# GENERAL RELATIVITY RAY TRACER

A Massively Parallel Free Software Alternative

ALEJANDRO GARCÍA MONTORO

Pablo Galindo Salgado, Alfonso Romero Sarabia
Universidad de Granada

December 2016

*Omnia sunt communia*

# ABSTRACT

Short summary of the contents in English. . . a great guide by Kent Beck how to write good abstracts can be found here:

https://plg.uwaterloo.ca/~migod/research/beckOOPSLA.html

*We have seen that computer programming is an art,*
*because it applies accumulated knowledge to the world,*
*because it requires skill and ingenuity, and especially*
*because it produces objects of beauty.*

— **knuth:1974 [knuth:1974]**


# ACKNOWLEDGMENTS

Put your acknowledgments here.

Many thanks to everybody who already sent me a postcard!

Regarding the typography and other help, many thanks go to Marco Kuhlmann, Philipp Lehman, Lothar Schlesier, Jim Young, Lorenzo Pantieri and Enrico Gregorio[1], Jörg Sommer, Joachim Köstler, Daniel Gottschlag, Denis Aydin, Paride Legovini, Steffen Prochnow, Nicolas Repp, Hinrich Harms, Roland Winkler, Jörg Weber, Henri Menke, Claus Lahiri, Clemens Niederberger, Stefano Bragaglia, Jörn Hees, and the whole LaTeX-community for support, ideas and some great software.

*Regarding LyX*: The LyX port was intially done by *Nicholas Mariette* in March 2009 and continued by *Ivo Pletikosić* in 2011. Thank you very much for your work and for the contributions to the original style.


1 Members of GuIT (Gruppo Italiano Utilizzatori di TeX e LaTeX)

# CONTENTS

# LIST OF FIGURES



# LIST OF TABLES

# LISTINGS

# ACRONYMS

TO DO LIST

# Part I

# INTRODUCTION

# Part II

# MATHEMATICS

# 1 LORENTZIAN VECTOR SPACES

This chapter follows the ideas on [23], which set up the basic background needed to understand what we will call later a spacetime, the main mathematical object in which we will develop our work.

Before going down that road we need to know what a Lorentzian vector space is and the basic concepts that we can build upon it.

## 1.1 BASIC DEFINITIONS

**Definition 1.1** (Lorentzian product and vector space)**.** Let $V$ be an $n$-dimensional vector space on $\mathbb{R}$, with $n \geqslant 2$.

A *Lorentzian product on* $V$ is a non-degenerate symmetric bilinear form on $V$ with index 1; that is, a bilinear form

$$g \colon V \times V \to \mathbb{R}$$

that satisfies the following:

1. $g$ is symmetric: $g(u, v) = g(v, u)$.
2. $g$ is non-degenerate: if $g(u, v) = 0 \ \ \forall v \in V$, then $\Rightarrow u = 0$.
3. $g$ has index 1: the maximum dimension of a subspace $W$ in which $g(u, u) \leqslant 0 \ \forall u \in W$ —where equality holds if and only if $u = 0$— is 1.

The vector space $V$ furnished with such a Lorentzian product is called a *Lorentzian vector space*.

As usual, we will say that two vectors are *orthogonal* on $(V, g)$ whenever $g(u, v) = 0$.

From now on until the end of this chapter, let us consider $(V, g)$ a Lorentzian vector space.

When such a Lorentzian product is added to a vector space, its elements can be classified depending on the value of the metric on them.

**Definition 1.2** (Classification of vectors on Lorentzian vector spaces)**.** A vector $v \in V$ is said to be:

- *spacelike* if $g(v, v) > 0$ or $v = 0$,
- *timelike* if $g(v, v) < 0$ or
- *null* or *lighlike* if $g(v, v) = 0$ with $v \neq 0$.

**Definition 1.3** (Light cone)**.** The *light cone of* $V$ is the subset of all null vectors.

The study of the Lorentzian vector spaces will be based on Lemma 1.4, which classifies, for each timelike vector, all elements of $(V, g)$ on two orthogonal subsets.

**Lemma 1.4.** *Let $v \in V$ be a timelike vector. Then, we can split the vector space as follows:*

$$V = \langle v \rangle \oplus v^{\perp},$$

*where* $v^{\perp} := \{u \in V / g(u, v) = 0\} = \langle v \rangle^{\perp}$.

*Furthermore,* $g_{|_{v^{\perp}}}$ *is positive definite and* $g_{|_{\langle v \rangle}}$ *is negative definite; i. e.,* $g_{|_{v^{\perp}}}$ *and* $-g_{|_{\langle v \rangle}}$ *are Euclidean products.*

*Proof.* As $\langle v \rangle$ is non degenerate, so it is $v^{\perp}$. Then, $V = \langle v \rangle + v^{\perp}$ is a direct sum. As the $v^{\perp}$ index is necessarily zero, $v^{\perp}$ is spacelike and, consequently

See [16, Lemma 5.26] for details. ■

*Remark* 1.5. The positive definite subspaces of a Lorentzian vector space are in fact Euclidean spaces; in particular, the Schwarz inequality holds in $v^{\perp}$, with $v$ a timelike vector:

$$|g(u, w)| \leqslant \sqrt{g(u, u)}\sqrt{g(w, w)} \quad \forall u, w \in v^{\perp}, \tag{1.1}$$

with equality if and only if $u$ and $w$ are linearly dependent.

Then, it is trivial that two timelike vectors are never orthogonal. An interesting result appears when we study what happens with two orthogonal null vectors.

From [24, Cor. 1.1.5], [16, p. 155], we can see:

**Proposition 1.6.** *Let $x, y \in V$ be two null vectors. Then*

*g(x,y)=0⇔x and y are linearly dependent.*

*Proof.* From Lemma 1.4 we can write $V = \langle v \rangle \oplus v^{\perp}$, with $g(v, v) = -1$. Therefore, there exist $a, b \in \mathbb{R} \setminus \{0\}$ and $u, w \in v^{\perp}$ such that:

$$x = av + u, \quad u \neq 0, \quad g(u, u) = a^2,$$
$$y = bv + w, \quad w \neq 0, \quad g(v, v) = b^2.$$

If $g(x, y) = 0$ then $|g(u, w)| = \sqrt{g(u, u)}\sqrt{g(w, w)}$ and therefore, from Equation 1.1, $u = kw$ for some $k \in \mathbb{R}$. In fact, $k = a/b$. The converse is trivial. ■

## 1.2 TIME CONES

From now on, $\mathcal{T}(V,g)$ will denote the set of all timelike vectors on $(V,g)$.

**Definition 1.7** (Time cone)**.** Let $v \in \mathcal{T}(V,g)$ be a timelike vector. The *time cone defined by* $v$ is the set

$$C(v) = \{u \in \mathcal{T}(V,g) \,:\, g(u,v) < 0\}.$$

It is clear that $C(v)$ is not empty, as $v$ itself lays on its own time cone.

Furthermore, as no two timelike vector can be orthogonal, $g(v,w) \neq 0$ for each $w \in \mathcal{T}(V,g)$; i. e., either $w \in C(v)$ or $w \in C(-v)$. This let us describe $\mathcal{T}(V,g)$ as a union of two time cones for each timelike vector.

$$\mathcal{T}(V,g) = C(v) \cup C(-v) \quad \forall v \in \mathcal{T}(V,g)$$

The following result characterizes when two timelike vectors lie in the same time cone

It is interesting to know when two timelike vectors lie in the same time cone. Lemma 1.8, based on [16, Lemma 5.29], gives us the answer.

**Lemma 1.8.** *Given* $u,v \in \mathcal{T}(V,g)$*, they belong to the same time cone if and only if* $g(u,v) < 0$*.*

*Proof.* By definition, if $g(u,v) < 0$, then $u \in C(v)$, and therefore $u$ and $v$ belong to the time cone defined by $v$.

Assume now that $u$ and $v$ belong to a same time cone defined by a $w \in V$ such that $g(w,w) = -1$. Using Lemma 1.4 we can write

$$\begin{aligned} u &= aw + y, \quad y \in w^{\perp}, \quad a = -g(u,w) > 0, \\ v &= bw + z, \quad z \in w^{\perp}, \quad b = -g(v,w) > 0. \end{aligned}$$

Taking into account

$$\sqrt{g(y,y)} < a \quad \text{and} \quad \sqrt{g(z,z)} < b,$$

we have

$$g(u,v) \leqslant -ab + |g(y,z)| \leqslant -ab + \sqrt{g(y,y)}\sqrt{g(z,z)} < -ab + ab = 0.$$

∎

**Corollary 1.9.** *Let* $w \in \mathcal{T}(V,g)$*,* $u,v \in C(w)$ *and* $a,b \in \mathbb{R}$ *with* $a,b \geqslant 0$ *and* $a^2 + b^2 \neq 0$*. Then, we have* $au + bv \in C(w)$*.*

*Proof.* As $g(au + bv, w) = ag(u,w) + bg(v,w) < 0$, Lemma 1.8 proves the result. ∎

This corollary let us conclude that time cones are convex subsets of $V$. Furthermore, it is important to realize that if $u \in C(v)$ then $C(u) = C(v)$.

Remark 1.5 showed that the Schwarz inequality holds on the orthogonal space of any timelike vector. This result is not true if we consider any pair of vectors, but what happens when we focus on timelike vectors? Some interesting results will appear.

## 1.3 WRONG-WAY INEQUALITIES

First of all, we can obtain an the so-called *wrong-way Schwarz inequality* —see [16, Prop. 5.30]— when considering only timelike vectors.

**Proposition 1.10** (Wrong-way Schwarz inequality). *Let $u, v \in \mathcal{T}(V, g)$ be a pair of timelike vectors. Then, its product satisfies the following:*

$$|g(u, v)| \geqslant \sqrt{-g(u, u)}\sqrt{-g(v, v)}$$

*with equality holding if and only if $u$ and $v$ are linearly dependent.*

*Proof.* Using Lemma 1.4, we know there exists $a \in \mathbb{R}$ and $x \in v^{\perp}$ such that $u = av + x$, and therefore

$$g(u, v)^2 = a^2 g(v, v)^2, \qquad g(u, u) = a^2 g(v, v) + g(x, x) < 0,$$

which gives

$$g(u, v)^2 = \{g(u, u) - g(x, x)\}g(v, v) \geqslant g(u, u)g(v, v).$$

The equality holds if and only if $g(x, x) = 0$, i. e., if and only if $x = 0$, which means $u = av$. ∎

We can obtain another wrong-way inequality. As a consequence of Proposition 1.10 and using Lemma 1.8, the so-called *wrong-way Minkowski inequality* is found —see [16, Cor. 5.31]—.

**Corollary 1.11** (Wrong-way Minkowski inequality). *Let $u, v \in \mathcal{T}(V, g)$ be two timelike vector that lie in the same time cone. The following inequality is hold:*

$$\sqrt{-g(u + v, u + v)} \geqslant \sqrt{-g(u, u)} + \sqrt{-g(v, v)}$$

*and equality holds if and only if $u$ and $v$ are linearly dependent.*

We can now formalize the notion of angle between two timelike vectors.

Indeed, let $u, v \in \mathcal{T}(V, g)$ be two timelike vectors that lie in the same time cone. From Lemma 1.8 and Proposition 1.10, we know that

$$-g(u, v) \geqslant \sqrt{-g(u, u)}\sqrt{-g(v, v)},$$

and therefore, there exists a unique $\theta \in \mathbb{R}$, $\theta \geqslant 0$, such that

$$\cosh\theta = \frac{-g(u, v)}{\sqrt{-g(u, u)}\sqrt{-g(v, v)}}.$$

$\theta$ is called the *hyperbolic angle* between $u$ and $v$.

## 1.4 TIME ORIENTATION

This last section on Lorentzian vector spaces introduces an important concept: the time orientation. We need some previous concepts before formalizing this idea.

**Definition 1.12** (Orthonormal basis of $(V, g)$)**.** Let $B = (v_1, \dots, v_n)$ be a basis of $V$. $B$ is said to be *orthonormal* when

- any two different vectors on $B$ are $g$-orthogonal,
- each $v_i$, with $i \in \{1, \dots, n-1\}$, is unitary spacelike and
- $v_n$ is unitary timelike.

Note that the matrix of $g$ on the basis $B$ is

$$M_B(g) = \begin{pmatrix} I_{n-1} & 0 \\ 0 & -1 \end{pmatrix}.$$

Let $B = (v_1, \dots, v_n)$ and $B' = (v'_1, \dots, v'_n)$ be two orthonormal basis of $(V, g)$. We say that $B$ and $B'$ define the same *time orientation* of $(V, g)$ when

$$g(v_n, v'_n) < 0;$$

that is, when both $v_n$ and $v'_n$ lie in the same time cone of $(V, g)$.

This defines two equivalence classes on the set of orthonormal basis: the class defined by $B = (v_1, .., v_{n-1}, v_n)$ and the class defined by $\tilde{B} = (v_1, .., v_{n-1}, -v_n)$.

**Definition 1.13** (Time orientation)**.** A *time orientation* on $(V, g)$ is each one of the two equivalence classes defined by $B$ and $\tilde{B}$.

Furthermore, we note that a time orientation on $(V, g)$ is given by each time cone of $(V, g)$.

Although the classical orientations on $V$ does not depend on $g$, it is important to realize that the notion of time orientation is a metric concept, that is, it is defined using the Lorentzian product $g$.

However, a time orientation does not change if we replace $g$ by a conformally related Lorentzian product $ag$, with $a \in \mathbb{R}$, $a > 0$.

## 1.5 TOPOLOGICAL REMARKS

We end this section explaining some topological remarks. If $V$ is an $n$-dimensional vector space and $B$ is a basis of $V$, then we have a linear isomorphism

$$b_B : V \longrightarrow \mathbb{R}^n,$$

defined by $b_B(v) = (a_1, .., a_n)$, where $(a_1, .., a_n)$ are the coordinates of $v$ in $B$. By using $b_B$ a topology $\mathfrak{T}_B$ can be defined in $V$ in such a way that the $\mathfrak{T}_B$-open subsets of $V$ are $b_B^{-1}(O)$, where $O$ is an open subset of $\mathbb{R}^n$. It is not difficult to see that $\mathfrak{T}_B = \mathfrak{T}_{B'}$ for any basis $B, B'$ of $V$. Endowed with this topology, the function $V \longrightarrow \mathbb{R}$, given by $v \mapsto g(v, v)$ is continuous. Hence $\mathcal{T}(V, g)$ is an open subset of $V$ and each time cone is also an open subset of $V$. Moreover, the set of nonzero spacelike vectors $\{v \in V : g(v, v) > 0\}$ is an open subset of $V$ and the light cone with the zero vector $\{v \in V : g(v, v) = 0\}$ is a closed subset of $V$. Finally, every null vector can be obtained as the limit of a sequence of timelike vectors as well of spacelike vectors (or both types of vectors). Thus, $\{v \in V : g(v, v) < 0\} \cup \{v \in V : g(v, v) > 0\}$ is a dense subset of $V$.

# 2 TENSOR ALGEBRA

Add references throughout the text.

This chapter covers some basic tools needed in the further development of this work. Although basic, its knowledge is crucial on a lot fields, and its interest for the study of the Geometry will unveil on the following chapters.

Nearly all definitions, results and ideas are based on [22, Chapters 6 and 9].

## 2.1 THE NOTION OF TENSOR

**Definition 2.1** (Multilinear map)**.** Let $V_1, V_2, \dots, V_r$ and $W$ be vector spaces over the same field $K$. A multilinear —r times linear— map from $V_1 \times V_2 \cdots \times V_r$ to $W$ is a map

$$T\colon V_1 \times V_2 \cdots \times V_r \longrightarrow W$$

that is linear in each of its components; i. e., that verifies the following conditions:

1. $$T(x_1, \dots, x_i + x_i', \dots, x_r) = T(x_1, \dots, x_i, \dots, x_r) + T(x_1, \dots, x_i', \dots, x_r),$$
2. $$T(x_1, \dots, a x_i, \dots, x_r) = a T(x_1, \dots, x_i, \dots, x_r),$$

for every $i \in \{1, 2, \dots, r\}$, where $x_j$ is an arbitrary vector in $V_j$ and $a \in K$.

Before going ahead with the definition of tensor, we must remember the concept of dual space.

Given a vector space $V$ over a field $K$, its *dual space* is the vector space defined as

$$V^* := \mathrm{Hom}_K(V, K);$$

that is, $V^*$ is the set of all linear maps $\varphi : V \to K$.

There are some interesting results concerning dual spaces that will be important in the understanding of the notion of tensor.

First of all, it is known that if $V$ is finite-dimensional, the dimensions of $V$ and $V^*$ are the same and, given a base of $V$, $B = \{v_1, \dots, v_n\}$, its *dual basis* is built as $B^* = \{\varphi^1, \dots, \varphi^n\}$[1], where

$$\varphi^i(v_j) = \delta^i_j.$$

Reflexiveness theorem is the worst translation ever

Furthermore, the reflexiveness theorem tells us that there exists a natural isomorphism between $V$ and its double dual space, $V^{**}$, when $V$ is finite-dimensional. This isomorphism assigns, to every vector $v \in V$, a function that maps every one-form into its evaluation on $v$:

$$\begin{aligned} \psi\colon V &\longrightarrow V^{**} \\ v &\longmapsto \psi_v\colon V^* \longrightarrow K \\ &\qquad\qquad \varphi \longmapsto \varphi(v). \end{aligned} \tag{2.1}$$

With the concepts of multilinear maps and dual spaces we can now build the definition of tensor, core concept of this section.

**Definition 2.2** (Tensor)**.** Let $V$ be a vector space over a field $K$, being $V^*$ its dual space. A tensor $r$ ($\geqslant 0$) times contravariant and $s$ ($\geqslant 0$) times covariant —i. e., a tensor of type $(r, s)$— is a multilinear map

$$T\colon \underbrace{V^* \times \cdots \times V^*}_{\text{r copies}} \times \underbrace{V \times \cdots \times V}_{\text{s copies}} \longrightarrow K.$$

Tensors of type $(0, s)$ are said to be *covariant*, whereas tensors of type $(r, 0)$ are called *contravariant*.

The *order* of the tensor is defined as the sum $r + s$.

**Example 2.3.** The following examples show how interesting the notion of tensor is, as it can include a vast selection of mathematical objects under the same concept; for example, we will see that both vectors and one-forms are tensors.

Let $V$ be an $n$-dimensional vector space and $V^*$ its dual space.

1. Let $\varphi \in V^*$; i. e., $\varphi\colon V \to K$ is a one-form. From Definition 2.2 it is clear that $\varphi$ is a tensor of type $(0, 1)$ over $V$.

2. Let $v \in V$ be a vector. Using the natural isomorphism between $V$ and its double dual space, the vector $v$ can be identified with $\psi_v\colon V^* \to K$, and thus we can understand $v$ as a tensor of type $(1, 0)$.

3. Consider now $f \in \operatorname{End}_K V$ and let $T_f\colon V^* \times V \to K$ be the map defined as $T_f(\varphi, v) := \varphi(f(v))$. It is clear that $T_f$ is a tensor of type $(1, 1)$. Moreover, if we consider $f$ to be the identity map $1_V$, then $T_{1_V}$ is the tensor that maps every pair of (vector, one-form) to the evaluation of the one-form on the vector; i. e., is the tensor associated to the natural isomorphism between $V$ and $V^{**}$.

1 From now on, Latin letters with subscripts will denote vectors, whereas Greek letters with superscripts will denote one-forms.

4. Common operations on several mathematical fields can also be seen as tensors. For example, the inner product can be defined as a tensor of type $(0,2)$ as follows:

$$\begin{aligned} T\colon V \times V &\to K \\ (v,w) &\mapsto \sum_{i=1}^{n} v_i w_i. \end{aligned}$$

   In general, every bilinear form is a $(0,2)$ tensor. This follows from the definition of bilinear form, which satisfies all conditions in Definition 2.2.

## 2.2 TENSOR BASIC OPERATIONS

### 2.2.1 *Tensor addition and product by a scalar*

Let $\mathcal{T}_{r,s}(V)$ be the set of all tensors of type $(r,s)$. It is clear that both $\mathcal{T}_{1,0}(V) = V^*$ and $\mathcal{T}_{0,1}(V) = V^{**}$ are vector spaces over K. A natural question arises: is $\mathcal{T}_{r,s(V)}$ a vector space for arbitrary $r$ and $s$? The following result gives us the answer we are looking for.

**Theorem 2.4.** *Let $T, T' \in \mathcal{T}_{r,s}(V)$ be two tensors of type $(r,s)$ and $a \in K$ a scalar. Consider the following operations:*

- $(T+T')(\varphi^1,\dots,\varphi^r,v_1,\dots,v_s) := T(\varphi^1,\dots,\varphi^r,v_1,\dots,v_s) + T'(\varphi^1,\dots,\varphi^r,v_1,\dots,v_s),$
- $(aT)(\varphi^1,\dots,\varphi^r,v_1,\dots,v_s) := aT(\varphi^1,\dots,\varphi^r,v_1,\dots,v_s),$

*where $\varphi^i \in V^*$ for all $i \in \{1,\dots,r\}$ and $v_j \in V$ for all $j \in \{1,\dots,s\}$.*

*The set $\mathcal{T}_{r,s}(V)$ with the preceding operations is a vector space.*

*Proof.* It is clear, from Definition 2.2, that $T+T', aT \in \mathcal{T}_{r,s}(V)$, given the linearity in each of the components of both $T$ and $T'$. These operations satisfy the following properties:

1. $(T+T')+T'' = T+(T'+T'')$.
2. There exists a *null tensor* $T_0 \in \mathcal{T}_{r,s}(V)$, defined as

$$T_0(\varphi^1,\dots,\varphi^r,v_1,\dots,v_s) = 0, \forall \varphi^i \in V^*, \forall v_i \in V,$$

   such that $T_0 + T = T + T_0 = T \ \forall T \in \mathcal{T}_{r,s}(V)$.

3. For each $T \in \mathcal{T}_{r,s}(V)$ there exists an *opposite tensor* $-T$ defined as

$$(-T)(\varphi^1,\dots,\varphi^r,v_1,\dots,v_s) = -T(\varphi^1,\dots,\varphi^r,v_1,\dots,v_s)$$

   that satisfies $T+(-T) = -T+T = 0$.

4. $T + T' = T' + T$.

This provides abelian goup structure to $\mathcal{T}_{r,s}(V)$. The following properties finally show that $\mathcal{T}_{r,s}(V)$ is a vector space:

1. $a(T + T') = aT + aT', \quad \forall a \in K, \ \forall T, T' \in \mathcal{T}_{r,s}(V)$.
2. $(a + b)T = aT + bT, \quad \forall a, b \in K, \ \forall T \in \mathcal{T}_{r,s}(V)$.
3. $(ab)T = a(bT), \quad \forall a, b \in K, \ \forall T \in \mathcal{T}_{r,s}(V)$.
4. $1T = T, \quad \forall T \in \mathcal{T}_{r,s}(V)$, where $1$ is the unity in $K$.

