

Introduction to General Relativity

Rodrigo Pereira Silva
University of São Paulo

Summer, 2023

During my investigations of General Relativity, I've noticed something that perpetuates along all the physical theories. Every physical theory has a physical motivation and background, but the tools necessary to work with it are mathematical. This mathematical ground plays a crucial role in General Relativity because it is heavy in the core of the theory.

Hence, from now on the beginning of these notes will be dedicated to the underlying principles of differential geometry and vector spaces, with some physical examples to get some motivation and clarification. We can talk about physics later.

The intuition behind General Relativity is that matter curves this entity called space-time, and deforms it. Therefore, spacetime is a deformed and wrinkled *thing*, and we have to do physics over it. To do physics over it, we need to define (and sometimes re-define) a series of properties that we usually expect from physical systems so we can calculate their properties, time-evolution, etc. Long story short: general relativity is an intrinsically geometric theory, so we need to learn geometry.

Of course all of this is very informal and intuitive, and in order to formalize these things we need to digress a little. Let's begin with the very first mathematical structure if we want to talk about geometry, which is a topology.

1 Topology

The least amount of structure we can have in a spacetime theory is a topology. So we shall define it.

Definition 1. Let M be a set. A topology \mathcal{O} is defined as a set $\mathcal{O} \subseteq \mathcal{P}(M)$ ¹ satisfying

- i. $\emptyset \in \mathcal{O}$, and $M \in \mathcal{O}$
- ii. $\forall V, U \in \mathcal{O} \Rightarrow V \cap U \in \mathcal{O}$
- iii. $\forall U_\alpha \in \mathcal{O} \Rightarrow \bigcup_{\alpha \in A} U_\alpha \in \mathcal{O}$, where A is a set of indices.

So this is a topology. Notice that a topology is built upon a set. So, for all topology \mathcal{O} there is an underlying set M .

Example 1. Take the simple set $M = \{1, 2, 3\}$. Now, define $\mathcal{O}_1 := \{\emptyset, \{1, 2, 3\}\}$. We shall proof that this is a topology:

¹ $\mathcal{P}(M)$ is called the superset of M , which is the set containing all the subsets of M .

- i. Trivial
- ii. Since $\emptyset \cap \{1, 2, 3\} = \emptyset \in \mathcal{O}_1$, this is satisfied.
- iii. Since $\emptyset \cup \{1, 2, 3\} = \{1, 2, 3\} \in \mathcal{O}_1$, this is also satisfied.

Now define $\mathcal{O}_2 := \{\emptyset, \{1\}, \{2\}, \{1, 2, 3\}\}$. We shall proof that this is not a topology:

- i. Trivial.
- ii. This is satisfied since $\{1, 2, 3\} \cap \{1\} = \{1\} \in \mathcal{O}_2$.
- iii. Since $\{1\} \cup \{2\} = \{1, 2\} \notin \mathcal{O}_2$, this is not a topology.

□

This example shows that although there are many sets $\mathcal{O} \subseteq \mathcal{P}(M)$, not all of them can be called a topology. Now, the next example introduces two very important cases (and they are important because they are completely useless):

Example 2. There are two topologies that can be created over any set M . They are:

- 1. $\mathcal{O}_{\text{discrete}} := \{\emptyset, \mathcal{P}(M)\}$
- 2. $\mathcal{O}_{\text{chaotic}} := \{\emptyset, M\}$.

□

Example 3. Another important topology is what we call the standard topology, or \mathcal{O}_{std} . Here's the implicit definition of it:

Definition 2. Let $M = \mathbb{R}^d$, and $\mathcal{O}_{\text{std}} \subseteq \mathcal{P}(\mathbb{R}^d)$. Let a softball be

$$B_r(p) := \left\{ (q_1, \dots, q_d) \mid \sum_{i=1}^d (q_i - p_i)^2 < r^2 \right\},$$

with $r \in \mathbb{R}^+$, $p \in \mathbb{R}^d$. We say that $\mathcal{U} \in \mathcal{O}_{\text{std}}$ iff

$$\forall p \in \mathcal{U} : \exists r \in \mathbb{R}^+ : B_r(p) \subseteq \mathcal{U}.$$

In other words, a set \mathcal{U} is a subset of the standard topology if, and only if, for any point p in \mathcal{U} , we can draw a circle of radius r around it, and every point within this circle is still in \mathcal{U} .

Remark 1. This intuition is just a helper to understand the definition better. When we talk about sets, there is no structure other than a collection of elements. Although we use the terminology "softball" and "radius" to make it more friendly, there is no notion of a radius, let say distance, within a pure set.

Now, a little bit of terminology:

- M is a set.
- \mathcal{O} is a topology, which is a set of open sets.
- (M, \mathcal{O}) is called a topological space.
- If $\mathcal{U} \in \mathcal{O}$, then \mathcal{U} is an open set.
- If $M \setminus A \in \mathcal{O}$, then A is a closed set.

1.1 Continuous maps

A map is defined as $f : M \longrightarrow N$. It means that f is a rule that takes elements from set M to set N . This question to be answered here is whether a map f is continuous or not. And the answer is that it depends on the topologies chosen based on M and N . So we shall define them as \mathcal{O}_M as the topology constructed based on set M , and \mathcal{O}_N as the topology constructed based on set N . Now, for the definition of continuity, we need first the definition of pre-image of a map:

Definition 3. Let $f : M \longrightarrow N$, and let $V \subseteq N$. Then, we define the pre-image of f as

$$\text{preim}_f(V) := \{m \in M | f(m) \in V\}.$$

Hence, the pre-image of a map are the set of all elements from M that are mapped into N . Finally, we can define continuous maps:

Definition 4. Let (M, \mathcal{O}_M) and (N, \mathcal{O}_N) be topological spaces. Then a map $f : M \longrightarrow N$ is called a *continuous map* with respect to \mathcal{O}_M and \mathcal{O}_N if

$$\forall V \in \mathcal{O}_N : \text{preim}_f(V) \in \mathcal{O}_M$$

This definition means that, a map is continuous iff the pre-images of all open sets in N are open sets in M . Let's see an example:

Example 4. Let $M = \{1, 2\}$ and $N = \{1, 2\}$. Then we define a map

$$\begin{aligned} f : M &\longrightarrow N \\ m &\longmapsto f(m) := m \end{aligned}$$

that is the identity map. The underlying topologies $\mathcal{O}_M := \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$, and $\mathcal{O}_N := \{\emptyset, \{1, 2\}\}$. To check whether or not this is a continuous map, we have to look at the pre-images:

$$\begin{aligned} \text{preim}_f(\emptyset) &= \emptyset \in \mathcal{O}_M, \\ \text{preim}_f(\{1, 2\}) &= M \in \mathcal{O}_M, \end{aligned}$$

then f is continuous.

