Treball final de grau

GRAU DE MATEMÀTIQUES

Facultat de Matemàtiques i Informàtica
Universitat de Barcelona

Constructing the space-time

Autor: Adrián Giral de las Heras

Director: Dr. Carlos Currás Bosch
Realitzat a: Departament de matemàtiques i informàtica

Barcelona, 27 de juny de 2018

Contents

0.1 Introduction

The main goal of this work is to construct the space-time, a topological space used for physical modelling. In 1905, Albert Einstein published a revolutionary article that changed the perception of space and time. He proposed that there is a maximum speed that anything can travel, the speed of light, and its value is constant for any observer. This caused interest to many scientists to study a new type of model based on Einstein's ideas. The phenomena happening on a velocity close to the speed of light contradict the theory of Newton's mechanics, which uses Euclidean spaces for modelling. The first space proposed to model these effects is called Minkowski space-time, which is $\mathbb{R}^4$ with a non Euclidean metric (this will be explained in detail in this work). The theory that studies physical phenomena based on Einstein's ideas, and modelled in the Minkowski space-time, is called the Theory of special relativity. However, this theory was not enough. The Minkowski space-time can only describe physical phenomena when there is no gravity involved. It took 10 years for Einstein to create the theory of general relativity, which is compatible with Newton's theory of gravitation. The intuitive idea is that, in presence of masses, the space-time is no longer "flat" ($\mathbb{R}^4$ has null curvature) and it becomes a "curved" topological space. Hence, in order to construct the space-time, we shall introduce the basics of differential geometry, which studies the concepts of differential calculus applied to curved spaces. Such spaces are called *differentiable manifolds*. Concepts as the curvature and the metric are defined using tensor algebra, hence, we will also study the basics of tensor calculus and algebra.

This work is aimed at readers with knowledge of linear algebra, topology, differentiable calculus and basic classical mechanics. It is common to find many mathematicians who struggle to understand theoretical physics. The methodology and notation can be very confusing, being tough for a mathematician to acknowledge abstract physical concepts. This work tries to be a guide for a mathematician to understand the basics of general relativity.

Chapter 1

Tensor calculus and algebra

Convention: The vector spaces will be over $\mathbb{R}$ and of finite dimension.

1.1 Tensor product of vector spaces

Definition 1.1. *Let E and F be two vector spaces of dimensions n and m respectively, we define $E \otimes F$ as a vector field such that given a map $\tau : E \times F \to E \otimes F$, the pair $(E \otimes F, \tau)$ satisfies:*

- *τ is bilinear.*
- *If $e_1, ..., e_n$ is a basis of E and $v_1, ..., v_m$ is a basis of F, $\{\tau(e_i, v_j)\}$ is a base of $E \otimes F$.*

Note that the dimension of $E \otimes F$ is $n \cdot m$. We can not know by this definition if such a pair exists and if it is unique. Let $(E \otimes F, \tau)$ be a pair that satisfies the properties above. Given an analogous pair $(E\hat{\otimes}F, \tau')$ that also satisfies such properties, we will say that f is an isomorphism between these pairs if it is an isomorphism between $E \otimes F$ and $E\hat{\otimes}F$ and makes the following diagram commutative

$$\begin{tikzcd}
E \otimes F \arrow[dr, "f"] & \\
E \times F \arrow[u, "\tau"] \arrow[r, "\tau'"] & E\hat{\otimes}F.
\end{tikzcd}$$

Proposition 1.2. *Two pairs $(E \otimes F, \tau)$ and $(E\hat{\otimes}F, \tau')$ satisfying the conditions of the definition 1.1 are always isomorphic.*

Proof. Let $e_1, ..., e_n$ be a basis of E and $v_1, ..., v_m$ be a basis of F. In order to make the last diagram commutative, f has to satisfy $f(\tau(e_i, v_j)) = \tau'(e_i, v_j)$. As $\tau(e_i, v_j)$ is a basis of $E \otimes F$ and $\tau'(e_i, v_j)$ is a basis of $E\hat{\otimes}F$, f is uniquely determined and it maps a basis to a basis. Therefore, f is an isomorphism. $\square$

Proposition 1.3. *There exist a vector space $E \otimes F$ and a map $\tau : E \times F \to E \otimes F$ satisfying the conditions of the definition 1.1*

Proof. Let $E \otimes F$ be any vector space of dimension $n \cdot m$. Let $\{\epsilon_{i,j}\}_{i=1...n;j=1...m}$ a basis for $E \otimes F$, and let $\{e_i\}$ and $\{v_j\}$ be bases for E and F respectively. We define the bilinear mapping τ such that $\tau(e_i, v_j) = \epsilon_{i,j}$. In order to satisfy the second condition of 1.1, given $\{e'_i\}$ and $\{v'_j\}$ any other bases for E and F, $\{\tau(e'_i, v'_j)\}$ must be also a basis for $E \otimes F$, which is an elemental exercise of linear algebra. $\square$

With these two propositions we conclude the existence and uniqueness (except for isomorphisms) of the pair $(E \otimes F, \tau)$.

Definition 1.4. *We define the **tensor product** of E and F as the vector space $E \otimes F$.*

Proposition 1.5. *Given a multilinear map $f : E_1 \times ... \times E_s \to F$, where $E_1, ..., E_s$ and F are vector spaces, there exists a unique linear mapping $f' : E_1 \otimes ... \otimes E_s \to F$ that makes the following diagram commutative:*

$$\begin{tikzcd}
E_1 \otimes ... \otimes E_s \arrow[dr, "f'"] & \\
E_1 \times ... \times E_s \arrow[u, "\tau"] \arrow[r, "f"] & F,
\end{tikzcd}$$

where τ is the mapping $\tau(x_1, ..., x_s) = x_1 \otimes ... \otimes x_s$.

Proof. Let $\{e_{i_1}\}_{i_1=1,...,n_1}$ be a basis of E_1. Let $\{v_{i_2}\}_{i_2=1,...,n_2}$ be a basis of E_2. ... Let $\{w_{i_s}\}_{i_s=1,...,n_s}$ be a basis of E_s. As $\{e_{i_1} \otimes ... \otimes w_{i_s}\}_{i_1=1,...,n_1,...,i_s=1,...,n_s}$ is a basis of $E_1 \otimes ... \otimes E_s$ the condition of commutativity determines f' uniquely. $\square$

1.2 Definition of tensor

Definition 1.6. *Let E be a vector space and E* its dual space, a tensor T of type (k,l) on E is a multilinear map*

$$T : \overbrace{E \times \cdots \times E}^{k} \times \overbrace{E^* \times \cdots \times E^*}^{l} \longrightarrow K \tag{1.1}$$

For example, if $w \in E^*$, and $v \in E$, $w(v) \in \mathbb{R}$, therefore, we can understand w as a (1,0) tensor. We can also think about the vector v as a (0,1) tensor considering it as a map from E* to $\mathbb{R}$ defined by $v(w) = w(v)$.

1.3 The tensors space

Let k and l be natural numbers. We can define the sum of two (k,l) tensors T and R as $(T + R)(v_1, ..., v_k, w^1, ..., w^l) = T(v_1, ..., v_k, w^1, ..., w^l) + R(v_1, ..., v_k, w^1, ..., w^l)$ and if $\lambda \in \mathbb{R}$ we define $(\lambda T)(v_1, ..., v_k, w^1, ..., w^l) = \lambda T(v_1, ..., v_k, w^1, ..., w^l)$. It is easy to check that the space of (k,l) tensors, denoted as $T_{(k,l)}(E)$, is a $\mathbb{R}$-vector space.

Let K be a (k,0) tensor on E. Using the proposition 1.5, one can deduce that there is a bijective correspondence between the space of multilinear maps $\times^k E = \overbrace{E \times \cdots \times E}^{k} \to \mathbb{R}$ and the space of linear maps $\otimes^k E = \overbrace{E \otimes \cdots \otimes E}^{k} \to \mathbb{R}$. Moreover, $(\otimes^k E)^* \backsimeq \otimes^k E^*$. Let $\{e_i\}_{i=1,\ldots,n}$ be a basis for E and $\{w^j\}_{j=1,\ldots,n}$ its dual basis. We can write any element w of $\otimes^k E^*$ as $w = \sum_{j_1,\ldots,j_k} \alpha_{j_1,\ldots,j_k} w^{j_1} \otimes \cdots \otimes w^{j_k}$ and any $v \in \otimes^k E$ as $v = \sum_{i_1,\ldots,i_k} \lambda^{i_1,\ldots,i_k} e_{i_1} \otimes \cdots \otimes e_{i_k}$. The element w acts on v by $w(v) = \sum_{i_1,\ldots,j_k} \lambda^{i_1,\ldots,i_k} \alpha_{j_1,\ldots,j_k} w^{j_1}(e_{i_1}) \cdots w^{j_k}(e_{i_k})$.

Let us generalize this to a (k,l) tensor T. Using T, we can define a linear map $\otimes^k E \otimes^l E^* \to \mathbb{R}$, this is, an element of $(\otimes^k E \otimes^l E^*)^* = \otimes^k E^* \otimes^l E$, and conversely. Therefore, there is a bijective correspondence between a (k,l) tensor and an element of $\otimes^k E^* \otimes^l E$, this is, $T_{(k,l)}(E) \cong \otimes^k E^* \otimes^l E$.

We will use these results in order to find a basis for $T_{(k,l)}(E)$. Let $\{e_i\}_{i=1,\ldots,n}$ be a basis for E and $\{w^j\}_{j=1,\ldots,n}$ its dual basis. We know that the elements $\{w^{j_1} \otimes \cdots \otimes w^{j_k} \otimes e_{i_1} \otimes \cdots \otimes e_{i_l}\}_{j_1,\ldots,i_l \in \{1,\ldots,n\}}$ form a basis for $\otimes^k E^* \otimes^l E \backsimeq T_{(k,l)}(E)$. The dimension of $T_{(k,l)}(E)$ is n^{k+l} and we can write the element $T \in T_{(k,l)}(E)$ on such basis as:

$$T = \sum_{i_1,\ldots,j_k} \lambda^{i_1,\ldots,i_l}_{j_1,\ldots,j_k} w^{j_1} \otimes \cdots \otimes w^{j_k} \otimes e_{i_1} \otimes \cdots \otimes e_{i_l} \tag{1.2}$$

In order to simplify the notation, from now on we shall remove the summation symbols. This is commonly known as Einstein's notation.

Definition 1.7. *We define the* ***tensor product of two tensors*** T ***and*** T'*, of type* (k,l) *and* (k',l') *respectively, as the* $(k+k', l+l')$ *tensor denoted by* $T \otimes T'$ *satisfying:*

$$(T \otimes T')((v,v'),(w,w')) = T(v,w) \cdot T'(v',w'),$$

where $(v,w) \in \times^k E \times^l E^*$ *and* $(v',w') \in \times^{k'} E \times^{l'} E^*$.

1.4 Identification of tensors with linear or multilinear maps

Proposition 1.8. *Denoting as* $\mathcal{L}(E;E)$ *the space of linear maps from E to E, the map* $\tau : T_{(1,1)} \longrightarrow \mathcal{L}(E;E)$*, defined as* $\tau(x,y')(v) = y'(v)x$*, satisfies the conditions of the definition 1.1.*

Proof. Indeed, it is clear that τ is bilinear. Let $\{e_i\}$ be a basis of E and $\{w^j\}$ be a basis of E^*. The set of elements $\{\epsilon_{ij}\}$ given by $\epsilon_{ij}(v) = w^j(v)e_i$ are a basis of $\mathcal{L}(E;E)$, which can be proved using elementary linear algebra. □

Using proposition 1.2, we can conclude that $\mathcal{L}(E,E)$ is isomorphic to $E \otimes E^*$, therefore we can identify these two vector spaces.

Let $\mathcal{L}(E, E; \mathbb{R})$ be the space of bilinear maps from $E \times E$ to $\mathbb{R}$. We can identify $\mathcal{L}(E, E; \mathbb{R})$ with $E^* \otimes E^*$ by defining the mapping $\tau : E^* \times E^* \to \mathcal{L}(E, E; \mathbb{R})$ such that $\tau(x', y')(v_1, v_2) = x'(v_1)y'(v_2)$. Using a similar argument to the last proof, we conclude that τ satisfies both conditions of 1.1. Analogously, we can identify $\otimes^k E^*$ to $\mathcal{L}(\overbrace{E, \cdots, E}^{k}; \mathbb{R})$.

Let us see now the identification of $E \otimes (\otimes^k E)$ with the space $\mathcal{L}(\overbrace{E, \cdots, E}^{k}; E)$. Let $\tau : E \times \overbrace{E^* \times \cdots \times E^*}^{k} \longrightarrow \mathcal{L}(\overbrace{E, \cdots, E}^{k}; E)$ be the mapping defined as $\tau(x, y_1', ..., y_k')(v_1, ..., v_k) = y_1'(v_1) \cdots y_k'(v_k)x$. Using similar arguments to the proof of the proposition 1.8, we conclude that τ satisfies the conditions of 1.1. This allows us to identify of $E \otimes (\otimes^k E)$ with $\mathcal{L}(\overbrace{E, \cdots, E}^{k}; E)$.

1.5 Symmetric and antisymmetric tensors

The concepts of symmetric and antisymmetric tensors will only be applicable to pure tensors, i.e, for $(0, l)$ or $(k, 0)$ tensors. Let us see, for example, the case of $(0, l)$ tensors. Let S_l be the the permutation group of the elements $\{1, ..., l\}$. Let $\sigma \in S_l$, and consider the multilinear map (which will be denoted again by σ), $\sigma : \overbrace{E \times \cdots \times E}^{l} \to \otimes^l E$ defined as $\sigma(x_1, ..., x_l) = x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(l)}$. Using proposition 1.5, there will exist a unique $\sigma' : \otimes^l E \to \otimes^l E$ making commutative the diagram

$$\begin{tikzcd}
\otimes^l E \arrow[dr, "\sigma'"] & \\
\overbrace{E \times \cdots \times E}^{l} \arrow[u, "\tau"] \arrow[r, "\sigma"] & \otimes^l E.
\end{tikzcd}$$

We define the symmetrisation S and the antisymmetrisation A as the linear mappings $\otimes^l \to \otimes^l$ given by

$$S(T) = \tfrac{1}{r!} \textstyle\sum_{\sigma \in S_l} \sigma'(T), \qquad A(T) = \tfrac{1}{r!} \textstyle\sum_{\sigma \in S_l} \epsilon(\sigma)\sigma'(T),$$

where $\epsilon(\sigma)$ is the signature of the permutation σ.

Definition 1.9. *A tensor $T \in \otimes^l E$ is **symmetric** if $S(T) = T$ and is antisymmetric if $A(T) = T$.*

To define these concepts for a $(k, 0)$ tensor is analogous, changing the space E by its dual space.