■

### 2.2.2 *Tensor product*

Now that we know that $\mathcal{T}_{r,s}(V)$ is a vector space, it will be important to study its dimension. Before going down that road, let us define the tensor product, concept upon which we will be able to build a basis for $\mathcal{T}_{r,s}(V)$.

**Definition 2.5** (Tensor product)**.** Let $T \in \mathcal{T}_{r,s}(V)$ and $T' \in \mathcal{T}_{r',s'}(V)$. The *tensor product*

$$T \otimes T' \colon \underbrace{V^* \times \cdots \times V^*}_{\text{r+r}' \text{ copies}} \times \underbrace{V \times \cdots \times V}_{\text{s+s}' \text{ copies}}$$

is defined as follows:

$$(T \otimes T')(\varphi^1, \ldots, \varphi^{r+r'}, v_1, \ldots, v_{s+s'}) :=$$
$$T(\varphi^1, \ldots, \varphi^r, v_1, \ldots, v_s) T(\varphi^{r+1}, \ldots, \varphi^{r+r'}, v_{s+1}, \ldots, v_{s+s'}),$$

where $\varphi^i \in V^*$ for all $i \in \{1, \ldots, r + r'\}$ and $v_j \in V$ for all $j \in \{1, \ldots, s + s'\}$.

It is easy to prove that $T \otimes T' \in \mathcal{T}_{r+r',s+s'}(V)$ for every $T \in \mathcal{T}_{r,s}(V)$ and $T' \in \mathcal{T}_{r',s'}(V)$. However, the proof is long and cumbersome to write, and it can be found in almost every elementary book on tensor algebra.

Add reference.

Furthermore, we can see that, given $T \in \mathcal{T}_{r,s}(V)$, $T' \in \mathcal{T}_{r',s'}(V)$ and $T'' \in \mathcal{T}_{r'',s''}(V)$:

$$(T \otimes T') \otimes T'' \in \mathcal{T}_{r+r'+r'',s+s'+s''}(V),$$
$$T \otimes (T' \otimes T'') \in \mathcal{T}_{r+r'+r'',s+s'+s''}(V)$$

and that the following equality holds:

$$(T \otimes T') \otimes T'' = T \otimes (T' \otimes T'').$$

In fact, the application

$$\begin{aligned} \mathcal{T}_{r,s}(V) \times \mathcal{T}_{r',s'}(V) &\longrightarrow \mathcal{T}_{r+r',s+s'}(V) \\ (T, T') &\longmapsto T \otimes T' \end{aligned}$$

is a bilinear map.

As a particular example, we can define the tensor product between two tensors of type $(1,0)$ and $(0,1)$ as the following tensor:[2]

$$v \otimes \varphi \in \mathcal{T}_{1,1}(V),$$

which maps every pair $(\psi, w)$ to the scalar $\psi(v)\varphi(w)$.

Upon this concept, and using the known properties of the dual vector space and its basis, we can state and prove a theorem that builds a basis for every $\mathcal{T}_{r,s}(V)$.

**Theorem 2.6.** *Let $V$ be an $n$-dimensional vector space over a field $K$. Let $\mathcal{B} = \{v_1, \dots, v_n\}$ be a basis of $V$ and $\mathcal{B}^* = \{\varphi^1, \dots, \varphi^n\}$ be its dual basis. Then,*

$$\mathcal{B}_T = \{v_{i_1} \otimes \cdots \otimes v_{i_r} \otimes \varphi^{j_1} \otimes \cdots \otimes \varphi^{j_s} \text{ , where every index moves independently from } 1 \text{ to } n\},$$

*is a basis of $\mathcal{T}_{r,s}(V)$. As a consequence,* $\dim_K \mathcal{T}_{r,s}(V) = n^{r+s}$.

*Proof.* In order to prove that $\mathcal{B}_T$ is a linear span of $\mathcal{T}_{r,s}(V)$, let us consider $T \in \mathcal{T}_{r,s}(V)$. Then, if

$$\begin{aligned} \psi^1 &= \sum_{i_1=1}^{n} a^1_{i_1} \varphi^{i_1}, \quad \cdots, \quad \psi^r = \sum_{i_r=1}^{n} a^r_{i_r} \varphi^{i_r} \text{ and} \\ w_1 &= \sum_{j_1=1}^{n} b^{j_1}_1 v_{j_1}, \quad \cdots, \quad w_s = \sum_{j_s=1}^{n} b^{j_s}_s v_{j_s}, \end{aligned}$$

we can write

$$T(\psi^1, \dots, \psi^r, w_1, \dots, w_s) = \sum_{\substack{i_1,\dots,i_s \\ j_1,\dots,j_r}} a^1_{i_1} \cdots a^r_{i_r} b^{j_1}_1 \dots b^{j_s}_s t^{i_1,\dots,i_r}_{j_1,\dots,j_s},$$

where $t^{i_1,\dots,i_r}_{j_1,\dots,j_s} := T(\varphi^{i_1}, \dots, \varphi^{i_r}, v_{j_1}, \dots, v_{j_s})$.

Following this reasoning —see [22] for the details—, we can finally write

$$T = \sum_{\substack{i_1,\dots,i_s \\ j_1,\dots,j_r}} t^{i_1,\dots,i_r}_{j_1,\dots,j_s} \left( v_{i_1} \otimes \cdots \otimes v_{i_r} \otimes \varphi^{j_1} \otimes \cdots \otimes \varphi^{j_s} \right),$$

which proves that $\mathcal{B}_T$ spans $\mathcal{T}_{r,s}(V)$.

2 Note that we are using the reflexiveness theorem to ease the notation, as we are writing $v$ —a letter used for vectors of $V$— to describe an element of $V^{**}$.

To prove that $\mathcal{B}_T$ is linearly independent, consider a linear combination of the elements in $\mathcal{B}_T$ that equals the null tensor,

$$\sum_{\substack{k_1,\dots,k_r \\ l_1,\dots,l_s}} a^{k_1,\dots,k_r}_{l_1,\dots,l_s} \left(v_{k_1} \otimes \cdots \otimes v_{k_r} \otimes \varphi^{l_1} \otimes \cdots \otimes \varphi^{l_s}\right) = T_0 \tag{2.2}$$

Evaluating both sides of Equation 2.2 on the corresponding elements of the bases $\mathcal{B}$ and $\mathcal{B}^*$ we obtain the equality

$$\sum_{\substack{k_1,\dots,k_r \\ l_1,\dots,l_s}} a^{k_1,\dots,k_r}_{l_1,\dots,l_s} \left(\delta^{i_1}_{k_1} \dots \delta^{i_r}_{k_r} \delta^{l_1}_{j_1} \dots \delta^{l_s}_{j_s}\right) = 0,$$

from which it is clear that $a^{i_1,\dots,i_r}_{j_1,\dots,j_s} = 0$ for every index. ∎

Two particularly important examples of the above theorem let us grasp the duality between vectors and one-forms.

Let $V$ be a vector space with dimension $n$, with basis $B = \{v_1, \dots, v_n\}$ and dual basis $B^* = \{\varphi^i, \dots, \varphi^n\}$.

We know that $\mathcal{T}_{1,0}(V) = V^{**}$, and using reflexiveness theorem we can write each vector $v \in V \simeq V^{**}$ as

$$v = \sum_{i=1}^{n} b^i v_i,$$

where each coordinate is the projection of the vector using the corresponding one-form: $b^i = \varphi^i(v)$.

Similarly, we can consider $\mathcal{T}_{0,1}(V) = V^*$. Each of its elements, $\psi \in V^*$, can be written in terms of the tensor basis $\mathcal{B}_{\mathcal{T}_{0,1}(V)} = B^*$:

$$\psi = \sum_{j=1}^{n} a_j \varphi^j,$$

Note here that the coordinates are retrieved evaluating the form in each element of the basis $B$: $a_j = \psi(v_j)$.

This shows that both evaluating a one-form on the elements of the basis $B$ and projecting a vector with the elements of the basis $B^*$ can be seen as dual operations: both of them gives the coordinates of the corresponding element on the corresponding basis.

The notion of tensor help us to understand this duality: we now see both elements as the *same* mathematical entity, where both operations are expressed in different bases with the corresponding change of coordinates.

## 2.3 CHANGE OF BASES ON $\mathcal{T}_{r,s}(V)$

We now want to study how the coordinates of a tensor change when we change the basis of $\mathcal{T}_{r,s}(V)$.

First of all, we must remember the relationship between the change of basis on a vector space V and the corresponding change of basis in its dual space, $V^*$.

Let $B = (v_1, \ldots, v_n)$ and $B' = (v'_1, \ldots, v'_n)$ be two bases of $V(K)$, and suppose that the change of coordinates is given by

$$v_j = \sum_{i=1}^{n} a^i_j v'_i.$$

It is known that —see [22, p. 162]—, if $B^* = (\varphi^1, \ldots, \varphi^n)$ and $B'^* = (\varphi'^1, \ldots, \varphi'^n)$ are the dual bases of the two preceding ones, the elements of the bases are related in the *opposite* way as before:

$$\varphi'^i = \sum_{j=1}^{n} a^i_j \varphi^j, \text{ where } a^i_j \in K. \tag{2.3}$$

Furthermore, we can see that

$$v'_j = \sum_{j=1}^{n} b^i_j v_j, \tag{2.4}$$

where $b^i_j$ is the element placed on the i-th row, j-th column of the inverse matrix of $(a^i_j)$.

This change of bases in V and $V^*$ is key to understand Proposition 2.7, which tells us how the coordinates of a tensor change when we change the basis of $\mathcal{T}_{r,s}(V)$.

**Proposition 2.7.** *Let* $B$, $B'$ *and* $B^*$, $B'^*$ *the bases of* V *and* $V^*$ *described above and consider two ordered bases of* $\mathcal{T}_{r,s}(V)$ *obtained from*

$$\mathcal{B}_T = \{v_{i_1} \otimes \cdots \otimes v_{i_r} \otimes \varphi^{j_1} \otimes \cdots \otimes \varphi^{j_s} \text{, where every index moves independently from } 1 \text{ to } n\}$$

*and*

$$\mathcal{B}'_T = \{v'_{i_1} \otimes \cdots \otimes v'_{i_r} \otimes \varphi'^{j_1} \otimes \cdots \otimes \varphi'^{j_s} \text{, where every index moves independently from } 1 \text{ to } n\}.$$

*Then, the analytic expression of the change of coordinates of a tensor* $T \in \mathcal{T}_{r,s}(V)$ *is the following:*

$$t'^{k_1,\ldots,k_r}_{l_1,\ldots,l_s} = \sum a^{k_1}_{i_1} \cdots a^{k_r}_{i_k} b^{j_1}_{l_1} \cdots b^{j_s}_{l_s} t^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}$$

*where the coefficients are the ones defined in Equation 2.3 and Equation 2.4.*

*Proof.* Considering the change of bases in V and $V^*$, the proof is direct, using the definition of the coordinates of a tensor $T \in \mathcal{T}_{r,s}(V)$:

$$\begin{aligned} t'^{k_1,\ldots,k_r}_{l_1,\ldots,l_s} &= T(\varphi'^{k_1}, \ldots, \varphi'^{k_r}, v'_{l_1}, \ldots, v'_{l_s}) = \\ &= T(\sum a^{k_1}_{i_1} \varphi^{i_1}, \ldots, \sum a^{k_r}_{i_r} \varphi i_r, \sum b^{j_1}_{l_1} v_{j_1}, \ldots, \sum b^{j_s}_{l_s} v_{j_s}) = \\ &= \sum a^{k_1}_{i_1} \cdots a^{k_r}_{i_k} b^{j_1}_{l_1} \cdots b^{j_s}_{l_s} T(\varphi^{k_1}, \ldots, \varphi^{k_r}, v_{l_1}, \ldots, v_{l_s}) = \\ &= \sum a^{k_1}_{i_1} \cdots a^{k_r}_{i_k} b^{j_1}_{l_1} \cdots b^{j_s}_{l_s} t^{i_1,\ldots,i_r}_{j_1,\ldots,j_s} \end{aligned}$$

∎

*Tensors of order 2*

We can study now the particular examples of tensors of order 2. From now on, let $V$ be an $n$-dimensional vector space, $V^*$ its dual space and $\mathcal{B} = (v_1, \dots, v_n)$ and $\mathcal{B}^* = (\varphi^1, \dots, \varphi^n)$ their respective bases.

Let us first consider $\mathcal{T}_{(2,0)}$, which is a vector space of dimension $n^2$. Let

$$\mathcal{B}_T = \{v_i \otimes v_j, \text{ where } i, j \in \{1, \dots, n\}\}$$

be one of its bases. We can now write a generic tensor $T \in \mathcal{T}_{2,0}$ as follows:

$$T = \sum_{i,j=1}^{n} t^{ij} v_i \otimes v_j,$$

where $t^{ij} = T(\varphi^i, \varphi^j)$.

Let $\mathcal{B}' = (v'_1, \dots, v'_n)$ and $\mathcal{B}'^* = (\varphi'^1, \dots, \varphi'^n)$ be two new bases of $V$ and $V^*$.

Following Proposition 2.7, we know that $T = \sum_{i,j=1}^{n} t^{ij} v_i \otimes v_j = \sum_{k,l=1}^{n} t'^{kl} v'_k \otimes v'_l$, where

$$t'^{kl} = \sum_{i,j=1}^{n} a_i^k a_j^l t^{ij}.$$

This expression can be written with matrix notation:

$$\left(t'^{kl}\right) = \mathbf{A} \left(t^{ij}\right) \mathbf{A}^t, \tag{2.5}$$

where $\mathbf{A} := \left(a_j^i\right)$ is the matrix whose elements are the ones defined in Equation 2.3; that is, the change of basis matrix in the dual space:

$$\mathbf{A} = M(1_V, \mathcal{B}'^*, \mathcal{B}^*)$$

We can follow a similar reasoning for $\mathcal{T}_{(0,2)}$, considering

$$\mathcal{B}_T = \{\varphi^i \otimes \varphi^j, \text{ where } i, j \in \{1, \dots, n\}\}$$

one of its bases. A generic tensor $T \in \mathcal{T}_{0,2}$ can be written in a similar way:

$$T = \sum_{i,j=1}^{n} t_{ij} \varphi^i \otimes \varphi^j,$$

where $t_{ij} = T(v_i, v_j)$.

The expression of the change of coordinates is now

$$t'_{kl} = \sum_{i,j=1}^{n} b_k^i b_l^j t_{ij},$$

where $\mathbf{B} := \left(b^i_j\right)$ is the inverse matrix of $\mathbf{A}$, as shown in Equation 2.4. This expression, written in matrix notation, is similar to Equation 2.5:

$$\left(t'_{kl}\right) = \mathbf{B}\left(t_{ij}\right)\mathbf{B}^t. \tag{2.6}$$

For $r = s = 1$, the study is even more interesting. Considering

$$\mathcal{B}_T = \{v_i \otimes \varphi^j, \text{ where } i, j \in \{1, \ldots, n\}\}$$

a generic tensor $T \in \mathcal{T}_{1,1}$ is written as follows:

$$T = \sum_{i,j=1}^{n} t^i_j v_i \otimes \varphi^j,$$

where $t^i_j = T(\varphi^i, v_j)$.

The expression of the change of coordinates is again similar to the previous ones:

$$t'^k_l = \sum_{i,j=1}^{n} a^k_i b^j_l t^i_j \quad \forall k, l \in \{1, 2, \ldots, n\}.$$

The matrix notation is now somewhat different:

$$(t'^k_l) = \mathbf{A}(t^i_j)\mathbf{A}^{-1}. \tag{2.7}$$

This last expression is very interesting, as it tells us that the *matrices* of the coordinates of $T$ in both bases of $\mathcal{T}_{1,1}(V)$ are similar, not congruent as in Equation 2.5 or Equation 2.6.

Remember now the third item in Example 2.3: it established a relationship between $\mathcal{T}_{1,1}(V)$ and $\mathrm{End}_K V$, that is now clear again: we see that Equation 2.7 is not only the change of coordinates of a tensor, but also the expression of the change of basis matrix of the corresponding endomorphism.

This is not surprising as both spaces are isomorphic, with a *natural* isomorphism between them:

**Theorem 2.8.** *The map*

$$\begin{aligned} \mathrm{End}_K V &\longrightarrow \mathcal{T}_{1,1}(V), \\ f &\longmapsto T_f \end{aligned}$$

*where* $T_f(\varphi, v) = \varphi(f(v)) \; \forall \varphi \in V^* \; \forall v \in V$, *is an isomorphism.*

*Proof.* It is obvious to see that $T_{f+f'} = T_f + T_{f'}$ and that $T_{af} = aT_f$, which tells us that the map is linear.

Suppose now that $T_f = T_0$; i.e., $\varphi(f(v)) = 0 \; \forall \varphi \in V^* \; \forall v \in V$. This implies immediately that $f(v) = 0 \; \forall v \in V$, as it is known that if $\psi(w) = 0$ for every $\psi \in V$, then $w = 0$. This shows that $f$ is the zero mapping, which proves that the map is one-to-one. ∎

Note that whereas all the three vector spaces studied —$\mathcal{T}_{2,0}$, $\mathcal{T}_{0,2}$ and $\mathcal{T}_{1,1}$— are isomorphic between them —they all have the same dimension: 2—, the natural isomorphism with $\mathrm{End}_K(V)$ is only found with $\mathcal{T}_{1,1}$.

*Remark* 2.9. This natural identification between tensors in $\mathcal{T}_{1,1}$ and operators is key in the physics literature, as one can talk either about operators over a vector space or about tensors $(1,1)$.

But how can we identify a tensor and its corresponding operator in practice? Let us see it: consider $f \in \mathrm{End}_K(V)$ defined, using a basis $\mathcal{B} = (v_1, \dots, v_n)$ of $V$, as

$$f(v_j) = \sum_{i=1}^{n} a_j^i v_i$$

and its corresponding tensor defined as

$$T_f = \sum_{i,j=1}^{n} t_j^i v_i \otimes \varphi^j.$$

A really quick proof shows that both coordinates $a_j^i$ and $t_j^i$ are exactly the same!

$$t_j^i = T_f(\varphi^i, v_j) = \varphi^i(f(v_j)) = a_j^i.$$

## 2.4 TENSOR CONTRACTION

We can study another way of obtaining new tensors from old ones. The operation we are going to define is motivated by the following property of the $(1,1)$ tensors.

Consider a tensor $T \in \mathcal{T}_{(1,1)}(V)$ and two basis of $\mathcal{T}_{(1,1)}(V)$: $\mathcal{B} = \{v_j \otimes \varphi^i\}$ and $\mathcal{B}' = \{v_j' \otimes \varphi'^i\}$, where both indices move independently from 1 to $n$. If the components of the tensor on the two different bases are the following:

$$t_j^i = T(\varphi^i, v_j) \quad t_j'^i = T(\varphi'^i, v_j'),$$

with the expression of coordinate change being

$$t'^k_l = \sum_{i,j=1}^{n} a_i^k b_l^j t_j^i \quad \forall k, l \in \{1, 2, \dots, n\}.$$

From the previous expression is straightforward to see that [22, p. 198] the sum of the coordinates with the same index is invariant under a change of basis. Indeed:

$$\sum_k t_k'^k = \sum_{i,j,k} a_i^k b_k^j t_j^i = \sum_{i,j} \left( \sum_k a_i^k b_k^j \right) t_j^i = \sum_{i,j} \delta_i^j t_j^i = \sum_i t_i^i.$$

Considering the tensor $T$ as an operator, it is clear that this invariant scalar (a $(0,0)$ tensor) is the trace. We have then a map

$$C\colon \mathcal{T}_{(1,1)}(V) \to \mathcal{T}_{(0,0)}(V)$$

defined as $C(T) = \sum_i T(\varphi^i, v_i)$.

The generalization of this map to other tensor types is what we call tensor contraction.

**Definition 2.10** (Tensor contraction)**.** Let $T \in \mathcal{T}_{(r,s)}(V)$, where $r, s > 0$. The contraction of $T$ with respect to the $i$-th contravariant slot and $j$-th covariant slot is the image of $T$ by the application

$$\begin{aligned} C^i_j\colon \mathcal{T}_{(r,s)}(V) &\to \mathcal{T}_{(r-1,s-1)}(V) \\ T &\mapsto C^i_j T, \end{aligned}$$

where $C^i_j T$ maps each tuple of one-forms and vectors

$$(\psi^1, \dots, \psi^{r-1}, w_1, \dots, w_{s-1})$$

to the value

$$\sum_k T(\psi^1, \dots, \underbrace{\varphi^k}_{i\text{-th slot}}, \dots, \psi^{r-1}, w_1, \dots, \underbrace{v_k}_{j\text{-th slot}}, \dots, w_{s-1}),$$

where $\{v_i\}$ is a basis of $V$ and $\{\varphi^i\}$ its dual basis.

It can be proved —see [22, Prop. 6.12]— that $C^i_j T$ is a tensor. As it does not depend on the choice of the basis, the concept of contraction is well-defined.

It is interesting to note that the coordinates of $C^i_j T$ are obtained making equal the $i$-th contravariant, $j$-th covariant coordinate and summing in that index —see [22, Prop. 6.14] for a more detailed explanation—:

$$C^i_j T = \left( \sum_m t^{k_1,\dots,k_{i-1},m,k_{i+1},\dots,k_r}_{l_1,\dots,l_{j-1},m,l_{j+1},\dots,l_s} \right)$$

## 2.5 ALTERNATIVE DEFINITION OF TENSOR: THE PHYSICS APPROACH

Equations 2.5, 2.6 and 2.7 can be interpreted otherwise, as they can be used to build an alternative definition of tensor.

This reinterpretation first needs a new concept to be developed, the multidimensional arrays, that will be used to represent tensors by their coordinates on a certain basis.

**Definition 2.11** (Multidimensional array)**.** Let $K$ be a field and $s, r$ and $n$ non-negative integers. A *multidimensional array* of type $(r, s)$

and order $n$ is an ordered set of $n^{r+s}$ K-scalars, which we will note as

$$\left(t^{k_1,\ldots,k_r}_{l_1,\ldots,l_s}\right),$$

with every index moving independently from 1 to $n$ and $t^{k_1,\ldots,k_r}_{l_1,\ldots,l_s} \in K$ for every $k_1, \ldots, k_r, l_1, \ldots, l_r \in \{1, \ldots, n\}$.

It is clear that multidimensional arrays are actually square matrices when $r = s = 1$. This shows that we are working not with a brand new definition but actually with a generalization of the well-known concept; we can generalize other ideas involving square matrices to multidimensional arrays.

For example, we can define an equivalence relation on the set of all multidimensional arrays of type $(r, s)$ and order $n$ as follows: we will say that $\left(t^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}\right)$ and $\left(t'^{k_1,\ldots,k_r}_{l_1,\ldots,l_s}\right)$ are similar when there is a matrix $\mathbf{A} = \left(a^j_i\right) \in Gl(n, K)$ such that

$$t'^{k_1,\ldots,k_r}_{l_1,\ldots,l_s} = \sum a^{k_1}_{i_1} \cdots a^{k_r}_{i_k} b^{j_1}_{l_1} \cdots b^{j_s}_{l_s} t^{i_1,\ldots,i_r}_{j_1,\ldots,j_s},$$

where $\mathbf{B} := \left(b^i_j\right)$ represents, as usual, the matrix $\mathbf{A}^{-1}$.

