ - U(H_{,Z\bar{Z}} dZ + H_{,\bar{Z}\bar{Z}} d\bar{Z})|^2.} \tag{48}$$

This is the line element generated by using transformation (44) on Minkowski space in complex coordinates (43).

Evidently $u = 0$ (or $U = 0$ respectively) is a common null hyperplane for the metrics (43) and (48). Following the 'cut and paste' idea, we can take one of those on each side of the hyperplane. It is possible to combine these two, by introducing the so-called *kink* or *ramp* function $U \cdot \Theta(U)$ (see figure 14 below)

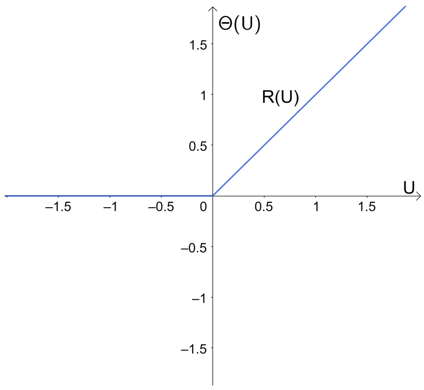

Figure 14: $R(U) = U \cdot \Theta(U)$, the so called ramp function.

and performing a trivial re-parametrisation ($u = U$, $r = R$ and $\zeta = Z$) for $U > 0$:

$$ds^2 = -2dUdR + 2|dZ - U\Theta(U)(H_{,Z\bar{Z}}dZ + H_{,\bar{Z}\bar{Z}}d\bar{Z})|^2. \tag{49}$$

The combination of these transformation can be expressed in the explicit combined form

$$u = U, \quad r = R - H\Theta(U) + U\Theta(U)H_{,Z}H_{,\bar{Z}}, \quad \zeta = Z - U\Theta(U)H_{,\bar{Z}}. \tag{50}$$

Tansformation (50) is clearly discontinuous for $u = 0$ and such that

$$(u = 0, r, \zeta, \bar{\zeta})_{\mathcal{M}^-} = (u = 0, r + H(\zeta, \bar{\zeta}), \zeta, \bar{\zeta})_{\mathcal{M}^+}, \tag{51}$$

which are precisely the *Penrose junction conditions* in complex coordinates (confer with (42)) for re-attaching the two halves with the warp function $H(\zeta, \bar{\zeta})$.

Using the properties $[U \cdot \Theta(U)]' = \Theta(U)$ and $\Theta^2(U) = \Theta(U)$ of the kink-function, we see that the only additional term generated when computing the differentials compared with (44) is $-H\delta(U)dU = -H(\zeta, \bar{\zeta})\delta(u)du$. Besides that the calculation is the same and therefore all other terms cancel and we obtain again the line element for an impulsive gravitational wave

$$\boxed{ds^2 = -2dudr - 2H(\zeta, \bar{\zeta})\delta(u)du^2 + 2d\zeta d\bar{\zeta}. \tag{52}}$$

4.3.3 Limit of a Suitable Solution

The metric for impulsive gravitational waves first arose out of a purely mathematical context in the work of Brinkmann in 1925 [6]. The first derivation in a physical context was given by Aichelburg and Sexl in 1971 in [3]. Applying a Lorentz boost to the so called Schwarzschild metric, which describes the gravitational field of a spherically symmetric mass, and then taking the ultra-relativistic limit, where c tends to one, yields an impulsive gravitational wave metric. The following section will give a discussion of their approach, referring also to [27].

The metric of a particle at rest with mass m in spherical coordinates (t, ρ, θ, ϕ) is given by the so-called Schwarzschild solution

$$ds^2 = -(1 - \frac{2m}{\rho})dt^2 + (1 - \frac{2m}{\rho})^{-1}d\rho^2 + \rho^2 d\Omega^2, \tag{53}$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$ (see for example [29, p. 124]). Because (53) is spherically symmetric we can employ isotropic coordinates. After such a transformation the line element will be of the form

$$ds^2 = -F(r)dt^2 + G(r)(dx^2 + dy^2 + dz^2), \tag{54}$$

where F and G are arbitrary functions of the new radial coordinate r and which we will now explicitly calculate:

Introducing the coordinate $\rho(r) = r(1 + \frac{m}{2r})^2$, its differential $d\rho = (1 - \frac{m^2}{4r^2})dr$ and the factor $A = \frac{m}{2r}$ we obtain

$$\begin{aligned} ds^2 &= -\frac{(1-\frac{m}{2r})^2}{(1+\frac{m}{2r})^2}dt^2 + \frac{(1+\frac{m}{2r})^2}{(1-\frac{m}{2r})^2}(1 - \frac{m^2}{4r^2})^2 dr^2 + r^2(1+\frac{m}{2r})^4 d\Omega^2 \\ &= -\frac{(1-A)^2}{(1+A)^2}dt^2 + \frac{(1+A)^2}{(1-A)^2}[(1-A)(1+A)]^2 dr^2 + r^2(1+A)^4 d\Omega^2 \\ &= -\frac{(1-A)^2}{(1+A)^2}dt^2 + (1+A)^4(dr^2 + r^2 d\Omega^2) \\ &= -\frac{(1-A)^2}{(1+A)^2}dt^2 + (1+A)^4(dx^2 + dy^2 + dz^2). \end{aligned} \tag{55}$$