1.6 Tensor contraction

Definition 1.10. *Let* $T = \lambda^{i_1,\ldots,i_l}_{j_1,\ldots,j_k} w^{j_1} \otimes \cdots \otimes w^{j_k} \otimes e_{i_1} \otimes \cdots \otimes e_{i_l}$ *be a (k,l) tensor on the basis described in the last subsection. For all* (α, β) *such that* $1 \leq \alpha \leq k$, $1 \leq \beta \leq l$, *we define the* (α, β) ***index contraction*** *as the linear map*

$$C^\alpha_\beta : \otimes^k E^* \otimes^l E \longrightarrow \otimes^{k-1} E^* \otimes^{l-1} E$$

such that $C^\alpha_\beta(T) = \delta_{i_\beta j_\alpha} \lambda^{i_1,\ldots,i_l}_{j_1,\ldots,j_k} w^{j_1} \otimes \cdots \otimes \hat{w}^{j_\alpha} \otimes \cdots \otimes w^{j_k} \otimes e_{i_1} \otimes \cdots \otimes \hat{e}_{i_\beta} \otimes \cdots \otimes e_{i_l}$, *where* δ_{ij} *is the Kronecker's delta and the hat on the elements means that they are removed.*

Note that the coefficients of $C^\alpha_\beta(T)$ will be $(C^\alpha_\beta(T))^{i_1,\ldots,i_{l-1}}_{j_1,\ldots,j_{k-1}} = \sum_{\gamma=1}^n \lambda^{i_1,\ldots,\gamma,\ldots,i_l}_{j_1,\ldots,\gamma,\ldots,j_k}$, where the superscript γ is on the α position and the subscript γ is on the β position.

Chapter 2

Differentiable manifolds

2.1 Definition of differentiable manifold

Definition 2.1. *A C^∞ **atlas** on a topological space M is a collection of pairs $\{(U_i, \varphi_i)\}_{i\in I}$ called local charts, where the U_i are open sets that cover M and, for each index i, φ_i is a homeomorphism of U_i onto an open subset of $\mathbb{R}^n$, that satisfies the following condition:*

For each pair $i, j \in I$ such that $U_i \cap U_j \neq \emptyset$, the map

$$\varphi_i(U_i \cap U_j) \xrightarrow{\varphi_j \circ \varphi_i^{-1}} \varphi_j(U_i \cap U_j) \tag{2.1}$$

is a C^∞ map of $\varphi_i(U_i \cap U_j) \in \mathbb{R}^n$ onto $\varphi_j(U_i \cap U_j) \in \mathbb{R}^n$.

We can define a C^r*atlas* just changing C^∞ by C^r in the last condition.

Definition 2.2. *Two n dimensional and C^∞ atlases are equivalent if their union is a n dimensional and C^∞ atlas.*

Definition 2.3. *A n dimensional and C^∞ **differentiable manifold** is a topological space M, Hausdorff, paracompact and equipped with an equivalence class of C^∞ atlases.*

2.2 Differenciable maps on manifolds

Definition 2.4. *Let M be a differentiable manifold and $\{(U_i, \varphi_i)\}_{i\in I}$ an atlas of M. A pair (U, φ), where U is an open set of M and φ a homeomorphism of U onto an open set of $\mathbb{R}^n$, is an admissible local chart of M if $\{(U, \varphi), (U_i, \varphi_i)\}_{i\in I}$ is an atlas of M.*

Definition 2.5. *Let M and N be two differentiable manifolds of dimensions m and n respectively. Let $p \in M$. A map $f : M \longrightarrow N$ is differentiable on p if there exist admissible local charts of M and N, (U, φ), (V, ψ), with $p \in U$, such that $f(U) \subset V$ and the map*

$$\varphi(U) \xrightarrow{\psi \circ f \circ \varphi^{-1}} \psi(V) \tag{2.2}$$

is differentiable on $\varphi(p)$. If f is differentiable $\forall p \in M$, we will say that f is differentiable.

Definition 2.6. *Let M be a differentiable manifold and let f be a differentiable map $f : M \to \mathbb{R}$. A map with this characteristics is called differentiable function.*

2.3 Tangent space

We want to find an equivalent concept of the differential of a smooth function on $\mathbb{R}^n$ applied on differentiable manifolds. As $\mathbb{R}^n$ is a differentiable manifold, the new concept of such derivative has to be compatible to the classic concept on euclidean spaces. In this chapter we will always consider M as a n dimensional differentiable manifold.

Definition 2.7. *Let $\gamma(t)$ be a curve on an open neighbourhood $I \subset \mathbb{R}$ of $a \in \mathbb{R}$ into M, this is, a differentiable map $\gamma : I \to M$. Let $p = \gamma(a)$ and let $\mathcal{F}(M)$ be the set of differentiable functions on M. The **tangent vector** of the curve on $\gamma(a) = p$ is the operator $\dot{\gamma}(a) : \mathcal{F}(U) \to \mathbb{R}$ defined by $f \mapsto (df(\gamma(t))/dt)_{t=a}$.*

Remark: The last expresion is well defined because $F : t \mapsto f(\gamma(t))$ is an ordinary differentiable function on t. To prove that, let (U, φ) be a local admissible chart containing $\gamma(a)$. $F(t) = f(\gamma(t)) = f(\varphi^{-1}(\varphi(\gamma(t))))$, and $\varphi \circ \gamma$ is differentiable by the definition of φ and γ, and $f \circ \varphi^{-1}$ is also differentiable because f is differentiable.

Let $\gamma(t)$ and $\alpha(s)$ be curves on M such that $\gamma(a) = \alpha(b) = p \in M$. Let us define the sum $(\dot{\gamma}(a) + \dot{\alpha}(b)) : \mathcal{F}(M) \to \mathbb{R}$ by $f \mapsto \dot{\gamma}(a)(f) + \dot{\alpha}(b)(f)$.

We also define the product of $\dot{\gamma}(a)$ by $\lambda \in \mathbb{R}$ as $(\lambda\dot{\gamma}(a))(f) = \lambda\dot{\gamma}(a)(f)$.

Theorem 2.8. *Using the same notation of above, the sum $\dot{\gamma}(a) + \dot{\alpha}(b)$ is the tangent vector of a curve on p, and so is $\lambda\dot{\gamma}(a)$.*

Proof. We want to find the curve which tangent vector on p is $\dot{\gamma}(a) + \dot{\alpha}(b)$. Let (U, φ) be a local chart where $p \in U \subset M$. We can suppose that $\varphi(p)$ is the origin of $\mathbb{R}^n$ because, if it is needed, we can change φ by $T \circ \varphi$, where T is a translation of $\mathbb{R}^n$. Let $\varphi(\gamma(t)) = (\gamma_1(t), ..., \gamma_n(t)) \in \mathbb{R}^n$ and $\varphi(\alpha(s)) = (\alpha_1(s), ..., \alpha_n(s)) \in \mathbb{R}^n$. We define the curve ψ into $\mathbb{R}^n$

$$\lambda \xrightarrow{\psi} ((\dot{\gamma}_1(a) + \dot{\alpha}_1(b))\lambda, ..., (\dot{\gamma}_n(a) + \dot{\alpha}_n(b))\lambda) \tag{2.3}$$

where $\dot{\gamma}_i(t)$ and $\dot{\alpha}_i(s)$ are ordinary derivatives. Let us consider the curve onto M, $\Psi(\lambda) = \varphi^{-1}(\psi(\lambda))$. Note that $\Psi(0) = p$. We are going to prove that Ψ is the curve that we want to find. Using the definition of tangent vector of Ψ on p, we have:

$$\dot{\Psi}(0)(f) = \tfrac{d}{d\lambda} f(\Psi(\lambda))_{\lambda=0} = \tfrac{d}{d\lambda} f \circ \varphi^{-1}((\dot{\gamma}_1(a) + \dot{\alpha}_1(b))\lambda, ..., (\dot{\gamma}_n(a) + \dot{\alpha}_n(b))\lambda)_{\lambda=0} =$$
$$\sum \partial_i (f \circ \varphi^{-1})_0 (\dot{\gamma}_n(a) + \dot{\alpha}_n(b)) = \tfrac{d}{dt} f \circ \varphi^{-1}(\gamma_1(t), ..., \gamma_n(t))_{t=a} + \tfrac{d}{ds} f \circ$$
$$\varphi^{-1}(\alpha_1(s), ..., \alpha_n(s))_{s=b} = \tfrac{d}{dt} f(\gamma(a)) + \tfrac{d}{ds} f(\alpha(b)) = \dot{\gamma}(a)(f) + \dot{\alpha}(b)(f).$$

To prove that $\lambda\dot{\gamma}(a)$ is the tangent vector of a curve on p, the procedure is analogous.

□

The last proposition shows that the set of tangent vectors of curves on $p \in M$ is a vector space.

Definition 2.9. *The set of tangent vectors of curves on $p \in M$ is called* ***tangent space****, represented by T_pM.*

Theorem 2.10. $Dim(T_pM) = Dim(M) = n$.

Proof. Using the same notation of the last proof, let us define the curve $\gamma : t \mapsto (0, ..., t, ..., 0)$(t on the i place and zeros on the rest). Let $\Gamma(t) = \varphi^{-1}(\gamma(t))$ be a curve on M, then the tangent vector of such curve on p applied on $f \in \mathcal{F}(M)$ is

$$\dot{\Gamma}(0)(f) = \frac{d}{dt} f \circ \varphi^{-1}(0, ..., t, ..., 0)_{t=0} = (\partial_i f \circ \varphi^{-1})_0 = (\partial_i f)_p. \tag{2.4}$$

Then, $(\partial_i)_p$ is the tangent vector of Γ on p. We will now prove that $(\partial_1)_p$, ...,$(\partial_n)_p$ form a basis of $T_p(M)$. Let $\alpha(s)$ be a smooth curve such that $\alpha(0) = p$ and let $\varphi(\alpha(s)) = (\alpha_1(s), ..., \alpha_n(s)) \in \mathbb{R}^n$, then

$$\dot{\alpha}(0)(f) = (\frac{df(\alpha(s))}{ds})_{s=0} = \sum (\partial_i f)_p \dot{\alpha}_i(0). \tag{2.5}$$

They are independent because if $\sum a_i (\partial_i)_p = 0$, applying such operator to the function α_i, we get $\sum a_i \delta_{ij} = 0 \Rightarrow a_i = 0$.

□

Definition 2.11. *Let $f \in \mathcal{F}(M)$, we define the* ***differential of f on*** *p as the operator $(df)_p \in T_p(M)^*$ defined by $(df)_p(v) = v(f)$ for all $v \in T_p(M)$.*

Note that, if the manifold is $\mathbb{R}^n$, $(df)_p \in T_p(\mathbb{R}^n)^*$, and this coincides with the classical concept of the directional derivative of a function on $\mathbb{R}^n$. From this definition, if $(U, \varphi = (x^1, ..., x^n))$ is a local chart containing p, and $x^1, ..., x^n$ are the coordinate functions on $\mathbb{R}^n$ of such a chart, $\{(dx^1)_p, ..., (dx^n)_p\}$ is the dual basis of $\{(\frac{\partial}{\partial x^1})_p, ..., (\frac{\partial}{\partial x^n})_p\}$.

2.4 Vector fields and 1-forms on manifolds

Definition 2.12. *The disjoint union of the tangent spaces of M, denoted by TM, is called* ***tangent bundle****. This is, $TM = \bigsqcup_{p \in M} T_pM = \bigcup_{p \in M} \{p\} \times T_pM$.*

Note that the tangent bundle is a differentiable manifold in its own right. To prove that, let us define the natural projection $\pi : TM \to M$ given by $\pi(p, v) = p$. Let $(U, \varphi = (x^1, ..., x^n))$ be a local chart of M. We can identify the tangent bundle of U with $U \times \mathbb{R}^n$ using the map $\phi : TU \to U \times \mathbb{R}^n$, $\phi(p, v_p) = (p; v^1, ..., v^n)$, where $v^1, ..., v^n$ are the components of the vector v_p on $\{(\frac{\partial}{\partial x^1})_p, ..., (\frac{\partial}{\partial x^n})_p\}$. It is evident that ϕ is a bijective mapping. We shall equip TU with the induced topology by ϕ^{-1}. Let us equip TM with the topology defined by; A subset of TM is an open set if and only if its intersection with any TU (being U an open set of M) is an open set of TU. One can check that TM is a paracompact and Hausdorff topological space. The dimension of TM is twice the dimension of M.

Definition 2.13. *A* ***vector field*** *V on a differentiable manifold M is a mapping $V : M \to TM$ such that $\pi \circ V$ is the identity map. V is* ***differentiable*** *(resp.* ***continuous****) if V is a differentiable map of manifolds (resp. continuous). Such a map is called a* ***section*** *of the tangent bundle.*

Note that if we have a chart $(U, \varphi = (x^1, ..., x^n))$, where $(x^1, ..., x^n)$ are the coordinate functions, then

$$V_p = V_1(p)(\frac{\partial}{\partial x^1})_p + ... + V_n(p)(\frac{\partial}{\partial x^n})_p \quad \forall p \in U \tag{2.6}$$

is differenciable on $p_0 \in U \iff V_1, ..., V_n$ are differentiable functions on p_0.

Definition 2.14. *The* ***Lie bracket*** *of two vector fields V and W of M, noted as $[V, W]$, is the only vector field that satisfies $[V, W]_p(f) = V_p(W(f)) - W_p(V(f))$, $\forall p \in M$, where f is any differentiable function defined on M.*

Let $\alpha, \beta \in \mathbb{R}$, $f, g \in \mathcal{F}(M)$ and X_1, X_2, X, Y_1, Y_2, Y and Z be smooth vector fields on M.

Properties:

- $[\alpha X_1 + \beta X_2, Y] = \alpha[X_1, Y] + \beta[X_2, Y]$
  $[X, \alpha Y_1 + \beta Y_2] = \alpha[X, Y_1] + \beta[X, Y_2]$
- $[X, Y] = -[Y, X]$
- $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$, also known as *Jacobi's identity*
- $[fX, gY] = fg[X, Y] + fX(g)Y - gY(f)X$, for any functions f and g.

Definition 2.15. *A* ***1-form*** *w on a smooth manifold M consists about assigning an element of T_pM^* to each $p \in M$.*

Having the same local chart as before, $w_p = w_1(p)(dx^1)_p + ... + w_n(p)(dx^n)_p$, $\forall p \in U$. We will say that the one-form is **differentiable** if w_i are differentiable functions on every local chart.

2.5 Tensors and tensor fields on manifolds; Riemannian manifold

Let M be a differentiable manifold, $p \in M$, and let $(U, \varphi = (x^1, ..., x^n))$ be a local chart such that $p \in U$. We can take $B = \{(\frac{\partial}{\partial x^1})_p, ..., (\frac{\partial}{\partial x^n})_p\}$ and $B^*= \{(dx^1)_p, ..., (dx^n)_p\}$ as basis for T_pM and T_pM^* respectively. A (k, l) tensor T_p on T_pM can be written as

$$T_p = t^{i_1,...,i_l}_{j_1,...,j_k}(\frac{\partial}{\partial x^{i_1}})_p \otimes ... \otimes (\frac{\partial}{\partial x^{i_l}})_p \otimes (dx^{j_1})_p \otimes ... \otimes (dx^{j_k})_p, \tag{2.7}$$

where $t^{i_1,...,i_l}_{j_1,...,j_k} = T_p((\frac{\partial}{\partial x^{i_1}})_p, ..., (\frac{\partial}{\partial x^{i_l}})_p, (dx^{j_1})_p, ..., (dx^{j_k})_p)$.