It is clear, from Proposition 2.7, that multidimensional arrays that represent the same tensor are always similar. The other implication is also true.

**Proposition 2.12.** *Let* $\left(t^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}\right)$ *and* $\left(t'^{k_1,\ldots,k_r}_{l_1,\ldots,l_s}\right)$ *be two similar multidimensional arrays of type* $(r, s)$ *and order* $n$*, and let* $V$ *be an* $n$*-dimensional vector space,* $V^*$ *its dual and* $B = (v_1, \ldots, v_n)$*,* $B^* = (\varphi^1, \varphi^n)$ *their corresponding bases.*

*If* $T$ *is the unique tensor of type* $(r, s)$ *over* $V$ *defined as*

$$T = \sum t^{i_1,\ldots,i_r}_{j_1,\ldots,j_s} v_{i_1} \otimes \cdots \otimes v_{i_r} \otimes \varphi^{j_1} \otimes \cdots \otimes \varphi^{j_s},$$

*then there exists a unique basis of V,* $B' = (v'_1, \ldots, v'_n)$ *—with its dual basis being* $B = (\varphi'^1, \ldots, \varphi'^n)$*— such that*

$$T = \sum t'^{k_1,\ldots,k_r}_{l_1,\ldots,l_s} v'_{k_1} \otimes \cdots \otimes v'_{k_r} \otimes \varphi'^{l_1} \otimes \cdots \otimes \varphi'^{l_s}$$

In order to prove Proposition 2.12, one can follow the same ideas used to check that similar matrices represent the same linear applications; see, for example, [22, p. 145].

Propositions 2.7 and 2.12 allows us to define a tensor in the following way:

**Definition 2.13** (Tensor — alternative version). A tensor of type $(r, s)$ and order $n$ over $V$ is an application that maps each ordered basis of $V$ to a multidimensional array of type $(r, s)$ and order $n$,

$$B = (v_1, \ldots, v_n) \mapsto \left(t^{i_1,\ldots,i_r}_{j_1,\ldots,j_s}\right),$$

such that if it maps

$$B' = (v'_1, \dots, v'_n) \mapsto \left( t'^{k_1,\dots,k_r}_{l_1,\dots,l_s} \right),$$

then

$$t'^{k_1,\dots,k_r}_{l_1,\dots,l_s} = \sum a^{k_1}_{i_1} \cdots a^{k_r}_{i_k} b^{j_1}_{l_1} \cdots b^{j_s}_{l_s} t^{i_1,\dots,i_r}_{j_1,\dots,j_s},$$

where $(a^1_i, a^2_i, \dots, a^n_i)$ are the coordinates of $v'_i$ on $B$ for all $i \in \{1, \dots, n\}$ and $\mathbf{B} := (b^i_j)$ is the inverse of $\mathbf{A} := (a^i_j)$.

The study on tensors of order 2 on page 20 becomes clearer with the preceding definition, as we see now that a $(1,1)$ tensor is an application that maps every ordered basis of $V$, $B = (v_1, \dots, v_n)$, to a square matrix of order $n$,

$$\begin{pmatrix} t^1_1 & t^1_2 & \cdots & t^1_n \\ t^2_1 & t^2_2 & \cdots & t^2_n \\ \vdots & \vdots & \ddots & \vdots \\ t^n_1 & t^n_2 & \cdots & t^n_n \end{pmatrix},$$

such that for every pair of different bases, the coordinate matrices are similar.

What is a vector, then, in this new language? It is nothing else — remember that we can identify vectors and tensors $(1,0)$— than an application that maps ordered bases onto ordered collections of $n$ numbers —given that $n$ is the dimension of the vector space—; what we always called coordinates of the vector are now seen as the result of applying a map —the vector itself— to a basis. Same exact concept, brand new vision.

## 2.6 TYPE-CHANGING

In physics, the usage of tensors of order 2 over $\mathbb{R}^3$ is made without explicitly saying nothing about its covariance or contravariance. This can be made only on those vector spaces that are adorned with a metric, which gives us the possibility to build isomorphisms between dual spaces and induce them between tensor spaces.

From now on, let us consider the metric vector space $(V, g)$, where $V$ is an $n$-dimensional vector space over $\mathbb{R}$ and $g\colon V \times V \to \mathbb{R}$ is a nondegenerate metric; i.e., $g$ is a nondegenerate symmetric $(0,2)$-tensor.

This addition to the naked vector space we have had until now will let us find an isomorphism between $V$ and its dual, $V^*$ without using bases; however, this is not a *natural* isomorphism as the one seen on Equation 2.1, as long as we need the metric in order to build it.

**Proposition 2.14.** *Let $v \in V$ and $\varphi \in V^*$ be a vector and a one-form.*

*Consider now the one-form $v^\flat \colon V \to \mathbb{R}$ defined by $v^\flat(w) = g(v, w)$ and the vector $\varphi^\sharp \in V$ implicitly defined by $g(\varphi^\sharp, w) = \varphi(w)$ $\forall w \in V$. Then:*

1. *$v^\flat \in V^*$ and $\varphi^\sharp \in V$.*
2. *$\flat \colon V \to V^*$ and $\sharp \colon V^* \to V$ are linear.*
   $v \mapsto v^\flat \qquad \varphi \mapsto \sharp^\flat$
3. *$\flat \circ \sharp = 1_{V^*}$ and $\sharp \circ \flat = 1_V$.*

This proposition, whose trivial proof can be found on [22, Proposition 9.30], let us define the so-called musical isomorphisms, that gives us the possibility to completely identify $V$ and $V^*$ —this identification depends on the metric $g$—:

**Definition 2.15** (Musical isomorphisms)**.** The isomorphisms $\flat \colon V \to V^*$ and $\sharp \colon V^* \to V$, defined on Proposition 2.14, are called *musical isomorphisms induced by g*. We call $\flat$ *flat* and $\sharp$ *sharp*.

These isomorphisms also allow us to translate geometrical objects from $V$ to $V^*$; particularly, we can see the induced metric $g^*$ on $V^*$:

**Proposition 2.16.** *The application $g^*$ that maps every $\varphi, \psi \in V^*$ on $g^*(\varphi, \psi) = g(\varphi^\sharp, \psi^\sharp)$ is a metric on $V^*$. Furthermore, $g^*$ is the only metric that allows $\flat$ and $\sharp$ to be isometries.*

Weird expression

*Proof.* Let us see that $g^*$ is a metric; i.e.that it is bilinear and symmetric: the fact that $g^*$ is bilinear comes from the facts that $g$ is bilinear and $\sharp$ is linear; using that $g$ is symmetric, we can write $g^*(\varphi, \psi) = g(\varphi^\sharp, \psi^\sharp) = g(\psi^\sharp, \varphi^\sharp) = g^*(\psi, \varphi)$, which proves that $g^*$ is symmetric.

$\flat$ is an isometry —and such is $\sharp$— given that

$$g^*(v^\flat, w^\flat) = g\left(\left(v^\flat\right)^\sharp, \left(w^\flat\right)^\sharp\right) = g(v, w). \tag{2.8}$$

Weird expression

If $g'$ is another metric that allows $\flat$ to be an isometry, then

$$g'(v^\flat, w^\flat) = g(v, w).$$

But using Equation 2.8 we see that $g^*(v^\flat, w^\flat) = g'(v^\flat, w^\flat)$. Given that $\flat$ is one-to-one we conclude that $g' = g^*$. ∎

The metrics $g$ and $g^*$ allow us to establish a connection between the coordinates of $v$ and $v^\flat$ and between the coordinates of $\varphi$ and $\varphi^\sharp$.

Let us consider $B = (v_1, \dots, v_n)$ and $B^* = (\varphi^1, \dots, \varphi^n)$ ordered bases of $V$ and its dual.

Let $(a^1, \dots, a^n)$ be the coordinates of $v \in V$ in $B$ and let $(b_1, \dots, b_n)$ be the coordinates of $v^\flat \in V^*$ in $B^*$. It is easy to see that

$$b_j = \sum_{k=1}^{n} g(v_j, v_k) a^k \quad \forall j \in \{1, 2, \dots, n\}. \tag{2.9}$$

Similary, if $(c_1, \dots, c_n)$ are the coordinates of $\varphi \in V^*$ in $B^*$ and $(d^1, \dots, d^n)$ are the coordinates of $\varphi^\sharp \in V$ in $B$, then we can write

$$d^j = \sum_{k=1}^{n} g^*(\varphi^j, \varphi^k) c_k \quad \forall j \in \{1, 2, \dots, n\}. \tag{2.10}$$

From this equations we can easily see that

$$M_{B^*}(g^*) = M_B(g)^{-1},$$

where $M_B(g)$ is the matrix associated to the bilinear form $g$ on the basis $B$.

Finally, we can see that if $B$ is an orthonormal basis of $(V, g)$, then $B^*$ is an orthonormal basis of $(V, g^*)$ and the previous expressions are even easier: $b_j = a^j$ and $d^j = c_j$.

Going back to the beginning of this section, we will see now how we could build isomorphisms between all tensors of order 2. Theorem 2.17 can be, of course, generalized in oder to find isomorphisms between all tensors of order $p \in \mathbb{N}$.

**Theorem 2.17.** *Let $T \in \mathcal{T}_{(0,2)}(V)$ be a $(0,2)$-tensor. If we consider $T'$ and $T''$, defined as*

$$\begin{aligned} T' \colon V^* \times V &\to \mathbb{R} \\ (\varphi, v) &\mapsto T(\varphi^\sharp, v) \end{aligned} \quad \textit{and} \quad \begin{aligned} T'' \colon V^* \times V^* &\to \mathbb{R} \\ (\varphi, \psi) &\mapsto T(\varphi^\sharp, \psi^\sharp) \end{aligned}$$

*then:*

1. $T' \in \mathcal{T}_{(1,1)}(V)$ *and* $T'' \in \mathcal{T}_{(2,0)}(V)$.
2. *The application that maps $T \mapsto T'$ —resp. $T \mapsto T''$— is an isomorphism between $\mathcal{T}_{(0,2)}(V)$ and $\mathcal{T}_{(1,1)}(V)$ —resp. between $\mathcal{T}_{(0,2)}(V)$ and $\mathcal{T}_{(2,0)}(V)$—.*

*Proof.* ■ DO IT

**Proposition 2.18.** *Let $B = (v_1, \dots, v_n)$ be an ordered basis of $(V, g)$ and let $B^* = (\varphi^1, \dots, \varphi^n)$ be its dual basis. Consider a tensor $T \in \mathcal{T}_{(0,2)}(V)$ that is defined as follows:*

$$T = \sum_{i,j}^{n} t_{ij} \varphi^i \otimes \varphi^j$$

*Then, if the tensors $T'$ and $T''$ from Theorem 2.17 have the coordinates*

$$T' = \sum_{i,j}^{n} t'^{i}_{j} v_i \otimes \varphi^j \quad \textit{and} \quad T'' = \sum_{i,j}^{n} t''^{ij} v_i \otimes v_j,$$

*the expressions that relate the tensors are given as*

$$t'^{i}_{j} = \sum_{k=1}^{n} g^*(\varphi^k, \varphi^i) t_{kj} \quad \textit{and} \quad t''^{ij} = \sum_{k,l=1}^{n} g^*(\varphi^k, \varphi^i) g^*(\varphi^l, \varphi^j) t_{kl}.$$

*Proof.* Given that

$$(\varphi^i)^\sharp = \sum_k d^k v_k = \sum_k (\sum_l g^*(\varphi^k, \varphi^l) c_l) v_k = \sum_k g^*(\varphi^k, \varphi^i) v_k,$$

it is easy to conclude the first expression:

$$\begin{aligned} t'^{i}_{j} &= T'(\varphi^i, v_j) = T((\varphi^i)^\sharp, v_j) = T(\sum_k g^*(\varphi^k, \varphi^i) v_k, v_j) = \\ &= \sum_{k=1}^n g^*(\varphi^k, \varphi^i) T(v_k, v_j) = \sum_k g^*(\varphi^k, \varphi^i) t_{kj} \end{aligned}$$

The second one is very similar:

$$\begin{aligned} t''^{ij} &= T''(\varphi^i, \varphi^j) = T((\varphi^i)^\sharp, (\varphi^j)^\sharp) = \\ &= T(\sum_k g^*(\varphi^k, \varphi^i) v_k, \sum_l g^*(\varphi^l, \varphi^j) v_l) = \\ &= \sum_{k,l=1}^n g^*(\varphi^k, \varphi^i) g^*(\varphi^l, \varphi^j) T(v_k, v_l) = \\ &= \sum_{k,l=1}^n g^*(\varphi^k, \varphi^i) g^*(\varphi^l, \varphi^j) t_{kl} \end{aligned}$$

■

*Remark* 2.19. From now on, in order to ease the notation, we will write the value of the metric in the elements of the basis as follows:

$$\begin{aligned} g_{ij} &:= g(v_i, v_j) \\ g^{ij} &:= g^*(\varphi^i, \varphi^j) \end{aligned}$$

*Remark* 2.20. Of course, these two expressions can be *inverted*, in order to obtain the coordinates of a $(0,2)$-tensor from the coordinates of tensors of type $(1,1)$ or $(2,0)$. Multiplying conveniently by the metric we obtain the inverted expressions. Consider the first expression, $t'^{i}_{j} = \sum_{k=1}^n g^*(\varphi^k, \varphi^i) t_{kj}$, multiply by $g_{il}$ and sum over $i$:

$$\begin{aligned} t'^{i}_{j} &= \sum_{k=1}^n g^{ki} t_{kj} \\ \sum_{i=1}^n g_{il} t'^{i}_{j} &= \sum_{k=1}^n \left( \sum_{i=1}^n g_{il} g^{ki} \right) t_{kj} \\ \sum_{i=1}^n g_{il} t'^{i}_{j} &= \sum_{k=1}^n \delta^k_l t_{kj} \\ \sum_{i=1}^n g_{il} t'^{i}_{j} &= t_{lj} \end{aligned}$$

We can obtain the inverted version of the second expression in a similar way, now multiplying by $g_{im}g_{jp}$ and summing over $i$ and $j$:

$$
\begin{aligned}
t''^{ij} &= \sum_{k,l=1}^{n} g^{ki}g^{lj}t_{kl} \\
\sum_{i,j=1}^{n} g_{im}g_{jp}t''^{ij} &= \sum_{i,j,k,l=1}^{n} \left(g_{im}g^{ki}\right)\left(g_{jp}g^{lj}\right)t_{kl} \\
\sum_{i,j=1}^{n} g_{im}g_{jp}t''^{ij} &= \sum_{k,l=1}^{n} \delta^k_m\delta^l_p t_{kl} \\
\sum_{i,j=1}^{n} g_{im}g_{jp}t''ij &= t_{mp}
\end{aligned}
$$

Fixing the indices to look nicer and leaving out the primes —which is somehow an abuse of the notation, but that cannot lead to any errors— we have the final expressions for both changes:

$$
t_{ij} = \sum_{k=1}^{n} g_{ki}t^k_j \quad \text{and} \quad t_{ij} = \sum_{k,l=1}^{n} g_{ki}g_{lj}t^{kl}
$$

Although we have seen Proposition 2.18 for tensors of order 2, the result is easily generalized to tensors of any other order. This provides us with a great tool to identify tensors of the same order through the musical isomorphisms.

This whole identification, which is in practice a technique consisting of multiplying by the metric in order to obtain the same tensor with a different type —as shown in Remark 2.20— is a well-known operation known as *raise and lower indices*. This operation will become extremely important on the differential geometry discussion, which will be key to this work.

But before going ahead, let us ease even more the identification between tensors. Let us consider again the example of tensors of order 2, but now with $B$ an orthonormal basis; we know from previous discussions that, in this case, $B^*$ is also orthonormal; i. e., the metric on the elements of the basis is

$$
g^{ij} = \delta^i_j.
$$

Then, the expressions for the change of coordinates, following the notation of Remark 2.20, are even easier:

$$
t^i_j = t_{ij} = t^{ij}
$$

We can finish this section with an interesting example on the use of the musical isomorphisms.

**Example 2.21.** Let $T \in \mathcal{T}_{1,3}(V)$ be a $(1,3)$-tensor. From Theorem 2.17 we know we can define $\tilde{T} \in \mathcal{T}_{(0,4)}(V)$ from $T$ as follows:

$$\tilde{T}(w_1, w_2, w_3, w_4) = T(w_1^\flat, w_2, w_3, w_4).$$

In terms of components, we can write $T$ using a basis of $\mathcal{T}_{1,3}(V)$, namely $\{v_m \otimes \varphi^i \otimes \varphi^j \otimes \varphi^k\}$, where every index moves independently from 1 to $n$. If we denote its coordinates as $t^m_{jkl}$, its expression is

$$T = \sum_{j,k,l,m}^{n} t^m_{jkl} v_m \otimes \varphi^i \otimes \varphi^j \otimes \varphi^k.$$

We can do the same with the tensor $\tilde{T}$ which will have its own coordinates $\tilde{t}_{ijkl}$ on a basis $\{\varphi^i \otimes \varphi^i \otimes \varphi^j \otimes \varphi^k\}$ of $\mathcal{T}_{0,4}(V)$:

$$T = \sum_{i,j,k,l}^{n} \tilde{t}_{ijkl} \varphi^i \otimes \varphi^i \otimes \varphi^j \otimes \varphi^k.$$

Let us finish this example with the most interesting equation we can obtain, the expression of the $\tilde{T}$ coordinates in terms of the coordinates of $T$. We just have to lower an index, as seen at Remark 2.20:

$$\tilde{t}_{ijkl} = \sum_{m=1}^{n} g_{im} t^m_{jkl}.$$

Using the Einstein summation convention and assuming the coordinates are the same, the expression is even cleaner:

$$t_{ijkl} = g_{im} t^m_{jkl}.$$

# 3 INTRODUCTION TO DIFFERENTIAL GEOMETRY

## 3.1 DIFFERENTIABLE MANIFOLDS

Roughly speaking, a manifold is a topological space that, locally, looks like the Euclidean space $\mathbb{R}^n$. This similitude is essential, and will let us control the manifold as if we were working in the Euclidean space; generally, the definitions concerning manifolds and the properties proved from them will be based on the known properties of $\mathbb{R}^n$.

The following definition specifies the formal concept of a topological manifold:

**Definition 3.1** (N-dimensional topological manifold)**.** Let $M^n$ be an $n$-dimensional topological space. The space $M^n$ is called a topological manifold if the following properties are satisfied:

1. $M^n$ is locally homeomorphic to $\mathbb{R}^n$.
2. $M^n$ is a Hausdorff space.
3. $M^n$ has a countable topological basis.

The first property states that, for every point $p \in M^n$, there exists an open neighbourhood $U \subset M^n$ of $p$ and a homeomorphism

$$h \colon U \to V$$

with $V \subset \mathbb{R}^n$ an open set.

One could think that the Hausdorff property is redundant, as the local homeomorphism may imply this topological characteristic. This is not true, and the usual counterexample is the line with two origins.

Let $M = \mathbb{R} \cup p$ be the union of the real line and a point $p \notin \mathbb{R}$. Define a topology in this space with $\mathbb{R} \subset M$ as an open set and the neighbourhoods of $p$ being the sets $(U \setminus \{0\}) \cup \{p\}$, where $U$ is a neighbourhood of $0 \in \mathbb{R}$. This space is locally Euclidean but not Hausdorff: the intersection of any two neighbourhoods of the points $0 \in \mathbb{R}$ and $p$ is non-empty.

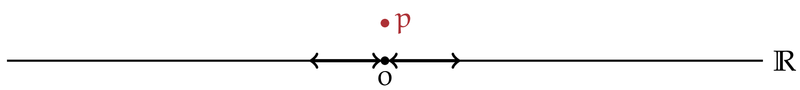


Figure 3.1: Line with two origins.

The last property of the definition will be proven key in our study, as it will let us define metrics on the manifold.

### 3.1.1 *Charts*

Is this word ok?

Maybe too much

Mmm... not sure about the word.

The main characteristic of the manifolds, its ressemblance to the Euclidean space, have to be exploited in order to understand the nature of the mathematical object.

The conceptual space where the manifolds live can be thought as the Plato's world of Ideas, where everything is pure but cannot be understood without studying particular examples.

The idea of the manifold will be understood, then, taking pieces of the manifold and lowering them to the real word; i. e., the Euclidean space, where we will be able to *physically* touch the manifold.

The essential tool to make this happen will be the coordinate charts. These tools are like prisms to see the manifold from the Euclidean perspective, and they will let us grasp the nature of the ideal concept of a manifold.

**Definition 3.2** (Coordinate chart)**.** A *coordinate chart* —or *coordinate system*— in a topological manifold $M^n$ is a homeomorphism $h\colon U \to V$ from an open subset of the manifold $U \subset M$ onto an open subset of the Euclidean space $V \subset \mathbb{R}^n$.

We call $U$ a *coordinate neighbourhood* in $M$.

One single chart may not cover the whole manifold. In order to completely understand it, we need a set of charts that describe it completely.

**Definition 3.3** (Coordinate atlas)**.** Let

$$A = \{h_\alpha \colon U_\alpha \to V_\alpha / \alpha \in I\}$$

be a set of coordinate charts in a topological manifold $M^n$, where $I$ is a family of indices and the open subsets $U_\alpha \subset M$ are the corresponding coordinate neighbourhoods.

$A$ is said to be an *atlas* of $M$ if every point is covered with a coordinate neighbourhood; i. e., if $\cup_{\alpha \in I} U_\alpha = M$.

### 3.1.2 *Differentiable structures*

The concept of manifold is quite general and includes a vast set of examples. We can impose, however, some properties on the smoothness of the manifold to restrict the objects we will work with.

This section introduces the notion of differentiable structure, whose definition is key in the later description of differentiable manifolds, the core concept of this chapter.

The first question in this study is the following: a chart describe perfectly a single piece of the manifold, but what happens when the domains of a pair of charts overlap? The following two definitions specify the concepts involved in this question.

**Definition 3.4** (Transition map)**.** Let $M^n$ be a manifold and $(U, \phi)$, $(V, \psi)$ a pair of coordinate charts in $M^n$ with overlapping domains, that is:

$$U \cap V \neq \emptyset$$

The homeomorphism between the open sets of the Euclidean space $\mathbb{R}^n$,

$$\psi \circ \phi^{-1} \colon \phi(U \cap V) \to \psi(U \cap V),$$

is called a *transition map*.

**Definition 3.5** (Smooth overlap)**.** Two charts $(U, \phi)$, $(V, \psi)$ are said to overlap smoothly if their domains are disjoint —i. e., if $U \cap V = \emptyset$— or if the transition map $\psi \circ \phi^{-1}$ is a diffeomorphism.