Note, that we changed to Cartesian coordinates as well and that the factor $(1 - \frac{2m}{\rho})$ transforms as

$$\begin{aligned} (1 - \frac{2m}{\rho}) &= 1 - \frac{2m}{r(1+\frac{m}{2r})^2} = \frac{r(1+\frac{m}{2r})^2 - 2m}{r(1+\frac{m}{2r})^2} \\ &= \frac{(4r^2 - 4mr + m^2)(4r)^{-1}}{(4r^2 + 4rm + m^2)(4r)^{-1}} = \frac{(2r-m)^2}{(2r+m)^2} = \frac{(1-\frac{m}{2r})^2}{(1+\frac{m}{2r})^2}. \end{aligned} \tag{56}$$

Now the line element is of the form (54).

Next, we deploy a Lorentz-boost (2) in x-direction ($dt' = \gamma(dt + vdx)$, $dx' = \gamma(dx + vdt)$, $dy' = y$ and $dz' = z$) such that the particle, i.e. the source of the field, is moving with constant velocity v in direction x.

This gives

$$ds^2 = (1+A)^4(-dt'^2 + dx'^2 + dy'^2 + dz'^2) + \left[(1+A)^4 - \left(\frac{1-A}{1+A}\right)^2\right] \frac{(dt' - vdx')^2}{1-v^2}, \tag{57}$$

as the following calculation shows:

$$(1+A)^4[-\gamma^2(dt+vdx)^2 + \gamma^2(dx+vdt)^2] + \left[(1+A)^4 - \left(\frac{1-A}{1+A}\right)^2\right] [\gamma(dt + vdx - vdx + v^2dt)]^2 \frac{1}{1-v^2} =$$

$$(1+A)^4[\gamma^2(-dt^2 + v^2dt^2 + dx^2 - v^2dx^2)] + \left[(1+A)^4 - \left(\frac{1-A}{1+A}\right)^2\right] [\gamma(dt - v^2dt)]^2 \frac{1}{1-v^2} =$$

$$(1+A)^4[\gamma^2(1-v^2)(-dt^2 + dx^2)] + \left[(1+A)^4 - \left(\frac{1-A}{1+A}\right)^2\right] \gamma^4(1-v^2)^2dt^2.$$

Because the squared gamma factor cancels with $(1-v^2)$ we obtain equation (57) as desired.

Taking the ultra-relativistic limit of the metric amounts to accelerating the particle to the speed of light, i.e. let v tend to 1. Additionally, we set $m = p\gamma^{-1}$ in the factor A, where p is a constant. In this way we are able to keep the total energy of the particle constant. For A we obtain:

$$\begin{aligned} A = \frac{m}{2r} &= \frac{m}{2(x^2+y^2+z^2)^{\frac{1}{2}}} = \\ &\frac{m}{2[(x'-vt')^2\gamma^2 + y'^2 + z'^2]^{\frac{1}{2}}} = \\ &\frac{p\gamma^{-1}}{2\gamma[(x'-vt')^2 + \gamma^{-2}(y'^2+z'^2)]^{\frac{1}{2}}} = \\ &\frac{p\gamma^{-2}}{2[(x'-vt')^2 + \gamma^{-2}(y'^2+z'^2)]^{\frac{1}{2}}} = \\ &\frac{p(1-v^2)}{2[(x'-vt')^2 + (1-v^2)(y'^2+z'^2)]^{\frac{1}{2}}}. \end{aligned} \tag{58}$$

It is convenient to work with the Taylor expansion of (57), which is,

$$ds^2 = (1+4A)(-dt'^2 + dx'^2 + dy'^2 + dz'^2) - 8A\frac{(dt' - vdx')^2}{(1-v^2)}. \tag{59}$$

The term $(1-v^2)$ in A guarantees $\lim_{v\to 1} A = 0$. Though, as can be seen in (59), this factor cancels in the right most term. Therefore, we need to check the limit

$$\lim_{v\to 1} \frac{4p}{[(x'-vt')^2 + (1-v^2)\rho^2]^{\frac{1}{2}}}, \tag{60}$$

where $\rho^2 = y'^2 + z'^2$.

For $t' = x'$ the term (60) clearly diverges. However, it is possible to create a workaround to solve this problem. We will discuss this in the following paragraph.