Example 2.16. A **metric tensor on** p is a symmetric and non degenerate (2,0) tensor on T_pM. With the same coordinates used in this section, we can express a metric tensor on p as $\mathcal{G}_p = g_{ij}\, dx^i \otimes dx^j$.

We can now generalize the concept of differentiable vector field applied to tensors.

Definition 2.17. *A (k,l)* ***tensor field*** *of type (k,l) on a differentiable manifold M consists about assigning a (k,l) tensor $T(p) \in \otimes^k T_pM \otimes^l T_pM^*$ to each $p \in M$.*

Just as in vector fields, we can represent a tensor field on U as:

$$T(p) = t^{i_1,...,i_l}_{j_1,...,j_k}(p)(\frac{\partial}{\partial x^{i_1}})_p \otimes ... \otimes (\frac{\partial}{\partial x^{i_l}})_p \otimes (dx^{j_1})_p \otimes ... \otimes (dx^{j_k})_p \quad \forall p \in U. \tag{2.8}$$

We will say that a tensor field is **differentiable** if $t^{j_1,...,j_k}_{i_1,...,i_l}(p)$ are differentiable functions on U.

Definition 2.18. *A* ***metric tensor field*** *is a $(2, 0)$ differentiable tensor field that consists about assigning a metric tensor $\mathcal{G}_p$ (defined in example 2.16) to every $p \in M$.*

Definition 2.19. *We define a* ***Riemannian manifold*** *as a pair $(M, \mathcal{G})$, where M is a smooth manifold and $\mathcal{G}$ is a positive defined metric tensor field on M. We will say that the manifold is* ***Pseudo-Riemannian*** *if $\mathcal{G}$ is not necessarily positive defined.*

Chapter 3

Connections

3.1 Definition of connection, geodesic curve

Definition 3.1. *Let $\mathcal{X}(M)$ be the set of smooth vector fields on M. A connection, or a covariant derivative operator, on M is a mapping*

$$\begin{array}{rcl} \mathcal{X}(M) \times \mathcal{X}(M) & \longrightarrow & \mathcal{X}(M) \\ (V, W) & \longmapsto & \nabla_V W \end{array}$$

such that:

- $\nabla_V (W_1 + W_2) = \nabla_V W_1 + \nabla_V W_2$.
- $\nabla_{V_1+V_2} W = \nabla_{V_1} W + \nabla_{V_2} W$.
- $\nabla_V (fW) = V(f)W + f \nabla_V W$, $f \in \mathcal{F}(M)$. *(Leibniz's rule)*
- $\nabla_{fV} W = f \nabla_V W$.

This is defined globally (on the entire manifold). It is proven in pages 122 and 123 of [1] that the constraint of ∇ to an open set $U \in M$ is also a connection on U. And conversely, let $\cup_{i \in I} U_i$ be an open cover of M and let ∇^{U_i} be a connection defined on U_i. We can define a connection ∇ on M if $\forall i, j \in I$ such that $U_i \cap U_j \neq \emptyset$, $\nabla^{U_i}|_{U_i \cap U_j} = \nabla^{U_j}|_{U_i \cap U_j}$. Using this results, if we know ∇ on every open set of an atlas of M, we can use local coordinates to calculate it.

Let $(U, \varphi = (x^1, ..., x^n))$ be a local chart and let $V = V^i \frac{\partial}{\partial x^i}$ and $W = W^j \frac{\partial}{\partial x^j}$ be two smooth vector fields on U, where W^j and V^i are differentiable functions on U. Using the definition of connection,

$$\nabla_V W = \nabla_V (W^j \tfrac{\partial}{\partial x^j}) = V(W^j) \tfrac{\partial}{\partial x^j} + W^j \nabla_{V^i \frac{\partial}{\partial x^i}} \tfrac{\partial}{\partial x^j} = V(W^j) \tfrac{\partial}{\partial x^j} + W^j V^i \nabla_{\frac{\partial}{\partial x^i}} \tfrac{\partial}{\partial x^j}.$$

As the field $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}$ is defined on U, we can express it as a linear combination of $\{\frac{\partial}{\partial x^k}\}$, therefore:

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma^k_{ij} \frac{\partial}{\partial x^k},$$

where the coefficients Γ^k_{ij} are differentiable functions on U, named **Christoffel symbols** of ∇ on (U, φ).

Definition 3.2. *We will say that the connection is **symmetric** or **torsion-free** if $\Gamma^k_{ij} = \Gamma^k_{ji}$. If ∇ is torsion-free, $\nabla_V W - \nabla_W V = [V, W]$.*

Note that

$$\nabla_V W|_p = (V_p(W^j) + W^j(p) V^i(p) \Gamma^k_{ij}(p)) \frac{\partial}{\partial x^j}|_p. \tag{3.1}$$

This means that the tangent vector $\nabla_V W|_p$ depends on V_p and on the restriction of W on a sufficiently small neighbourhood of p. Moreover, from this result we can deduce that $\nabla_V W|_p$ depends on V_p and on the value of W on a small smooth curve containing p such that the tangent vector on p is V_p. Therefore, given any smooth curve $\gamma : I \to M$ and $V \in \mathcal{X}(M)$, we can define $\nabla_{\dot\gamma} V$ because $\forall t \in I$, $\dot\gamma(t) \in T_{\gamma(t)}M$ and V is defined along γ.

Definition 3.3. *A differentiable vector field $V \in \mathcal{X}(M)$ is called **parallel** if $\nabla_{\dot\gamma} V = 0$.*

Let $a \in I$ and $v \in T_{\gamma(a)}M$. Using Picard's theorem, we deduce that there only exists a unique differentiable vector field V satisfying:

$$\nabla_{\dot\gamma} V = 0$$
$$V_{\gamma(a)} = v$$

The field V along the curve γ is commonly known as **parallel transport** of v along γ.

Definition 3.4. *(Geodesic curve) A geodesic curve with respect to a connection ∇ is a curve γ satisfying $\nabla_{\dot\gamma}\dot\gamma = 0$.*

Let $t \mapsto \gamma(t)$ be a parametrisation of γ. Using a local chart $(U, \varphi = (x^1, ..., x^n))$ and the equation (3.1), we can write $\nabla_{\dot\gamma}\dot\gamma = 0$ as

$$\frac{d^2 x^i}{dt^2} + \Gamma^i_{jk} \frac{dx^j}{dt} \frac{dx^k}{dt} = 0, \tag{3.2}$$

where Γ^i_{jk} are the Christoffel symbols relative to the basis $\{(\frac{\partial}{\partial x^1}), ..., (\frac{\partial}{\partial x^n})\}$.

3.2 Covariant derivative of a tensor field; Levi-Civita connection

In order to define the Levi-Civita connection, we shall generalize the concept of covariant derivative of a vector field to a tensor field.

Definition 3.5. *Let T be a (k,l) tensor field defined on M and ∇ a connection. $\forall V \in \mathcal{X}(M)$, $\nabla_V T$ is a (k,l) tensor field on M given by:*

$$\nabla_V T(V_1, ..., V_k; w^1, ..., w^l) = V(T(V_1, ..., w^l)) - \sum_{i=1}^k T(V_1, ..., \nabla_V V_i, ..., V_k; w^1, ..., w^l) - \sum_{j=1}^l T(V_1, ..., V_k; w^1, ..., \nabla_V w^j, ..., w^l).$$

From this definition, we can deduce that given a one-form w, the covariant derivative of w with respect to V is the one-form:

$$\nabla_V w(W) = V(w(W)) - w(\nabla_V W).$$

We can also infer that given $f \in \mathcal{F}(M)$, $\nabla_V f = V(f)$.

Given $p \in M$ and $v = V_p$, the mapping

$$\begin{aligned} T_p M &\longrightarrow \otimes^l T_p M \otimes^k T_p M^* \\ v &\longmapsto (\nabla_V T)_p \end{aligned}$$

is linear. We proved that we can identify such a mapping with a $(k+1, l)$ tensor. We define as ∇T_p such a tensor, and ∇T will be the tensor field that associates ∇T_p to each $p \in M$. Having a local chart $(U, \varphi = (x^1, ..., x^n))$, ∇T is expressed as

$$\nabla T_p = \nabla_i T^{i_1, ..., i_l}_{j_1, ..., j_k}(p) (\frac{\partial}{\partial x^{i_1}})_p \otimes ... \otimes (\frac{\partial}{\partial x^{i_l}})_p \otimes (dx^{j_1})_p \otimes ... \otimes (dx^{j_k})_p \otimes (dx^i)_p,$$

where $\nabla_i T^{i_1, ..., i_l}_{j_1, ..., j_k}$ are the function components of ∇T. Such components functions applied on p will be the matrix of the linear mapping above in the basis $\{\frac{\partial}{\partial x^i}\}$ of $T_p M$ and $\{\frac{\partial}{\partial x^{i_1}} \otimes ... \otimes \frac{\partial}{\partial x^{i_l}} \otimes dx^{j_1} \otimes ... \otimes dx^{j_k}\}$ of $\otimes^l T_p M \otimes^k T_p M^*$.

Definition 3.6. *A* ***parallel tensor field*** *is a tensor field T that satisfies $\nabla_V T = 0$, $\forall V \in \mathcal{X}(M)$,i.e, $\nabla T = 0$.*

Definition 3.7. *Given a Riemannian (or pseudo-Riemannian) manifold $(M, \mathcal{G})$, a connection ∇ is a* ***Levi-Civita connection*** *if it is torsion-free and $\nabla \mathcal{G} = 0$.*

Theorem 3.8. *There exist a unique Levi-Civita connection for each pseudo-Riemannian manifold $(M, \mathcal{G})$.*

Proof. Let us write the condition $\nabla \mathcal{G} = 0$ thrice, permuting the fields V, W and Z,

$$V(\mathcal{G}(W, Z)) = \mathcal{G}(\nabla_V W, Z) + \mathcal{G}(W, \nabla_V Z)$$

$$Z(\mathcal{G}(V, W)) = \mathcal{G}(\nabla_Z V, W) + \mathcal{G}(V, \nabla_Z W)$$

$$W(\mathcal{G}(Z, V)) = \mathcal{G}(\nabla_W Z, V) + \mathcal{G}(Z, \nabla_W V).$$

Summing the first two equations and subtracting the third one, using that ∇ is torsion free and $\mathcal{G}$ symmetric, we get:

$$V(\mathcal{G}(W, Z)) + Z(\mathcal{G}(V, W)) - W(\mathcal{G}(Z, V)) =$$
$$\mathcal{G}([V, W], Z) + \mathcal{G}([Z, W], V) + \mathcal{G}([V, Z], W) + 2\mathcal{G}(\nabla_Z V, W).$$

Therefore:

$$\mathcal{G}(\nabla_Z V, W) = \tfrac{1}{2}(V(\mathcal{G}(W,Z)) + Z(\mathcal{G}(V,W)) - W(\mathcal{G}(Z,V)) -$$
$$\mathcal{G}([V,W],Z) - \mathcal{G}([Z,W],V) - \mathcal{G}([V,Z],W)). \tag{3.3}$$

This expression let us know the scalar product of $(\nabla_Z V)_p$ with any other W_p, determining a unique $\nabla_Z V$ on each $p \in M$. One can easily prove the existence of ∇ by showing that the field $\nabla_Z V$ determined by (3.3) satisfies the properties of a connection. □

From now on, every connection ∇ that we use will be the Levi-Civita connection. The equation (3.3) is commonly known as Levi-Civita equation. Given a local chart $(U, \varphi = (x^1, ..., x^n)$, let us see how the Christoffel symbols are determined in the basis $(\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^n})$. Let $\mathcal{G}_{ij} = \mathcal{G}(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})$. Using the constraint of the connection in U, let us write the Levi-Civita equation (3.3) for the fields $\frac{\partial}{\partial x^i}$, $\frac{\partial}{\partial x^j}$ and $\frac{\partial}{\partial x^k}$. As $[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0$, the equation is written as

$$\Gamma^r_{ij}\mathcal{G}_{rk} = \tfrac{1}{2}\{\tfrac{\partial}{\partial x^i}(\mathcal{G}_{jk}) + \tfrac{\partial}{\partial x^j}(\mathcal{G}_{ki}) - \tfrac{\partial}{\partial x^k}(\mathcal{G}_{ij})\}.$$

Therefore, designing as $(\mathcal{G}^{ij})$ to the inverse matrix of $(\mathcal{G}_{ij})$,

$$\Gamma^r_{ij} = \frac{1}{2}\mathcal{G}^{rk}\{\frac{\partial}{\partial x^i}(\mathcal{G}_{jk}) + \frac{\partial}{\partial x^j}(\mathcal{G}_{ki}) - \frac{\partial}{\partial x^k}(\mathcal{G}_{ij})\}. \tag{3.4}$$

3.3 The Riemann curvature tensor field

Definition 3.9. *The Riemann curvature tensor field, or* ***Riemann tensor****, is the (3,1) tensor field given by:*

$$R(V,W,Z) = \nabla_V\nabla_W Z - \nabla_W\nabla_V Z - \nabla_{[V,W]}Z \tag{3.5}$$

where V, W and Z are smooth vector fields.

If $\mathcal{X}(M)$ is the space of smooth vector fields on M, it is easy to show that $R : \mathcal{X}(M) \times \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$, $(V,W,Z) \mapsto \nabla_V\nabla_W Z - \nabla_W\nabla_V Z - \nabla_{[V,W]}Z$, is a $\mathcal{F}(M)$-multilinear mapping, therefore R is a (3,1) tensor field. Also, from the definition, R is antisymmetric on the first two components. From now on we will use $R(V,W)Z$ instead of $R(V,W,Z)$.