The description of two charts that overlap smoothly can be naturally extended to the concept of smooth atlas, that will make possible to do calculus on the manifold.

**Definition 3.6** (Smooth coordinate atlas)**.** An atlas $A$ is said to be smooth if every pair of charts in $A$ overlap smoothly.

But what happens if we define two different atlases in the manifold? Will the calculus depend on this choice? Fortunately we can find, for each manifold, one particular atlas that contain every other atlas defined there. It is formally described in the following definition and its uniqueness is proved in Proposition 3.8.

**Definition 3.7** (Complete atlas)**.** A *complete atlas* —or *maximal atlas*— on $M^n$ is a smooth atlas that contains each coordinate chart in $M^n$ that overlaps smoothly with every coordinate chart in $M^n$.

**Proposition 3.8** (Complete atlas uniqueness)**.** *Let $M^n$ be a topological manifold.*

- *Every smooth atlas on $M^n$ is contained in a complete atlas.*
- *Two smooth atlas on $M^n$ determine the same complete atlas if and only if its union is a smooth atlas.*

*Proof.* Let $A$ be a smooth atlas on $M^n$ and define $A'$ as the set of all $n$-dimensional coordinate charts that overlaps smoothly with every chart on $A$. We are going to see that $A'$ is a complete atlas.

It is trivial to see that $A'$ is an atlas, since $A \subset A'$ and $A$ is an atlas. The smoothness of the atlas is a consequence of the fact that smoothness is a local property ■ Finish.

**Definition 3.9** (Differentiable manifold)**.** A *differentiable manifold* is a pair $(M, A)$, where $M$ is a topological manifold and $A$ is a complete atlas.

**Example 3.10.** The concept of differentiable manifold is, probably, the most important idea throughout all this work. Let us see then some examples in order to better understand that these spaces we will going to work with are not that abstract —although they can be—.

1. The Euclidean space $\mathbb{R}^n$ is a differentiable manifold considering the identity map as its atlas.
2. Every *smooth surface*[1] of $\mathbb{R}^3$ is an example of a differentiable manifold. As a subset of $\mathbb{R}^3$, the local homeomorphism, the Hausdorff property and the countable basis are trivial. Furthermore, the definition of smooth surface gives us for free the complete atlas.
3. The sphere $S^n$ is an $n$-dimensional differentiable manifold. As an atlas we can consider the union of the two stereographic projections onto $\mathbb{R}^n$ from the north and south poles.

### 3.1.3 *Differentiable maps on manifolds*

The concept of differentiable maps on manifolds is the first one in which we are going to generalize concepts from the Euclidean space using the local homeomorphism.

The idea is simple: we know how to build differentiable maps between open sets of $\mathbb{R}^n$, so we are going to define differentiability between manifolds going through the images of the coordinate neighbourhoods of the points.

As the differentiability is a local concept, being the manifolds locally Euclidean is enough to generalize it.

**Definition 3.11.** Let $F\colon M \to N$ be a map between two differentiable manifolds: $M$ and $N$. $F$ is said to be *differentiable* or *smooth* if the following conditions are satisfied:

1. There is a chart $(U, \varphi)$ for every point $p \in M$ and another one, $(V, \psi)$ for its image, $F(p) \in N$, such that $p \in U$, $F(p) \in V$ and $F(U) \subset V$.
2. The map $\psi \circ F \circ \varphi^{-1} : \varphi(U) \to \psi(V)$ is differentiable in the usual sense.

This definition includes also the case in which $M$, $N$ or even both of them are the euclidean spaces $\mathbb{R}^m$ and $\mathbb{R}^n$. There is no ambiguity between this and the euclidean definition of smoothness, as one can

1 We consider the definition of smooth surface seen in a basic course of curves and surfaces: a subset of $\mathbb{R}^3$ such that every point is covered by the image of a differentiable map whose restriction to an open subset containing the point is an homeomorphism and whose differential is a monomorphism.

take the identity map as coordinate chart when one of the manifolds is an euclidean space and the usual definition will be found.

From this definition it is trivial to prove that, if a family of smooth maps cover a manifold with the maps being equal where their images overlap, a unique smooth function that is equal to each individual map on its image can be built.

Furthermore, it is easy to see that the identity of a manifold, the coordinate charts and the composition of smooth functions are smooth. Smoothness also implies continuity.

As well as the definition of smoothness, the definition of diffeomorphism can be generalized to manifolds, being Definition 3.12 its formal expression.

**Definition 3.12** (Diffeomorphism)**.** A function $f\colon M \to N$ between two manifolds is said to be a *diffeomorphism* if it is a smooth bijective map with its inverse being also smooth.

When there exists such map, $M$ and $N$ are said to be diffeomorphic.

### 3.1.4 *Tangent space*

Once we know what a differentiable function is, the next step we need to take in order to set up a proper place to do calculus on manifolds is to define the differential.

First, let us remember some concepts about regular surfaces on $\mathbb{R}^3$. Let $S, S'$ be two regular surfaces on $\mathbb{R}^3$ and let $f\colon S \to S'$ be a differentiable map between them. The differential of $f$ on $p \in S$ was defined as a function that transforms tangent vectors to the first surface into tangent vectors to the second,

$$(df)_p\colon T_pS \to T_{f(p)}S'.$$

What can we learn from this? Our goal is to define the differential of a differentiable map between *manifolds*. It would be ideal that it generalizes the notion we already have about differentials on surfaces, so it is mandatory to first generalize the concept of tangent plane.

The tangent plane to a regular surface on one of its points $p$ is, as we know, the vector subspace of all the tangent vectors to the point. This vector space was shown to be isomorphic to the space of directional derivatives on $p$. Instead of trying to generalize the concept of tangent vector, the idea we will follow is to extend the notion of directional derivatives, building the new *tangent plane*-like space from these.

The usual directional derivative is a linear map that satisfies the Leibniz rule, so we are going to define a tangent vector to a manifold as an axiomatization of this concept: the derivation.

From now on, we will note the set of all the smooth real-valued functions on a manifold $M$ as $\mathcal{F}(M)$:

$$\mathcal{F}(M) := \{f\colon M \to \mathbb{R} / f \text{ is smooth}\}$$

**Definition 3.13** (Derivation)**.** Let $p$ be a point on a manifold $M$. A *derivation* at $p$ is a map

$$D_p \colon \mathcal{F}(M) \to \mathbb{R}$$

that is linear and leibnizian; i. e., that satisfies the following properties:

1. $D_p(af + bg) = aD_p(f) + bD_p(g)$, where $a, b \in \mathbb{R}$ and $f, g \in \mathcal{F}(M)$.
2. $D_p(fg) = D_p(f)g(p) + f(p)D_p(g)$, where $f, g \in \mathcal{F}(M)$.

Taking into account the one-to-one correspondence between tangent vectors and derivations on the euclidean case —the directional derivative is actually a derivation—, the idea of the generalization of tangent vector on Definition 3.14 is more clear now.

**Definition 3.14** (Tangent vector)**.** Let $M$ be a manifold and $p \in M$ one of its points. A *tangent vector to M on p* is a derivation at $p$.

It is trivial to see that the directional derivative is a tangent vector to the well-known manifold $\mathbb{R}^n$. Being this *derivation — tangent vector* duality clear, it is now natural to arrive to Definition 3.15.

**Definition 3.15** (Tangent space)**.** Let $M$ be a manifold and $p \in M$ one of its points. The *tangent space to* $M$ *at* $p$, noted as $T_pM$, is the set of all tangent vectors to $M$ on $p$; i. e., the family of derivations at $p$.

*Remark* 3.16. $T_pM$ is a vector space with the usual definitions of function addition and product by a scalar, and if $x \colon U \to M$ is a chart that covers $p$, $\{\frac{\partial}{\partial x^1}\big|_p, \dots, \frac{\partial}{\partial x^n}\big|_p\}$[2] is its associated basis on $T_pM$. See [4, p. 8] for details.

As for every vector space, we can define its dual version.

**Definition 3.17** (Cotangent space)**.** Let $M$ be a manifold and $p \in M$ one of its points. The *cotangent space to* $M$ *at* $p$, denoted as $T_pM^*$ is the dual space of the vector space $T_pM$.

The elements $\omega \in T_pM^*$ are called *one-forms* or *covectors* on $p$.

The extension of the idea of differential is now straightforward: we have just to remember how the differential on the euclidean case can be defined from derivations and repeat the nearly exact same definition on manifolds.

**Definition 3.18** (Differential or pushforward)**.** Let $M$ and $N$ be two manifolds and let $F \colon M \to N$ be a smooth map.

2 We note by $\frac{\partial}{\partial x^i}\big|_p \colon \mathcal{F}(M) \to \mathbb{R}$ the function that maps every smooth function $f$ to the value of its derivative with respect to the $i$-th coordinate evaluated at $p$.

Consider, for each $p \in M$, the function

$$\begin{aligned} dF\colon T_pM &\to T_{F(p)}M \\ X &\mapsto dF(X), \end{aligned}$$

that maps each tangent vector to $M$ at $p$, $X$, to a tangent vector to $N$ at $F(p)$, $F_*X$, defined as follows:

$$\begin{aligned} dF(X)\colon \mathcal{F}(M) &\to \mathbb{R} \\ f &\mapsto X(f \circ F). \end{aligned}$$

The function $dF$ is the differental of $F$ at $p$, which is also known as the *pushforward* of $p$ by $F$.

On the $\mathbb{R}^3$ surfaces scenario, it is not odd to define tangent vectors using their close relation with the curves on the surface. In order to obtain a better understanding of the manifolds tangent space, let us see what a curve on a manifold is and how a tangent vector on a point can be identified with them.

**Definition 3.19** (Curve on a manifold)**.** Let $M$ be a manifold and $I \subset R$ an open set on $\mathbb{R}$. A *curve on* $M$ is a continuous map

$$\gamma\colon I \to M.$$

Every smooth curve is differentiable in the manifold sense, and having understood the duality between derivations and tangent vectors, we can naturally obtain the tangent vector to a curve on an instant $t_0 \in I$ by applying the definition we just saw.

**Proposition 3.20** (Tangent vector to a curve)**.** *The tangent vector to a curve* $\gamma\colon I \to M$ *on an instant* $t_0 \in I$*, noted as* $\gamma'(t_0) \in T_{\gamma(t_0)}M$ *is the pushforward of* $t$ *by* $\gamma$*; i. e., the tangent vector to* $M$ *defined as*

$$\begin{aligned} \gamma'(t_0)\colon \mathcal{F}(M) &\to \mathbb{R} \\ f &\mapsto \frac{d}{dt}(f \circ \gamma)(t_0) \end{aligned}$$

Proposition 3.20 tells us how to assign a vector from the tangent space of a manifold $M$ to every curve $\gamma$ on it, but is there a curve that could be assigned to every tangent vector on $M$?; i. e., is every element of the tangent space to $M$ the tangent vector of a curve? The following result answers this question.

**Theorem 3.21.** *Let* $p$ *be a point on a manifold* $M$*. There exists, for every* $X \in T_pM$*, a smooth curve on* $M$ *whose tangent vector is* $X$*.*

*Proof.* If $\varphi$ is the manifold chart that covers $p$ and $X = (X^1, \dots, X^n)$ are the coordinates of an element of the tangent space, then we can define

$$\gamma(t) = \varphi^{-1}(tX^1, \dots, tX^n)$$

in such a way that it is smooth on $\gamma(0) = p$ and that its tangent vector is $\gamma'(0) = X$. ∎

## 3.2 VECTOR FIELDS

In our journey to understand geometry on manifolds, one key step is to generalize what we called directional derivative in the euclidean spaces. The directional derivative of a function on a point gives us information on how the function changes when moving in the given direction; the concept of geodesic will need of this idea, but first we have to set up some definitions and technical results.

Let us start, then, by generalizing the concept of vector field to manifolds. As in the euclidean sense, we can define a *vector field* on a manifold $M$ as a correspondence $X$ that maps every point $p$ on the manifold to a vector $X(p)$ in the tangent space $T_pM$.

To formalize this concept, we should first define the set of the tangent spaces at every point of the manifold, which is the target set of the map we just described. This definition can be found on [16, p. 26] and [4, p. 13]

**Definition 3.22** (Tangent bundle)**.** Let $M$ be a smooth manifold and let $A = \{(U_\alpha, h_\alpha)\}$ be a smooth atlas on $M$.

Consider now the set

$$TM = \bigcup_{p \in M} T_pM,$$

where the projection $\pi \colon TM \to M$ maps every tangent vector $v$ to $p$, the manifold point such that $v \in T_pM$.

We can furnish $TM$ with the atlas $A' = \{(\pi^{-1}(U_\alpha), h'_\alpha)\}$, where $h'_\alpha$ defines the coordinates of every point $v \in TM$, as the union of the coordinates of $p(= \pi(v))$ in $U_\alpha$ with the coordinates of $v$ in the associated basis of $T_pM$; i. e., if $(x^1, \ldots, x^n)$ are the coordinate functions that assigns every point $p \in M$ to its coordinates on $\mathbb{R}^n$, the coordinates of the elements of $TM$ are

$$(x^1 \circ \pi, \ldots, x^n \circ \pi, \frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}),$$

where $\frac{\partial}{\partial x^i} \colon \pi^{-1}(U_\alpha) \to \mathbb{R}$ are the coordinate functions of the tangent space, given by $\frac{\partial v}{\partial x^i} = v(x^i)$.

$TM$ is called the tangent bundle of $M$.

It can be proved —see [4, Example 2.1] or [16, pp. 26, 27]— that $A'$ is, indeed, an atlas and, therefore, that the tangent bundle of every smooth manifold of dimension $n$ is in turn a smooth manifold of dimension $2n$.

*Remark* 3.23 (Cotangent bundle). The dual version of the tangent bundle, the *cotangent bundle*, can also be defined, and has the expected properties. It is defined as

$$T^*M = \bigcup_{p \in M} T_pM^*.$$

This recently defined space is the target set of what we described as a vector field. Definition 3.24 formalize this idea.

**Definition 3.24** (Vector field)**.** A *vector field* $X$ in a smooth manifold $M$ is a map

$$X\colon M \mapsto TM$$

that maps every point $p$ on the manifold to a vector $X(p) \in T_pM$, often noted as $X_p$.

From now on, we will note the set of smooth vector fields on $M$ as $\mathfrak{X}(M)$.

Analogously, one can define the set of smooth covector fields, that is, the set of functions assigning a one-form to each point of the manifold:

$$\mathfrak{X}^*(M) = \left\{ \begin{array}{l} \theta\colon M \to T^*M \\ \qquad p \mapsto \theta_p \in T_pM^* \end{array} \middle/ \theta \text{ is smooth} \right\}$$

Another interesting way to look at the vector fields, shown on [4, p. 23], consists on considering again the idea of the vectors as directional derivatives: a vector field $X$ on $p$ is then a map that receives a smooth function $f$ on $M$ and gives us another function on $M$, noted as $Xf$ and defined as follows:

$$(Xf)(p) = X_p(f).$$

One can define an interesting operation on vector fields considering them as derivations: the bracket operation.

**Definition 3.25** (Bracket operation)**.** Let $V$ and $W$ be vector fields on $M$.

The bracket operation on $V$ and $W$, noted as $[V, W]$ is the vector field defined as

$$[V, W] = VW - WV,$$

which is an application that maps every $f \in \mathcal{F}(M)$ to the function $V(Wf) - W(Vf)$.

The proof that $[V, W]$ is indeed a vector field can be found at [4, p. 24].

Going ahead with the generalization of euclidean concepts to the manifolds, we can define what a vector field along a curve is:

**Definition 3.26** (Vector field along a curve)**.** Let $c\colon I \to M$ be a curve on a manifold defined on the open subset $I \subset \mathbb{R}$. A *vector field along the curve* $c$, $V$, is a function

$$V\colon I \mapsto TM$$

that maps every instant $t \in I$ to a vector $X(c(t)) \in T_{c(t)}M$, where $X$ is a vector field.

**Example 3.27.** One of the most interesting examples of vector fields along a curve is the velocity of the curve itself. Let $\gamma\colon I \to \mathcal{M}$ be a smooth curve on M. The application that maps every instant to the tangent vector to the curve at that instant,

$$\begin{aligned}\gamma'\colon I &\to TM\\ t &\mapsto \gamma'(t),\end{aligned}$$

where $\gamma'(t)$ is defined on Proposition 3.20, is a vector field along $\gamma$.

## 3.3 PARTITIONS OF UNITY

A collection $\mathcal{L}$ of subsets of a space S is locally finite provided each point of S has a neighbourhood that meets only finitely many elements of $\mathcal{L}$. Let $\{f_\alpha\colon \alpha \in A\}$ be a collection of smooth functions on a manifold M such that $\{\operatorname{supp} f_\alpha\colon \alpha \in A\}$ is locally finite. Then the sum $\sum_\alpha f_\alpha$ is a well-defined smooth function on M , since on some neighbourhood of each point all but a finite number of $f_\alpha$ are identically zero.

**Definition 3.28** (Smooth partition of unity)**.** A *smooth partition of unity* on a manifold M is a collection $\{f_\alpha\colon \alpha \in A\}$ of functions $f_\alpha \in \mathfrak{F}(M)$ such that

1. $0 \leqslant f_\alpha \leqslant 1 \quad \forall \alpha \in A$.
2. $\{\operatorname{supp} f_\alpha\colon \alpha \in A\}$ is locally finite.
3. $\sum_\alpha f_\alpha = 1$.

The partition is said to be *subordinate* to an open covering $\mathfrak{C}$ of M provided each set $\operatorname{supp} f_\alpha$ is contained in some element of $\mathfrak{C}$.

Partitions of unity are an indispensable tool for assembling locally defined objects into a global object (or decomposing a global object into a sum of local objects). For such purposes partitions of unity with "small" supports are needed.

## 3.4 TENSOR FIELDS

Just as we introduce the concept of vector fields on Definition 3.24, a similar concept arises now: the tensor fields.

We are going to introduce this concept following the line of reasoning on [16, Ch. 2].

**Definition 3.29** (Tensor field)**.** A tensor field $A$ on a manifold M r times contravariant and $s$ times covariant is a classic tensor on the vector space $\mathfrak{X}(M)$, whose scalar field is $\mathfrak{F}(M)$, the set of smooth real valued functions on M:

$$A\colon \underbrace{\mathfrak{X}^*(M) \times \cdots \times \mathfrak{X}^*(M)}_{\text{r copies}} \times \underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{\text{s copies}} \to \mathfrak{F}(M).$$

As usual, we denote by $\mathcal{T}_{(r,s)}(M)$ the set of all tensor fields of type $(r, s)$ on a manifold $M$.

It is interesting to study that the name we gave to this concept is not random: indeed, we can see a tensor field $A$ as a proper *field*, in which every point of the manifold is mapped to a tensor.

The basis of this idea comes from the following result, whose proof can be studied on [16, Ch. 2, Proposition 2].

**Proposition 3.30.** *Let $p \in M$ be a point on a manifold $M$ and $A \in \mathcal{T}_{(r,s)}(M)$ a tensor field.*

*Let $\theta^i$ and $\bar{\theta}^i$ be covector fields for every $i \in \{1, \dots, r\}$, and such that*

$$\theta^i_{|_p} = \bar{\theta}^i_{|_p} \quad \forall i \in \{1, \dots, r\}.$$

*Similarly, let $X_i$ and $\bar{X}_i$ be vector fields for every $i \in \{1, \dots, s\}$, and such that*

$$X_{i|_p} = \bar{X}_{i|_p} \quad \forall i \in \{1, \dots, s\}.$$

*Then,*

$$A(\theta^1, \dots, \theta^r, X_1, \dots, X_s)(p) = A(\bar{\theta}^1, \dots, \bar{\theta}^r, \bar{X}_1, \dots, \bar{X}_s)(p).$$

This result let us consider each tensor field $A \in \mathcal{T}_{(r,s)}(M)$ as the following field on $M$, which is denoted in the same exact way:

$$\begin{aligned} A \colon M &\to \mathcal{T}_{(r,s)}(M) \\ p &\mapsto A_p \colon \underbrace{T_pM^* \times \cdots \times T_pM^*}_{\text{r copies}} \times \underbrace{T_pM \times \cdots \times T_pM}_{\text{s copies}} \to \mathbb{R} \end{aligned}$$

where the tensor $A_p$, now defined on the tangent and cotangent space, is the following mapping:

$$(\alpha^1, \dots, \alpha^r, x_1, \dots, x_s) \xmapsto{A_p} A(\theta^1, \dots, \theta^r, X_1, \dots, X_s),$$

where $\theta^i$ is any covector field such that $\theta^i_{|_p} = \alpha^i$ and $X_i$ is any vector field such that $X_{i|_p} = x_i$ for every $i \in \{1, \dots, n\}$.

The operation involving tensors, the definition of the tensor components, the tensor contraction and all other classic results thoroughly studied in chapter 2 hold also here for the tensor fields on a manifold.

## 3.5 CONNECTIONS ON MANIFOLDS

In this section we will define a connection on a manifold, which in turn will give us the tools to generalize the concept of directional derivative arriving to the definition of covariant derivative.

Finally, when adding a metric to the smooth manifold, an interesting connection will appear naturally: the Levi-Civita connection, whose interest on the study of the geodesic will be shown.

### 3.5.1 *Affine connections*

The following definitions and results can be found at [4, Ch. 2, Section 2] and [16, pp. 59-67].

**Definition 3.31** (Affine connection)**.** An affine connection $\nabla$ on a smooth manifold $M$ is a map

$$\nabla \colon \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M),$$

noted as $(X, Y) \xrightarrow{\nabla} \nabla_X Y$, that satisfies the following properties:

1. $\nabla_{fX+gY} Z = f\nabla_X Z + g\nabla_Y Z$,
2. $\nabla_X(Y + Z) = \nabla_X Y + \nabla_X Z$,
3. $\nabla_X(fY) = f\nabla_X Y + (Xf)Y$,

where $X, Y, Z \in \mathfrak{X}(M)$ and $f, g \in \mathcal{F}(M)$.

On [4, Chapter 2, Remark 2.3] we can see how the last property of Definition 3.31 let us show that the affine connection is a local concept. Consider a coordinate system $(x^1, \ldots, x^n)$ around $p$ and describe the vector fields $X, Y$ as follows:

$$X = \sum_i x^i X_i, \qquad Y = \sum_j y^j X_j,$$

where $X_i = \frac{\partial}{\partial x^i}$. Then, we can write

$$\nabla_x Y = \sum_i x^i \nabla_{X^i} \left( \sum_j y^j X_j \right) = \sum_{ij} x^i y^j \nabla_{X^i} X_j + \sum_{ij} x^i (X_i y^j) X_j.$$

As $(\nabla_{X^i} X_j)_p \in T_p M$, and using that $\{X_1(p), \ldots, X_n(p)\}$ is a basis of $T_p M$, we can write the coordinate expression of $\nabla_{X^i} X_j$ as follows:

$$\nabla_{X^i} X_j = \sum_k \Gamma^k_{ij} X_k,$$

where the functions $\Gamma^k_{ij}$ are necessarily differentiable. Finally, we can write

$$\nabla_X Y = \sum_k \left( \sum_{ij} x^i y^j \Gamma^k_{ij} + X(y^k) \right) X_k.$$

This shows that $\nabla_X Y(p)$ depends on $x^i(p)$, $y^k(p)$ and the derivatives $X(y^k)(p)$.