It is possible to seperate a term in (60), that does not contribute to the curvature of the sapcetime, as we shall see shortly:

$$\begin{aligned} &\lim_{v\to 1} \frac{4p}{[(x'-vt')^2+(1-v^2)\rho^2]^{\frac{1}{2}}} = \\ \lim &\left(\frac{4p}{[(x'-vt')^2+(1-v^2)\rho^2]^{\frac{1}{2}} - [(x'-vt')^2+(1-v^2)]^{\frac{1}{2}}} \right) + \frac{4p}{|t'-x'|}. \end{aligned} \tag{61}$$

As explained in [27] the limit of the bracketed term in (61) tends to $-4p \ln \rho^2 \delta(x'-t') = -8p \ln \rho \delta(x'-t')$.

The limit of the line element (59) then amounts to

$$\lim_{v\to 1} ds^2 = 8p \ln \rho \delta(x'-t')(dt'-dx')^2 - \frac{4p}{|t'-x'|}(dt'-dx')^2 - dt' + dx'. \tag{62}$$

Now, we can deploy another transformation to get rid of the separated term, which does not contribute to the curvature. Setting

$$\begin{aligned} \bar{x} &= x' - 2p \ln(|t'-x'|) \\ \bar{t} &= t' - 2p \ln(|t'-x'|), \end{aligned} \tag{63}$$

we obtain

$$\begin{aligned} dt' - dx' &= d\bar{t} - d\bar{x} \\ dt' + dx' &= d\bar{t} + d\bar{x} - \frac{4p}{|t'-x'|}(dt'-dx') \end{aligned} \tag{64}$$

and

$$-dt'^2 + dx'^2 = -d\bar{t}^2 + d\bar{x}^2 + \frac{4p}{|\bar{t}-\bar{x}|}(d\bar{t}-d\bar{x})^2. \tag{65}$$

Thus, the line element of a particle with rest mass m in the ultra-realtivistic limit $v \to 1$ amounts to

$$ds^2 = -d\bar{t}^2 + d\bar{x}^2 + d\bar{y}^2 + d\bar{z}^2 + 8p \ln \rho \delta(\bar{t}-\bar{x})(d\bar{t}-d\bar{x})^2. \tag{66}$$

Using light cone coordinates, this gives

$$\boxed{ds^2 = 8p \ln \rho \delta(u) du^2 - du dr + dy'^2 + dz'^2. \qquad (67)}$$

Now it is easily seen that (67) is of the form (37) with $f(x, y) = 8pln(\sqrt{y^2 + x^2})$. Hence, the metric generated by a particle with rest mass m and moving with the speed of light has the form of an impulsive pp-wave.

4.4 What if It Hits? – Impulsive pp-Waves Colliding with an Observer

Concluding, we will give a short discussion of the effects a bypassing impulsive gravitational wave might have on an observer. Or, to put it differently, what happens if such a wave hits an oblivious space ship travelling through the galaxy?

Just above, in chapter 4.3.3, we have seen, that massive bodies moving with the speed of light could generate impulsive pp-waves, propagating through spacetime. Therefore, consider the wave hitting the space ship at the point $(0, r, x, y)$. What then happens is that the r-coordinate is instantly changed to $r \to r + H(x, y)$ as given in (42). Moreover, the r-direction of the trajectory of the ship changes, as well as its x and y-directions. Conferring with the Christoffel symbols of (37), we see that the derivatives of the function f in each coordinate except u govern the manner of change in directions.

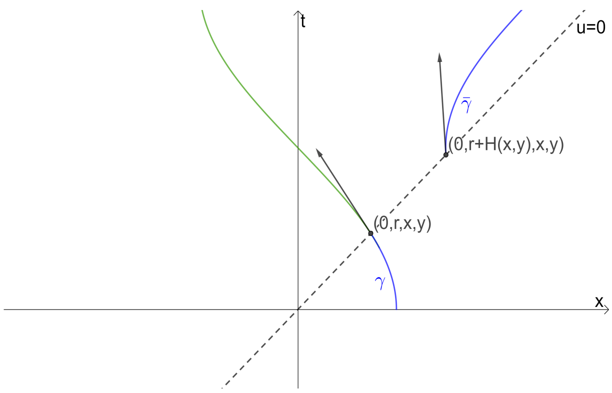

Figure 15: The curve γ depicts the trajectory of a spacecraft before it is hit by a gravitational wave moving on the hyperplane $u = 0$ and $\bar{\gamma}$ shows the trajectory after the collision. The attached vectors are tangential to the curves and thus indicate the velocity.

To sum up, if a space ship travelling straight through the universe, is hit by an impulsive gravitational wave, it would undergo a jump in the r-direction and abruptly change its path of flight in the directions r, x and y. The wave changed the curvature of spacetime and thus the ship's trajectory. What an unforseen event this would be for the passengers?