Now let us define the inverse identity function, which is of course is the identity function. But now, notice that

$$\begin{aligned} f : N &\longrightarrow M \\ n &\longmapsto f(n) := n \end{aligned}$$

Thinking about the underlying topologies, let's see the pre-images:

$$\text{preim}_{f^{-1}}(\{1\}) = \{1\} \notin \mathcal{O}_N,$$

then f^{-1} is not continuous. □

Henceforth, we see that even the identity function is not continuous, and the whole concept of continuity of maps is based on the underlying topologies defined for the topological spaces.

1.1.1 Composition of two continuous maps

Let $M \overset{f}{\rhd} N \overset{g}{\rhd} P$, but also the composition of these two functions,

$$\begin{aligned} f \circ g &: N \longrightarrow P \\ m &\longmapsto (g \circ f)(m) := g(f(m)) \end{aligned}$$

Now that things are defined, we can enunciate the theorem:

Theorem 1. *Let $f : M \longrightarrow N$ and $g : N \longrightarrow P$ be two continuous functions. Then, the composition $f \circ g$ is also continuous.*

The proof is not difficult:

Proof. Let $V \in \mathcal{O}_P$ be a subset of the target set of the composite function, then:

$$\begin{aligned} \text{preim}_{g \circ f}(V) &:= \{m \in M \mid (g \circ f)(m) \in V\} \\ &= \{m \in M \mid f(m) \in \text{preim}_g(V)\} \\ &= \text{preim}_f(\underbrace{\text{preim}_g(V)}_{\in \mathcal{O}_N}) \\ &\quad \underbrace{\hspace{10em}}_{\in \mathcal{O}_M} \end{aligned} \tag{1.1}$$

Henceforth, the composition is continuous. ■

1.2 Inheriting a Topology

Sometimes it is useful to define a topology from an already existing topological space. One particularly important way to inherit a topology when it comes to spacetime physics is described in this section. Let M be a set with a topology already given, \mathcal{O}_M . Consider now a subset $S \subseteq M$. The question is: can one construct on S a topology from \mathcal{O}_M on M ? The answer is yes, and here's how:

Definition 5. The topology $\mathcal{O}|_S \subseteq \mathcal{P}(S)$ is said to be a subset topology when constructed like

$$\mathcal{O}|_S := \{\mathcal{U} \cap S \mid \mathcal{U} \in \mathcal{O}_M\}.$$

A result to consider is that, if we have a mapping f taking elements from M (\mathcal{O}_M) into N (\mathcal{O}_N) that is continuous, and we have a restricted function defined as

$$f|_S := S \subseteq M \longrightarrow N$$

then we can claim that this function is continuous if it was constructed as it was shown above (that is, if the topology from S is the topology $\mathcal{O}|_S$).

2 Topological Manifolds

Now we can restrict even more our universe of possibilities. For spacetime physics, we may focus on a set of topological spaces (M, \mathcal{O}) that can be charted, in a very similar manner on how the Earth is charted in an atlas.

Definition 6. A topological space (M, \mathcal{O}) is called a d -dimensional topological manifold if

$$\forall p \in M : \exists \mathcal{U} \in \mathcal{O} : \exists x : \mathcal{U} \longrightarrow x(\mathcal{U}) \subseteq \mathbb{R}^d$$

where:

- x is invertible: $\exists x^{-1} : x(\mathcal{U}) \longrightarrow \mathcal{U}$
- x is continuous.
- x^{-1} is continuous.

Notice that the requirement of continuity for x is addressable since we have a topology for both the target and the domain set of the map x : the set $\mathcal{U} \subseteq M$, which has topology \mathcal{O} , and \mathbb{R}^d has topology \mathcal{O}_{std} .

This is fascinating. This definition is saying that, for all points in our set M , there is a subset $\mathcal{U} \in \mathcal{O}$, and there is a map x that takes elements from this neighborhood \mathcal{U} to an Euclidean \mathbb{R}^d space. This accounts for the many figures showing a weird, wobbly picture with a map taking surfaces from this picture to a nice, flat space. This is what we want, for short: a topological space that, if you take a close look and focus your attention in a tiny part of it, it resembles \mathbb{R}^d . And this is good because that's where we know how to make physics.

A bit of terminology: the tuple (\mathcal{U}, x) is called a chart. And it is called a chart because there is a nice and pictorial way to look at these topological manifolds. Imagine that you are creating a map of the Earth; imagine what you need to do in order to accomplish this task: you have to take real-world objects (in here they lay in our M set) and you have to make pictures of them, so you apply a map x that creates this picture with certain characteristics. The key point is: there are many ways to create pictures of a real place, and these pictures do not affect at all the real place. This is precisely the relation of the chart produced by some *specific* map x : you are making a massive choice by discarding all the other possible maps and, nevertheless, your real world (your topological space (M, \mathcal{O})) remains unaffected.

Definition 7. An atlas \mathcal{A} is a collection of charts:

$$\mathcal{A} := \{(\mathcal{U}_{(\alpha)}, x_{(\alpha)}) | \alpha \in A\}$$

if

$$\bigcup_{\alpha \in A} \mathcal{U}_{(\alpha)} = M.$$

Henceforth, an atlas is a collection of charts, as long as if union of the pre-images of x reconstruct again the real world.

Definition 8. A chart map is the map $x(p) = (x^1(p), x^2(p), \dots, x^d(p))$ where each $x^i(p) := \mathcal{U} \longrightarrow \mathbb{R}$ is called a coordinate map.

2.1 Chart Transition Maps

Imagine two charts (\mathcal{U}, x) , and (\mathcal{V}, y) with overlapping regions, that is, $\mathcal{U} \cap \mathcal{V} \neq \emptyset$. Now, take any point within the intersection of these sets. We can map them to \mathbb{R}^d using either x and y . The question here is: how can we make a transition of coordinates from x to y , for example? And the answer is: given $p \in x(\mathcal{U} \cap \mathcal{V}) \subseteq \mathbb{R}^d$, take $x^{-1}(p) \in \mathcal{U} \cap \mathcal{V}$, and then apply the map y in it to go to $y(\mathcal{U} \cap \mathcal{V}) \subseteq \mathbb{R}^d$. The whole process is then $y \circ x^{-1} := y(x^{-1}(p))$. This is called a chart transition map.

Often it is desirable (or indeed the only way) to define properties of real-world objects by judging suitable conditions not on the real-world object itself, but on a chart-representative of that real world object.

3 Multilinear Algebra

It's not the case that we will equip our spacetime set with a vector space structure. Indeed, we have so far equipped it with a space topology structure. However, the tangent spaces (which shall be defined in its moment) $T_p M$ to a smooth manifold associated with a set M has a vector space structure. And all the underlying vector and tensor manipulation done in General Relativity is done actually within this space. Long story short: we are going to talk about tensors now. But first, we need to define what a vector space is.