This is a somewhat technical definition, but as shown in Proposition 3.32, it provides us with the concept of covariant derivative, which will be shown to be a generalization of the directional derivative on $\mathbb{R}^n$.

**Proposition 3.32** (Covariant derivative)**.** *Let $M$ be a smooth manifold with an affine connection $\nabla$ and let $c\colon I \to M$ be a smooth curve. Then there is a unique function that maps each vector field $V$ along $c$ onto another vector field along $c$, called* covariant derivative of $V$ along $c$, *and noted as $\frac{DV}{dt}$, that satisfies the following properties:*

1. $\frac{D}{dt}(V+W) = \frac{DV}{dt} + \frac{DW}{dt}$.
2. $\frac{D}{dt}(fV) = \frac{df}{dt}V + f\frac{DV}{dt}$.
3. *If $V$ is described as $V(t) = X(c(t))$, where $X \in \mathfrak{X}(M)$, then*
$$\frac{DV}{dt} = \nabla_{\frac{dc}{dt}} X.$$

*where $W$ is another vector field along $C$ and $f \in \mathcal{F}(M)$.*

Proposition 3.32 gives us an actual derivation on vector fields along smooth curves. The concept of connection, whose definition may appear artificial at first, shows now its interest: it provides us with a way of derivating vectors along curves; i.e., we have now the possibility to consider the concept of *acceleration* on curves on manifolds.

*Proof of Proposition 3.32.* Assuming the existence of such a map, considering the local coordinates of $V$ and using the properties that define the covariant derivative, one can prove that it is unique.

Let $x\colon U \subset \mathbb{R}^n \to M$ be a coordinate chart that assigns the local expression $(x^1(t), \dots, x^n(t))$ to the curve $c$. If we note $X_i = \frac{\partial}{\partial x^i}$ and write the field $V$ locally as $V = \sum_j v^j X_j$, applying all three properties of the covariant derivative we conclude that

$$\frac{DV}{dt} = \sum_j \frac{dv^j}{dt} X_j + \sum_{i,j} \frac{dx^i}{dt} v^j \nabla_{X_i} X_j, \tag{3.1}$$

that is, the covariant derivative is unique.

On the other hand, defining the covariant derivative as in Equation 3.1, its existence and the satisfaction of the defined properties is straightforward. ∎

The remaining technical details of the previous reasoning can be found on [4, p. 43], from where this proof was taken.

**Definition 3.33** (Parallel vector field)**.** Let $M$ be a smooth manifold furnished with an affine connection $\nabla$. A vector field $V$ along a curve $c\colon I \to M$ is called *parallel* whenever $\frac{DV}{dt} = 0$ for every $t \in I$.

**Proposition 3.34** (Parallel transport)**.** *Let $M$ be a smooth manifold furnished with an affine connection $\nabla$. Let $c\colon I \to M$ be a smooth curve on $M$ and $V_0$ a tangent vector to $M$ on $c(t_0)$; i.e., $V_0 \in T_{c(t_0)}M$.*

*Then, there exists a unique parallel vector field $V$ along $c$ such that $V(t_0) = V_0$. We call $V$ the* parallel transport of $V(t_0)$ along $c$.

*Proof.* See [4, Proposition 2.6 on section 2.2] ∎

Write it if there's time

### 3.5.2 *Semi-Riemannian connections*

This section follows the line of reasoning on [4, Ch. 2, Section 3].

Introduction stating that a metric is now needed.

Until now, the geometry of

**Definition 3.35** (Metric compatible connection)**.** Let $M$ be a smooth manifold furnished with an affine connection $\nabla$ and a semi-Riemannian metric $g$.

The connection is said to be compatible with the metric when, for every smooth curve $c$ and for every pair of vector fields $P, P'$ along $c$, the product of the vector fields is constant:

$$g(P, P') = constant$$

This definition will let us compute the derivative of the vector fields product using the Leibniz rule, as shown in the following proposition.

**Proposition 3.36** (Derivative of the metric on vector fields)**.** *Let $M$ be a semi-Riemannian manifold furnished with a connection $\nabla$ compatible with the metric $g$. Let $c\colon I \to M$ be a smooth curve and $V$ and $W$ two vector fields along $c$. Then, the Leibniz rule is satisfied when computing the derivative of the vector fields product; that is:*

$$\frac{d}{dt} g(V, W) = g(\frac{DV}{dt}, W) + g(V, \frac{DW}{dt}), t \in I.$$

Write if time

*Proof.* See [4, Ch. 2, Section 3, Proposition 3.2]. ∎

In fact, we can redefine what a connection compatible with a metric is using Proposition 3.36.

**Corollary 3.37** (Compatible connection redefinition)**.** *An affine connection on a semi-Riemannian manifold $M$ is compatible with a metric $g$ if and only if*

$$X(g(Y, Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z), \quad X, Y, Z \in \mathfrak{X}(M).$$

*Proof.* The first implication is straightforward, let us see the other one.

Consider that $\nabla$ is compatible with the metric. Let $p \in M$ and define $c\colon I \subset \mathbb{R} \to M$ to be a smooth curve satisfying

$$c(t_0) = p, \quad t_0 \in I \quad \text{and} \quad \left.\frac{dc}{dt}\right|_{t=t_0} = X(p).$$

Then,

$$X(p)\,(g(Y, Z)) = \frac{d}{dt} g(Y, Z)\Big|_{t=t_0} = g(\nabla_{X(p)} Y, Z)_p + g(Y, \nabla_{X(p)} Z)_p,$$

which proves the result, as $p$ is an arbitrary point. ∎

**Definition 3.38** (Symmetric connection)**.** An affine connection $\nabla$ on a smooth manifold $M$ is said to be symmetric if

$$\nabla_X Y - \nabla_Y X = [X, Y] \quad \forall X, Y \in \mathfrak{X}(M).$$

The name chosen for this property may not be clear at first sight, but it is not in vain: if we consider a coordinate system $(U, x)$, the symmetric connections satisfy

$$\nabla_{X_i} X_j - \nabla_{X_j} X_i = [X_i, X_j] = 0,$$

which is equivalent to the symmetric expression

$$\Gamma^k_{ij} = \Gamma^k_{ji}.$$

Note that $\Gamma^k_{ij}$ and $\Gamma^k_{ji}$ are not equal in general. This happens only when $X_i = \frac{\partial}{\partial x^i}$ as we assume here. In this case, the Schwarz lemma shows us that $[X_i, X_j] = 0$; i. e., that $\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i}$.

We can now state the main theorem for this section, which, paraphrasing Barret O'Neill [16, p. 60], it is said to be the miracle of semi-Riemannian geometry.

**Theorem 3.39** (Levi-Civita connection)**.** *There exists, for every semi-Riemannian manifold $M$, a unique affine connection $\nabla$ satisfying the following properties:*

1. *$\nabla$ is symmetric.*
2. *$\nabla$ is compatible with the semi-Riemannian metric.*

*The connection $\nabla$ is known as the* Levi-Civita connection.

*Proof.* Assume that such connection exists; then, if $g$ is the metric on $M$, the following equalities hold:

$$Xg(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \tag{3.2}$$

$$Yg(Z, X) = g(\nabla_Y Z, X) + g(Z, \nabla_Y X) \tag{3.3}$$

$$Zg(X, Y) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y) \tag{3.4}$$

$$\tag{3.5}$$

Summing equalities 3.2 and 3.3, substracting 3.4 and using the symmetry of $\nabla$:

$$\begin{aligned} &Xg(Y, Z) + Yg(Z, X) - Zg(X, Y) = \\ =&g([X, Z], Y) + g([Y, Z], X) + g([X, Y], Z) + 2g(Z, \nabla_Y X) \end{aligned}$$

Therefore, we can extract an equality that proves that $\nabla$ is uniquely determined by the metric:

$$\begin{aligned} g(\nabla_Y X, Z) = -\frac{1}{2}\big(&Xg(Y, Z) + Yg(Z, X) - Zg(X, Y) - \\ &- g([X, Z], Y) - g([Y, Z], X) - g([X, Y], Z)\big). \end{aligned} \tag{3.6}$$

The proof of its existence is straightforward when defining the connection with Equation 3.6, commonly known as the *Koszul formula*. ∎

*Remark* 3.40 (Christoffel symbols, covariant and usual derivatives). When considering a coordinate system $(U, x)$, the functions $\Gamma^k_{ij}$ defined before with the expression $\nabla_{x^i} X_j = \sum_k \Gamma^k_{ij} X_k$ are known as the *coefficients of the connection* $\nabla$ on $U$ or, more commonly, as the *Christoffel symbols of the connection*.

From Equation 3.6 we can write

$$\sum_l \Gamma^l_{ij} g_{lk} = \frac{1}{2}\left(\frac{\partial}{\partial x^i} g_{jk} + \frac{\partial}{\partial x^j} g_{ki} - \frac{\partial}{\partial x^k} g_{ij}\right).$$

Taking into account the fact that $g_{km}$ has an inverse, namely $g^{km}$, and using the Einstein summation convention, we obtain the classic expression of the Christoffel symbols of the Levi-Civita connection:

$$\Gamma^m_{ij} = \frac{1}{2} g^{km}\left(\frac{\partial}{\partial x^i} g_{jk} + \frac{\partial}{\partial x^j} g_{ki} - \frac{\partial}{\partial x^k} g_{ij}\right) \tag{3.7}$$

We can reformulate Equation 3.1 in terms of the Christoffel symbols, obtaining the covariant derivative classical expression using again the Einstein notation:

$$\frac{DV}{dt} = X_k\left(\frac{dv^k}{dt} + \Gamma^k_{ij} v^j \frac{dx^i}{dt}\right). \tag{3.8}$$

There is still one question left: is the covariant derivative an actual generalization of the directional derivative on the euclidean space $\mathbb{R}^n$? If we inspect Equation 3.8 we realize that the only difference between directional and covariant derivatives is the term where the Christoffel symbols appear. But it is straightforward to see that these terms vanishes when the metric $g$ is euclidean; from Equation 3.7, and using that $g^{ij} = \delta^i_j$ for every $i, j \in \{1, \ldots, n\}$:

$$\begin{aligned}
\Gamma^m_{ij} &= \frac{1}{2} g^{km}\left(\frac{\partial}{\partial x^i} g_{jk} + \frac{\partial}{\partial x^j} g_{ki} - \frac{\partial}{\partial x^k} g_{ij}\right) = \\
&= \frac{1}{2}\sum_k g^{km}\left(\frac{\partial}{\partial x^i} g_{jk} + \frac{\partial}{\partial x^j} g_{ki} - \frac{\partial}{\partial x^k} g_{ij}\right) = \\
&= \frac{1}{2}\sum_k \delta^k_m\left(\frac{\partial}{\partial x^i} \delta^j_k + \frac{\partial}{\partial x^j} \delta^k_i - \frac{\partial}{\partial x^k} \delta^i_j\right) = \\
&= 0
\end{aligned}$$

We can conclude what we state at the beginning of this section: when restricted to the euclidean spaces, the covariant derivative and the directional derivative agree.

## 3.6 GEODESICS

From now on, let $M$ be a semi-Riemannian manifold.

### 3.6.1 *Basic definition*

**Definition 3.41** (Geodesic)**.** Let $\gamma\colon I \to M$ be a curve on M. $\gamma$ is said to be a *geodesic on* M when

$$\frac{D}{dt}\left(\frac{d\gamma}{dt}\right) = 0 \quad \forall t \in I,$$

that is, the vector field $\gamma'$ is parallel.

It is a common abuse of notation, using that the covariant derivative is an actual derivation, to write that a curve $\gamma$ on M is a geodesic when $\gamma'' = 0$.

One milestone on the development of this work is to obtain the equations that characterise geodesics: these kind of equations will let us apply a numerical algorithm in order to obtain positions of particles moving on manifolds. We will see later that a spacetime can be modelled as a 4-dimensional manifold whose geodesics will tell us the path light follows.

As our primary objective is to study the movement of light, this first small step in understanding what a geodesic is and what equation it satisfies is very important.

*Remark* 3.42 (Differential equations satisfied by a geodesic)*.* Therefore, let us see the local equations satisfied by a geodesic $\gamma$ on on a coordinate system $(U, x)$ around $\gamma(t_0)$, $t_0 \in I$.

Let $\gamma(t) = (x^1(t), \dots, x^n(t))$ be the coordinates of the curve on U. Using the expression of the covariant derivative and assuming $\gamma$ is a geodesic; i. e., that its velocity vector field is parallel, we can write:

$$0 = \frac{\partial}{\partial x^k}\left(\frac{d^2x^k}{dt^2} + \Gamma^k_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt}\right).$$

We conclude that the differential equations system of order 2 given by

$$\frac{d^2x^k}{dt^2} + \Gamma^k_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt} = 0 \quad k = 1, \dots, n \tag{3.9}$$

describes a necessary condition for the curves on M to be geodesics.

### 3.6.2 *Variational characterization of geodesics*

When working with geodesics, its natural, elegant definition, although really interesting for theoretical purposes, is not practical. This section aims to find a characterization that let us study geodesics.

First of all, we need to define some more concepts on curves.

**Definition 3.43** (Variation [4, Ch. 9, Definition 2.1])**.** Let $c\colon [0,1] \to M$ be a piecewise smooth curve on a semi-Riemannian manifold M. A *variation of* c is a continuous map

$$f\colon (-\varepsilon, \varepsilon) \times [0,1] \to M$$

that satisfies:

1. $f(0, t) = c(t), \quad t \in [0, 1]$.
2. There is a partition of $[0, 1]$, $0 = t_0 < t_1 < \cdots < t_{k+1} = 1$, such that $f_{|(-\varepsilon,\varepsilon)\times[t_i,t_{i+1}]}$ is smooth for every $i = 0, 1, \ldots, k$.

The variation is said to be *proper* if $f(s, 0) = c(0)$ and $f(s, 1) = c(1)$ for every $s \in (-\varepsilon, \varepsilon)$.

**Definition 3.44** (Variational curve of a c)**.** Let $f$ be a variation of a piecewise smooth curve $c$ on a semi-Riemannian manifold $M$, and fix an arbitrary $s \in (-\varepsilon, \varepsilon)$.

The curve

$$\begin{aligned} f_s \colon [0, 1] &\to M \\ t &\mapsto f_s(t) := f(s, t) \end{aligned}$$

c or f?

is called the variational curve of c.

The family of curves $\{f_s\}_{s\in(-\varepsilon,\varepsilon)}$ trivially includes $f_0 = c$, and so it can be considered as a set of close curves to $c$ if $\varepsilon > 0$ is sufficiently small.

Furthermore, tt is clear that the variation is proper if every curve on this family have the same initial point $c(0)$ and the same final point $c(1)$.

Curva transversal de la variación.

**Definition 3.45** (Transversal curve of variation)**.** Let $f$ be a variation of a piecewise smooth curve $c$ on a semi-Riemannian manifold $M$, and fix an arbitrary $t \in [0, 1]$.

The curve

$$\begin{aligned} f_t \colon (-\varepsilon, \varepsilon) &\to M \\ s &\mapsto f_t(s) := f(s, t) \end{aligned}$$

f or c?

is called the sectional curve of f .

The velocity vector of $f_t$ is

$$V(t) := \frac{\partial f}{\partial s}(0, t),$$

which defines a piecewise smooth vector field along $c$: the so-called *variational field* of $f$.

Furthermore, if $f$ is a proper variation, the velocity satisfies $V(0) = 0$ and $V(1) = 0$.

We can define now a quantity minimized by geodesics, in order to find a characterization with which we can work. In a semi-Riemannian manifold, this quantity will be the energy.

**Definition 3.46** (Energy of a variation)**.** Let $f$ be a variation of a smooth curve $c$ on a manifold $M$

The *energy of* $f$ is defined a

$$\begin{aligned} E_f \colon (-\varepsilon, \varepsilon) &\to \mathbb{R} \\ s \mapsto E_f(s) &:= \int_0^1 g\left(\frac{\partial f}{\partial t}(s,t), \frac{\partial f}{\partial t}(s,t)\right) dt, \end{aligned}$$

where $g$ is a metric on $M$ and $\frac{\partial f}{\partial t}(s,t)$ denotes the velocity of $f_s$ on the instant $t$.

Some interesting properties of the energy function are studied on the following proposition, which will help us to find the characterization we are looking for.

**Proposition 3.47.** *Let* $c\colon [0,1] \to M$ *a piecewise smooth curve and let* $f\colon (-\varepsilon,\varepsilon) \times [0,1] \to M$ *be a proper variation of* $c$.

*The energy of* $f$, $E_f\colon (-\varepsilon,\varepsilon) \to \mathbb{R}$ *satisfies:*

$$\frac{1}{2}E_f'(0) = -\int_0^1 g\left(V(t), \frac{D}{dt}\left(\frac{dc}{dt}\right)\right) dt - \sum_{i=1}^{k} g\left(V(t_i), \frac{dc}{dt}(t_i^+) - \frac{dc}{dt}(t_i^-)\right),$$

*where* $V$ *is the variational field of* $f$ *and*

$$\frac{dc}{dt}(t_i^+) = \lim_{\substack{t\to t_i \\ t>t_i}} \frac{dc}{dt}(t), \qquad \frac{dc}{dt}(t_i^-) = \lim_{\substack{t\to t_i \\ t<t_i}} \frac{dc}{dt}(t)$$

*Proof.* ■

Add proof or reference.

The characterization of geodesics we are looking for comes as a consequence of Proposition 3.47.

**Proposition 3.48.** *A piecewise smooth curve* $c\colon [0,1] \to M$ *is a geodesic if and only if*

$$\frac{dE_f}{ds}(0) = 0 \quad \textit{for every } f\textit{, proper variation of } c.$$

*Proof.* ■

Add proof or reference.

Geodesics are, in short, the critical points of the energy of every proper variation.

## 3.7 CURVATURE

**Lemma 3.49.** *Let* $M$ *be a semi-Riemannian manifold with Levi-civita connection* $\nabla$. *The function*

$$\begin{aligned} R\colon \mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) &\to \mathfrak{X}(M) \\ (X,Y,Z) &\mapsto R(X,Y)Z, \end{aligned}$$

*defined as*

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z,$$

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

*Proof.* To see the curvature $R$ as a $(1,3)$ tensor consider the function defined as

$$\begin{aligned} \tilde{R} \colon \mathfrak{X}(M)^* \times \mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) &\to \mathfrak{X}(M) \\ (\Omega, X, Y, Z) &\mapsto \Omega\left(R(X,Y)Z\right) \end{aligned}$$

It is straightforward to see that $\tilde{R}$, that will be denoted also as $R$, is multilinear and can factor out its scalars. ∎

The Riemannian curvature can be considered as an $\mathbb{R}$-multilinear function ay any $p \in M$ working on individual tangent vectors.

*Remark* 3.50 (Curvature tensor on tangent vectors). If $u, v, w \in T_pM$ and $U, V, W \in \mathfrak{X}(\Omega)$, where $\Omega$ is some open neighbourhood of $p$ such that

$$U_p = u, \qquad V_p = v, \qquad W_p = w,$$

then $(R(U,V)W)_p$ does not depend on the local extensions of the tangent vectors at $p \in M$ [16, p. 38], and therefore we can define

$$\begin{aligned} R \colon T_pM \times T_pM \times T_pM &\to T_pM \\ (u, v, w) &\mapsto R(u,v)w := (R(U,V)W)_p \end{aligned}$$

In particular, for any $u, v \in T_pM$ we have a linear operator

$$R(u,v) \colon T_pM \to T_pM$$

sending each $w \in T_pM$ to $R(u,v)w \in T_pM$, which is called the *curvature operator* defined by $u, v \in T_pM$.

Let us compute the expression of the curvature tensor components. Consider that its coordinate expression is:

$$R\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)\frac{\partial}{\partial x^k} = \sum_l R^l_{ijk}\frac{\partial}{\partial x^l}. \tag{3.10}$$

Using the curvature tensor definition and the Schwarz lemma, that tells us that $[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0$, the tensor has the expression

$$R\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)\frac{\partial}{\partial x^k} = \nabla_{\frac{\partial}{\partial x^i}}\nabla_{\frac{\partial}{\partial x^j}}\frac{\partial}{\partial x^k} - \nabla_{\frac{\partial}{\partial x^j}}\nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^k}.$$

Let us compute the first term in details; the second one will have the same expression with swapped indices $i$ and $j$. Using the Christoffel symbols, we have

$$\nabla_{\frac{\partial}{\partial x^j}}\frac{\partial}{\partial x^k} = \sum_h \Gamma^h_{jk}\frac{\partial}{\partial x^h},$$

from where we can write

$$\nabla_{\frac{\partial}{\partial x^i}}\left(\nabla_{\frac{\partial}{\partial x^j}}\frac{\partial}{\partial x^k}\right)=\sum_h \nabla_{\frac{\partial}{\partial x^i}}\Gamma^h_{jk}\frac{\partial}{\partial x^h}=$$
$$=\sum_h \frac{\partial}{\partial x^i}\Gamma^h_{jk}\frac{\partial}{\partial x^h}+\sum_h \Gamma^h_{jk}\nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^h}=$$
$$=\sum_h \frac{\partial}{\partial x^i}\left(\Gamma^h_{jk}\right)\frac{\partial}{\partial x^h}+\sum_{h,p}\Gamma^h_{jk}\Gamma^p_{ih}\frac{\partial}{\partial x^p}.$$

If we swap the indices $h$ and $p$ on the last addend, the expression of this one term is somewhat more clear:

$$\nabla_{\frac{\partial}{\partial x^i}}\nabla_{\frac{\partial}{\partial x^j}}\frac{\partial}{\partial x^k}=\sum_h\left(\frac{\partial}{\partial x^i}\Gamma^h_{jk}+\sum_p \Gamma^p_{jk}\Gamma^h_{ip}\right)\frac{\partial}{\partial x^h} \tag{3.11}$$

From Equation 3.10, using Equation 3.11 and its analogous with $i$ and $j$ indices swapped, we can finally write the expression of the curvature tensor components:

$$R^l_{ijk}=\frac{\partial}{\partial x^i}\Gamma^l_{jk}-\frac{\partial}{\partial x^j}\Gamma^l_{ik}+\sum_p\left(\Gamma^p_{jk}\Gamma^l_{ip}-\Gamma^p_{ik}\Gamma^l_{jp}\right)$$

The following identities are the symmetries of $R$:

**Proposition 3.51.** *If $u, v, w, z \in T_pM$, then*

1. $R(u,v) = -R(v,u)$ *as operators on* $T_pM$.
2. $g(R(u,v)w,z) = -g(R(u,v)z,w)$.
3. $R(u,v)w + R(v,w)u + R(w,u)v = 0$ (first Bianchi identity).
4. $g(R(u,v)w,z) = g(R(w,z)u,v)$.