Let us now find an expression of R in local coordinates. Given a local chart of M, $(U, \varphi = (x^1, ..., x^n))$,

$$R(\tfrac{\partial}{\partial x^i}, \tfrac{\partial}{\partial x^j})\tfrac{\partial}{\partial x^k} = R^l_{kij}\tfrac{\partial}{\partial x^l}.$$

We want to find the expression for R^l_{kij} in function of the Christoffel symbols. Using that $[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0$,

$$R(\tfrac{\partial}{\partial x^i}, \tfrac{\partial}{\partial x^j})\tfrac{\partial}{\partial x^k} = \nabla_{\frac{\partial}{\partial x^i}}(\Gamma^s_{jk}\tfrac{\partial}{\partial x^s}) - \nabla_{\frac{\partial}{\partial x^j}}(\Gamma^s_{ik}\tfrac{\partial}{\partial x^s}) = (\tfrac{\partial\Gamma^s_{jk}}{\partial x^i} - \tfrac{\partial\Gamma^s_{ik}}{\partial x^j})\tfrac{\partial}{\partial x^s} + (\Gamma^s_{jk}\Gamma^l_{is} - \Gamma^s_{ik}\Gamma^l_{js})\tfrac{\partial}{\partial x^l} =$$
$$(\tfrac{\partial\Gamma^l_{jk}}{\partial x^i} - \tfrac{\partial\Gamma^l_{ik}}{\partial x^j} + \Gamma^s_{jk}\Gamma^l_{is} - \Gamma^s_{ik}\Gamma^l_{js})\tfrac{\partial}{\partial x^l}.$$

Therefore:

$$R^l_{kij} = \frac{\partial \Gamma^l_{jk}}{\partial x^i} - \frac{\partial \Gamma^l_{ik}}{\partial x^j} + \Gamma^s_{jk}\Gamma^l_{is} - \Gamma^s_{ik}\Gamma^l_{js} \tag{3.6}$$

Proposition 3.10. *The Riemann curvature tensor field satisfies the following identity:*

$$\mathcal{G}(R(V,W)Z,Y) = -\mathcal{G}(R(V,W)Y,Z)$$

Proof. We have

$$\mathcal{G}(R(V,W)Z,Y) = \mathcal{G}((\nabla_V\nabla_W - \nabla_W\nabla_V - \nabla_{[V,W]})Z,Y).$$

Moreover, as $\nabla\mathcal{G} = 0$,

$$\mathcal{G}(\nabla_V\nabla_W Z,Y) = V(\mathcal{G}(\nabla_W Z,Y) - \mathcal{G}(\nabla_W Z,\nabla_V Y) =$$
$$VW\mathcal{G}(Z,Y) - V\mathcal{G}(Z,\nabla_W Y) - W\mathcal{G}(Z,\nabla_V Y) + \mathcal{G}(Z,\nabla_W\nabla_V Y).$$

Doing the same with $\mathcal{G}(\nabla_W\nabla_V Z,Y)$ and substituting in the first equation of the proof, we obtain $\mathcal{G}(R(V,W)Z,Y) = -\mathcal{G}(R(V,W)Y,Z)$. □

3.4 Second Bianchi identity

Let M be a pseude-Riemannian manifold. The Riemann curvature tensor field R is a $(3,1)$ tensor field. Therefore, given a vector field V, $\nabla_V R$ will also be a $(3,1)$ tensor field. Hence, for all $p \in M$, we can identify R_p with a linear map $T_pM \times T_pM \times T_pM \longrightarrow T_pM$.

Proposition 3.11. $\nabla_V R$ *satisfies*

$$\nabla_V R(W,Z,Y) = \nabla_V(R(W,Z)Y) - R(\nabla_V W,Z)Y - R(W,\nabla_V Z)Y - R(W,Z)\nabla_V Y.$$

Proof. Identifying the $(3,1)$ tensor field R with a mapping $T_pM \times T_pM \times T_pM \longrightarrow T_pM$, we can write $R(W,Z)Y = C^1_1C^2_2C^3_3(W \otimes Z \otimes Y \otimes R)$, where C^i_i is the index contraction defined in section 1. Therefore, $\nabla_V(R(W,Z)Y) = C^1_1C^2_2C^3_3(\nabla_V W \otimes Z \otimes Y \otimes R) + C^1_1C^2_2C^3_3(W \otimes \nabla_V Z \otimes Y \otimes R) + C^1_1C^2_2C^3_3(W \otimes Z \otimes \nabla_V Y \otimes R) + C^1_1C^2_2C^3_3(W \otimes Z \otimes Y \otimes \nabla_V R)$, which lead us to the equation that we want to prove. □

Proposition 3.12. *(second Bianchi identity) The Riemann curvature tensor field satisfies the following identity:*

$$(\nabla_V R)(W,Z,Y) + (\nabla_W R)(Z,V,Y) + (\nabla_Z R)(V,W,Y) = 0.$$

We will write this identity as $\mathcal{P}_{V,W,Z}(\nabla_V R)(W,Z,Y) = 0$

Proof. $\mathcal{P}_{V,W,Z}(\nabla_V R)(W,Z,Y) =$
$$\mathcal{P}_{V,W,Z}\{\nabla_V(R(W,Z)Y) - R(\nabla_V W,Z)Y - R(W,\nabla_V Z)Y - R(W,Z)\nabla_V Y\} =$$
$$\mathcal{P}_{V,W,Z}\{\nabla_V(R(W,Z)Y) - R(\nabla_V W,Z)Y - R(W,\nabla_V Z)Y - R(W,Z)\nabla_V Y\}.$$

As ∇ is torsion-free, we can write this expression as

$$\mathcal{P}_{V,W,Z}\{\nabla_V(R(W,Z)Y) - R(\nabla_V W,Z)Y - R(W,\nabla_Z V)Y - R(W,[V,Z])Y - R(W,Z)\nabla_V Y\} =$$
$$\mathcal{P}\nabla_V(R(W,Z)Y) - \mathcal{P}R(\nabla_V W,Z)Y + \mathcal{P}R(\nabla_Z V,W)Y - \mathcal{P}R(W,[V,Z])Y - \mathcal{P}R(W,Z)\nabla_V Y,$$

where $\mathcal{P}R(V,W)Z = \mathcal{P}R(W,Z)V + \mathcal{P}R(Z,V)W + \mathcal{P}R(V,W)Z$. However, as $R(V,W)Z = \nabla_V\nabla_W Z - \nabla_W\nabla_V Z - \nabla_{[V,W]}Z$, we can consider $R(V,W)$ as the operator $[\nabla_V,\nabla_W] - \nabla_{[V,W]}$. Therefore,

$$\nabla_V(R(W,Z)Y) - R(W,Z)\nabla_V Y = [\nabla_V, R(W,Z)]Y = [\nabla_V,[\nabla_W,\nabla_Z]]Y - [\nabla_V,\nabla_{[W,Z]}]Y,$$

and

$$R(W,[V,Z]) = [\nabla_W,\nabla_{[V,Z]}] - \nabla_{[W,[V,Z]]}.$$

Substituting above these results, and using Jacobi's identity, we prove the theorem. □

3.5 Musical isomorphism

We can use the metric $\mathcal{G}$ to define the canonical isomorphism $b : T_pM \longrightarrow T_pM^*$ defined by $b(v)(w) = \mathcal{G}(v,w)$. We have

$$b(\frac{\partial}{\partial x^i}) = \mathcal{G}_{ij}dx^j,$$

where $\mathcal{G}_{ij}$ are the components of the matrix of $\mathcal{G}$, defined as $g_{ij} = g(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j})$. The inverse isomorphism, denoted by $\sharp$, is given by

$$\sharp(dx^i) = \mathcal{G}^{ij}\frac{\partial}{\partial x^i},$$

where $\mathcal{G}^{ij}$ are the components of the inverse matrix of $\mathcal{G}$. Using $\sharp$, we can define the following isomorphism $\sharp'$ as

$$\begin{aligned}\sharp' : \otimes^k T_pM^* &\longrightarrow T_pM \otimes (\otimes^{k-1}T_pM^*)\\ w^1\otimes ...\otimes w^k &\longmapsto \sharp(w^1)\otimes w^2\otimes ...\otimes w^k.\end{aligned}$$

$\sharp'$ on each point determines an isomorphism of tensor fields called **musical isomorphism**. Physicists usually refer to "raise and lower indices" when using the musical isomorphism. The inverse of $\sharp'$ is denoted as b'.

3.6 Divergence of a tensor field

Let T be a (k,l) tensor field on a pseudo-Riemannian manifold (M,g) such that $l > 0$. We define the *divergence* of T as

$$div(T) = C_1^{k+1}(\nabla T),$$

where C is the index contraction that contracts the first index of T_pM with the last index of T_pM^*. Note that $div(T)$ is a $(k,l-1)$ tensor field. This definition is compatible with the classic definition of the divergence of a vector field in $\mathbb{R}^n$. Indeed, if V is a vector field on a smooth manifold, $div(V) = \nabla_i V^i$. If the manifold is $\mathbb{R}^n$ with the Euclidean metric,

$\nabla_i V^i = \partial_i V^i$.

We want to define the divergence of a tensor field for $l = 0$ and $k \neq 0$ on each $p \in M$. If T is a $(k, 0)$ tensor field, $\sharp' T$ is a $(k-1, 1)$ tensor field, and we define

$$divT = div(\sharp' T).$$

3.7 Ricci tensor field and scalar curvature

Definition 3.13. *Let R be the Riemann tensor field of a smooth manifold M. The **Ricci Tensor field** is the (2,0) tensor field defined by $Ric = C_1^1(R)$.*

Let us use the usual local chart, and let R_{ij} be the components of the Ricci tensor, i.e., $Ric = R_{ij} dx^i \otimes dx^j$. The Riemann tensor in this basis will be $R = R^l_{ijk} \frac{\partial}{\partial x^l} \otimes dx^i \otimes dx^j \otimes dx^k$. Therefore:

$$Ric = C_1^1(R) = C_1^1(R^l_{ijk} \frac{\partial}{\partial x^l} \otimes dx^i \otimes dx^j \otimes dx^k) = R^l_{ljk} dx^j \otimes dx^k,$$

hence, we have the equality:

$$R_{ij} = R^l_{lij}. \tag{3.7}$$

Definition 3.14. *The **scalar curvature** is the function R given by $R = C(\sharp' Ric)$, where $\sharp'$ is the musical isomorphism, and C is the only possible contraction.*

From now on we will refer R to the scalar curvature (not to the Riemann curvature tensor field). Given a local chart,

$$R = \mathcal{G}^{ij} R_{ij}.$$

Chapter 4

Basics of special relativity

4.1 Introduction

At the end of the XIX century, it was believed that Newton's mechanics described the concepts of speed and force for any inertial (non-accelerating) frame of reference. This was based on the theory of the Galilean relativity. Such a theory predicts that any inertial frame of reference is valid to describe the movement of a particle. This was based on the Galilean transformation of the components of space and time between frames of reference. Let us see how this transformation works. For instance, consider two inertial frames of reference F and F', and let $(r,t) \in \mathbb{R}^3 \times I$ and $(r',t') \in \mathbb{R}^3 \times I$ be the space and time components of a particle in F and F' respectively. If F' moves with a constant speed v from F, the Galilean transformation is given by

$$\begin{aligned} r' &= r - vt \\ t' &= t \end{aligned}$$

being (t,r) and (t',r') both equivalent.

Remark 4.1. Note that these expressions transform the origin of F to the origin of F' at $t = 0$. This means that these transformations are refereed to two frames of reference whose origins coincide at one instant. If the origins do not coincide, we use the corresponding translation in order to use the transformations above. With no lack of generality , from now on we will only consider transformations between frames of references whose origins are invariant.

One can deduce from Newton's mechanics that there is not a maximum speed. This is because what ever the speed of a particle is in F, we can always have a frame of reference F' in where the speed of the particle is greater. Experimentally, it is impossible to exceed a certain speed when accelerating charges using a voltage. If we assume that all inertial frames of reference are equally valid and we find one in which there is a limit speed, we conclude that the Galilean transformations are wrong. One can also deduce that such a transformation is not right because it doesn't satisfy the Maxwell equations for the electromagnetism. Therefore, there was a need to develop a new theory to correct these contradictions. Such a theory is the Einstein's theory of relativity.

4.2 Postulates; Lorentz transformation

The postulates used by Einstein to develop the theory of special relativity were:

- First postulate (principle of relativity): The laws of physics are the same in all inertial frames of reference.
- Second postulate (invariance of c): The speed of light in free space has the same value c in all inertial frames of reference.

The Galilean transformation doesn't satisfy the second postulate. Therefore, it is needed to use a new method to transform the components (r, t) in F to (r', t') in a frame F' that moves with a constant vector speed v from F. Such a method is the linear transformation called *Lorentz transformation* and it is given by

$$ct' = \gamma(ct - \tfrac{v\cdot r}{c})$$
$$r' = r + (\gamma - 1)\tfrac{r\cdot v}{v\cdot v}v - \gamma v t$$

where $\gamma = \frac{1}{\sqrt{1-\frac{v\cdot v}{c^2}}}$ (Lorentz factor). For the derivation of the Lorentz transformation see [2] sections 14.3 and 14.6. In these expressions, "$\cdot$" represents the scalar product in $\mathbb{R}^3$ with respect to the Euclidean metric. Such a transformation will only make sense if $|v| < c$, hence, an additional postulate is needed, that restricts the value of $|v|$ as seen above. Note that for $v \ll c$, $\gamma(v) \approx 1$, and the Lorentz transformation in this case will be approximately $ct' \approx ct$ and $r' \approx r - vt$, i.e., the Galilean transformation.

4.3 Consequences of the Lorentz transformation

Let F and F' be inertial frames of reference such that F' moves parallel to the x axis of F at a constant speed v such that $0 < v < c$. Let P be a motionless point that in F is written as $P = (p, 0, 0)$. The coordinates of P in F' at t' are given by:

$$(x', y', z') = (\tfrac{p-vt}{\sqrt{1-v^2/c^2}}, 0, 0)$$
$$t' = \tfrac{-(v/c^2)p+t}{\sqrt{1-v^2/c^2}}.$$

Isolating t in the second equation and substituting it in to the first equation, $(x', y', z') = (p\sqrt{1-v^2/c^2} - vt', 0, 0)$. Let Q be a different motionless point that in F is $Q = (q, 0, 0)$. Following the same procedure, Q described in F' at the instant t' is $(q\sqrt{1-v^2/c^2} - vt', 0, 0)$. Note that the distance between P and Q in F is $|p - q|$ and the distance in F' is given by $|p-q|\sqrt{1-v^2/c^2}$. As $\sqrt{1-v^2/c^2} < 1$, the distance measured by an observer of F' is shorter than the distance measured by an observer of F. This phenomenon is known as *length contraction*.

Using the same frames of reference of above, assume two different events happening in the same point P of F, $P = (p^1, p^2, p^3)$, at different instants t_0 and t_1. These events will be seen in F' at the times t'_0 and t'_1 given by

$$t'_i = \frac{-(v/c^2)p^1 + t_i}{\sqrt{1 - v^2/c^2}}, \qquad i = 0, 1.$$

Therefore, $t'_1 - t'_0 = \frac{t_1 - t_0}{\sqrt{1-v^2/c^2}}$ and $(t'_1 - t'_0) > (t_1 - t_0)$. This effect is known as *time dilation*. To put an example, this means that if someone travels on a high speed spaceship, his clock will go slower than the clock of someone staying on Earth. This is because, from Earth, the spaceship reaches different points of space A and B at t_A and t_B respectively, while in the spaceship the events (t_A, A) and (t_B, B) happen in the same place. Therefore, the observer in the spaceship is in the frame of reference F, and the Earth is the frame of reference F'.