3.1 Vector Spaces

We shall promptly start with its definition.

Definition 9. A \mathbb{R} -vector space (V, \oplus, \odot) is a tuple with a set V , and two operations

$$\begin{aligned}\oplus : V \times V &\longrightarrow V, \\ \odot : \mathbb{R} \times V &\longrightarrow V,\end{aligned}$$

such that these operations satisfy:

- i. $v \oplus w = w \oplus v, \forall v, w \in V$
- ii. $(u \oplus v) \oplus w = u \oplus (v \oplus w), \forall v, w, u \in V$
- iii. $\exists 0 \in V : \forall v \in V : v \oplus 0 = v$
- iv. $\forall v \in V : \exists (-v) \in V : v \oplus (-v) = 0$
- v. $\lambda \odot (\mu \odot v) = (\lambda \cdot \mu) \odot v, \forall \lambda, \mu \in \mathbb{R}$
- vi. $(\lambda + \mu) \odot v = \lambda \odot v + \mu \odot v, \forall \lambda, \mu \in \mathbb{R}, \forall v \in V$
- vii. $\lambda \odot (v + w) = \lambda \odot v + \lambda \odot w, \forall \lambda \in \mathbb{R}, \forall v, w \in V$
- viii. $1 \odot v = v, \forall v \in V$

Notice that there is a clear distinction between the sum on the real numbers $+$ from the sum in the vector space \oplus ; and the multiplication on the real numbers \cdot and the multiplication in the vector space \odot . When elements from the real numbers interact between them, we have to use ordinary sum and multiplication. However, when they interact with a vector, then the structure defined within the vector space is needed.

Remark 2. The elements of a vector space is often referred to, informally, as a vector. It is important to notice that a set is not a vector space: it is necessary to define the operations \oplus and \odot such that it obeys the definition given above.

Example 5. Let $P := \left\{ p : [-1, +1] \longrightarrow \mathbb{R} \mid p(x) = \sum_{n=0}^N p_n \cdot x^n, p_n \in \mathbb{R} \right\}$, which is the set of polynomials of fixed degree.

The question is whether or not an element of this set P is a vector, and the answer is a resounding no. There is nowhere any \cdot or $+$ operation defined here. However, if we define

$$\begin{aligned} \oplus : P \times P &\longrightarrow P \\ (p, q) &\longmapsto p \oplus q : (p \oplus q)(x) := p(x) + q(x) \end{aligned}$$

and

$$\begin{aligned} \odot : \mathbb{R} \times P &\longrightarrow P \\ (\lambda, p) &\longmapsto \lambda \odot p : (\lambda \odot p)(x) := \lambda \cdot p(x) \end{aligned}$$

Now, we have indeed a vector space (because it is already known that these operations defined above satisfy the conditions in the definition of a vector space). \square

Now, our next piece of information in the journey to define what a tensor is, is to define what a linear map from a vector space is.

3.2 Linear Maps

Let's begin with the definition, which is somehow very similar to the definition from linear maps in topological spaces.

Definition 10. Let (V, \oplus_V, \odot_V) and (W, \oplus_W, \odot_W) be two vector spaces. Then a map

$$\phi : V \longrightarrow W$$

is called linear if

- i. $\phi(v \oplus_V w) = \phi(v) \oplus_W \phi(w)$
- ii. $\phi(\lambda \odot_V v) = \lambda \odot_W \phi(v)$

Notice, from the definition, that the addition and s-multiplication defined in V and W don't need to be the same. On the left-hand side of the definitions above, we are at vector space (V, \oplus_V, \odot_V) , while on the left side we were already taken to the target set W , with its \oplus_W addition and \odot_W s-multiplication. And also an important remark is that, in order to define or not the linearity of a map, we need to know the thorough structure of the vector spaces involved, e.g. we need to define addition and s-multiplication.

Example 6. Using the same set P of polynomials of fixed degree, we can define a differentiation operator

$$\begin{aligned} \delta : P &\longrightarrow P \\ p &\longmapsto \delta(p) := p' \end{aligned}$$

so δ is the derivative of the polynomial p . The usual rules of differentiation (sum rule and multiplication by a constant) provides us the proof that this operator is indeed linear according to the definition of a linear operator. \square

Now, a bit of notation to make things a little less cluttered:

Notation. The indication of a linear map can be expressed as simply

$$\phi : V \rightarrow W.$$

Next, a theorem:

Theorem 2. Let (V, \oplus_V, \odot_V) , (W, \oplus_W, \odot_W) , (U, \oplus_U, \odot_U) be vector spaces. Now take the linear maps ϕ and ψ such that

$$V \xrightarrow{\phi} W \xrightarrow{\psi} U$$

Then

$$\phi \circ \psi : V \longrightarrow U$$

is linear.

An immediate example of this theorem is the composition of two first-order derivatives, which is a second-order derivative. Second-order derivatives are as linear as any-order derivative.

3.3 Vector space of Homomorphisms

If we have two vector spaces (V, \oplus_V, \odot_V) , (W, \oplus_W, \odot_W) , (U, \oplus_U, \odot_U) , we can define the set of all linear maps between them, which is named the set of homomorphisms between V and W , and it is simply defined as

$$\text{Hom}(V, W) := \{\phi : V \rightarrow W\}.$$

The funny thing is that we can make turn this set into a vector space itself. It suffices to define the addition and multiplication for this set:

$$\begin{aligned} \oplus_H : \text{Hom}(V, W) \times \text{Hom}(V, W) &\longrightarrow \text{Hom}(V, W) \\ (\phi, \psi) &\longmapsto \phi \oplus_H \psi \end{aligned}$$

where this new vector space $\text{Hom}(V, W)$ can inherit the addition and multiplication from its “parent” vector spaces, that is:

$$(\phi \oplus_H \psi)(v) = \phi(v) \oplus_W \psi(v)$$

And for s-multiplication, we have

$$\begin{aligned} \odot : \mathbb{R} \times \text{Hom}(V, W) &\longrightarrow \text{Hom}(V, W) \\ (\lambda, v) &\longmapsto \lambda \odot_H v \end{aligned}$$

where

$$(\lambda \odot_H \phi)(v) := \lambda \odot_W \phi(v).$$

therefore, $(\text{Hom}(V, W), \oplus_H, \odot_H)$ is a vector space.