These symmetries of $R$ lead to a less obvious symmetry of its covariant derivative $\nabla R$, called the *second Bianchi identity*. First, let us recall that $\nabla R$ is the $(1,4)$-tensor field defined by

$$\begin{aligned}(\nabla_U R)(X,Y)Z =& \nabla_U(R(X,Y)Z) - R(\nabla_U X, Y)Z - \\ &- R(X,\nabla_U Y)Z - R(X,Y)\nabla_U Z\end{aligned}$$

for all $U, X, Y, Z \in \mathfrak{X}(M)$.

**Definition 3.52** (Ricci tensor)**.** The Ricci tensor Ric of a semi-Riemannian metric $g$ is defined as the following contraction of the Riemann curvature tensor $R$ of $g$:

$$\mathrm{Ric}(X,Y) := \mathrm{trace}(V \mapsto R(V,X)Y),$$

that is, $\mathrm{Ric}(X,Y) = \sum_i \omega^i(R(X_i,X)Y)$; where $\{X_i\}$ is a local basis of vector fields and $\{\omega^i\}$ is its dual basis of one-forms.

Therefore, for a coordinate system $(x_1, \dots, x_n)$ we have

$$\mathrm{Ric} = \sum_{j,k} \mathrm{Ric}\left(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^j}\right) dx^j \otimes dx^k,$$

where the components $\mathrm{Ric}(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k})$ of Ric are given by

$$\mathrm{Ric}(\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}) = \sum_i dx^i \left(R(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})\frac{\partial}{\partial x^k}\right) = \sum_i R^i_{ijk}.$$

Equivalently,

$$\mathrm{Ric}(X, Y) = \sum_{i,h} g^{ih} g(R(\frac{\partial}{\partial x^i}, X)Y, \frac{\partial}{\partial x^h}).$$

**Definition 3.53** (Scalar curvature)**.** The *scalar curvature* $S$ of $g$ is the contraction of the $(1,1)$ tensor field $g$-equivalent to the Ricci tensor. In terms of coordinates:

$$S = \sum_{i,j} g^{ij} (\mathrm{Ric})_{ij} = \sum_{i,j,k} g^{ij} R^k_{kij}$$

From the second identity of Bianchi we get [16, p. 88]

$$\nabla S = 2\,\mathrm{div}\,\widehat{\mathrm{Ric}},$$

where $\nabla S$ is the gradient of te scalar curvature, $\widehat{\mathrm{Ric}}$ is the contravariant tensor field $g$-equivalent to Ric and div is the divergence acting on symmetric $(2,0)$ tensor fields.

## 3.8 KILLING VECTORS FIELDS

Add references to Oneill

**Definition 3.54** (Integral curve)**.** Let $M$ be a smooth manifold and consider a vector field $X \in \mathfrak{X}(M)$. An integral curve of $X$ is a curve

$$\gamma \colon I \to M,$$

where $I$ is an open interval of $\mathbb{R}$ with $0 \in I$, such that

$$\gamma'(t) = X_{\gamma(t)} \quad \forall t \in I, \tag{3.12}$$

that is, the vector field $X$ assigns to every point of the curve its own velocity at that instant.

**Lemma 3.55.** *Let $p_0 \in M$ be a point on a smooth manifold. There exists a unique (maximal) integral curve of $X \in \mathfrak{X}(M)$, $\gamma \colon I \to M$, such that*

$$\gamma(0) = p_0. \tag{3.13}$$

We are now able to see that an integral curve starting at a point on $M$ can be seen as the solution to the differential equation described by Equation 3.12 restricted to the initial conditions given by Equation 3.13.

**Lemma 3.56.** *Consider a vector field* $X \in \mathfrak{X}(M)$ *and a point* $p_0 \in M$. *Then, there exists* $U$, *an open neighbourhood of* $p_0$ *at* $M$, *a number* $\varepsilon > 0$ *and a smooth function*

$$\varphi\colon (-\varepsilon, \varepsilon) \times U \to M,$$

*such that for any* $q \in U$, *the map*

$$t \mapsto \varphi(t, q)$$

*is the integral curve of* $X$ *starting at* $q$; *i. e.,* $\varphi$ *is the general solution of the differential equation 3.12.*

*The solution* $\{\varphi_t\}_{t\in(-\varepsilon,\varepsilon)}$ *is known as* a local flow of $X$.

**Definition 3.57** (Killing vector field)**.** Let $(M, g)$ be a semi-Riemannian manifold, and consider a vector field $X \in \mathfrak{X}(M)$.

$X$ is said to be a Killing vector field if every local flow $\{\varphi_t\}$ of $X$ satisfies that

$$\varphi_t\colon U \to \varphi_t(U)$$

is an isometry.

When $X$ is a Killing vector field, $g$ does not change under the action of any flow of $X$; i. e., $\varphi_t^* g = g$.

Before finding an interesting characterization of the Killing vector fields, it is necessary to introduce an important concept: the Lie derivative.

**Definition 3.58** (Lie derivative)**.** Let $(M, g)$ be a semi-Riemannian manifold and let $X \in \mathfrak{X}(M)$ be a vector field on $M$. The *Lie derivative of* $g$ *with respect to* $X$ is a function acting on pairs of smooth vector fields $Y, Z \in \mathfrak{X}(M)$ defined by:

$$(L_X g)(Y, Z) = Xg(Y, Z) - g([X, Y], Z) - g(Y, [X, Z]) \tag{3.14}$$

**Proposition 3.59** (Killing vector fields characterisation)**.** *A vector field* $X \in \mathfrak{X}(M)$ *on a semi-Riemannian manifold* $(M, g)$ *is a Killing vector field if and only if*

$$L_X g = 0,$$

*where* $L_X$ *is the Lie derivative with respect to* $X$.

We can rewrite the Lie derivative to better work with it. Using Equation 3.2 and the identity $[X, Y] = \nabla_X Y - \nabla_Y X$, Equation 3.14 is equivalent to

$$(L_X g)(Y, Z) = g(\nabla_Y X, Z)) + g(Y, \nabla_Z X). \tag{3.15}$$

Using this definition of Lie derivative and Proposition 3.59, we can now conclude that $X \in \mathfrak{X}(M)$ is a Killing vector field if and only if

$$g(\nabla_Y X, Z) + g(Y, \nabla_Z X) = 0 \quad \forall X, Y \in \mathfrak{X}(M). \tag{3.16}$$

**Proposition 3.60.** *Let $X \in \mathfrak{X}(M)$ be a Killing vector field on a semi-Riemannian manifold $(M, g)$ and considet $\gamma$ a geodesic on $M$. Then,*

$$g(\gamma'(t), X_{\gamma(t)}) = c,$$

*where $c \in \mathbb{R}$ depends only on $\gamma$.*

*Proof.* Indeed, using that $\frac{D\gamma'}{dt} = 0$ and Equation 3.16:

$$\frac{d}{dt} g\left(\gamma'(t), X_{\gamma(t)}\right) = g\left(\frac{D\gamma'}{dt}, X_{\gamma(t)}\right) + g\left(\gamma'(t), \nabla_{\gamma'(t)} X\right) = 0.$$

∎

# 4

# THE NOTION OF SPACETIME

In this section we will follow [23]

## 4.1 LORENTZIAN MANIFOLDS

A *Lorentzian metric* on an $n(\geqslant 2)$-dimensional manifold $M$[1] is a symmetric 2-covariant tensor field $g$ such that

$$g_p : T_pM \times T_pM \longrightarrow \mathbb{R}$$

is a Lorentzian product for all $p \in M$. A *Lorentzian manifold* is a pair $(M, g)$ consisting of an $n(\geqslant 2)$-dimensional manifold $M$ and a Lorentzian metric $g$ on $M$.

It should be noticed that if a manifold $M$ admits a symmetric 2-covariant tensor field $g$ such that $g_p$ is non-degenerate for all $p \in M$, then $g$ has a Levi-Civita connection $\nabla$. This assertion follows from the classical Koszul formula (see Theorem 3.39) which defines $\nabla$ just from the non-degeneracy property (equivalently, note also that the Christofell symbols $\Gamma^k_{ij}$ may be defined using only the non-degeneration property). Therefore, from the connectedness of $M$, for each two points $p_0, p_1 \in M$ there exists a piecewise smooth curve

Add definition of piecewise smooth curve somewhere

$$\gamma : [a, b] \longrightarrow M,$$

$a, b \in \mathbb{R}$, $a < b$ such that $\gamma(a) = p_0$ and $\gamma(b) = p_1$. Therefore, we have the corresponding parallel transport

$$P^{\gamma}_{a,b} : T_{p_0}M \longrightarrow T_{p_1}M$$

which is a linear isometry between $(T_{p_0}M, g_{p_0})$ and $(T_{p_1}M, g_{p_1})$.

Consequently, $\text{index}(g_{p_0}) = \text{index}(g_{p_1})$, and we may speak of the *index of* $g$.

A non-degenerate symmetric 2-covariant tensor field $g$ is called a *semi-Riemannian metric*, thus, $g$ is Riemannian if its index is zero and Lorentzian if its index is 1 and $\dim M \geqslant 2$.

A semi-Riemannian metric of index $s$ such that $0 < s < \dim M$ is said to be *indefinite*.

Thus, a *semi-Riemannian* (resp. *Riemannian*, *indefinite Riemannian*) *manifold* is a pair $(M, g)$, where $g$ is a semi-Riemannian (resp. Riemannian, indefinite Riemannian) metric.

1 Unless otherwise is specified, a manifold will be assumed to be of class $C^\infty$, connected and with a countable basis in its topology.

## 4.2 TIME ORIENTATION

Now we will explain the concept of time orientation of a Lorentzian manifold.

Let $(M, g)$ be a Lorentzian manifold and denote by $C_p(M, g)$ the set consisting of the two time cones of $(T_pM, g_p)$, $p \in M$. Put

$$C(M, g) = \bigcup_{p \in M} C_p(M, g).$$

A *time orientation* on $(M, g)$ is a map

$$\tau : M \longrightarrow C(M, g)$$

such that $\tau(p) \in C_p(M, g)$, i.e. $\tau(p)$ is a time cone of $(T_pM, g_p)$, and such that for each $p_0 \in M$ there exist an open neighbourhood $U$ of $p_0$ and $X \in \mathfrak{X}(U)$ which satisfies

$$X_p \in \tau(p), \quad \text{for} \quad \text{all} \quad p \in U.$$

If a Lorentzian manifold $(M, g)$ admits a time orientation, it is called *time orientable*. A time orientable Lorentzian manifold $(M, g)$ produces two *time oriented Lorentzian manifolds* $(M, g, \tau)$ and $(M, g, \tau')$, where $\tau'(p)$ is the opposite cone of $\tau(p)$ in $(T_pM, g_p)$. A 4-dimensional time oriented Lorentzian manifold is called a *spacetime*.

The following result characterizes the existence of a time orientation [16, Lemma 5.32]

**Proposition 4.1.** *A Lorentzian manifold $(M, g)$ is time orientable if and only if there exists $Y \in \mathfrak{X}(M)$ such that $g(Y, Y) < 0$.*

*Proof.* If such a vector field exists, then we can choose $\tau(p)$ as the time cone of $(T_pM, g_p)$ such that $Y_p \in \tau(p)$ for all $p \in M$. Conversely, let $\tau$ be a time orientation on $(M, g)$. For each $p_0 \in M$ there exist a neighbourhood $U^{p_0}$ and $X_U \in \mathfrak{X}(U^{p_0})$ such that $(X_U)_p \in \tau(p)$ for all $p \in U^{p_0}$.

Let $\{f_\alpha\}$ be a smooth partition of unity subordinate to the open covering $\{U^p : p \in M\}$; i.e. $\{\text{supp}(f_\alpha)\}$ is locally finite, $f_\alpha \geqslant 0$, $\sum_\alpha f_\alpha = 1$ and $\text{supp}(f_\alpha) \subset U_\alpha$ for some $U_\alpha$ of the covering of $M$ (see section 3.3).

The vector field

$$Y := \sum f_\alpha X_{U_\alpha}$$

is then well-defined and for each $p \in M$ there exists an open neighbourhood $V(p)$ such that $V \cap \text{supp}(f_\alpha) = \emptyset$ for all $\alpha \neq i_1, .., i_k$; therefore

$$Y_{|V} = f_{i_1} X_{U_{i_1}} + .. + f_{i_k} X_{U_{i_k}},$$

with $\sum_j f_{i_j} = 1$.

Then, using the convexity of time cones, Corollary 1.9, the vector field $Y$ is timelike everywhere. ∎

**Example 4.2.** Let us see some examples that clarify the concepts exposed.

1. Let $\mathbb{L}^n$ be the $n$-dimensional Lorentz-Minkowski space, i.e. $\mathbb{L}^n$ is $\mathbb{R}^n$ endowed with the Lorentzian metric $g = dx_1^2 + ... + dx_{n-1}^2 - dx_n^2$, where $(x_1, .., x_n)$ is the usual coordinate system of $\mathbb{R}^n$. The coordinate vector field $\partial/\partial x_n$ is unitary timelike and hence, Proposition 4.1, $\mathbb{L}^n$ is time orientable.

2. Let $\mathbb{S}_1^n$ be the $n$-dimensional De Sitter space; i.e. $\mathbb{S}_1^n = \{p \in \mathbb{L}^{n+1} : g(p,p) = 1\}$, where $g$ denotes the Lorentzian metric of $\mathbb{L}^{n+1}$. For each $p \in \mathbb{S}_1^n$, we have $T_p\mathbb{S}_1^n = \{v \in \mathbb{L}^{n+1} : g(p,v) = 0\}$ and denote by $g_p$ the restriction of $g$ to $T_p\mathbb{S}_1^n$, which is Lorentzian because $\mathbb{L}^{n+1} = T_p\mathbb{S}_1^n \oplus \langle p \rangle$, the direct sum is also $g$-orthogonal and $p$ is spacelike. Observe that a vector field on $\mathbb{S}_1^n$ can be contemplated as a smooth map

   $$X : \mathbb{S}_1^n \longrightarrow \mathbb{L}^{n+1}$$

   such that at each point $p \in \mathbb{S}_1^n$ we have $X_p$ is $g$-orthogonal to $p$. Thus, if we put $p = (y,t) \in \mathbb{S}_1^n$, $y \in \mathbb{R}^n$, $t \in \mathbb{R}$, then $X_p = (\frac{t}{1+t^2}y, 1)$ is a well-defined timelike vector field on $\mathbb{S}_1^n$. Therefore, Proposition 4.1, the Lorentzian manifold $\mathbb{S}_1^n$ is time orientable.

3. Let $\mathbb{H}_1^n$ be the $n$-dimensional anti De Sitter space; i.e. $\mathbb{H}_1^n = \{p \in \mathbb{R}^{n+1} : g'(p,p) = -1\}$, where $g' = dx_1^2 + ... + dx_{n-1}^2 - dx_n^2 - dx_{n+1}^2$ and $(x_1, .., x_n, x_{n+1})$ is the usual coordinate system of $\mathbb{R}^{n+1}$. The semi-Riemannian metric $g'$ on $\mathbb{R}^{n+1}$ has index 2 and $\mathbb{R}_2^{n+1}$ will denote $(\mathbb{R}^{n+1}, g')$. For each $p \in \mathbb{H}_1^n$, we have $T_p\mathbb{H}_1^n = \{v \in \mathbb{R}^{n+1} : g'(p,v) = 0\}$ and denote by $g'_p$ the restriction of $g'$ to $T_p\mathbb{H}_1^n$, which is Lorentzian because $\mathbb{R}_2^{n+1} = T_p\mathbb{H}_1^n \oplus \langle p \rangle$, the direct sum is also $g'$-orthogonal and $p$ satisfies $g'(p,p) = -1$. The vector field

   $$X : \mathbb{H}_1^n \longrightarrow \mathbb{R}_2^{n+1}$$

   given by $X_p = (0, t, -s)$ for $p = (y, s, t)$, $y \in R^{n-1}$, $s, t \in \mathbb{R}$, is timelike everywhere.

   Therefore, Proposition 4.1, the Lorentzian manifold $\mathbb{H}_1^n$ is time orientable.

The following result gives a geometric characterization of time orientability [24, p. 255]

**Corollary 4.3.** *A Lorentzian manifold* $(M, g)$ *is time orientable if and only if for any piecewise smooth curve* $\gamma : [a, b] \longrightarrow M$ *such that* $\gamma(a) = \gamma(b) = p$, *we have*

$$g(P^{\gamma}_{a,b}(v), v) < 0$$

*for all* $v \in \mathcal{T}(T_pM, g_p)$, *for all* $p \in M$.

*Proof.* Assume $(M, g)$ is time orientable and consider $X \in \mathfrak{X}(M)$ such that $g(X, X) < 0$. Changing $X$ to $-X$, if necessary, we may assume $g(X_p, v) < 0$. Let $Y$ be a vector field along $\gamma$ such that $\frac{DY}{dt} = 0$ and $Y(a) = v$.

Note that we have $Y(b) = P^{\gamma}_{a,b}(v)$. Consider the function $f : [a, b] \longrightarrow \mathbb{R}$ given by

$$f(t) = g(X_{\gamma(t)}, Y(t))$$

which is continuous and never vanishes because $X_{\gamma(t)}$ and $Y(t)(= P^{\gamma}_{a,t}(v))$ are timelike. Therefore $f(t) < 0$ for all $t \in [a, b]$ and, in particular, $f(b) < 0$. This means, taking into account Lemma 1.8, that $Y(b)$ and $X_{\gamma(b)}$ lie in the same time cone.

Conversely, let us consider, for two arbitrary points $p$ and $q$ of $M$, two piecewise smooth curves $\alpha$ and $\beta$ from $p$ to $q$. We want to show that for any $v \in \mathcal{T}(T_pM, g_p)$ the parallel transported vectors $P^{\alpha}(v)$, $P^{\beta}(v)$ lie in the same time cone of $(T_qM, g_q)$.

In order to achieve this conclusion we construct a piecewise smooth curve $\gamma : [a, b] \longrightarrow M$ from $\alpha$ and $\beta$ in a standard way such that $\gamma(a) = \gamma(b) = p$. Note that

$$g(P^{\alpha}(v), P^{\beta}(v)) = g((P^{\beta})^{-1}P^{\alpha}(v), v) = g(P^{\gamma}_{a,b}(v), v) < 0,$$

which means that $P^{\alpha}(v)$ and $P^{\beta}(v)$ lie in the same time cone of $(T_qM, g_q)$.

Therefore, we have a well-defined way to chose a time cone $\tau(p)$ at any $p \in M$. Finally, we will to show the smoothness. Given $p_0 \in M$ and the time cone $\tau(p_0)$ consider $v \in \tau(p_0)$. Let $X$ be a vector field which extends $v$; i.e. such that $X_{p_0} = v$. Note that $X$ remains timelike in some (connected) open neighbourhood $U$ of $p_0$. For each $q \in U$ we construct a piecewise smooth curve $\alpha : [a, b] \longrightarrow U$ satisfying $\alpha(a) = p_0$, $\alpha(b) = q$ and consider the function $h : [a, b] \longrightarrow \mathbb{R}$ given by

$$h(t) = g(X_{\alpha(t)}, P^{\alpha}_{a,t}(v))$$

which is continuous and never vanishes. Therefore $h(t) < 0$ for all $t \in [a, b]$ and, in particular, $h(b) < 0$. This means, taking into account Lemma 1.8, that $X_q$ and $P^{\alpha}_{a,b}(v)$ lie in the same time cone of $(T_qM, g_q)$ and thus $X_q \in \tau(q)$ for all $q \in U$. ∎

Now assume each closed piecewise smooth curve is null homotopic by means of a piecewise smooth homotopy.

In this case, Corollary 4.3 says that $(M, g)$ must be time orientable. But it is known that this fact holds true whenever $M$ is assumed to be simply connected. Therefore, we have

**Corollary 4.4.** *If $M$ is simply connected and $g$ is a Lorentzian metric on $M$, then the Lorentzian manifold $(M, g)$ must be time orientable.*

A well-known non time orientable Lorentzian manifold is the following Lorentzian cylinder [24, Example 1.2.3]

**Example 4.5.** Let $g$ be the Lorentzian metric on $\mathbb{R}^2$ given by

$$g\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial x}\right)_{(x,y)} = -g\left(\frac{\partial}{\partial y}, \frac{\partial}{\partial y}\right)_{(x,y)} = \cos 2y,$$
$$g\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)_{(x,y)} = \sin 2y.$$

Observe that

$$\det\begin{pmatrix} \cos 2y & \sin 2y \\ \sin 2y & -\cos 2y \end{pmatrix} = -1 < 0$$

everywhere, which implies that $g$ is Lorentzian. The map $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, defined by $f(x,y) = (x, y+\pi)$, is clearly an isometry of $(\mathbb{R}^2, g)$. Put $M := \mathbb{R}^2/\mathbb{Z}$, where the action of $\mathbb{Z}$ on $\mathbb{R}^2$ is defined via $f$ as follows

$$\big(m, (x,y)\big) \mapsto f^m(x,y) = (x, y+m\pi).$$

Then $M$ is a cylinder and the metric $g$ may be induced to a Lorentzian metric $\tilde{g}$ in $M$. We want to show that $(M, \tilde{g})$ is not time orientable. If we choose a time cone at $(0,0)$ then along the axis $x = 0$ it changes its position in the counter-clockwise rotation sense. Note that $(0,0)$ and $(0,\pi)$ represent the same point of $M$ but the time cones at these points are not compatible with the equivalence relation in $\mathbb{R}^2$ induced by $f$. Note that $Y = -\sin y \frac{\partial}{\partial x} + \cos y \frac{\partial}{\partial y}$ is a timelike vector field on $(\mathbb{R}^2, g)$ (of course, $(\mathbb{R}^2, g)$ is time orientable from Corollary 4.4) which satisfies $Y_{(0,0)} = \frac{\partial}{\partial y}\,|_{(0,0)}$ and $Y_{(0,\pi)} = -\frac{\partial}{\partial y}\,|_{(0,\pi)}$.

Taking into account that $df_{(0,0)} Y_{(0,0)} = -Y_{(0,\pi)}$, $Y$ cannot be induced on $M$. On the other hand, assume there exists $\tilde{X} \in \mathfrak{X}(M)$ such that $\tilde{g}(\tilde{X}, \tilde{X}) < 0$ and let $X \in \mathfrak{X}(\mathbb{R}^2)$, $g(X,X) < 0$, which projects onto $\tilde{X}$. Necessarily $df_{(x,y)} X_{(x,y)} = X_{(x,y+\pi)}$ and $g(Y_{(x,y)}, X_{(x,y)}) \neq 0$ for all $(x,y) \in \mathbb{R}^2$.