Let us now assume two different events happening in different points of F, $P = (p^1, p^2, p^3)$ and $Q = (q^1, q^2, q^3)$, at the same time t_0, i.e., simultaneous events in F. These two events will be seen by an observer of F' at the instants t'_1 and t'_2 given by

$$t'_1 = \frac{-(v/c^2)p^1 + t_0}{\sqrt{1-v^2/c^2}}, \qquad t'_2 = \frac{-(v/c^2)q^1 + t_0}{\sqrt{1-v^2/c^2}}.$$

Therefore, if $p^1 \neq q^1$, these events will not be simultaneous for an observer of F'. Contrary to Newton's mechanics, in special relativity **simultaneity is not absolute**.

Let us see another important consequence of the Lorentz transformation that will be useful to define a metric in space-time.

Definition 4.2. *Given an inertial frame of reference F, and two points of $\mathbb{R}^4$ in F, (t_1, r_1) and (t_2, r_2), we define the* ***space-time interval*** *as $c^2(t_2 - t_1)^2 - |r_2 - r_1|^2$, where $|\cdot|$ is the Euclidean norm in $\mathbb{R}^3$.*

Let (t'_1, r'_1) and (t'_2, r'_2) be the same two points in any other inertial frame of reference F'. One can prove by using the Lorentz transformation that $c^2(t_2 - t_1)^2 - |r_2 - r_1|^2 = c^2(t'_2 - t'_1)^2 - |r'_2 - r'_1|^2$ (we will leave this as an exercise for the reader) . Therefore, the value of is independent of the chosen frame of reference.

4.4 Minkowski space-time

In this section we will describe the first model of space-time. We want to keep constant the distance between two events when changing the frame of reference. Using the invariance of the space-time interval, let us define the following scalar product $\mathcal{G}$ in $\mathbb{R}^4$. Given an inertial frame of reference, consider two different times t_x and t_y, and let $x = (ct_x, x^1, x^2, x^3)$ and $y = (ct_y, y^1, y^2, y^3)$ be two different events, where c is the speed of light. We define

$$\mathcal{G}(x, y) = -c^2 t_x t_y + x^1 y^1 + x^2 y^2 + x^3 y^3.$$

Note that ct is a longitude magnitude, being measured in meters as are the other components.

Proposition 4.3. *If f is the Lorentz transformation between two inertial frames of reference, then*

$$\mathcal{G}(x,y) = \mathcal{G}(f(x), f(y)).$$

Proof. Using that $\mathcal{G}$ is bilinear and symmetrical with respect of the two arguments, we can write

$$\mathcal{G}(x,y) = \tfrac{1}{2}\{\mathcal{G}(x+y, x+y) - \mathcal{G}(x,x) - \mathcal{G}(y,y)\}.$$

We know that $c^2(t_2-t_1)^2 - |r_2-r_1|^2 = -\mathcal{G}((ct,r),(ct,r)) = -\mathcal{G}(f(ct,r), f(ct,r))$, therefore $\mathcal{G}(x+y,x+y)$, $\mathcal{G}(x,x)$ and $\mathcal{G}(y,y)$ are invariant when applying f. $\square$

Definition 4.4. *The **Minkowski space-time** $\mathcal{M}$ is a pseudo-Riemannian manifold with null Riemann curvature and isometric to the manifold $(\mathbb{R}^4, \mathcal{G})$. The metric tensor field $\mathcal{G}$ can be written in a Cartesian coordinate system as*

$$\mathcal{G} = -dx^0 \otimes dx^0 + dx^1 \otimes dx^1 + dx^2 \otimes dx^2 + dx^3 \otimes dx^3.$$

and in matrix form as

$$\mathcal{G} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The isometry is given by the identification $(x^0, x^1, x^2, x^3) \mapsto (ct, x, y, z)$. Using such identification, the tensor field $\mathcal{G}$ can be written as $\mathcal{G} = -c^2 dt \otimes dt + dx \otimes dx + dy \otimes dy + dz \otimes dz$.

Let F be an inertial frame of reference in the Minkowski space-time $\mathcal{M}$, we define the set C of F as

$$C := \{(ct, x, y, z) \in \mathcal{M} | -c^2t^2 + x^2 + y^2 + z^2 < 0\}.$$

C has two connected components C^+ and C^-. We define C^+ as the connected component of C contained in the semi-space $t > 0$ (the future). C is the set of events that can have a relation of causality with the event in the origin of F. A vector of $\mathbb{R}^4$ is called *time-like* if belongs to C. The boundary of C, ∂C, is called *light cone*, and it is the set of events connected with the origin by a unique light ray (see figure 4.1).

Proposition 4.5. *Let F and F' be inertial frames of reference of $\mathcal{M}$ and v the speed of F' relative to F. If f is the Lorentz transformation between F and F', then $f(C^+) = C^+$.*

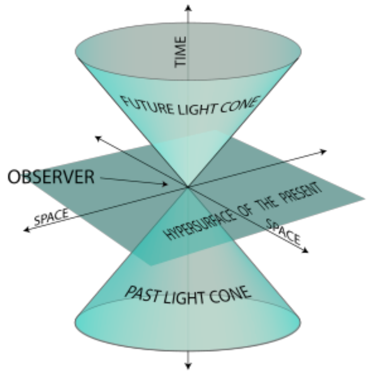

Figure 4.1: Light cone of a two dimensional space

Proof. From the proposition 4.1, we deduce that $f(C) = C$. As f is linear, f is continuous and it will transform the connected component C^+ to a connected component of C. Therefore, it is sufficient to prove that $f(w) \in C^+$ for any time-like vector $w \in C^+$. Let $w = (1,0,0,0)$ and $f(w)$ its Lorentz transformed vector, given by

$$f(w) = \begin{cases} \gamma(1 - \frac{v\cdot(0,0,0)}{c}) \\ \\ (0,0,0) + (\gamma - 1)\frac{(0,0,0)\cdot v}{v\cdot v}v - \gamma v\frac{1}{c}). \end{cases}$$

Therefore, $f(w) = \gamma(1, \frac{1}{c}v)$, which belongs to C^+. $\square$

Characterisation of the space-time interval

Given an inertial frame of reference in the Minkowski space-time, let O and A be two different events (O is placed in the origin of the frame of reference), we can characterise the space-time interval $\mathcal{G}(\vec{OA},\vec{OA})$ between O and A by the following:

- If $\mathcal{G}(\vec{OA},\vec{OA}) < 0$, the interval is called *time-like*.
  -It is negative in all inertial frames of reference.
  -It does not exist an inertial frame of reference in where O and A are simultaneous. Therefore, it is not possible to have an inertial frame of reference such that A happens before O.
  -It is possible to find a frame of reference in which O and A happen in the same place.

- If $\mathcal{G}(\vec{OA},\vec{OA}) > 0$, the interval is called *space-like*.
  -It is positive in all inertial frames of reference.
  -There exists an inertial frame of reference in which O and A are simultaneous.
  -There does not exist a frame of reference in which O and A happen in the same place.

- If $\mathcal{G}(\vec{OA}, \vec{OA}) = 0$, the interval is called *light-like*.
  -O and A are connected by a unique light ray. They both belong to C.
  -O and A can not be simultaneous events in any inertial frame of reference.
  -There does not exist a frame of reference in where O and A happen in the same place (we can not have a frame of reference travelling at c).

4.5 Proper time

Given an inertial frame of reference F, let us consider a straight line l in $\mathbb{R}^4$ with time-like direction vector v. Such a line can be parametrised as

$$\begin{cases} r = vt + b \\ \\ t = t \end{cases}$$

with $||v|| < c$. For an observer of F, the line l will represent the movement of a particle with uniform rectilinear motion and constant speed v. As F is inertial, we can have another inertial frame of reference F' moving at a constant speed v from F such that the particle is static and in the origin of F'. Let P_A and P_B be different points of l such that P_A is previous to P_B. For an observer of F', these events will be represented by the coordinates $(ct_A, 0, 0, 0)$ and $(ct_B, 0, 0, 0)$, with $t_A < t_B$. The distance between P_A and P_B, calculated in F', is given by

$$d(P_A, P_B) = \sqrt{\mathcal{G}(P_B - P_A, P_B - P_A)} = ic(t_B - t_A).$$

Therefore, the time interval $(t_B - t_A)$ measured in the frame of reference in which the particle is motionless is $d(P_A, P_B)/ic$. However, $(t_B - t_A)$ can be calculated from any inertial frame of reference because $d(P_A, P_B)$ is invariant when applying the Lorentz transformation.

Consider that the particle reaches the event $P_1 = (ct_1, p)$, that in F_1, it is placed on l, and it changes drastically its speed to a constant velocity $w \neq v$. Let $P_A = (ct_A, A)$ and $P_B = (ct_B, B)$ be events represented in F such that $t_A < t_1 < t_B$. An observer of F can represent the motion of such a particle in $\mathbb{R}^4$ by two segments ending at P_1 forming an angle that have time-like director vectors. Suppose that an observer is travelling with the particle. From A to p he will be in an inertial frame of reference and the time interval will be $d(P_A, P_1)/ic$. At p, when the speed changes, its frame of reference will no longer be inertial. An instant later, his frame of reference will be inertial again from p to B, and the time measured from him between these two events will be $d(P_1, P_B)/ic$. We postulate that the time measured by such an observer (that is not inertial) does not suffer any jump in time at p. Therefore, the time measured by his clock between A and B is

$$\tfrac{1}{ic}(d(P_A, P_1) + d(P_1, P_B)).$$

Suppose now that the particle is moving arbitrarily with respect to F, and without any drastic change of speed. Let us consider its trajectory in F given by a smooth curve γ of $\mathbb{R}^4$

and instant speed v such that $||v|| < c$, i.e, with a positive time-like tangent vector on each point of γ. Let P_A and P_B be two different points on such a curve. We want to determine the time interval measured by an observer moving with the particle, known as *proper time* of the particle. As this observer is not necessarily in an inertial frame of reference, an additional postulate is needed. Let us consider $n-1$ points of γ, $P_1, P_2, ..., P_{n-1}$ situated in numerical order between P_A and P_B. Let γ' be the line formed by the segments $P_A P_1$, $P_1 P_2$,..., $P_{n-1} P_B$. Such a line describes the motion of a particle that changes of inertial frame of reference at the points $P_1, P_2, ..., P_{n-1}$. Using the postulate defined above, the time measured by a clock moving with trajectory γ' is

$$\tfrac{1}{ic}(d(P_A, P_1) + d(P_1, P_2) + ... + d(P_{n-1}, P_B)).$$

It is evident that the greater the n is, the closer γ' is to γ. We postulate: given two observers, the more approximate their trajectories in the Minkowski space-time are, the more approximate their measured time interval becomes. Using these additional postulates, we want to define the proper time of the particle on the trajectory γ between the events P_A and P_B tending n to infinity. The last expression will become

$$\tfrac{1}{ic}(\textit{length of } \gamma \textit{ between } P_A \textit{ and } P_B).$$

Definition 4.6. *The time between two events, P_A and P_B, measured by an observer whose motion is represented in the Minkowski space-time by the piecewise smooth curve γ, with positive time-like tangent vector on each point (except the points in where γ is not differentiable), is the length of γ between P_A and P_B divided by ic.*

4.6 4-velocity, 4-momentum and 4-force

In classical mechanics, given any inertial frame of reference, it is assumed that the total momentum of a system of particles, defined as $\vec{p} = \sum_i m \frac{d\vec{r_i}}{dt}$ ($\vec{r_i} \in \mathbb{R}^3$ is the position vector of the particle i), is conserved (momentum conservation law). Redefining the definition of momentum in special relativity by changing the term $\frac{d\vec{r_i}}{dt}$ by $\frac{d(ct,\vec{r_i})}{dt}$, one deduces that this conservation law is not satisfied for some inertial frame of references (see [2] section 15.1). This contradicts the first postulate of special relativity. Therefore, we need to find a good definition of momentum that satisfies a conservation law in the Minkowski space-time and for all inertial frames of references.

Definition 4.7. *Let $\gamma(\tau)$ be the trajectory of a particle in the Minkowski space-time parametrised by its proper time τ (also named particle's life). We define the* ***4-velocity*** *of the particle at the instant τ_0 as $\frac{d\gamma(\tau)}{d\tau}|_{\tau=\tau_0}$. The* ***4-momentum*** *of the particle at τ_0 is given by $m\frac{d\gamma(\tau)}{d\tau}|_{\tau=\tau_0}$, where m is the mass of the particle.*

Let us analyse these definitions in an inertial frame of reference F. The particle's life will be parametrised as

$$\begin{cases} \vec{r} = (\gamma^1(t), \gamma^2(t), \gamma^3(t)) \\ \\ ct = ct \end{cases},$$

where $(\gamma^1(t), \gamma^2(t), \gamma^3(t))$ are the coordinates of the particle in F, and t is the time measured from F. We can express t in function of τ, therefore, the last expression can be written as

$$\begin{cases} \vec{r} = (\gamma^1(t(\tau)), \gamma^2(t(\tau)), \gamma^3(t(\tau))) \\ \\ ct = ct(\tau) \end{cases}.$$

The 4-velocity $\vec{v}$ of the particle will be

$$\vec{v} = \frac{d\gamma}{d\tau} = \frac{d\gamma}{dt}\frac{dt}{d\tau} = (c, \frac{d\gamma^1}{dt}, \frac{d\gamma^2}{dt}, \frac{d\gamma^3}{dt})\frac{dt}{d\tau} = (c, v^1, v^2, v^3)\frac{dt}{d\tau},$$

where (v^1, v^2, v^3) is the usual velocity in F. Using the definition of proper time,

$$\tau(t) = \frac{1}{ic}\int_0^t \sqrt{\mathcal{G}(\frac{d\gamma}{d\xi}, \frac{d\gamma}{d\xi})}d\xi,$$

therefore,

$$\begin{aligned} \frac{d\tau}{dt} = \frac{1}{ic}\sqrt{\mathcal{G}(\frac{d\gamma}{dt}, \frac{d\gamma}{dt})} &= \frac{1}{ic}\sqrt{\mathcal{G}((c, v^1, v^2, v^3), (c, v^1, v^2, v^3))} \\ &= \frac{\sqrt{c^2 - v^2}}{c} = \sqrt{1 - \frac{v^2}{c^2}}, \end{aligned}$$

where $v = \sqrt{(v^1)^2 + (v^2)^2 + (v^3)^2}$. Hence, if $\vec{v}$ is the 4-velocity and $\vec{\mathcal{P}}$ the 4-momentum, in F they will be expressed as

$$\vec{v} = (c, v^1, v^2, v^3)\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \text{ and } \vec{\mathcal{P}} = (c, v^1, v^2, v^3)\frac{m}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

The **4-force** or **Minkowski force** is defined as $\vec{\mathcal{F}} = \frac{d\vec{\mathcal{P}}}{d\tau}$ (analogously to $F = \frac{d\vec{P}}{dt}$ of classical mechanics).

Chapter 5

Relativistic continuum mechanics; the stress energy tensor field

Continuum mechanics is a branch of mechanics such that materials are analysed and modelled as if they could be divided infinite times. In other words, a model of an object is made assuming that the substance of the object completely fills the space it occupies, i.e, the object is a continuum. It is well known that this is not true on a very small scale, thus, this theory will only be applicable on a large scale. In this section we will only emphasise on the aspects needed for our purpose, giving a brief, intuitive explanation of the other aspects.