Example 7. Let $\text{Hom}(P, P)$, where P is the set of fixed degree polynomials. Then, $\delta \in \text{Hom}(P, P)$, as well as $\delta \circ \delta$. Also, any composition involving some multiplication and addition is an element of $\text{Hom}(P, P)$. \square

3.3.1 Dual spaces

The reason why this is a subsection of the Homomorphisms section is because this is a special case of Homomorphism, where we construct a set of all linear maps from V to \mathbb{R} . That's it. Formally, we call this the dual space V^* and we define it as follows:

Definition 11. Let (V, \oplus_V, \odot_V) and $(\mathbb{R}, +, \cdot)$ be vector spaces. Then the dual space V^* is

$$V^* := \{\phi : V \rightarrow \mathbb{R}\} = \text{Hom}(V, \mathbb{R}).$$

And since this is just a special case from Homomorphisms, it is immediate that $\text{Hom}(V, \mathbb{R})$ is also a vector space. A bit of terminology: an element of V^* is called a covector.

Example 8. Let $I : P \rightarrow \mathbb{R}$, therefore $I \in P^*$. We can define I as, for example,

$$I(p) := \int_0^1 dx p(x),$$

being the integral operator, which is known to be linear. \square

Finally, we have all the ingredients necessary to introduce the notion of tensors.

3.4 Tensors

If you have a vector space and you consider its dual space, then you have all the building blocks necessary to construct your tensors, which consist of multilinear maps.

Definition 12. Let (V, \oplus_V, \odot_V) be a vector space. An (r, s) -tensor T over V is a multilinear map

$$T : \underbrace{V^* \times \dots \times V^*}_r \times \underbrace{V \times \dots \times V}_s \rightarrow \mathbb{R}$$

Let's clear out what multilinearity means with an example:

Example 9. If T is a $(1, 1)$ -tensor, it means that it can have as input two entries: one covector and one vector, so

$$T(\phi + \psi, v) = T(\phi, v) + T(\psi, v)$$

where $\phi \in V^*$ and $v \in V$. To be multilinear means that

$$T(\phi, v + w) = T(\phi, v) + T(\phi, w)$$

and

$$T(\phi, \lambda v) = \lambda T(\phi, v).$$

Finally,

$$T(\phi + \psi, v + w) = T(\phi, v) + T(\phi, w) + T(\psi, v) + T(\psi, w).$$

\square

Now, an example of a tensor.

Example 10. Define the (0,2)-tensor

$$g : \mathbb{P} \times \mathbb{P} \rightarrow \mathbb{R}$$

where we can take

$$(p, q) \mapsto \int_0^1 dx p(x)q(x).$$

□

A nice piece of information is that vector and covectors are themselves tensors. The first result for covectors is immediate: $\phi \in V^* \iff \phi : V \rightarrow \mathbb{R}$, then it is immediate that ϕ is a (1, 0)-tensor.

Now, for the vector, we have $v \in V = (V^*)^* \iff v : V^* \rightarrow \mathbb{R}$, so v is a (0, 1)-tensor. Notice that the equality casually thrown there $V = (V^*)^*$ is not trivial at all, and only works for finite dimensional vector spaces. Of course we still don't know what a dimension is, but we can define it when we talk about a vector basis.

3.5 Bases

Basis are not very good for you. When you introduce a basis into your calculations, every result you obtain is conditioned to this basis. It means that when you choose a basis, you make an absolutely massive choice and collapse it into one: the basis you have chosen. This situation is very similar with choosing a chart for your topological space. With that being said, let's talk about basis.

Definition 13. Let (V, \oplus_V, \odot_V) be a vector space. A subset $B \subset V$ is called a basis if

$$\forall v \in V \exists! \text{ finite } F \subset B : \exists! (v^1, \dots, v^n) : v = \sum_{i=1}^n v^i f_i.$$

where $F = \{f_k\}_{k=1}^n$.

With this definition in hands, we can define what is the dimension of a vector space:

Definition 14. If a vector space (V, \oplus_V, \odot_V) has a basis $B \subset V$ with finitely many elements, say d many, then we say that $\dim V := d$.

Now, a little remark on the components of a vector within a vector space with a chosen basis:

Remark 3. Let (V, \oplus_V, \odot_V) be a finite dimensional vector space. Having chosen a basis e_1, \dots, e_n of (V, \oplus_V, \odot_V) , we may uniquely associate

$$v \mapsto (v^1, \dots, v^n)$$

which are called the components of v with respect to a chosen basis. These components are such that

$$v = v^1 e_1 + \dots + v^n e_n.$$

3.5.1 A basis for the Dual Space

Having chosen the basis (e_1, \dots, e_n) for V , we can also choose a basis $(\epsilon^1, \dots, \epsilon^n)$ for V^* . There is no reason why we couldn't choose completely independent basis for each set, but this introduces as double as much arbitrariness. Henceforth, it is desirable (and more economical) to choose the basis of the dual space based on the basis chosen for the vector space, and let them be attached somehow.

So, if (e_1, \dots, e_n) have been chosen, it is a nice practice to create a basis for the dual space as

$$\epsilon^a(e_b) = \delta_b^a = \begin{cases} 1, & a = b \\ 0, & a \neq b \end{cases}$$

This uniquely determines the choice of ϵ 's from the choice of e 's. In the case where the basis of the dual space V^* satisfies this, we call this basis the *dual basis*.

Example 11. Let P be the set of polynomials of fixed degree, equipped with the addition and s-multiplication, so we talk about a vector space. We choose the degree to be fixed at 3. Now, the basis to this vector space (e_0, e_1, e_2, e_3) is given by $(1, x, x^2, x^3)$, so $e_a(x) := x^a$. Now, we can create the dual basis for this vector space as being

$$\epsilon_a := \frac{1}{a!} \left. \partial_x^a \right|_{x=0},$$

where ∂_x^a is the a th derivative with respect to x , and after the derivative is calculated, you evaluate it at $x = 0$. It is immediately verifiable that this definition satisfies the requirements to be a dual basis. \square

Now that we have talked about a basis, we can take a look at the components of a tensor.

3.5.2 Components of a tensor

The components of a tensor are defined as follows:

Definition 15. Let T be an (r, s) -tensor on a finite dimensional vector space V , a basis $\{e_1, \dots, e_n\}$ of V , and a dual basis $\{\epsilon^1, \dots, \epsilon^n\}$ for V^* . Then, define the $(r + s)^{\dim V}$ many real numbers

$$T_{j_1, \dots, j_s}^{i_1, \dots, i_r} \in \mathbb{R} := T(\epsilon^{i_1}, \dots, \epsilon^{i_r}, e_{j_1}, \dots, e_{j_s})$$

where $i_1, \dots, i_r, j_1, \dots, j_s \in \{1, \dots, \dim V\}$.

Henceforth, the components of a (r, s) -tensor are defined based on the basis of the vector space V and the dual vector space V^* , which is a natural way to define it.