Therefore, either $g(Y,X) > 0$ or $g(Y,X) < 0$ everywhere. But this is incompatible with

$$g\left(Y_{(0,\pi)}, X_{(0,\pi)}\right) = -g\left(df_{(0,0)} Y_{(0,0)}, df_{(0,0)} X_{(0,0)}\right) =$$
$$= -g\left(Y_{(0,0)}, X_{(0,0)}\right).$$

Previous example shows a (connected) orientable manifold $M$ which admits a Lorentzian metric $\tilde{g}$ such that $(M, \tilde{g})$ is not time orientable. It is possible to have a time orientable Lorentzian manifold $(N, g)$ where $N$ is not (topologically) orientable. Even more, it is also easy to construct a non time orientable Lorentzian manifold $(P, g')$ such that $P$ is not (topologically) orientable.

As in the non orientable case, a Lorentzian manifold $(M, g)$ which is not time orientable admits a double Lorentzian covering manifold

$(\hat{M}, \hat{g})$ which is time orientable. Note that $(\hat{M}, \hat{g})$ and $(M, g)$ have the same local geometry, but the first one possesses a globally defined timelike vector field and the second one does not.

## 4.3 ONE DIMENSIONAL DISTRIBUTIONS

It is classical that, by using a partition of the unity on a (paracompact) manifold $M$, we can always construct a Riemannian metric on $M$. But, the same procedure does not work in the Lorentzian case. In fact, although we can consider a Lorentzian metric on each coordinate open subset of $M$, it may be not possible to glue the locally defined Lorentzian metrics, as in the Riemannian case, to produce a Lorentzian metric defined on the whole manifold $M$. Therefore, it is natural to ask when a manifold admits a Lorentzian metric.

Bad reference

The answer is the well-known result [**greub72**].

**Proposition 4.6.** *An $n(\geqslant 2)$-dimensional manifold $M$ admits a Lorentzian metric if and only if it admits a $1$-dimensional distribution.*

*Proof.* First consider a Lorentzian metric $g$ on $M$, and let $g_R$ be an arbitrarily chosen Riemannian metric on $M$.

A $(1,1)$-tensor field $P$ on $M$ can be defined by setting, for each $u \in T_P M$, $P(u)$ the unique vector of $T_p M$ such that

$$g_R\big(P(u), v\big) = g(u, v)$$

for all $v \in T_p M$, $p \in M$. Clearly, $P$ is $g_R$-selfadjoint and, therefore, at any point $p \in M$, there exists a $g_R$-orthonormal basis of $T_p M$ consisting of eigenvectors of $P$. Observe that none of the eigenvalues is zero, $n-1$ are positive and one is negative.

Now, making use of $P$ we will define a 1-dimensional distribution on $M$ [14]; i.e., an assignment to each $p \in M$ of a 1-dimensional subspace $\mathfrak{D}_p$ of $T_p M$ such that each $p \in M$ has an open neighbourhood $U$ and a vector field $X \in \mathfrak{X}(U)$ such that

$$\mathfrak{D}_p = \langle X_q \rangle \quad \forall q \in U.$$

Put $\mathcal{D}_p$ the eigenspace associated to the negative eigenvalue of $P$ at $p$, then $\mathfrak{D}$ defines a $1$-dimensional distribution (or line field) on $M$. It should be noted that $\mathfrak{D}$ clearly depends on the arbitrary Riemannian metric $g_R$.

Conversely, if a $1$-dimensional distribution $\mathfrak{D}$ on $M$ is given, fix an arbitrary Riemannian metric $g_R$ on $M$. We know that there exist an open covering $\{U_\alpha\}$ of $M$ and vector fields $X_\alpha \in \mathfrak{X}(U_\alpha)$ such that, locally,

$$\mathfrak{D} = \langle X_\alpha \rangle, \quad \text{with} \quad g_R(X_\alpha, X_\alpha) = 1.$$

By putting

$$g_L(u, v) := g_R(u, v) - 2\, g_R\big(u, X_\alpha(p)\big)\, g_R\big(v, X_\alpha(p)\big),$$

for any tangent vectors $u, v \in T_pM$ with $p \in U_\alpha$, it is easily seen that $g_L$ does not depend on $\alpha$ and therefore, it is a Lorentzian metric on all $M$. ∎

*Remark* 4.7. Instead of a 1-dimensional distribution if we have $X \in \mathfrak{X}(M)$ such that $X_p \neq 0$ for all $p \in M$, we can construct a Lorentzian metric $g_L$, starting from a Riemannian metric $g_R$ on $M$, as follows

$$g_L(u,v) := g_R(u,v) - 2\,\frac{g_R(u,X_p)g_R(v,X_p)}{g_R(X_p,X_p)}$$

where $u, v \in T_pM$, $p \in M$.

Proposition 4.6 can be generalized to obtain **[greub72]** Bad reference.

Any non-compact manifold admits a non-vanishing vector field. In fact, it can be taken as the gradient, with respect to any Riemannian metric, of a smooth function with no critical points.

Thus any $n$-dimensional (with $n \geqslant 2$) non-compact manifold admits a Lorentzian metric.

On the other hand, an $n$-dimensional (with $n \geqslant 2$) compact manifold $M$ admits a 1-dimensional distribution if and only if its Euler-Poincaré characteristic $\chi(M)$ is zero. Therefore, any $(2n+1)$-dimensional compact orientable manifold admits a Lorentzian metric.

The existence of a 1-dimensional distribution on a manifold is closely related to the existence of a non-vanishing vector field. In fact, it is a standard topological result that

**Proposition 4.8.** *An $n(\geqslant 2)$-dimensional compact manifold $M$ admits a non-vanishing vector field if and only if $\chi(M) = 0$.*

On a simply connected manifold (compact or not), every 1-dimensional distribution on $M$ arises from a global non-vanishing vector field $X \in \mathfrak{X}(M)$. However, a 1-dimensional distribution cannot be lifted in general to a global non-vanishing vector field as the following example shows **[greub72]**. Bad reference

*Remark* 4.9 (Lie groups). Let us remember that a *Lie group* is a group $G$ that is also a differentiable manifold, and such that the map

$$\begin{aligned} G \times G &\to G \\ (a,b) &\mapsto ab^{-1} \end{aligned}$$

is differentiable—see [14, p. 38]. Review this reference

We denote by $L_a$ (resp. $R_a$) the left (resp. right) translation of $G$ by the element $a \in G$, that is:

$$\begin{aligned} L_a\colon G &\to G \\ x &\mapsto ax \end{aligned}$$

and

$$\begin{aligned} R_a \colon G &\to G \\ x &\mapsto xa. \end{aligned}$$

Both $L_a$ and $R_a$ are diffeomorphisms on G.

A vector field X on G is called left invariant (resp. right invariant) if $(dL_a)_b X_b = X_{L_a(b)}$ (resp. $(dR_a)_b X_b = X_{R_a(b)}$) for all $a, b \in G$.

Given $v \in T_e G$ there exists a unique left invariant vector field X on G such that $X_e = v$. In fact, $X_a := (dL_a)_e(a)$ $\forall a \in G$ (analogously for a right invariant vector field). Note that $v \neq 0$ implies $X_a \neq 0$ for all $a \in G$. Even more, if $\{v_1, \dots, v_n\}$ is a basis of $T_e G$, then there exists a set of left invariant vector fields $\{X_1, \dots, X_n\}$ such that $X_i(w) = v_i$, $1 \leqslant i \leqslant n$, and $\{X_1(a), \dots, X_n(a)\}$ is a basis of $T_a G$ for all $a \in G$.

Therefore, G admits a global basis of $\mathfrak{X}(G)$. In this case, G is said to be *parallelizable* [14, Ch. 1, Sec. 4].

Review this reference

Consider the special orthogonal group of order 3, SO(3), and put $M = \mathbb{S}^1 \times SO(3)$. Both $\mathbb{S}^1$ and SO(3) are Lie groups.

M is a 4-dimensional compact Lie group. Moreover it is parallelizable, and therefore every vector field $X \in \mathfrak{X}(M)$ can be contemplated as a smooth map

$$X : M \to \mathbb{R}^4$$

and, by fixing a diffeomorphism $\psi : \mathbb{RP}^3 \to SO(3)$, a 1-dimensional distribution $\mathfrak{D}$ can be seen as a smooth map

$$\mathfrak{D} : M \to SO(3).$$

In particular, the canonical projection on the second factor $\mathfrak{D}_2$ defines a natural 1-dimensional distribution on $M = \mathbb{S}^1 \times SO(3)$. If we assume that $\mathfrak{D}_2$ lifts to a vector field X without any zero, then, taking into account that $\mathbb{R}^4 - \{0\}$ is simply connected, one easily shows that SO(3) would be also simply connected, which is not true. Hence $\mathfrak{D}_2$ cannot be lifted to a global vector field on $\mathbb{S}^1 \times SO(3)$.

# 5

# EINSTEIN FIELD EQUATIONS

## 5.1 EINSTEIN TENSOR OF A METRIC

**Definition 5.1** (Einstein tensor)**.** Let $(M, g)$ be a four-dimensional spacetime. The symmetric 2-covariant tensor field

$$G := \mathrm{Ric} - \frac{1}{2} S g,$$

where Ric is the Ricci tensor of $g$ and $S$ the scalar curvature of $g$, is called the *Einstein tensor of* $g$.

It can be proved that $G$, the Einstein tensor of $g$, satisfies the following:

Reference?

1. $\mathrm{trace}_g\, G = -S$ and so $\mathrm{Ric} = G - (1/2)(\mathrm{trace}_g G) g$, where $\mathrm{trace}_g\, G$ denotes the contraction of the $(1, 1)$-tensor field $g$-equivalent to $G$. Therefore, to have $G$ is equivalent to have Ric, and both tensors have the same physical information.

2. If for any symmetric 2-covariant tensor field $T$, we define

   $$G_k := T + k(\mathrm{trace}_g T) g,$$

   for a fixed $k \in \mathbb{R}$, then the map $T \longmapsto G_k$ is involutive if and only if $k = 0$ or $k = -\frac{1}{2}$.

3. $\mathrm{Ric} = \lambda\, g$ if and only if $G = -\lambda\, g$. A spacetime such that Ric is proportional to the metric tensor $g$, or equivalently, $G$ is proportional to $g$, is called an *Einstein spacetime*.

## 5.2 STRESS-ENERGY TENSORS

**Definition 5.2** (Stress-energy tensor)**.** A *stress-energy tensor* on a spacetime $M$ is a symmetric 2-covariant tensor field $T$ such that

$$T(V, V) \geqslant 0 \tag{5.1}$$

for any timelike vector $V \in T_p M$ $\forall p \in M$.

Note that a continuity argument (see section 1.5) shows that inequality 5.1 also holds true if $V$ is a lightlike vector on $T_p M$.

It is commonly argued that the mathematical way to express that gravity attracts on average is

$$\mathrm{Ric}(V, V) \geqslant 0 \tag{5.2}$$

for any timelike tangent vector $V \in T_p M$ and for all $p \in M$; that is, the Ricci tensor of a physically realistic spacetime must be an stress-energy tensor. This assumption restricts the familiy of Lorentzian metrics that can be physically significant.

## 5.3 EINSTEIN FIELD EQUATIONS

Consider a stress-energy tensor $T$ on a spacetime $(M, g)$ such that

1. The function
$$T - \frac{1}{2}(\mathrm{trace}_g\, T)g \tag{5.3}$$
is also a stress-energy tensor.

2. $T$ satisfies the conservation law
$$\mathrm{div}\, \widehat{T} = 0, \tag{5.4}$$
where $\widehat{T}$ is the 2-contravariant tensor field $g$-equivalent to $T$.

We say that the spacetime $(M, g)$ obeys the Einstein field equation with respect to $T$ if

$$G(g) = T, \tag{5.5}$$

where $G$ is the Einstein tensor of $g$, that is,

$$\mathrm{Ric} - \frac{1}{2} S g = T. \tag{5.6}$$

Which one?

Note that we can also write, using previous formula , that

$$T - \frac{1}{2}(\mathrm{trace}_g\, T)g = \mathrm{Ric}. \tag{5.7}$$

Therefore, assumption 5.3 agrees with the timelike convergent condition, which express, using the Ricci tensor, that gravitational effects are on average attractive.

On the other hand, assumption 5.4 is mandatory, as we know that the Einstein tensor $G$ satisfies

$$\mathrm{div}\, \hat{G} = 0.$$

Einstein field equation (Equation 5.5 or Equation 5.6) postulated how matter and radiation in a region of the universe can be described by a Lorentzian metric $g$.

In fact, Equation 5.5 is similar to the Poisson equation,

$$\Delta\phi = k\rho,$$

where $k > 0$ is a universal constant, $\rho$ is the function describing the density of matter and $\phi$ is the potential function.

This equation postulated, in the pre-relativistic physics, how matter can be described by a potential function.

Note that if we know $\phi$, then the corresponding gravitational field is

$$-\nabla\phi,$$

where $\nabla$ denotes the usual gradient for functions on an open subset of the Euclidean space $\mathbb{R}^3$.

Then, the gravitational force, $F$, acting on a mass $m$, is obtained as follows:

$$F = -m\nabla\phi.$$

We see then that, roughly speaking, $F$ had the same role in the old physics as the curvature tensor has now in Relativity.

Imagine now the trivial stress-energy tensor

$$T = 0.$$

If $(M, g)$ obeys the Einstein field equation with respect to $T = 0$, then, using Equation 5.7, we realize that

$$\mathrm{Ric} = 0.$$

Assume now that $(M, g)$ obeys the Einstein field equation

$$\mathrm{Ric} = 0. \tag{5.8}$$

Then, it is trivial that $T = 0$.

We can conclude that the mathematical way of expressing the absence of matter and radiation on the spacetime is the one shown on Equation 5.8, which is called the *vacuum Einstein field equation*.

# 6 KERR SPACETIME

The Kerr metric is

$$g_{\mu\nu} = \begin{pmatrix} -\alpha^2 + \varpi^2\omega^2 & 0 & 0 & -\varpi^2\omega \\ 0 & \rho^2/\Delta & 0 & 0 \\ 0 & 0 & \rho^2 & 0 \\ -\varpi^2\omega^2 & 0 & 0 & \varpi^2 \end{pmatrix},$$

with inverse

$$g^{\mu\nu} = \begin{pmatrix} -1/\alpha^2 & 0 & 0 & -\omega/\alpha^2 \\ 0 & \Delta/\rho^2 & 0 & 0 \\ 0 & 0 & 1/\rho^2 & 0 \\ -\omega/\alpha^2 & 0 & 0 & (1/\varpi^2) - (\omega^2/\alpha^2) \end{pmatrix},$$

where

$$\omega = \frac{2ar}{\Sigma^2}, \quad \varpi = \frac{\Sigma \sin\vartheta}{\rho}, \quad \alpha = \frac{\rho\sqrt{\Delta}}{\Sigma}, \quad \rho = \sqrt{r^2 + a^2\cos^2\vartheta},$$
$$\Delta = r^2 - 2r + a^2, \text{ and } \Sigma = \sqrt{(r^2 + a^2)^2 - a^2\Delta\sin^2\vartheta}. \tag{6.1}$$

# 7 EQUATIONS OF MOTION

This chapter aims to find the most computationally stable equations of motion for $\gamma$, a free falling lightlike particle moving on a Kerr spacetime.

The classical equations derived from the definition of geodesic in terms of the Christoffel symbols,

$$\frac{d^2x^k}{dt^2} + \Gamma^k_{ij}\frac{dx^i}{dt}\frac{dx^j}{dt} = 0 \quad k = 1, \dots, n,$$

are not interesting enough to later implement a numeric algorithm that integrates them: they do not directly provide conserved quantities that can be useful in our computations and, furthermore, we would have to deal with a second order system of differential equations.

In our journey to find such a system that fits our needs, we will find the classical Hamiltonian formulation [3, Sec. 33.5], that gives us a first order system with conserved quantities but with numerically unstable equations; this version would force us to deal with the zeros of a pair of square roots.

Finally, we will be able to get rid of those square roots, simply by defining another equivalent Hamiltonian that will lead us to a first order system with conserved quantities and which will be numerically-friendly [13].

From now on, and valid throughout this chapter, let $\gamma$ be a lightlike particle, whose tangent vector components, expressed in Boyer-Lindquist coordinates, are

$$\mathbf{v}^\alpha = (\dot{t}, \dot{r}, \dot{\vartheta}, \dot{\varphi}), \tag{7.1}$$

whereas its covariant equivalent version, the momentum, is noted as

$$\mathbf{p}_\alpha = \mathbf{v}_\alpha = (p_t, p_r, p_\vartheta, p_\varphi). \tag{7.2}$$

The relation between the two quantities is obtained through the operation of lower and raising indices using the metric. Therefore, we have the two equations

$$\mathbf{v}^\alpha = g^{\beta\alpha}\mathbf{p}_\beta, \tag{7.3}$$

$$\mathbf{p}_\alpha = g_{\beta\alpha}\mathbf{v}^\beta. \tag{7.4}$$

## 7.1 CLASSICAL HAMILTONIAN FORMULATION

It is known [3, Sec. 33.5] that the geodesic equation $\gamma'' = 0$ is equivalent to the system

$$\frac{dx^\mu}{d\lambda} = \frac{\partial \mathcal{H}}{\partial p_\mu}, \qquad \frac{dp_\mu}{d\lambda} = -\frac{\partial \mathcal{H}}{\partial x^\mu}.$$

, where $\lambda$ is an affine parameter such that $d/d\lambda =$ and where

$$\mathcal{H} = \frac{1}{2} g^{\mu\nu} p_\mu p_\nu$$

is the usual Hamiltonian: a half of the square of the velocity.

Define integral of motion

This Hamiltonian formalism let us obtain four first integrals of motion[3, pp. 898-899], that come as a result from the stationary and axial symmetry of the geometry of Kerr spacetimes, from the nearly miraculous work of Carter when computing the constant named after him and from the modulus of the momentum:

- The killing vector $\partial_t$ raises the conserved quantity

  $$g_{\alpha\beta}(\partial_t)^\alpha u^\beta = (\partial_t)^\alpha p_\alpha = p_t = -E,$$

- The killing vector $\partial_\phi$ raises the conserved quantity

  $$g_{\alpha\beta}(\partial_\phi)^\alpha u^\beta = (\partial_\phi)^\alpha p_\alpha = p_\phi = L_z,$$

- The Carter's constant, which can be written as

  $$Q = p_\vartheta^2 + \cos^2\vartheta \left( a^2 \left(\mu^2 - E^2\right) + \frac{L_z^2}{\sin^2\vartheta} \right). \tag{7.5}$$

- The modulus of the tangent vector $p_\alpha$, which is

  $$p_\alpha p_\beta g^{\alpha\beta} = -\mu^2, \tag{7.6}$$

  that is,

  $$\mu^2 + \frac{p_\vartheta^2 + p_r^2\Delta}{\rho^2} + \frac{L_z^2}{\varpi^2} = \frac{E - L_z\omega}{\alpha^2}$$

Using the constants of motion, the relation in Equation 7.3 and the components notation used in equations 7.1 and 7.2, we obtain four equations:

$$\dot{t} = \frac{E}{\alpha^2} - \frac{L_z\omega}{\alpha^2} \tag{7.7}$$

$$\dot{r} = \frac{p_r\Delta}{\rho^2} \tag{7.8}$$

$$\dot{\vartheta} = \frac{p_\vartheta}{\rho^2} \tag{7.9}$$

$$\dot{\varphi} = \frac{E\omega}{\alpha^2} + L_z\left(\frac{1}{\varpi^2} - \frac{\omega^2}{\alpha^2}\right) \tag{7.10}$$

From this direct computation, we can express all four equations in terms of the constants $E$, $L_z$, $Q$ and $\mu$, which yields the following result [3, p. 899].

**Theorem 7.1.** *A free falling lightlike particle $\gamma$ satisfies:*

$$\rho^2 \dot{t} = -(aE\sin^2\vartheta - L_z) + \frac{r^2+a^2}{\Delta}P \tag{7.11}$$

$$\rho^2 \dot{r} = \sqrt{R} \tag{7.12}$$

$$\rho^2 \dot{\vartheta} = \sqrt{\Theta} \tag{7.13}$$

$$\rho^2 \dot{\varphi} = -(aE - \frac{L_z}{\sin^2\vartheta}) + \frac{a}{\Delta}P, \tag{7.14}$$

*where $R$ and $\Theta$ are the functions defined as follows:*

$$R := P^2 - \Delta\left(r^2\mu^2 + Q + (L_z - aE)^2\right), \tag{7.15}$$

$$\Theta := Q - \cos^2\vartheta\left(\frac{L_z^2}{\sin^2\vartheta} + \omega^2\left(\mu^2 - E^2\right)\right), \tag{7.16}$$

*with $P$ an auxiliary function defined as*

$$P := E(r^2 + a^2) - aL_z. \tag{7.17}$$

*Proof.* We will only prove equations 7.12 and 7.13, as the computation of the other two is really similar and does not contribute anything interesting. Furthermore, only the $\dot{r}$ and $\dot{\vartheta}$ equations will prove to be useful on the next section of this chapter.

*$\dot{r}$ Equation*

Using Equation 7.6, we can find the value of $p_r$:

$$p_r^2 = \left(\frac{(E - L_z\omega)^2}{\alpha^2} - \frac{L_z^2}{\varpi^2} - \frac{p_\vartheta^2}{\rho^2} - \mu^2\right)\frac{\rho^2}{\Delta}.$$

We can now substitute $p_r$ for its value in the squared version of Equation 7.8:

$$\dot{r}^2 = \Delta\left(\underset{\substack{\uparrow \\ \text{from } p_\vartheta = \rho^2\dot{\vartheta}}}{-\dot{\vartheta}^2} + \frac{E^2}{\alpha^2\rho^2} - \frac{\mu^2}{\rho^2} - \frac{2L_zE\omega}{\alpha^2\rho^2} + \frac{L_z^2\omega^2}{\alpha^2\rho^2} - \frac{L_z^2}{\rho^2\varpi^2}\right) \tag{7.18}$$

If we write Equation 7.18 in the form

$$\rho^2\dot{r} = \sqrt{\widehat{R}},$$

the function $\widehat{R}$ has the expression

$$\begin{aligned}\widehat{R} &= \rho^4\left(-\frac{\Theta}{\rho^4} + \frac{(E^2 - L_z\omega)^2}{\alpha^2\rho^2} - \frac{\mu^2}{\rho^2} - \frac{L_z^2}{\rho^2\varpi^2}\right) = \\ &= \Delta\left(-\Theta + \frac{(E - L_z\omega)^2\rho^2}{\alpha^2} - \mu^2\rho^2 - \frac{L_z^2\rho^2}{\varpi^2}\right).\end{aligned} \tag{7.19}$$

Notice that

$$\frac{\rho^2}{\alpha^2} = \frac{\Sigma^2}{\Delta}, \qquad \frac{\rho^2}{\varpi^2} = \frac{\rho^4}{\Sigma^2 \sin^2\vartheta}.$$

Let us now substitute $\Theta$ by its definition (**??**) on Equation 7.19, from which we obtain:

$$\begin{aligned}
\widehat{R} =& \Delta\Bigg(-Q + \cos^2\vartheta\left(\frac{L_z^2}{\sin^2\vartheta} + a^2(\mu^2 - E^2)\right) + \frac{\Sigma^2}{\Delta}(E - L_z\omega)^2 - \\
& -\mu^2\rho^2 - \frac{L_z^2\rho^4}{\Sigma^2\sin^2\vartheta}\Bigg) = \\
=& \Delta\Bigg(-Q + \cos^2\vartheta\left(\frac{L_z^2}{\sin^2\vartheta} + a^2(\mu^2 - E^2)\right) + \\
& + \frac{\Sigma^2}{\Delta}\left(E^2 + L_z^2\left(\frac{2ar}{\Sigma^2}\right)^2 - 2EL_z\frac{2ar}{\Sigma^2}\right) - \mu^2\rho^2 - \\
& - \frac{L_z^2}{\sin^2\vartheta}\frac{\rho^4}{(r^2+\omega)^2 - a^2\Delta\sin^2\vartheta}\Bigg)
\end{aligned} \tag{7.20}$$

This can be simplified in order to get a more readable expression, although the work will be somewhat cumbersome.