List of Figures

I took measures to contact the holders of the property rights to acquire their approval regarding the usage of respective pictures. If you still find a violation of property rights, please contact me.

References

[1] B. P. Abbott et al. "GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence". *Phys. Rev. Lett.* 116 (2016).

[2] B. P. Abbott et al. "Observation of Gravitational Waves from a Binary Black Hole Merger". *Phys. Rev. Lett.* 116 (2016).

[3] P. C. Aichelburg and R. U. Sexl. "On the gravitational field of a massless particle". *General Relativity and Gravitation* 2 (1971), pp. 303–312.

[4] C. Bär. *Elementary Differential Geometry*. Cambridge University Press, 2010.

[5] M. Blau. "Plane Waves and Penrose Limits". Lecture Notes. Institut de Physique, Université de Neuchâtel, 2011. URL: `http://www.blau.itp.unibe.ch/lecturesPP.pdf` (visited on 03/10/2020).

[6] H.W. Brinkmann. "Einstein spaces which are mapped conformally on each other". *Mathematische Annalen* 94 (1925), pp. 119–145.

[7] C. Denson Hill and P. Nurowski. *How the green light was given for gravitational wave search*. 2016. arXiv: `1608.08673`.

[8] A. Einstein. "Zur Elektrodynamik bewegter Körper". *Annalen der Physik und Chemie* 322 (1905), pp. 891–921.

[9] F. Embacher. *Elemente der theoretischen Physik: Band 1, Klassische Mechanik und Spezielle Relativitätstheorie Eine Einführung für das Lehramts- und Bachelorstudium*. Vieweg Studium. Vieweg+Teubner Verlag, 2010.

[10] R. Eötvös. *Über die Anziehung der Erde auf verschiedene Substanzen*. Vol. 8. R. Friedländer und Sohn, 1890.

[11] A. Fölsing and E. Osers. *Albert Einstein: A Biography*. Penguin, 1998.

[12] G. Galilei and M.A. Finocchiaro. *Galileo on the World Systems: A New Abridged Translation and Guide*. University of California Press, 1997.

[13] J.B. Griffiths and J. Podolský. *Exact Space-Times in Einstein's General Relativity*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2009.

[14] J. C. Maxwell. *Proceedings of the Royal Society of London*. 13. Taylor and Francis, 1864.

[15] A. A. Michelson and E. W. Morley. "On the relative motion of the Earth and the luminiferous ether". *American Journal of Science* 34.203 (1887), pp. 333–345.

[16] H. Minkowski. *Raum und Zeit*. Teubner, 1909.

[17] C.W. Misner, K.S. Thorne, and J.A. Wheeler. *Gravitation*. Gravitation. W. H. Freeman, 1973.

[18] I. Newton. *Philosophiae naturalis principia mathematica*. J. Societatis Regiae ac Typis J. Streater, 1687.

[19] B. O'Neill. *Semi-Riemannian Geometry With Applications to Relativity*. Pure and Applied Mathematics. Elsevier Science, 1983.

[20] R. Penrose. “A Remarkable property of plane waves in general relativity”. Rev. Mod. Phys. 37 (1965), pp. 215–220.

[21] R. Penrose. *Structure of space-time*. De Witt, C. M. and Wheeler, J. A. (New York: Benjamin), 1972.

[22] J. Podolský, C. Sämann, R. Steinbauer, and R. Švarc. “The global existence, uniqueness and C^1-regularity of geodesics in nonexpanding impulsive gravitational waves”. *Classical and Quantum Gravity* 32 (2014).

[23] J. Podolský, R. Steinbauer, and R. Švarc. “Gyratonic pp waves and their impulsive limit”. *Phys. Rev. D* 90 (2014).

[24] Press Release. *The Nobel Prize in Physics*. 1993. URL: `https://www.nobelprize.org/prizes/physics/1993/press-release/` (visited on 12/04/2019).

[25] Press Release. *The Nobel Prize in Physics*. 2017. URL: `https://www.nobelprize.org/prizes/physics/2017/press-release/` (visited on 12/04/2019).

[26] B. Rossi and D. B. Hall. “Variation of the Rate of Decay of Mesotrons with Momentum”. *Phys. Rev.* 59 (1941), pp. 223–228.

[27] R. Steinbauer. “Colombeau-Theorie und ultrarelativistischer Limes”. Diploma Thesis. University of Vienna, 1995.

[28] Angelique Van Ombergen et al. “The effect of spaceflight and microgravity on the human brain”. *Journal of Neurology* 264 (2017).

[29] R.M. Wald. *General Relativity*. University of Chicago Press, 1984.

[30] N.M.J. Woodhouse. *General Relativity*. Springer Undergraduate Mathematics Series. Springer London, 2006.