One really useful application of having a basis is that, once the components and the basis of the tensor are known, one can reconstruct the entire tensor.

Example 12. Let T be a $(1, 1)$ -tensor. Then, each component of the tensor is defined as

$$T_j^i := T(\epsilon^i, e_j)$$

then

$$\begin{aligned}
T(\phi, v) &= T\left(\sum_{i=1}^{\dim V} \phi_i \epsilon^i, \sum_{j=1}^{\dim V} v^j e_j\right) \\
&= \sum_{i=1}^{\dim V} \sum_{j=1}^{\dim V} \phi_i v^j \underbrace{T(\epsilon^i, e_j)}_{T_j^i} \\
&= \phi_i v^j T(\epsilon^i, e_j),
\end{aligned}$$

□

In the last line of this example I introduced the Einstein summation convention, which says that if you have an index upstairs and the same index downstairs in different components, there's an implicit summation occurring. This is possible simply because elements from basis V are labeled with upstairs indices, while elements from basis V^* are labeled with downstairs indices.

4 Differentiable Manifolds

Topological manifolds gives us the idea of continuity of curves that are in a topological space. Now we want somehow to associate a velocity of a curve as well. The idea of talking about differentiability of a curve is then very natural, and we want to address it now. A first question to be asked is whether or not the structure already given to a topological manifold is enough for us to talk about differentiability, and the answer is no.

In order to talk about the differentiation of a curve in a topological manifold, we need to talk about charts, simply because we don't know how to differentiate a map that takes real numbers and throw them into a topological manifold, and that is precisely what a parameterized curve does:

$$\gamma : \mathbb{R} \longrightarrow M.$$

The usual notion of differentiation lies in maps that takes elements from \mathbb{R} into elements of \mathbb{R}^d , and in order to accomplish this in the context of topological manifolds, first we need to talk about charts. Of course charts are fantasies of the real world; it would be catastrophic if my notion of differentiability depended on the type of chart I chose to represent my manifold, because differentiability is a very intrinsic notion of a real-world object. To see how problematic it is, imagine a real-world sphere that I chose to represent with a differentiable chart on a Monday, but at Friday I chose to represent with a non-differentiable chart. This means that the same object—the real-world sphere—created a sharp edge during the week. All of a sudden, physics is not consistent anymore. This silly story serves to motivate us to define a kind of compatibility between charts, in a manner that they should not disagree in overlapping regions.

4.1 Compatible charts

Let \mathcal{U} and \mathcal{V} be two overlapping regions ($\mathcal{U} \cap \mathcal{V} \neq \emptyset$) of our manifold (M, \mathcal{O}) . We can choose for each of these regions maps x and y such that they form charts, so we have

(\mathcal{U}, x) and (\mathcal{V}, y) . Now, while x can map an overlapping region of the real-world into a chart,

$$\mathbb{R}[\gamma\mathcal{U} \cap \mathcal{V}]xx(\mathcal{U} \cap \mathcal{V}) \subseteq \mathbb{R}^d$$

y can do the same thing:

$$\mathbb{R}[\gamma\mathcal{U} \cap \mathcal{V}]yy(\mathcal{U} \cap \mathcal{V}) \subseteq \mathbb{R}^d$$

the thing is: we can work with the composite functions $x \circ \gamma$ and $y \circ \gamma$ and their notions of differentiability because we know the domain and the image of these functions, but we need somehow to make sure that the notions established in the differentiability level by the map x are not broken by the ones established by the map y .

The way to tight these concepts together is to examine the chart transition map, given by

$$y \circ x^{-1} : \mathbb{R}^d \longrightarrow \mathbb{R}^d,$$

since it gives us the opportunity to study the desired composition

$$y \circ \gamma = (y \circ x^{-1}) \circ (x \circ \gamma).$$

Notice here that $x \circ \gamma : \mathbb{R} \longrightarrow \mathbb{R}^d$, and we know how to work with this. However, the first term of this composition $y \circ x^{-1} : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is, in general, only continuous (by the definition of a topological manifold). Hence, we have the composition of a (generally) continuous function with a differentiable function, and the result is (generally) a continuous function. So if we take any two charts (\mathcal{U}, x) , (\mathcal{V}, y) and try to address the differentiability problem with no further constraints, we will fail.

Until this point, we are using any imaginable charts (\mathcal{U}, x) , (\mathcal{V}, y) on the topological manifold (M, \mathcal{O}) . In other words, we are taking (\mathcal{U}, x) , (\mathcal{V}, y) from the *maximal atlas* \mathcal{A} of (M, \mathcal{O}) . The concept of maximal atlas is simple: it is the set of all possible charts from a topological manifold. This is not what we want, and the set of charts need to be more restricted. Indeed, the answer is not really complicated: we just need to choose the charts that are compatible with respect to that operation. This seems a little circular, but we can finally define the compatibility of two charts.

Definition 16. Two charts (\mathcal{U}, x) , (\mathcal{V}, y) of a topological manifold are called \clubsuit -compatible if either:

- (a) $\mathcal{U} \cap \mathcal{V} = \emptyset$, or
- (b) $\mathcal{U} \cap \mathcal{V} \neq \emptyset$ and the chart transition maps

$$\begin{aligned} y \circ x^{-1} : x(\mathcal{U} \cap \mathcal{V}) &\subseteq \mathbb{R}^d \longrightarrow y(\mathcal{U} \cap \mathcal{V}) \subseteq \mathbb{R}^d \\ x \circ y^{-1} : y(\mathcal{U} \cap \mathcal{V}) &\subseteq \mathbb{R}^d \longrightarrow x(\mathcal{U} \cap \mathcal{V}) \subseteq \mathbb{R}^d \end{aligned}$$

have the \clubsuit -property.

Definition 17. An atlas \mathcal{A}_\clubsuit is a \clubsuit -compatible atlas if any two charts in \mathcal{A}_\clubsuit are \clubsuit -compatible.

Definition 18. A \clubsuit -manifold is a triple $(M, \mathcal{O}, \mathcal{A}_\clubsuit)$.

Let's take the obviously ridiculous \clubsuit out of the way and say what this symbol can be in Table ???. We see that, by imposing more and more conditions on the transition functions, you get more and more restricted. Indeed, we didn't have to add any other extra structure in order to talk about differentiability of curves; what we had to do was to make a choice (hence, to add information into our already existing structure) in order to "clean" our atlas with bad-behaving charts.

\clubsuit	Property known	Description
C^0	$C^0(\mathbb{R} \rightarrow \mathbb{R})$	Continuous maps with respect to \mathcal{O}_{std}
C^1	$C^1(\mathbb{R} \rightarrow \mathbb{R})$	Differentiable (once) and the result is continuous
\vdots	\vdots	\vdots
C^∞	$C^\infty(\mathbb{R} \rightarrow \mathbb{R})$	Smooth manifold

Now, an important theorem:

Theorem 3. *Any $C^{k \geq 1}$ -atlas $\mathcal{A}_{C^{k \geq 1}}$ contains a C^∞ -atlas.*

What it means: if you can guarantee that you have an atlas with at least one C^1 transition function, then you can remove charts from it until you get an atlas with only smooth charts.