Let us start by simplifying the previous expression by actually making the product by the factorised $\Delta$ and the inner $\Sigma^2$:

$$\begin{aligned}
\widehat{R} =& -Q\Delta + \Delta\cos^2\vartheta\left(\frac{L_z^2}{\sin^2\vartheta} + a^2(\mu^2 - E^2)\right) + \Sigma^2E^2 + \\
& + \frac{L_z^2(2ar)^2}{\Sigma^2} - 2EL_z 2ar - \mu^2\rho^2\Delta - \frac{L_z^2}{\sin^2\vartheta}\frac{\Delta\rho^4}{\Sigma^2}
\end{aligned} \tag{7.21}$$

Rearranging Equation 7.21, we obtain

$$\begin{aligned}
\widehat{R} =& \frac{\overbrace{L_z^2\left(4a^2r^2 - \frac{\Delta\rho^4}{\sin^2\vartheta}\right)}^{(\dagger)}}{\Sigma^2} - 4arEL_z - \mu^2\rho^2\Delta + \Sigma^2E^2 - Q\Delta + \\
& + \Delta\cos^2\vartheta\left(\frac{L_z^2}{\sin^2\vartheta} + a^2\left(\mu^2 - E^2\right)\right).
\end{aligned} \tag{7.22}$$

Let us focus now on $(\dagger)$:

$$\begin{aligned}
(\dagger) =& \frac{L_z^2\left(4a^2r^2 - \frac{\Delta\rho^4}{\sin^2\vartheta}\right)}{\Sigma^2} = \frac{L_z^2\left(4a^2r^2\sin^2\vartheta - \rho^4\Delta\right)}{\sin^2\vartheta\left((r^2+a^2)^2 - a^2\Delta\sin^2\vartheta\right)} \\
=& \frac{\left\{L_z^2\left(4a^2r^2\sin^2\vartheta + \left(r^2+2r-a^2\right)\left(r^2+\omega^2\cos^2\vartheta\right)^2\right)\right\}(\ddagger)}{\sin^2\vartheta\left((r^2+a^2)^2 - a^2\Delta\sin^2\vartheta\right)}
\end{aligned} \tag{7.23}$$

Let us try to simplify (‡), the numerator of (†), first:

$$
\begin{aligned}
(\ddagger) = & L_z^2 \Big( -r^6 + 2r^6 + r^4 \left(-a^2 - 2a^2 \cos^2\vartheta\right) + r^3 \left(4a^2\cos^2\vartheta\right) + \\
& + r^2 \left(-2a^4\cos^2\vartheta - a^4\cos^4\vartheta + 4a^2\sin^2\vartheta\right) + r\left(2a^4\cos^4\vartheta\right) - \\
& - a^6\cos^4\vartheta \Big) = L_z^2 \left( \left(r^2+a^2\right)^2 - a^2\left(r^2-2r+a^2\right)\sin^2\vartheta \right) \cdot \\
& \cdot \left( -\frac{a^2}{2} + 2r - r^2 - \frac{1}{2}a^2\cos^2\vartheta + \frac{1}{2a^2\sin^2\vartheta} \right) = \\
= & L_z^2\Sigma^2 \left( -\frac{a^2}{2} + 2r - r^2 - \frac{1}{2}a^2\cos^2\vartheta + \frac{1}{2a^2\sin^2\vartheta} \right) \qquad (7.24)
\end{aligned}
$$

Then, the term (†) becomes:

$$
(\dagger) = \frac{L_z^2\left(-\frac{a^2}{2} + 2r - r^2 - \frac{1}{2}a^2\cos^2\vartheta + \frac{1}{2}a^2\sin^2\vartheta\right)}{\sin^2\vartheta},
$$

and so the function $\widehat{R}$ can be now written as:

$$
\begin{aligned}
\widehat{R} = & \frac{L_z^2\left(-\frac{a^2}{2} + 2r - r^2 - \frac{1}{2}a^2\cos^2\vartheta + \frac{1}{2}a^2\sin^2\vartheta\right)}{\sin^2\vartheta} - \\
& - 4arEL_z - \mu^2\rho^2\Delta + \Sigma^2E^2 - Q\Delta + \\
& + \Delta\cos^2\vartheta\left(\frac{L_z^2}{\sin^2\vartheta} + a^2\left(\mu^2 - E^2\right)\right). \qquad (7.25)
\end{aligned}
$$

Substituting $\Sigma$ and $\Delta$ by their values, defined on equations 6.1, we obtain

$$
\begin{aligned}
\widehat{R} = & \frac{a^2L_z^2}{2} - \Delta Q - 4aL_zrE + a^4E^2 + 2a^2r^2E^2 + r^4E^2 - a^2r^2\mu^2 + \\
& + 2r^3\mu^2 - r^4\mu^2 + \\
& + \cos^2\vartheta\left(-a^4E^2 + 2a^2rE^2 - a^2r^2E^2\right) + \\
& + \sin^2\vartheta\left(-a^4E^2 + 2a^2rE^2 - a^2r^2E^2\right) + \\
& + \cot^2\vartheta\left(\frac{a^2L_z^2}{2} - 2L_z^2r + L_z^2r^2\right) + \\
& + \csc^2\vartheta\left(-\frac{a^2L_z^2}{2} + 2L_z^2r - L_z^2r\right). \qquad (7.26)
\end{aligned}
$$

We can simplify the last two pairs of terms using that $\sin^2\vartheta + \cos^2\vartheta = 1$ and that $\csc^2\vartheta - \cot^2\vartheta = 1$:

$$
\begin{aligned}
\widehat{R} = & \frac{a^2L_z^2}{2} - \Delta Q - 4aL_zrE + a^4E^2 + 2a^2r^2E^2 + r^4E^2 - a^2r^2\mu^2 + \\
& + 2r^3\mu^2 - r^4\mu^2 + \left(-a^4E^2 + 2a^2rE^2 - a^2r^2E^2\right) + \\
& + \left(-\frac{a^2L_z^2}{2} + 2L_z^2r - L_z^2r\right). \qquad (7.27)
\end{aligned}
$$

Factoring out common terms in the last two addends, using the definition of $\Delta$ and Equation 7.17, we simplify a little bit more:

$$
\begin{aligned}
\widehat{R} = {} & \frac{a^2L_z^2}{2} - \Delta Q - 4aL_zrE + a^4E^2 + 2a^2r^2E^2 + r^4E^2 - a^2r^2\mu^2 + \\
& + 2r^3\mu^2 - r^4\mu^2 + a^2E^2\left(-a^2 + 2r - r^2\right) + L_z^2\left(-\frac{a^2}{2} + 2r - r^2\right) = \\
= {} & \frac{a^2L_z^2}{2} - \Delta Q - 4aL_zrE + a^4E^2 + 2a^2r^2E^2 + r^4E^2 - a^2r^2\mu^2 + \\
& + 2r^3\mu^2 - r^4\mu^2 - a^2E^2\Delta + L_z^2\Delta + \frac{a^2L_z^2}{2} = \\
= {} & \overbrace{a^2L_z^2 + a^4E^2 + 2a^2E^2r^2 + r^4E^2 - 2L_zr^2Ea - 2a^3L_zE}^{P^2} + \\
& + 2L_zr^2Ea + 2a^3L_zE - \Delta Q - 4aL_zrE - a^2E^2\Delta - L_z^2\Delta - \\
& - a^2r^2\mu^2 + 2r^3\mu^2 - r^4\mu^2 - \Delta Q = \\
= {} & P^2 + \mu^2\overbrace{\left(-r^4 + 2r^3 - a^2r^2\right)}^{\Delta r^2} - L_z^2\Delta - (a^2E^2)\Delta + 2L_zEar^2 + \\
& + 2a^2L_zE - 4aL_zrE - Q\Delta = \\
= {} & P^2 + r^2\mu^2\Delta + L_z^2\Delta - a^2E^{\Delta} - 2L_zaE\left(2r - r^2 - a^2\right) - Q\Delta = \\
= {} & P^2 + r^2\mu^2\Delta - L_z^2\Delta - a^2E^2\Delta - 2L_zaE\Delta - Q\Delta = \\
= {} & P^2 - \Delta\left(r^2\mu^2 + Q + (L_z - aE)^2\right) \qquad (7.28)
\end{aligned}
$$

This proves that $\widehat{R} = R$, and therefore proves Equation 7.12.

*$\dot{\vartheta}$ Equation*

Using the definition of the Carter constant (Equation 7.5), we can directly find the value of $p_\vartheta$

$$
p_\vartheta^2 = Q + a^2E^2\cos^2\vartheta - a^2\mu^2\cos^2\vartheta - \frac{L_z^2}{\sin^2\vartheta}\cos^2\vartheta = \qquad (7.29)
$$

$$
= Q - \cos^2\vartheta\left(a^2\left(\mu^2 - E^2\right)‘\frac{L_z^2}{\sin^2\vartheta}\right). \qquad (7.30)
$$

Therefore, Equation 7.9 is directly equivalent to the following one:

$$
\rho^2\dot{\vartheta} = \sqrt{\Theta}, \qquad (7.31)
$$

where $\Theta$ is defined in Equation 7.16. This finally proves Equation 7.13.
■

## 7.2 VARIATIONAL FORMULATION

The study of the variational characterization of geodesics gave us an interesting result, from which we can change the problem of finding a geodesic by an equivalent one: to find the solution of a variational problem.

Proposition 3.48 states that $\gamma$ is a geodesic if and only if

$$\frac{dE_f}{ds}(0) = 0.$$

This lead us to understand geodesics —we could even define them that way— as the critical points of the energy, and yields a variational problem whose Lagrangian depends on the proper variation $f(s,t)$.

Using the characterization from Definition 3.46, the variational problem equivalent to finding the equations of motion for $\gamma$ has the following Lagrangian:

$$\mathcal{L} = g\left(\frac{df}{dt}, \frac{df}{dt}\right) \tag{7.32}$$

This Lagrangian can now be expressed in terms of $\gamma$ components. We can then write the usual form of the Lagrangian:

$$\mathcal{L} = \frac{1}{2}\mathbf{v}^{\mu}\mathbf{v}_{\mu} = \tag{7.33}$$

$$= \frac{1}{2}\left(\dot{t}\left(-\dot{\varphi}^2\omega\varpi^2 + \dot{t}\left(-\alpha^2 + \omega^2\varpi^2\right)\right) + \right. \tag{7.34}$$

$$\left. + \dot{r}^2\frac{\rho^2}{\Delta} + \dot{\vartheta}\rho^2 + \dot{\varphi}\left(\dot{\varphi}r^2 - \dot{t}\omega\varpi^2\right)\right) \tag{7.35}$$

This is similar to the formalism we developed in the previous section, but we can now consider the Hamiltonian version of this formulation, which appear simply by applying the Legendre transform:

$$\mathcal{H} = \sum p_i q_i - \mathcal{L}.$$

Our goal now is to recover the system described at [13, Eq. (A.15)]. In order to do that, we can rewrite $\mathcal{H}$ as follows:

$$\mathcal{H} = \frac{p_r^2\Delta}{2\rho^2} + \frac{p_\vartheta^2}{2\rho^2} + \mathfrak{f}, \tag{7.36}$$

where $\mathfrak{f}$ is the function consisting on the remaining terms of $\mathcal{H}$.

Although $\mathfrak{f}$ is completely defined and can be written as is, we are using the fact that $\mathcal{H} = \frac{-\mu^2}{2}$, and write it using the remaining terms.

Let us first work a little bit more on $\mathcal{H}$, rewriting Equation 7.36. First of all, we realize that we can write the definition of the components of $\mathbf{p}_\alpha$ from Equation 7.4:

$$p_t = -\dot{t}\alpha^2 - \dot{\varphi}\omega\varpi^2 + \dot{t}\omega^2\varpi^2 \tag{7.37}$$

$$p_r = \frac{\dot{r}\rho^2}{\Delta} \tag{7.38}$$

$$p_\vartheta = \dot{\vartheta}\rho^2 \tag{7.39}$$

$$p_\varphi = \dot{\varphi}\varpi^2 - \dot{t}\omega\varpi^2 \tag{7.40}$$

$$\tag{7.41}$$

Now, using these equations and the ones obtained in Theorem 7.1, we can rewrite $\mathcal{H}$ as follows:

$$\begin{aligned}
\mathcal{H} &= \frac{p_r^2\Delta}{2\rho^2} + \frac{p_\vartheta^2}{2\rho^2} + \mathfrak{f} \underset{\substack{\uparrow \\ 7.38,\ 7.39}}{=} \frac{\left(\frac{\dot{r}\rho^2}{\Delta}\right)^2 \Delta}{2\rho^2} + \frac{\left(\dot{\vartheta}\rho^2\right)^2}{2\rho^2} + \mathfrak{f} = \\
&= \dot{r}^2\frac{\rho^2}{2\Delta} + \dot{\vartheta}^2\rho^2 + \mathfrak{f} \underset{\substack{\uparrow \\ 7.12,\ 7.13}}{=} \frac{R}{\rho^4}\frac{\rho^2}{2\Delta} + \frac{\Theta}{\rho^4}\rho^2 + \mathfrak{f} = \\
&= \frac{R}{2\rho^2\Delta} + \frac{\Theta}{\rho^2} + \mathfrak{f}.
\end{aligned}$$

Now, using that $\mathcal{H} = \frac{-\mu^2}{2}$, we can express $\mathfrak{f}$ easily:

$$\mathfrak{f} = -\frac{R + \Delta\Theta}{2\Delta\rho^2} - \frac{\mu^2}{2}.$$

The final version of the Hamiltonian is:

$$\mathcal{H} = \frac{p_r^2\Delta}{2\rho^2} + \frac{p_\vartheta^2}{2\rho^2} - \frac{R + \Delta\Theta}{2\Delta\rho^2} - \frac{\mu^2}{2}. \tag{7.42}$$

From the general Hamilton's equations:

$$\dot{p}_i = -\frac{\partial\mathcal{H}}{\partial q_i}, \quad \dot{q}_i = \frac{\partial\mathcal{H}}{\partial p_i},$$

which in this case read as follows

$$\begin{aligned}
\dot{r} &= \frac{\partial\mathcal{H}}{\partial p_r}, \quad \dot{\vartheta} = \frac{\partial\mathcal{H}}{\partial p_\vartheta}, \quad \dot{\varphi} = \frac{\partial\mathcal{H}}{\partial p_\varphi} \\
\dot{p}_r &= -\frac{\partial\mathcal{H}}{\partial r}, \quad \dot{p}_\vartheta = -\frac{\partial\mathcal{H}}{\partial\vartheta},
\end{aligned}$$

we obtain the expected first order numerically stable system:

$$\dot{r} = \frac{\Delta}{\rho^2}p_r \tag{7.43}$$

$$\dot{\vartheta} = \frac{1}{\rho^2}p_\vartheta \tag{7.44}$$

$$\dot{\varphi} = \frac{\partial}{\partial p_\varphi}\left(\frac{R + \Delta\Theta}{2\Delta\rho^2}\right) \tag{7.45}$$

$$\dot{p}_r = -\frac{\partial}{\partial r}\left(-\frac{\Delta}{2\rho^2}p_r^2 - \frac{1}{2\rho^2}p_\vartheta^2 + \left(\frac{R + \Delta\Theta}{2\Delta\rho^2}\right)\right) \tag{7.46}$$

$$\dot{p}_\vartheta = -\frac{\partial}{\partial\vartheta}\left(-\frac{\Delta}{2\rho^2}p_r^2 - \frac{1}{2\rho^2}p_\vartheta^2 + \left(\frac{R + \Delta\Theta}{2\Delta\rho^2}\right)\right). \tag{7.47}$$

Part III

# COMPUTER SCIENCE

# 8 ANALYSIS OF THE PROBLEM

# 9 SOLUTION DESIGN

## 9.1 ALGORITHM DESCRIPTION

### 9.1.1 *Initial Conditions Computation*

### 9.1.2 *Raytracing*

## 9.2 PARALLELIZATION TECHNIQUES

### 9.2.1 *Classic Parallelization*

### 9.2.2 *General-Purpose Computing on Graphics Processing Units*

# 10 SOLUTION IMPLEMENTATION

## 10.1 TECHNOLOGIES USED

### 10.1.1 *Python*

### 10.1.2 *CUDA*

### 10.1.3 *PyCUDA*

### 10.1.4 *Documentation: Sphinx and Doxygen*

## 10.2 ALGORITHM IMPLEMENTATION

### 10.2.1 *Initial Conditions Computation*

### 10.2.2 *Raytracing*

## 10.3 CUDA PARALLELIZATION

### 10.3.1 *Device Setup*

### 10.3.2 *CUDA Kernel*

### 10.3.3 *Fine-Tuning*

# 11 DOCUMENTATION DEVELOPMENT

# 12 RESULTS

## 12.1 COMPUTATIONAL RESULTS

## 12.2 SCIENTIFIC RESULTS

# Part IV

# CONCLUSIONS

Change whole names by initials.

# REFERENCES


[1] William M. Boothby. *An introduction to differentiable manifolds and riemannian geometry*. Second. Vol. 120. Pure and applied mathematics. Academic Press, 1986.

[2] Theodor Bröcker and Klaus Jänich. *Introducción a la topología diferencial*. Trans. by Juan Vázquez. First. AC, 1977.

[3] K.S. Thorne C.W. Misner and J.A. Wheeler. *Gravitation*. San Francisco: Freeman, 1973.

[4] Manfredo Perdigão do Carmo. *Geometria Riemanniana*. Projeto Euclides. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 1979.

[5] Chi kwan Chan. *A massive parallel ODE integrator for performing general relativistic radiative transfer using ray tracing*. https://github.com/chanchikwan/gray. 2016.

[6] Chi kwan Chan, Dimitrios Psaltis and Feryal Ozel. 'GRay: a Massively Parallel GPU-Based Code for Ray Tracing in Relativistic Spacetimes'. In: *The Astrophysical Journal* 777.1 (2013). DOI: 10.1088/0004-637X/777/1/13. eprint: arXiv:1303.5057.

[7] José Ángel Fernández Roldán. 'Formas diferenciales con aplicaciones a la física'. Universidad de Salamanca, 2013.

[8] Ernst Hairer, Syvert P. Nørsett and Gerhard Wanner. *Solving Ordinary Differential Equations I. Nonstiff Problems*. Vol. 8. Springer Series in Computational Mathematics. Berlin New York: Springer-Verlag Berlin Heidelberg, 1993. ISBN: 978-3-540-56670-0.

[9] Ernst Hairer and Gerhard Wanner. *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*. Vol. 14. Springer Series in Computational Mathematics. Berlin New York: Springer-Verlag Berlin Heidelberg, 1996. ISBN: 978-3-642-05220-0.

[10] S.W. Hawking. 'The Occurrence of Singularities in Cosmology. III. Causality and Singularities'. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 300.1461 (1967), pp. 187–201.

[11] S.W. Hawking and G.F.R. Ellis. *The Large Scale Structure of Space-time*. Cambridge University Press, 1973.

[12] P.E. Ehrlich J.K. Beem and K.L. Easley. *Global Lorentzian Geometry*. New York: Dekker, 1996.

[13] Oliver James, Eugénie von Tunzelmann, Paul Franklin and Kip S Thorne. 'Gravitational lensing by spinning black holes in astrophysics, and in the movie Interstellar'. In: *Classical and Quantum Gravity* 32.6 (2015), p. 065001. URL: http://stacks.iop.org/0264-9381/32/i=6/a=065001.

[14] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*. Vol. 1. Wiley Inters. Publ., 1963.

[15] Katsumi Nomizu. *Fundamentals of Linear Algebra*. Second. Chelsea Publishing Company, 1979.

[16] Barrett O'Neill. *Semi-Riemannian Geometry. With applications to Relativity*. Ed. by Samuel Eilenberg and Hyman Bass. First. Pure and applied Mathematics. Academic Press, 1983.

[17] Barrett O'Neill. *The Geometry of Kerr Black Holes*. Wellesley, MA: A K Peters, 1995.

[18] F.J. Palomo and A. Romero. 'Certain Actual Topics on Modern Lorentzian Geometry'. In: *Handbook of Differential Geometry*. Vol. 2. Elsevier/North-Holland, Amsterdam, 2006, pp. 513–546.

[19] A. Romero. 'Fundamentos Matemáticos de la Relatividad General: Cosmología'. In: *Publicaciones del Deparamento de Matemáticas, Universidad de Murcia*. Gris 14. 1996, pp. 81–93.

[20] A. Romero. 'Geometría y Relatividad. Una introducción a la geometría básica de la teoría'. In: *II Jornadas: Lecciones de Matemáticas*. 14. Rev. S.A.E.M. Thales Epsilon. 1998, pp. 305–320.

[21] A. Romero. 'Geometría de Lorentz: de lenguaje a herramienta básica en Relatividad General'. In: *Conferencies FME, UPC, Curs Einstein*. Vol. 2. Rev. S.A.E.M. Thales Epsilon. 2006, pp. 127–151.

[22] Alfonso Romero. *Álgebra lineal y geometría*. Second. La Madraza, 1986.

[23] Alfonso Romero. *An introduction to certain topics on Lorentzian geometry*. Vol. 1. Topics in Modern Differential Geometry. Simon Stevein Transactions on Geometry, 2010, pp. 195–220.

[24] R. K. Sachs and H. Wu. *General Relativity for mathematicians*. Vol. 48. Graduate Texts in Math. Springer-Verlag, 1977.

[25] R. K. Sachs and H. Wu. 'General relativity and cosmology'. In: *Bulletin of American Mathematical Society* 83.6 (1977), pp. 1101–1164.

[26] Jason Sanders and Edward Kandrot. *CUDA by example: an introduction to general-purpose GPU programming*. First. Addison-Wesley, 2010.

# DECLARATION

Put your declaration here.

*Granada, December 2016*

Alejandro García Montoro

## COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured here:

http://postcards.miede.de/

*Final Version* as of 8th December 2016 (`classicthesis`).