5.1 Mass density; proper density

Consider a continuum medium in motion (for example a fluid) in a region $U \in \mathbb{R}^3$ with respect to an inertial frame of reference F. The density function $\rho : I \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ of the medium is related to the mass of a measurable region $V \subset U$ at the instant $t \in I$ by

$$m(V,t) = \int_V \rho(x,t)\, dV.$$

We consider that the density function is differentiable. Let x_0 be a point of U and t_0 be an instant in time. Let V_δ be the ball of center x_0 and radium δ. The density function can be expressed as

$$\rho(x_0, t_0) = \lim_{\delta \to 0} \frac{m(V_\delta, t_0)}{vol(V_\delta)}.$$

Let P be a particle of the medium. Let τ be its proper time and τ_0 be any instant of it. Let F' be an inertial frame of reference such that at $t' = 0$ its origin coincides with the particle at τ_0 and the speed of P is null at that instant. As the particle is not necessarily moving at a constant speed, P can have a non-null speed in F' at any other time $\tau_1 \neq \tau_0$. An inertial frame of reference with these characteristics is called *proper frame of P at τ_0*. Measuring the mass density of the medium from F' at the origin of F' and at $t' = 0$, we get a number σ called *proper density of the medium of P at τ_0*.

5.2 Body force density and stress tensor field

Consider, as before, a continuum medium in motion in a region $U \in \mathbb{R}^3$. Let V be a region of U. Following the classical dynamics of Newton and Euler, the motion of V is produced by the action of applied forces which are assumed to be of two kinds: *surface forces* $\vec{\mathbf{F}}_C$ and *body forces* $\vec{\mathbf{F}}_B$. The surface forces are those which the rest of the medium exerts on V, and body forces are those originating outside of the medium (for example, if V is affected by a gravitational field). For further details, see [3] section 4.1. Thus, the total force $\mathcal{F}$ applied on V at the instant t can be expressed as:

$$\mathcal{F}(V,t) = \vec{\mathbf{F}}_B(V,t) + \vec{\mathbf{F}}_C(V,t).$$

Definition 5.1. *The body force density $\vec{f}(x,t)$, which depends differentially of x and t, is a vector field defined as*

$$\vec{f}(x,t) = \lim_{\delta \to 0} \frac{\mathbf{F}_B(V_\delta,t)}{vol(V_\delta)},$$

where V_δ is the ball of center x and radium δ.

$\vec{f}(x,t)$ is, intuitively, the body force exerted on the point x at the instant t. Its magnitude is force per unit of volume.

We want to define the body force density in relativity. Let F be an inertial frame of reference. Given an event $(ct_0, x_0) \in \mathbb{R}^4$, consider a particle of the medium in motion and placed on x_0 at the instant t_0. Let F' be an inertial frame of reference such that the particle coincides with the origin of F' at $t' = 0$, and its instant speed at that time is null in F' (proper frame of the particle at that time). Let H be the hyperplane that in F' is orthogonal to the ct' axis (by the Minkowski metric) and contains (ct_0, x_0) (intuitively, H is a picture of the space at t_0, taken by an observer of F'). Let $\vec{f} = (f_1, f_2, f_3)$ be the body force density at the origin of F' at $t' = 0$. The vector $(0, f_1, f_2, f_3)$ will be contained in H. From now on we will denote such a vector of $\mathbb{R}^4$ as $\vec{f}'$.

Definition 5.2. *With the considerations above, for each (ct, x) we have defined a vector $\vec{f}'(t,x) \in \mathbb{R}^4$ called **Minkowski body force density**. The function $\vec{f}'$ is supposed to be differentiable on its components.*

Cauchy's principle *Let S^2 be the set of unit vectors of $\mathbb{R}^3$. There exists a differentiable mapping*

$$\begin{array}{ccc} U \times S^2 \times I & \longrightarrow & \mathbb{R}^3 \\ (x, \vec{u}, t) & \longmapsto & \vec{T}(x, \vec{u}, t) \end{array}$$

such that for all region $V \subset U$ with a smooth boundary, and also for all tetrahedron $V \subset U$ (even if its boundary is not smooth), the surface force on V at the instant t is given by

$$\vec{\mathbf{F}}_C(V,t) = \int_{\partial V} \vec{T}(x, \vec{n}(x), t) ds. \tag{5.1}$$

where, for all $x \in \partial V$, $\vec{n}(x)$ is the normal unit vector to ∂V, and ds is the surface element of ∂V. Let us assume that the integral above makes sense when V is a tetrahedron (even if $\vec{n}(x)$ is not defined on the edges). It is also supposed that $\vec{T}(x,\vec{u},t) = -\vec{T}(x,-\vec{u},t)$.

Let us see the significance of $\vec{T}$. Given $x \in U$ and $\vec{u} \in S^2$, let s be a small open neighbourhood of x on a plane perpendicular to $\vec{u}$ containing x. Intuitively, $\vec{T}(x,\vec{u},t)$ is the force that the part of the medium at one side of s exerts to the other side with direction $\vec{u}$, therefore, it is a force per unit area. Using Newton's third law (action reaction), one can easily intuit why the condition $\vec{T}(x,\vec{u},t) = -\vec{T}(x,-\vec{u},t)$ is needed.

The mapping $\vec{T}(x,\vec{u},t)$, given by the Cauchy's principle, is defined for vectors belonging to S^2. Let us extend such a differentiable map for vectors of $\mathbb{R}^3$ by

$$T(x,\vec{u},t) = \begin{cases} 0 & \text{if } \vec{u} = 0 \\ \\ ||\vec{u}||\vec{T}(x, \frac{\vec{u}}{||u||}, t) & \text{if } \vec{u} \neq 0 \end{cases}.$$

Cauchy proved that $T(x,\vec{u},t)$ depends linearly on $\vec{u}$ (Cauchy's Theorem) (see [2], theorem 12.4). We proved in chapter one that given a vector space E, the space of linear mappings from E to E can be identified to $E^* \otimes E$. Therefore, the linear mapping $\vec{u} \mapsto T(x,\vec{u},t)$ defines a tensor of type $(1,1)$.

Definition 5.3. *For each $t \in I$, $T(x,\vec{u},t)$ determines a $(1,1)$ tensor field $\mathcal{T}$ on U named **Cauchy stress tensor**.*

We want to find an analogous concept in the Minkowski space-time. Let F be an inertial frame of reference in where the events are represented by $(ct,x) \in \mathbb{R}^4$. Consider a particle of the medium placed in x_0 at the instant t_0. Let F' be an inertial frame of reference such that its origin coincides with the particle at $t' = 0$ and the speed of the particle with respect to F at that instant is null (proper frame of the particle at t_0). Let $\mathcal{T}$ be the Cauchy stress tensor at the instant $t' = 0$ according to the observers of F'. Let $\mathcal{T}_{(t_0,x_0)}$ be the tensor field applied on the origin of F' at $t' = 0$. By virtue of the Cauchy's theorem, $\mathcal{T}_{(t_0,x_0)}$ is a linear transformation. Consider the hyperplane H containing (t_0,x_0) and orthogonal to the ct' axis by the Minkowski metric. H represents the space of F' at $t' = 0$ (informally, H is a picture of the space at the instant t' taken by an observer of F') . With this considerations, $\mathcal{T}_{(t_0,x_0)}$ can be understood as a linear transformation of H. We want to extend such a transformation to $\mathbb{R}^4$, which we will also design by $\mathcal{T}_{(t_0,x_0)}$. Let us impose that if $v \in \mathbb{R}^4$ is an orthogonal vector of H, $\mathcal{T}_{(t_0,x_0)}(v) = 0$. Thus, we have defined a tensor $\mathcal{T}_{(t,x)}$ defined on each $(t,x) \in \mathbb{R}^4$. Note that this tensor field only depends on the medium, because the frame of reference F' used to define $\mathcal{T}_{(t,x)}$ is uniquely determined by the medium at each $(t,x) \in \mathbb{R}^4$.

Definition 5.4. *Using the notation above, and considering that the tensor field $\mathcal{T}$ is differentiable, we name $\mathcal{T}$ as the **stress tensor field**.*

Note that $\mathcal{T}$ can be also taken as a $(0,2)$ tensor field using the musical isomorphism.

5.3 Stress-energy tensor field

In this section we will define the stress-energy tensor field, which will be crucial to locally determine the metric of space-time in presence of masses.

In classical continuum mechanics, the equations of motion of a continuum medium, also known as Cauchy momentum equations, are given by

$$\frac{\partial(\rho\vec{v})}{\partial t} + div(\rho\vec{v}\otimes\vec{v} - \mathcal{T}) = \vec{f},$$

where ρ is the density, $\vec{f}$ the body force density, $\vec{v}$ is the speed vector field of the medium and $\mathcal{T}$ is the stress tensor field that in this expression is of type $(0,2)$. Considering that the mass of the medium is constant in time, we can deduce the following equation, also known as *continuity equation* (see [2] section 12.2)

$$\frac{\partial\rho}{\partial t} + div(\rho\vec{v}) = 0.$$

These last expressions depend on the chosen frame of reference. We want to find an expression in the Minkowski space-time that does not depend on the frame of reference. It will also be necessary the compatibility with classical mechanics, i.e, if the speed of the medium is a lot smaller than c, we should get the Cauchy momentum equations.

Consider the following equation in the Minkowski space-time:

$$div(\sigma\vec{u}\otimes\vec{u} - \mathcal{T}) = \vec{f}', \tag{5.2}$$

where σ is the proper density, $\mathcal{T}$ is the $(0,2)$ stress tensor field, $\vec{f}'$ is the Minkowski body force density and $\vec{u}$ is the speed vector field of the medium. In order to show the compatibility of the equation (5.2) with the classical theory, assume that all of the particles of the medium move at a small speed. Therefore, on each point of the fluid, and at each instant, we have a proper frame F' moving from F with a constant speed $v \ll c$. We can approximate this by assuming that the frames of reference F' coincide with F. The proper density σ will be approximately ρ measured in F. Using components in F, the equation (5.2) will be written as

$$\frac{\partial}{\partial x^\alpha}(\rho v^i v^\alpha - \mathcal{T}^{i\alpha}) + \frac{\partial}{\partial ct}(\rho v^i v^0 - \mathcal{T}^{i0}) = f^i,$$

where $\alpha \in \{1,2,3\}$ and $i \in \{0,1,2,3\}$. Considering that all proper frames coincide with F, $\mathcal{T}^{j0} = 0$ (see section 5.2). Furthermore, as in this case the speed of the medium is small compared with c, the Lorentz factor γ is almost 1 Therefore, $v^0 = c$ and the 0 component will be

$$c\left(\frac{\partial\rho}{\partial t} + div(\rho u)\right) = 0,$$

where u is the three dimensional speed vector field defined by the three space components of $\vec{u}$. Note that dividing the equation by c, one gets the continuity equation. The other three components will give us the Cauchy momentum equations.

Definition 5.5. *We define the **stress-energy tensor field** as*

$$T = \frac{1}{c^2}(\sigma\vec{u} \otimes \vec{u} - \mathcal{T}). \tag{5.3}$$

Using this definition, the equation (5.2) is written as

$$c^2 divT = \vec{f}'.$$

Note that we can use the Minkowski metric to convert T to a $(1,1)$ tensor field, which can be understood as a linear mapping (see section 1.4). Let p be a point in the Minkowski space-time and $u = \vec{u}(p)$. We want to prove that, considering T as a linear mapping, u is an eigenvector with eigenvalue $-\sigma$. Suppose that F is an inertial frame of reference in which the vector u is the director vector of the time axis, i.e, the coordinates of u are $(c, 0, 0, 0)$ in F. If the index 0 is the time component, and the space components are $1, 2$ and 3, the components of $T(u)$ in F are given by

$$T(u)^i = T^i_k u^k = cT^i_0 = \tfrac{\sigma}{c}(u^i u_0 - \mathcal{T}^i_0).$$

The components $\mathcal{T}^i_0$ are all 0 in F (see section 5.2). Therefore,

$$T(u)^i = \tfrac{\sigma}{c}(u^i u_0) = \tfrac{\sigma}{c}(u^i \mathcal{G}_{00} u_0) = -\sigma u^i.$$

One can see that T does not have any other time-like eigenvector except of u multiplied by any scalar. The vector field $\sigma\vec{u}$ is called *4-momentum density vector field*. Suppose that we know T. It is possible to extract such a vector field by calculating the only time-like eigenvector of norm *ic* and multiplying it by the opposite of its eigenvalue. Note that we can also extract the information of the surface forces by knowing the stress-energy tensor field.

Chapter 6

The space-time; An introduction to general relativity

In the case of special relativity, the effect of gravity doesn't exist, being incomplete and not compatible with Newton's theory of gravity, which predicts that a mass placed in space creates a gravitational field, attracting the rest of the masses. This incompatibility is what pushed Einstein to create the theory of general relativity, which is compatible with Newton's theory of gravitation. Before formalising the basics of general relativity, we will see an example that will help us to understand and develop such a theory.