4.2 Diffeomorphisms

Consider the map $\phi : M \rightarrow N$ where M and N are naked maps, that is, that don't have any further structure. Then, the structure-preserving maps are the bijections (which are the invertible maps). Otherwise, the map is not invertible and we can't fully recover one set from the other via the map ϕ .

Example 13. Consider the two sets $\{1, 2, 3\}$ and $\{a, b\}$. There is no possible map that preserves the structure of these two sets. We can map $(1, a)$ and $(2, b)$, but there's no pair for the element 3. Hence, we would have to map 3 to either a or b , making the map not invertible, whence not preserving set structure. \square

We now many sets that have bijections between them: the set of natural numbers \mathbb{N} have a bijection with the integers \mathbb{Z} , as well as with the rationals \mathbb{Q} , but not with the reals \mathbb{R} . We can now define what it means to be isomorphic:

Definition 19. M and N are set theoretically isomorphic, and is denoted as $M \cong_{\text{set}} N$ if there is a bijection $\phi : M \rightarrow N$.

There is a multitude of "-morphisms", and all of them preserve a structure somehow of a set (not a naked set, but a set with more structure). The next case we talk about are homeomorphic sets.

Definition 20. Consider two topological spaces (M, \mathcal{O}_M) , (N, \mathcal{O}_N) . If there is a bijection $\phi : M \rightarrow N$, and both ϕ and ϕ^{-1} are continuous, they are said to be topologically isomorphic, or $(M, \mathcal{O}_M) \cong_{\text{top}} (N, \mathcal{O}_N)$. This topological isomorphism is called homeomorphism.

For a vector space, we can do a similar definition:

Definition 21. Consider two vector spaces (V, \oplus_V, \odot_V) and (W, \oplus_W, \odot_W) . These vector spaces are said to be isomorphic if there is a bijection $\phi : V \rightarrow W$ that is linear in both directions.

Before the star of the section, which are diffeomorphic sets, we first need to define what are C^∞ maps:

Definition 22. Consider $\mathcal{U} \subset (M, \mathcal{O}_M)$ and $\mathcal{V} \subset (N, \mathcal{O}_N)$. Consider now

$$\mathbb{R}^d \llbracket x \mathcal{U} \subseteq M \llbracket \phi \mathcal{V} \subseteq N \llbracket y \mathbb{R}^e$$

If the transition map from \mathbb{R}^d to \mathbb{R}^e , that is $y \circ \phi \circ x^{-1}$, is C^∞ , then we say that the maps are C^∞ .

Now, we can define what a diffeomorphism is:

Definition 23 (Diffeomorphisms). Two C^∞ manifolds $(M, \mathcal{O}_M, \mathcal{A}_M)$, $(N, \mathcal{O}_N, \mathcal{A}_N)$ are said to be topologically diffeomorphic if there is a bijection $\phi : M \rightarrow N$ such that ϕ and ϕ^{-1} are both C^∞ maps.

The idea here is that two diffeomorphic manifolds are somehow equivalent. They don't really have a shape yet, but we can guarantee at this point that they are smooth. A spear and a soccer ball are not diffeomorphic, but a soccer ball and any other smooth shape you can think of are diffeomorphic, and they are guaranteed to be diffeomorphic because if we find one C^1 map, then the theorem shows that we can remove charts from the atlas until the atlas is an \mathcal{A}_∞ atlas.

One last spooky fact is that there is uncountably finitely many number of C^∞ -manifolds that one can make out of a given C^0 -manifold (if any) (up to diffeomorphism).

5 Tangent Spaces

The question that leads this section is: “what is the velocity of a curve γ at a point p ”? We shall define it.

5.1 Velocity

We shall begin with a straight auxiliary definition.

Definition 24. Let $(C^\infty, \oplus, \odot)$ be a vector space, where

$$C^\infty := \{f : M \rightarrow \mathbb{R} \mid f \text{ is smooth}\}$$

equipped with

- i. $(f \oplus g)(p) := f(p) + g(p)$
- ii. $(\lambda \odot g)(p) := \lambda \cdot g(p)$

Now we can properly define a velocity:

Definition 25 (Velocity). Consider a smooth manifold $(M, \mathcal{O}, \mathcal{A})$, a curve $\gamma : \mathbb{R} \longrightarrow M$ at least C^1 , and suppose $\gamma(\lambda_0) = p$. The velocity of γ at p is the linear map

$$v_{\gamma,p} : C^\infty(M) \llbracket \sim \mathbb{R}$$

and this velocity is defined as

$$v_{\gamma,p}(f) := (f \circ \gamma)'(\lambda_0)$$

It is not straightforward to define the velocity as a linear map that depends on a smooth function f , but the intuition is not far: imagine that you are running along a path γ in the real world (which in our models is described as the smooth manifold), and you experience a temperature around you, which is described by this smooth function f that takes your position in the manifold and spits a real number (which is the measurement of the temperature at that location). The velocity is then how this temperature in the world is changing at that particular point p . The definition of the velocity is hence the directional derivative.

5.2 Tangent vector space

Definition 26 (Tangent Space). For each point $p \in M$ we can define the set “tangent space to the point p ” as

$$T_p M := \{v_{\gamma,p} | \gamma \text{ smooth curves}\}$$

Remark 4. $T_p M$ can be made into a vector space.

Proof. The strategy here is to define \oplus and \odot such that when we take two vectors from $T_p M$ and add them together, the result is still in $T_p M$, and the same goes for multiplication. Since the velocities are linear maps from the (supposedly) vector space into \mathbb{R} , we say that their addition and multiplication is part of $\text{Hom}(C^\infty, \mathbb{R})$. Then, we define addition as

$$\begin{aligned} \oplus : T_p M \times T_p M &\longrightarrow \text{Hom}(C^\infty, \mathbb{R}) \\ (v_{\gamma,p} \oplus v_{\delta,p})(f) &:= v_{\gamma,p}(f) + v_{\delta,p}(f), \end{aligned}$$

and s-multiplication as

$$\begin{aligned} \odot : \mathbb{R} \times T_p M &\longrightarrow \text{Hom}(C^\infty, \mathbb{R}) \\ (\alpha \odot v_{\gamma,p})(f) &:= \alpha \cdot v_{\gamma,p}(f). \end{aligned}$$

Now we need to show that:

- i. $\exists \tau$ curve such that $\alpha \odot v_{\gamma,p} = v_{\tau,p}$.
- ii. $\exists \sigma$ curve such that $v_{\gamma,p} \oplus v_{\gamma,p} = v_{\sigma,p}$

We begin with the first item on the list by defining a curve

$$\begin{aligned} \tau : \mathbb{R} &\longrightarrow M \\ \lambda &\longmapsto \tau(\lambda) := \gamma(\alpha\lambda + \lambda_0) = (\gamma \circ \mu_\alpha)(\lambda), \end{aligned}$$

where $\mu_\alpha : \mathbb{R} \rightarrow \mathbb{R}$ such that $r \mapsto \alpha \cdot r + \lambda_0$, and we claim that this curve does the trick. Now,

$$\tau(0) = \gamma(\lambda_0) = p$$

so

$$v_{\tau,p} := (f \circ \tau)'(0) = (f \circ \gamma \circ \mu_\alpha)'(0) = \alpha(f \circ \gamma)'(\lambda_0) = \alpha \cdot v_{\gamma,p}.$$

Now, for the sum, we have to commit a sin: we pick a chart (\mathcal{U}, x) and apply it to the two curves we define: our γ that passes through point p at parameter value λ_0 , and another δ that we say passes through point p at parameter value λ_1 . We claim that the curve

$$\sigma_x(\lambda) := x^{-1}((x \circ \gamma)(\lambda_0 + \lambda) + (x \circ \delta)(\lambda_1 + \lambda) - (x \circ \gamma)(\lambda_0))$$

does the trick. What we do is: we calculate the values of the curves γ and δ at the point p and return it to the manifold via x^{-1} , so we are applying a chart, but we return from this chart. Of course this is still ill-defined since it is chart-dependent, but we'll see in the end that the result is chart-independent. The first step is to show that the curve σ_x passes through point p , which we assume that occurs at $\lambda = 0$:

$$\begin{aligned} \sigma_x(0) &= x^{-1}((x \circ \gamma)(\lambda_0) + (x \circ \delta)(\lambda_1) - (x \circ \gamma)(\lambda_0)) \\ &= x^{-1}((x \circ \gamma)(\lambda_1)) \\ &= \gamma(\lambda_1) = p. \end{aligned}$$

Now, we calculate the value of $v_{\sigma_x,p}$ and check if it is indeed equal to $v_{\gamma,p} + v_{\delta,p}$:

$$\begin{aligned} v_{\sigma_x,p}(f) &:= (f \circ \sigma_x)'(0) \\ &= ((f \circ x^{-1}) \circ (x \circ \sigma_x))'(0). \end{aligned}$$

In this second line, we can insert the identity map to make things easier. Now, since $f \circ x^{-1} : \mathbb{R}^d \rightarrow \mathbb{R}$ and $x \circ \sigma_x : \mathbb{R} \rightarrow \mathbb{R}^d$, we use the multidimensional chain rule to assess that:

$$v_{\sigma_x,p}(f) = (x^i \circ \sigma_x)'(0) \cdot (\partial_i(f \circ x^{-1}))(x(\sigma_x(0))).$$

Notation makes things a little convoluted, but what is happening here is that we are taking the inner derivative first (the term with x^i in it) and evaluating at the point 0, and we have to do it component-wise because the entries of x are in \mathbb{R}^d , and then we are taking the outer derivative (the term with x^{-1} in it), but since the domain is in \mathbb{R}^d , we have to take the partial derivative with respect to the i -th component, and then we evaluate at the point $x(\sigma_x(0)) = x(p)$, the chart-valued p . Now, plugging the definition of σ_x in the expression above yields:

$$\begin{aligned} v_{\sigma_x,p}(f) &= (x \circ x^{-1}((x \circ \gamma)(\lambda_0 + \lambda) + (x \circ \delta)(\lambda_1 + \lambda) - (x \circ \gamma)(\lambda_0)))^{i'} \cdot (\partial_i(f \circ x^{-1}))(x(p)) \\ &= ((x \circ \gamma)^{i'}(\lambda_0) + (x \circ \delta)^{i'}(\lambda_1)) \cdot (\partial_i(f \circ x^{-1}))(x(p)) \\ &= (\partial_i(f \circ x^{-1}))(x(p)) \cdot (x \circ \gamma)^{i'}(\lambda_0) + (\partial_i(f \circ x^{-1}))(x(p)) \cdot (x \circ \delta)^{i'}(\lambda_1) \\ &= (f \circ x^{-1} \circ x \circ \gamma)'(\lambda_0) + (f \circ x^{-1} \circ x \circ \delta)'(\lambda_1) \\ &= (f \circ \gamma)'(\lambda_0) + (f \circ \delta)'(\lambda_1) \\ &= v_{\gamma,p}(f) + v_{\delta,p}(f) \end{aligned}$$

as we wished to proof, and there is no dependency on the chart. ■

This proof is not just a mathematical throw up, this is actually to be used within the calculations of general relativity.

5.2.1 Components of a vector with respect to a chart

To know the components of a vector, we need to know the basis of the tangent space. Let $(\mathcal{U}, x) \in \mathcal{A}_{\text{smooth}}$. Let $\gamma : \mathbb{R} \rightarrow \mathcal{U}$, and $\gamma(0) = p$. Calculate $v_{\gamma,p} := (f \circ \gamma)'(0)$:

$$\begin{aligned} v_{\gamma,p}(f) &:= (f \circ \gamma)'(0) \\ &= ((f \circ x^{-1}) \circ (x \circ \gamma))'(0) \\ &= (x^i \circ \gamma)'(0) \cdot (\partial_i(f \circ x^{-1}))(x(p)) \end{aligned}$$

The first thing to notice is that the index on ∂_i tells us which entry to derive by. That is, it makes no reference whatsoever to x , but simply says “what ever the i -th entry is, derive by that.”² Now this is a lot of writing and so we introduce some new notation in order to simplify it: we define:

$$\left(\frac{\partial f}{\partial x^i} \right)_p := \partial_i(f \circ x^{-1})(x(p)), \quad \dot{\gamma}_x^i(0) := (x^i \circ \gamma)'(0),$$

so we can write the velocity as

$$v_{\gamma,p}f = \dot{\gamma}_x^i(0) \cdot \left(\frac{\partial}{\partial x^i} \right)_p f$$

and, as a map, it is

$$v_{\gamma,p} = \dot{\gamma}_x^i(0) \cdot \left(\frac{\partial}{\partial x^i} \right)_p.$$

Definition 27 (Chart-induced basis). We call

$$\left(\frac{\partial}{\partial x^i} \right)_p$$

the i -th element of the basis of T_pM .

Definition 28 (Components w.r.t. chart induced basis). We call

$$\dot{\gamma}_x^i(0)$$

the i -th component of a vector in T_pM .

Notice that these elements are all dependent of the chart (\mathcal{U}, x) chosen at the beginning, so we say that this basis is a chart-induced basis of T_pM . However, we need to show that the elements $\partial/\partial x^i$ are linearly independent.

².This is analogous to the fact that given $f : \mathbb{R} \rightarrow \mathbb{R}$ we define $f' : \mathbb{R} \rightarrow \mathbb{R}$ completely independently of what variable we're using. So $\frac{df}{dx}$ is not a general expression, but is a notation choice once we have decided that x is our variable.

Proof. We need to proof that

$$\lambda_i \left(\frac{\partial}{\partial x^i} \right)_p = 0 \iff \lambda_i = 0 \quad \forall i \in \{1, \dots, d\}.$$

The trick here is to apply it to a function that we choose to be x^j , then

$$\begin{aligned} \lambda^i \left(\frac{\partial}{\partial x^i} \right)_p (x^j) &= \lambda_i \partial_i (x^j \circ x^i)'(x(p)) \\ &= \lambda^i \delta_i^j = \lambda^j = 0 \end{aligned}$$

So it means that every scalar λ_j yields zero, henceforth, these elements are linearly independent. ■

Corollary 1. $\dim T_p M = d = \dim M$.

The last thing to point out in this subsection is some terminology. If $X \in T_p M : \exists \gamma : \mathbb{R} \rightarrow \mathbb{R} : X = v_{\gamma, p}$, and also $\exists (X^1, \dots, X^d) \in \mathbb{R} : X = X^i \partial / \partial x^i$. This is a generalization of the chart-induced basis: we can choose other basis, as well, and a chart-induced basis is not the only way to do it. We can perfectly create a basis without picking a chart.

5.2.2 Change of vector components under a change of chart

A vector does not change under a change of chart simply because it is defined independently of a chart. What changes are the *components* of a vector. The coordinate system we choose to look at our physical problem does not change the physics of the problem. All these concepts tight together with our initial idea that both charts and bases are just fantasies of the real world (or our description of real world): they help us to make calculations because they have more structure and more tools to work with, but the “reality” is independent of these fantasies.

Let (\mathcal{U}, x) and (\mathcal{V}, y) be overlapping charts and $p \in \mathcal{U} \cap \mathcal{V}$. Let $X \in T_p M$, then we can write this vector using the chart x and, independently, write this vector using chart y :

$$X_{(y)}^j \left(\frac{\partial}{\partial y^j} \right)_p = X = X_{(x)}^i \left(\frac{\partial}{\partial x^i} \right)_p \quad (5.1)$$

The question is: how these two relate to one another? (and we know they do, since the object X in the middle is invariant, so there needs to be a way to go from the left-most side to the right-most side). In order to study this change of components formula, consider:

$$\begin{aligned} \left(\frac{\partial}{\partial x^i} \right)_p f &:= \partial_i (f \circ x^{-1})(x(p)) \\ &= \partial_i ((f \circ y^{-1}) \circ (y \circ x^{-1}))(x(p)) \\ &= \partial_i (y^j \circ x^{-1})(x(p)) \cdot (\partial_j (f \circ y^{-1}))(y(p)) \\ &=: \left(\frac{\partial y^j}{\partial x^i} \right)_p \cdot \left(\frac{\partial}{\partial y^j} \right)_p f \end{aligned}$$

Hence, we have that

$$X_{(x)}^i \left(\frac{\partial y^j}{\partial x^i} \right)_p \cdot \left(\frac{\partial}{\partial y^j} \right)_p = X_{(y)}^j \left(\frac{\partial}{\partial y^j} \right)_p,$$

which yields, plugging back into Eq. (??), to

$$\left[X_{(x)}^i \left(\frac{\partial y^j}{\partial x^i} \right)_p - X_{(y)}^j \right] \left(\frac{\partial}{\partial y^j} \right)_p = 0.$$

Remember that the term outside of the square brackets is a base to the vector space induced by chart (\mathcal{V}, y) , so everything in square brackets is zero, which finally yields

$$X_{(y)}^j = X_{(x)}^i \left(\frac{\partial y^j}{\partial x^i} \right)_p. \quad (5.2)$$

These are the components of the same vector (real object) when we change from one chart (fantasy) to another.

5.3 Cotangent spaces

It is pretty trivial, with all that has been presented, that we consider the dual space of the $T_p M$, which is denoted as $T_p^* M$:

Definition 29 (Cotangent space). Let $T_p M$ be the tangent space to the point $p \in M$. The dual of this space is called the cotangent space

$$T_p^* M := \{ \phi : T_p M \rightarrow \mathbb{R} \}$$

Example 14. Let $f \in C^\infty(M)$, we can define

$$(\mathrm{d}f)_p : T_p M \rightarrow \mathbb{R}$$

such that

$$X \mapsto (\mathrm{d}f)_p(X) := Xf \in \mathbb{R}.$$

which means that $(\mathrm{d}f)_p \in T_p^* M$. The map $(\mathrm{d}f)_p$ is called the gradient of f at the point $p \in M$. We can calculate the components of this gradient. To calculate the components of a $(0, 1)$ -tensor, we do as we defined earlier, which is to apply the tensor in the coordinates of the cotangent space:

$$\begin{aligned} ((\mathrm{d}f)_p)_j &:= (\mathrm{d}f)_p \left(\left(\frac{\partial}{\partial x^j} \right)_p \right) \\ &= \left(\frac{\partial f}{\partial x^j} \right)_p \\ &= \partial_j (f \circ x^{-1})(x(p)). \end{aligned}$$

□

Theorem 4. Consider a chart (\mathcal{U}, x) and chart maps $x^i : \mathcal{U} \rightarrow \mathbb{R}$. Then,

$$\{(\mathrm{d}x^1)_p, \dots, (\mathrm{d}x^d)_p\}$$

is the dual basis of the dual space.

Proof. Simply apply the gradient to the elements of the basis of $T_p M$ and we get

$$(\mathrm{d}x^a)_p \left(\left(\frac{\partial}{\partial x^b} \right)_p \right) = \left(\frac{\partial x^a}{\partial x^b} \right)_p = \delta_b^a,$$

which is precisely what we want. ■

5.3.1 Change of components of a covector under a change of chart

The principle is really similar to the change of coordinates of a cotangent space once we change the chart. For example, let $\omega \in T_p^*M$, then

$$\omega_{(y)j}(\mathrm{d}y^j)_p = \omega = \omega_{(x)i}(\mathrm{d}x^i)_p,$$

and the rule we need to use is

$$\omega_{(y)j} = \left(\frac{\partial x^i}{\partial y^j} \right)_p \omega_{(x)i}. \tag{5.3}$$