6.1 The spinning disc

Let F be an inertial frame of reference and suppose that there is a disc of radium R contained in the plane $z = 0$ such that its center coincides with the origin of F. Suppose that at the instant $t = 0$ the disc starts spinning with a constant angular speed ω such that $\omega R < c$. Using a polar coordinate system, consider a point of the disc placed in $(r_0, \theta_0, 0)$ at $t = 0$. We can parametrise the trajectory of P by the curve

$$\begin{cases} x(t) = r_0 \cos(\omega t + \theta_0) \\ y(t) = r_0 \sin(\omega t + \theta_0) \\ z(t) = 0 \\ t = t \end{cases}, \tag{6.1}$$

where (x, y, z) are the Cartesian space coordinates of P with respect to F. As in this example the coordinate z will always be zero, we shall remove it from now on and consider this example in a two dimensional space. Therefore, we can represent the motion of P in the Minkowski space-time by the curve $\gamma(t) = (ct, x(t), y(t))$. The norm of $\dot{\gamma}(t)$ by the Minkowski metric will be $||\dot{\gamma}(t)|| = \sqrt{\omega^2 r^2 - c^2}$. As $||\dot{\gamma}(t)||$ does not depend on t, the proper time of P is given by

$$\tau = \frac{1}{ic}(\sqrt{\omega^2 r^2 - c^2})t = \sqrt{1 - \frac{r^2\omega^2}{c^2}}t.$$

Considering this expression and supposing that there is a clock stuck on P, one deduces that the further that P is from the center of the disc, the slower the time measured by the clock will go. Parametrising the trajectory of P by its proper time, we have

$$\gamma(\tau) = \begin{cases} x(\tau) = r_0 \cos(\frac{\omega}{\alpha_0}\tau + \theta_0) \\ \\ y(\tau) = r_0 \sin(\frac{\omega}{\alpha_0}\tau + \theta_0) \\ \\ ct(\tau) = \frac{c\tau}{\alpha_0} \end{cases},$$

where $\alpha_0 = \sqrt{1 - \frac{r_0^2\omega^2}{c^2}}$. The 4-acceleration will be

$$\frac{d^2\gamma(\tau)}{d\tau^2} = -\frac{\omega^2 r_0}{\alpha_0^2}(0, \cos(\frac{\omega}{\alpha_0}\tau + \theta_0), \sin(\frac{\omega}{\alpha_0}\tau + \theta_0)).$$

This means that a particle of mass m stuck on P will experiment a 4-force directed to the center of the disc and of norm $m\frac{\omega^2 r_0}{\alpha_0^2}$. (centripetal force)

Consider another frame of reference F' spinning with respect to F also with an angular speed ω, and such that its origin coincides with the origin of F (and with the center of the disc). Note that F' is not inertial and the disc will be static in such a frame of reference. For the observers of F' (intuitively, the inhabitants of the disc), the trajectory of P in space will only be a single constant point given by (r_0, θ_0). Let us express the Minkowski metric using the coordinates (ct, r, θ). Using the equations (6.1), we have

$$\begin{cases} dx = \cos(\omega t + \theta)dr - \frac{r\omega\sin(\omega t+\theta)}{c}cdt - r\sin(\omega t + \theta)d\theta \\ dy = \sin(\omega t + \theta)dr + \frac{r\omega\cos(\omega t+\theta)}{c}cdt + r\cos(\omega t + \theta)d\theta \\ cdt = cdt. \end{cases}$$

Therefore,

$$\mathcal{G} = -c^2dt^2 + dx^2 + dy^2 = dr^2 + r^2d\theta^2 + \frac{2r^2\omega}{c}cdtd\theta - (1 - \frac{r^2\omega^2}{c^2})c^2dt^2. \tag{6.2}$$

A point P of the disc will be denoted in F' as (r_0, θ_0). It will have a trajectory in the Minkowski space-time given by $\gamma(t) = (ct, r_0, \theta_0)$, with tangent vector $\dot{\gamma}(t) = (c, 0, 0)$. Using (6.3), we can calculate the norm of $(c, 0, 0)$, which will be $\sqrt{r_0^2\omega^2 - c^2}$. Therefore, the proper time of P gives

$$\tau = \frac{1}{ic}(\sqrt{\omega^2 r_0^2 - c^2})t = \sqrt{1 - \frac{r_0^2\omega^2}{c^2}}t.$$

γ parametrised with τ will be $\gamma(\tau) = (\frac{c\tau}{\alpha}, r_0, \theta_0)$, where $\alpha = \sqrt{1 - \frac{r_0^2\omega^2}{c^2}}$. This means that $\dot{\gamma}^1 = \dot{\gamma}^2 = 0$ and $\dot{\gamma}^0 = c/\alpha$. Therefore $\ddot{\gamma} = 0$. When we choose a non inertial frame of reference, fictitious forces occur. For example, the Coriolis force is a fictitious force that acts on objects that are in motion relative to a rotating frame of reference. It would be false to say that there is no force acting on P using that $\ddot{\gamma} = 0$. If P is static in F', there must be a real force keeping it motionless. Such a force is the one we found analysing this problem in the frame F, which is inertial. We want to find an expression that allows us to extract the 4-acceleration found from F in any frame of reference (even if it is not inertial). Let us consider the acceleration given by $\nabla_{\dot{\gamma}}\dot{\gamma}$, where ∇ is the Levi-Civita connection. The coordinates of $\nabla_{\dot{\gamma}}\dot{\gamma}$ will be

$$(\nabla_{\dot{\gamma}}\dot{\gamma})^i = \ddot{\gamma}^i + \Gamma^i_{jk}\dot{\gamma}^j\dot{\gamma}^k = \tfrac{c^2}{\alpha^2}\Gamma^i_{00}$$

Let us calculate the Christoffel symbols using the Levi-Civitta equations given by

$$\mathcal{G}_{il}\Gamma^i_{jk} = \tfrac{1}{2}\left(\tfrac{\partial \mathcal{G}_{jl}}{\partial x^k} + \tfrac{\partial \mathcal{G}_{kl}}{\partial x^j} - \tfrac{\partial \mathcal{G}_{jk}}{\partial x^l}\right).$$

The metric $\mathcal{G}$ in the point (ct, r_0, θ_0) is written as

$$\mathcal{G} = \begin{pmatrix} \frac{r_0^2\omega^2}{c^2} - 1 & 0 & \frac{r_0^2\omega}{c} \\ 0 & 1 & 0 \\ \frac{r_0^2\omega}{c} & 0 & r_0^2 \end{pmatrix}.$$

Therefore,

$$\mathcal{G}_{il}\Gamma^i_{00} = \tfrac{\partial \mathcal{G}_{0l}}{\partial x^0} - \tfrac{1}{2}\tfrac{\partial \mathcal{G}_{00}}{\partial x^l}.$$

However, the only non null derivatives of $\mathcal{G}$ are the derivatives with respect the coordinate $x^1 = r$. Hence, the only non null $\mathcal{G}_{il}\Gamma^i_{00}$ will occur when $l = 1$. Therefore,

$$\mathcal{G}_{i1}\Gamma^i_{00} = -\tfrac{1}{2}\tfrac{\partial \mathcal{G}_{00}}{\partial r} = -\tfrac{r_0\omega^2}{c^2}.$$

Therefore, the Christoffel symbols Γ^i_{00} are given by

$$\Gamma^i_{00} = \mathcal{G}^{i1}\mathcal{G}_{i1}\Gamma^i_{00},$$

where $\mathcal{G}^{ij}$ are the coefficients of the inverse matrix of $\mathcal{G}$,

$$\mathcal{G}^{-1} = \begin{pmatrix} -1 & 0 & \frac{\omega}{c} \\ 0 & 1 & 0 \\ \frac{\omega}{c} & 0 & \frac{1}{r_0^2} - \frac{\omega^2}{c^2} \end{pmatrix}.$$

The only non null Γ^i_{00} is when $i = 1$ and we have

$$\nabla_{\dot{\gamma}}\dot{\gamma} = \tfrac{r_0\omega^2}{\alpha^2}(0, -1, 0).$$

Given this acceleration, we can calculate the 4-force by $F = m\nabla_{\dot{\gamma}}\dot{\gamma}$, which will give us a force directed to the center of the disc and with norm $\frac{mr_0\omega^2}{\alpha^2}$. Note that this is the same result that we got using an inertial frame of reference and the classical derivative. Let us see what happens if we leave a freely moving particle P of mass m at the point (r_0, θ_0) of the disc. If $\gamma(t) = (ct, r(t), \theta(t))$ is the trajectory describing the particle's life such that $\gamma(0) = (0, r_0, \theta_0)$, assuming that $\nabla_{\dot{\gamma}}\dot{\gamma}$ is the real acceleration, the condition that no forces will act on P will be written as $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$, i.e, γ is a geodesic curve. The equation of a geodesic curve is given by

$$\ddot{\gamma}^i + \Gamma^i_{jk}\dot{\gamma}^j\dot{\gamma}^k = 0.$$

As the particle is motionless at $t = 0$, $\dot{\gamma}(0) = (c, 0, 0)$. Therefore, the geodesic equation at $t = 0$ is written as

$$\ddot{\gamma}^i + \Gamma^i_{00}c^2 = 0.$$

If an inhabitant of the disc (an observer of F') refers the fictitious acceleration $\ddot{\gamma}^i$ to its proper time $\tau = \alpha t$ (not to the time t of F), we have

$$\frac{d^2\gamma^i}{d\tau^2} = \frac{1}{\alpha^2}\ddot{\gamma}^i = -\frac{c^2}{\alpha^2}\Gamma^i_{00}.$$

The inhabitant of the disc may think that there is a force acting on P with the same norm of the centripetal force and opposite direction (centrifugal force). This will only happen because he refers the movement using the coordinates of F', which is not inertial. An observer of F will see that there is no force acting on the particle, and its movement will be caused by the motion of the rotating disc (not because of the effect of a force). Imagine that we close an observer of F' in a small box. If such an observer leaves a particle Q of mass m freely in motion inside the box, he might think that there is a gravitational field attracting Q. Indeed, if he leaves a second particle of mass $M \neq m$ freely in motion, its acceleration will be the same as it was for the first particle. This is a characteristics of gravitational fields. Therefore, the observer inside a small box can not know if the motion of the particle is created by a "fictitious" force due to the motion of the box with respect to the inertial observers, or by a "real" force created by a gravitational field.

Let us see what happens with the space geometry of F' (not the space-time geometry). When an observer of F' measures the longitude of the circumference $r = r_0$, the result will not be $2\pi r_0$. Before formalise rigorously this result, we shall give first an intuitive explanation. Let us divide the circumference $r = r_0$ to n arcs denoted as $\alpha'_1, \alpha'_2, ..., \alpha'_n$ such that n is sufficiently large to confuse each α'_i with a straight segment. Hence, we will consider that each α'_i is straight. If we do the same in F, we will have n straight segments $\alpha_1, \alpha_2, ..., \alpha_n$. The measure of each α_i will be the measure of α'_i multiplied by $\sqrt{1 - \omega^2 r_0^2/c^2}$. Hence, if the perimeter of $r = r_0$ in F is $2\pi r_0$, in F' will be $\frac{2\pi r_0}{\sqrt{1-\omega^2 r_0^2/c^2}}$. If an observer of F' measures the radium of the circumference, as it is perpendicular to the movement, there will not be length contraction (see section 4.3). Therefore, the radium will remain the same between F and F'. One concludes that the relation between the circumferences centred in the origin and their diameter is π in F and greater than π in F'. Hence, the geometry of the space in F' is no longer Euclidean.

Let us formalise with more rigour the results above. Consider a non inertial observer placed in $P_0 = (r_0, \theta_0)$. Given a fixed time $t = t' = 0$, where t is measured in F and t' in F', consider an inertial frame of reference F'' such that its origin coincides with P_0 at $t'' = 0$ and the speed of the observer is null at that instant, i.e, the proper frame of P_0 at $t' = 0$. Indeed, for each point $P = (r, \theta)$, motionless from F', we have a proper frame of P at $t' = 0$. Let us design such an inertial frame of reference as F''_P. Every F''_P moves with respect to the others, hence, the distances will vary depending on which inertial frame of reference we are. Using the first postulate of the special relativity, we can only measure distances in an inertial frame of reference. In order to give a definition of "measuring distances" in F' that is compatible with every measurement made in an inertial frame of reference, the only alternative that we have is thinking that the distances in F' are given by a metric such that the metric of the tangent space of a point P is the metric of F''_P. Note that these considerations above are refereed to the space of F' (not the space-time) because the time is fixed at $t' = 0$. Every point of the disc $P_0 = (r_0, \theta_0)$ origins a curve γ of the Minkowski space-time with tangent vector $\dot{\gamma}(0) = (c, 0, 0)$. The frame of reference F''_{P_0} will have $\dot{\gamma}(0)$ as director vector of the time axis, and the perpendicular plane at $t'' = 0$ will be the space axes. Let us see what the condition of perpendicularity with $\dot{\gamma}(0)$ means. Given a vector (cT, R, Θ) which origin is on $(0, r_0, \theta_0)$, the perpendicularity with $(c, 0, 0)$ is written as

$$\begin{pmatrix} c & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{r_0^2\omega^2}{c^2} - 1 & 0 & \frac{r_0^2\omega}{c} \\ 0 & 1 & 0 \\ \frac{r_0^2\omega}{c} & 0 & r_0^2 \end{pmatrix} \begin{pmatrix} cT \\ R \\ \Theta \end{pmatrix} = 0.$$

Therefore, we can express T as a function of Theta by

$$T = \frac{r_0^2\omega}{c^2 - r_0^2\omega^2}\Theta.$$

The vectors with origin $\gamma(0)$ which are perpendicular to $(c, 0, 0)$ are $(c\frac{r_0^2\omega}{c^2 - r_0^2\omega^2}\Theta, R, \Theta)$. The scalar product of two vectors that are perpendicular to $\dot{\gamma}(0)$ is given by

$$\begin{pmatrix} c\frac{r_0^2\omega}{c^2 - r_0^2\omega^2}\Theta & R & \Theta \end{pmatrix} \begin{pmatrix} \frac{r_0^2\omega^2}{c^2} - 1 & 0 & \frac{r_0^2\omega}{c} \\ 0 & 1 & 0 \\ \frac{r_0^2\omega}{c} & 0 & r_0^2 \end{pmatrix} \begin{pmatrix} c\frac{r_0^2\omega}{c^2 - r_0^2\omega^2}\Theta' \\ R' \\ \Theta' \end{pmatrix} =$$

$$\begin{pmatrix} 0 & R & \frac{(r_0^2\omega)^2}{c^2 - r_0^2\omega^2}\Theta + r_0^2\Theta \end{pmatrix} \begin{pmatrix} c\frac{r_0^2\omega}{c^2 - r_0^2\omega^2}\Theta' \\ R' \\ \Theta' \end{pmatrix} =$$

$$RR' + (\frac{(r_0^2\omega)^2}{c^2 - r_0^2\omega^2} + r_0^2)\Theta\Theta' = RR' + \frac{r_0^2}{1 - \frac{r_0^2\omega^2}{c^2}}\Theta\Theta'.$$

The last expression is the metric of the space (not the space-time) of F''_{P_0}. Therefore, the metric that induces each one of the metrics of F''_P is the non euclidean metric for F' given by

$$dr^2 + \frac{r^2}{1 - \frac{r^2\omega^2}{c^2}} d\theta^2.$$

6.2 Einstein's equivalence principle and the theory of gravitation

In the last example, we deduced that what causes the movement of free particles in a small box on the spinning disc can be confused between a gravitational field and a force produced by its motion. This concept is known as *Einstein's equivalence principle*, which states that in any sufficiently small region of space, the effects of gravitation are the same as those from acceleration. However, we saw that given the free particle's life $\gamma(\tau)$, whatever the reason of its movement is, its trajectory will satisfy $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$, where ∇ is the Levi-Civitta connection with respect to the metric tensor field. We also saw that the metric (6.2) depends on the angular velocity of the disc. Indeed, if the disc is not spinning, the metric is written as

$$\mathcal{G} = -c^2 dt + dr^2 + r^2 d\theta^2,$$

which is the usual Minkowski metric in polar coordinates. Hence, one can deduce that the rotating movement of the disc deforms the Minkowski metric. Moreover, we saw that the space given at any time is also deformed when $\omega \neq 0$. Using the idea that the effect of a gravitational field is equal to the effect of the disc's movement (Einstein's equivalence principle), we can conceive that a gravitational field creates a deformation of the metric tensor field. When there is no gravity involved, a free particle's life is represented as a geodesic curve of the Minkowski space-time parametrised by its proper time (a geodesic curve is a straight line in $(\mathbb{R}^4, \mathcal{G})$). The tangent vector at each point of the particle's life will be a time-like vector aiming towards the future (contained in C^+). When a gravitational field is involved, we shall change the Minkowski metric to another one, and the events will no longer be represented in $\mathbb{R}^4$, being the new space-time a more general four dimensional differentiable manifold M. However, the particle's life will still be a geodesic curve of the new space-time with the new metric, also parametrised by its proper time. Let $\gamma(\tau)$ be the free falling particle's life in space-time, and let (ct, x^1, x^2, x^3) be coordinates in a local chart of space-time such that the vector $\frac{\partial}{\partial t}$ aims to the future (we will study with more details the concept of time orientation). The condition that its trajectory in space-time is a geodesic curve is written as

$$\ddot{\gamma}^i + \Gamma^i_{jk}\dot{\gamma}^j\dot{\gamma}^k = 0.$$

The observers whose life is represented by the coordinates (ct, x^1, x^2, x^3) will think that the particle is accelerating at $\ddot{\gamma}^i = -\Gamma^i_{jk}\dot{\gamma}^j\dot{\gamma}^k$, and they will attribute this acceleration to a 4-force created by a gravitational field. But this is not a real force, it is a fictitious force caused by the special election of the observers. In the expression of such a fictitious force, the Christoffel symbols are involved, and, in the expression of the Christoffel symbols, the first derivatives of the components of the metric $\mathcal{G}$ appear. Therefore, the (fictitious) gravitational force depends on the first derivatives of the metric components. To give an analogy with classical mechanics, as any force produced by a potential V is given by the first derivatives of V, in the theory of general relativity the metric $\mathcal{G}$ will substitute the gravitational potential V.

Let us see another 1-dimensional example built using Einstein's equivalence principle, which will help us to develop Einstein's gravitation theory. Consider an observer in a small spaceship placed on the surface of the Earth. Such an observer experiences the vertical force produced by the Earth's gravitational field. If he drops a mass from a certain altitude, the mass will experience an acceleration of $-g$. The equivalence principle allows us to consider such an observer in the spaceship accelerating at g from an inertial frame of reference in a zero-gravity space. Indeed, if he drops a mass, the observer will see the mass accelerating at $-g$. We will use this example to repostulate the equivalence principle. If we decrease the acceleration of the spaceship by g in both cases, the apparent gravitational field should be the same in both frames of reference (the origin of the frame of reference is in the center of the spaceship). In the case of the spaceship accelerating at g, decreasing this value by g, the spaceship will have null acceleration and its speed will remain constant. Therefore, if the observer at the center of the spaceship drops a mass freely moving, it will not experience any apparent force and it will "float" (the spaceship is travelling in a zero gravity space). In the case of the spaceship standing motionless on the Earth, decreasing the value of its acceleration (which was 0) by g, we have a spaceship accelerating at $-g$, i.e, free falling, and using the equivalence principle, we deduce that if the spaceship is very small, the observer will experience that he is in a zero-gravity space. Remember that the trajectory of a free falling particle's life is a geodesic curve. Let us repostulate the equivalence principle as: Any observer whose life is a geodesic curve will experience no gravity in a small neighbourhood. But, how can we describe mathematically the equivalence principle?

We have seen that given a free falling particle's life $\gamma(\tau)$, for every τ_0, we can have an inertial frame of reference F_{τ_0} such that its origin coincides with $\gamma(\tau_0)$ and the particle at the instant τ_0 is static in F_{τ_0} (proper frame of reference of the particle at τ_0). As the particle's life is a geodesic curve of the space-time, by using the equivalence principle, it is logic to assume that at every point of $\gamma(\tau)$, we can have an inertial frame of reference F_τ such that the effects of gravity are null in a small neighbourhood of space and time. Therefore, one can deduce that the induced metric of F_τ to the tangent space $T_{\gamma(\tau)}M$ can be represented through a change of coordinates in the tangent space as the Minkowski metric (note that the tangent space of any point of the space-time is $\mathbb{R}^4$).

6.3 Definition of space-time

Definition 6.1. *A **Lorenz manifold** is a pseudo-Riemannian manifold $(M, \mathcal{G})$ such that M is a four dimensional differentiable manifold and $\mathcal{G}$ is a metric tensor field such that for any $p \in M$, there exists a basis e_0, e_1, e_2, e_3 of T_pM such that the scalar products $\mathcal{G}(e_i, e_j)$ are given by the matrix*

$$\mathcal{G} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Similarly with what we defined in the Minkowski space-time, a vector $v \in T_pM$ is called time-like if $\mathcal{G}(v, v) < 0$, space-like if $\mathcal{G}(v, v) > 0$ and light-like if $\mathcal{G}(v, v) = 0$. C_p is the cone formed by the time-like vectors of T_pM. C_p will have two connected components. We will say that a Lorenz manifold $(M, \mathcal{G})$ is orientable with respect to time if for all $p \in M$ we can choose a connected component of C_p, which we will denote as C_p^+, that satisfies the following condition: There exists a smooth vector field $V \in \mathcal{X}(M)$ such that, for each local chart (U, φ) of M, $V_p \in C_p^+$ for all $p \in U$. The election of the connected component of C_p that satisfies the last condition is called a temporal orientation of M.

Definition 6.2. *(Space-time) We define the **space-time** as a Lorenz manifold that is orientable with respect to time, and with an election of a temporal orientation.*

6.4 Einstein's field equation

So far, we have seen the definition of space-time and its properties. But, how do we determine the metric tensor field in a region of the space-time? Einstein deduced an equation that allows us to determine such a tensor field. Let us see the deduction of this equation. In the frame of the classical mechanics, the matter in a continuum medium is described by the mass density σ. According to Newton's theory, this matter will create a gravitational field, whose potential V is related with σ by the Poisson equation

$$\Delta V = 4\pi K\sigma,$$

where Δ is the Laplacian operator and K is the gravitational constant. But, how can we write this equation based on the theory of gravitation proposed by Einstein? In general relativity, we have seen that the potential V is substituted by the metric $\mathcal{G}$. But if we replace V, which is a scalar, with $\mathcal{G}$, which is a tensor field, we will also have to substitute σ by a tensor field of the same type as $\mathcal{G}$. The relativistic equation would have to be an equation of tensor fields. In chapter 5 we have defined the stress-energy tensor field T, which is of type $(2, 0)$ (same type of $\mathcal{G}$), and contains the information of the matter, including σ. Such a tensor field satisfies the continuity equation

$$c^2 divT = \vec{f}',$$

where $\vec{f}'$ is the Minkowski body force density. Let us assume that the matter of the medium creates the gravitational field and that there are no other sources generating body forces (such as an electromagnetic field). In the gravitational theory developed in section 6.2, the gravitational forces do not exist, hence, there will be no body forces involved and the continuity equation will be written as

$$divT = 0.$$

Therefore, a matter in absence of electric charges (which create electromagnetic fields) will be described by the stress-energy tensor T that satisfies $divT = 0$. Hence, trying to build an analogous expression to the Poisson equation, we are searching for an equation that relates $\mathcal{G}$ with T and that does not depend on the chosen coordinates system. The second member will be, then, kT, where k is a constant. The first member would have to be a tensor field G' that only depends on $\mathcal{G}$ (ΔV only depends on V). Given a local chart, G' would have to contain the second derivatives of $\mathcal{G}_{ij}$, analogously to ΔV, which contains the second derivatives of V. Therefore, the equation that we are looking for is given by

$$G' = kT. \tag{6.3}$$

In order to find the tensor field G', we must search for those whose divergence is null. Indeed, by virtue of the continuity equation, $div(kT) = 0$. Einstein proposed the tensor field

$$G = Ric - \frac{1}{2}R\mathcal{G}, \tag{6.4}$$

where Ric is the Ricci tensor field and R the scalar curvature. G is known as the *Einstein's tensor field*. Let us prove that $divG = 0$.

Theorem 6.3. *$Div(Ric) = \frac{1}{2}dR$, where dR is the differential of the scalar curvature.*

Proof. As usual, we will chose a local chart $(U, \varphi = (x^1, ..., x^n))$. Using the second Bianchi's identity,

$$\nabla_r R^i_{jkh} + \nabla_h R^i_{jrk} + \nabla_k R^i_{jhr} = 0.$$

From the definition of Riemann curvature, the third term will be equal to $-\nabla_k R^i_{jrh}$. Note that in the last expression, for each i, we have an independent equation. Taking $r = i$ and summing with respect to i, we have

$$\nabla_i R^i_{jkh} + \nabla_h R^i_{jik} - \nabla_k R^i_{jih} = 0.$$

Therefore, we deduce

$$\mathcal{G}^{jk}\nabla_i R^i_{jkh} + \mathcal{G}^{jk}\nabla_h R^i_{jik} - \mathcal{G}^{jk}\nabla_k R^i_{jih} = 0. \tag{6.5}$$

In order to continue this proof, we need the following lemma:

Lemma 6.4. $\mathcal{G}^{jk}\nabla_i R^i_{jkh} = -\mathcal{G}^{ir}\nabla_i R^k_{rkh}$.

Proof. $R^i_{jkh} = \delta_{mi}R^m_{jkh} = \mathcal{G}^{ir}\mathcal{G}_{rm}R^m_{jkh} = \mathcal{G}^{ir}\mathcal{G}(\frac{\partial}{\partial x^i}, R(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^h})\frac{\partial}{\partial x^j})$.
Using proposition 3.10, the last expression can be written as

$$-\mathcal{G}^{ir}\mathcal{G}(\frac{\partial}{\partial x^j}, R(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^h})\frac{\partial}{\partial x^i}) = -\mathcal{G}^{ir}\mathcal{G}_{jm}R^m_{rkh}.$$

Therefore,

$$R^i_{jkh} = -\mathcal{G}^{ir}\mathcal{G}_{jm}R^m_{rkh} \Rightarrow \mathcal{G}^{jk}R^i_{jkh} = -\mathcal{G}^{ir}R^m_{rkh}.$$

Let us apply $\nabla_{\frac{\partial}{\partial x^i}}$ to the tensor K with components $K^i_h = \mathcal{G}^{jk}R^i_{jkh}$, and sum with respect to i; we will obtain the expression

$$\nabla_{\frac{\partial}{\partial x^i}} K^i_h = \nabla_i(\mathcal{G}^{jk})R^i_{jkh} + \mathcal{G}^{jk}\nabla_i R^i_{jkh}.$$

By how we defined the Levi-Civitta connection, we know that $\nabla_i \mathcal{G}_{jk} = 0$. We want to prove that $\nabla_i \mathcal{G}^{jk} = 0$ (($\mathcal{G}^{jk}$) is the inverse matrix of ($\mathcal{G}_{jk}$)). Using that $\mathcal{G}\mathcal{G}^{-1} = I$, we have

$$(\nabla_i \mathcal{G})\mathcal{G}^{-1} + \mathcal{G}\nabla_i \mathcal{G}^{-1} = 0,$$

therefore, as $(\nabla_i \mathcal{G}) = 0$, the first member is zero and we deduce that $\nabla_i \mathcal{G}^{-1} = 0$. Hence, as $\nabla_i \mathcal{G}^{jk} = 0$, by applying ∇_i on both sides of the equation $\mathcal{G}^{jk} R^i_{jkh} = -\mathcal{G}^{ir} R^m_{rkh}$, we prove the lemma. □

Continuation of the theorem's proof

Applying the lemma to the identity (6.5), we obtain

$$-\mathcal{G}^{ir}\nabla_i R^k_{rkh} + \mathcal{G}^{jk}\nabla_h R^i_{jik} - \mathcal{G}^{jk}\nabla_k R^i_{jih} = 0.$$

By the definition of Ricci tensor field, the equation above can be written as

$$-\mathcal{G}^{ir}\nabla_i R_{rh} + \mathcal{G}^{jk}\nabla_h R_{jk} - \mathcal{G}^{jk}\nabla_k R_{jh} = 0.$$

Note that the first and the third members are equal. Using that $\nabla_i \mathcal{G}^{ir} = 0$, the identity can be written as

$$-2(div(Ric))_h + \nabla_h R = 0.$$

As R is a function, $\nabla_h R = \partial R / \partial x^h$. The equation above implies that $2div(Ric) = dR$. □

Corollary 6.5. *$divG = 0$, where G is the Einstein's tensor field.*

Proof. We have to prove that $div(Ric) = \frac{1}{2} div(R\mathcal{G})$, hence, using the last theorem, it is sufficient to prove that $div(R\mathcal{G}) = dR$. In a local chart $(U, \varphi = (x^1, ..., x^n))$, we have

$$(div(R\mathcal{G}))_h = \nabla_i (R\mathcal{G}^{ik}\mathcal{G}_{kh}) = (\nabla_i R)\mathcal{G}^{ik}\mathcal{G}_{kh} = \nabla_h R = \frac{\partial R}{\partial x^h}.$$

□

Note that given a local chart, the Ricci curvature tensor field contains derivatives of the Christoffel symbols, hence, it depends on the second derivatives of $\mathcal{G}_{ij}$ (see equation 3.6). Einstein proposed that the equation that had to substitute the Poisson equation is $kT = G$. This expression is known as Einstein's field equation. One could question if G is the only tensor field satisfying $divG = 0$ and (determining a local chart) if it is the only one that is expressed in function of the second derivatives of $\mathcal{G}_{ij}$. Given $\Lambda \in \mathbb{R}$, if we add $\Lambda\mathcal{G}$ with G, we obtain the tensor field

$$Ric - \tfrac{1}{2}R\mathcal{G} + \Lambda\mathcal{G},$$

which divergence is also null because $div\mathcal{G} = 0$ (remember that $\nabla\mathcal{G} = 0$) and it contains the second derivatives of $\mathcal{G}_{ij}$. Poincare proved that the only tensor fields satisfying these conditions where $Ric - \frac{1}{2}R\mathcal{G} + \Lambda\mathcal{G}$ for all $\Lambda \in \mathbb{R}$. Therefore, the equation substituting the Poisson equation will be

$$G + \lambda\mathcal{G} = kT,$$

where k, and Λ are constants. By doing an approximation of this equation to a non relativistic frame, and comparing with Newton's theory of gravitation, one deduces that $k = 8K\pi/c^2$, where K is the gravitational constant, and $\Lambda = 0$ (see [2] section 18.5). Therefore, the Einstein's field equation is written as

$$Ric - \tfrac{1}{2}R\mathcal{G} = \frac{8K\pi}{c^2}T.$$

Bibliography

[1] Carlos Currás Bosch, *Geometria diferencial: varietats diferenciables i varietats de Riemann*, Universitat de Barcelona; 63, (2003).

[2] Joan Girbau, *Geometria diferencial i relativitat*, Manuals de la Universitat Autònoma de Barcelona; 10, (1993).

[3] Xavier Oliver Olivella i Carlos Agelet de Saracíbar Bosch, *Mecánica de medios continuos para ingenieros*, Edicions UPC (2000).

[4] Bert Janssen, *Breve repaso de la relatividad especial*, Universidad de Granada (2005) http://www.ugr.es/ bjanssen/text/repaso.pdf

[5] Carroll S.M. *An introduction to General Relativity. Space time and geometry*, Addison Wesley (2004)