General Relativity Fall 2019
Lecture 5: manifolds and tangent vectors

Yacine Ali-Haïmoud
September 17th 2019

MANIFOLDS

In simple words, a manifold is a **set that locally "looks like"** $\mathbb{R}^n$. A little more precisely, it is a set made of patches that look like $\mathbb{R}^n$, smoothly sewn together. Before giving the formal mathematical definition, we need to define a few preliminary concepts:

• An **open ball** of $\mathbb{R}^n$ around a point $y \in \mathbb{R}^n$, with radius r, is the set $\mathcal{O}(y; r) \equiv \{x \in \mathbb{R}^n$ such that $|x - y| < r\}$.
• An **open set** of $\mathbb{R}^n$ is the union of an arbitrary number of open balls: $U \subset \mathbb{R}^n$ is open if $\forall\ y \in U, \exists\ r > 0$ such that $\mathcal{O}(y; r) \subset U$.

• A map $\phi : A \to B$ is said to be **one-to-one** or **injective** (these are interchangeable) is no two elements of A have the same image: $\phi(x) = \phi(y)$ if *and only if* $x = y$.
• A map $\phi : A \to B$ is said to be **onto** or **surjective** if all the elements of B can be written as the image of (at least) one element of A: $\forall\ y \in B, \exists\ x \in A$ such that $y = \phi(x)$.
• A map $\phi : A \to B$ is **bijective** if it is both **injective and surjective**, in other words, if it is **invertible**. We denote the inverse map $\phi^{-1} : B \to A$. Note that an injective map $\phi : A \to B$ is always bijective when viewing it as a map from A to $\phi(A) = \{\phi(x)$ such that $x \in A\}$.

Let us now consider a general set $\mathcal{M}$ which is not $\mathbb{R}^n$.

• A **chart** or **coordinate system** on a subset $U \subset \mathcal{M}$ is a **one-to-one** map $\phi : U \to \mathbb{R}^n$ such that the image set, $\phi(U) \subset \mathbb{R}^n$ is **open**. The restricted map $\phi : U \to \phi(U)$ is thus bijective, and we call $\phi^{-1} : \phi(U) \subset \mathbb{R}^n \to \mathcal{M}$ its inverse..
• An **atlas** on a set $\mathcal{M}$ is a collection of charts and associated domains, $\{(U_i, \phi_i)\}$, such that:
(*i*) The domains U_i **cover** $\mathcal{M}$, i.e. $\underset{i}{\text{Union}}(U_i) = \mathcal{M}$,
(*ii*) Overlapping charts are compatible, i.e. if two charts have overlapping domains $U_1 \cap U_2 \neq \emptyset$, then the maps $\phi_2 \circ \phi_1^{-1}$ and $\phi_1 \circ \phi_2^{-1}$, both from $\mathbb{R}^n$ to $\mathbb{R}^n$, are **infinitely differentiable** (symbol: $\mathcal{C}^\infty$), see Fig. 1.

This allows us to finally define a **manifold**, as a set $\mathcal{M}$ with a **maximal atlas**, i.e. one that contains every possible compatible chart. Note that n (the dimension of $\mathbb{R}^n$) must be the **same everywhere** on the manifold.

Examples of sets which are not manifolds (see Fig. 2): a 1-dimensional line attached to a 2-dimensional surface, as this would require 2 different dimensions. A self-intersecting line, which would not "look like $\mathbb{R}^n$ at the intersection point.

You can think of manifolds as smooth $n-$dimensional surfaces of $\mathbb{R}^p$, with $p > n$. Indeed, **Whitney's embedding theorem** states that a manifold can always be mapped to a smooth n-dimensional surface of $\mathbb{R}^{2n}$ ("embedded" in $\mathbb{R}^{2n}$). This is not necessarily helpful, however, unless you have good visualisation skills in 4 dimensions or higher. Moreover, a manifold has well-defined **intrinsic** properties regardless of its embedding.

In general, a manifold cannot be covered with a single chart. Consider for instance a circle: you can describe all of it *except for a point* by an angle $\theta \in (0, 2\pi)$. To cover the full circle, you need at least 2 charts. Note that you cannot just use $[0, 2\pi)$, as this would not be an open set of $\mathbb{R}$. The same holds for a shpere.

We know what a $\mathcal{C}^\infty$ function of $\mathbb{R}^n \to \mathbb{R}$, but it is not a priori obvious how to define this notion for functions $f : \mathcal{M} \to \mathbb{R}$. We can do so as follows: at any given point $p \in \mathcal{M}$, find a chart $\{(U, \phi)\}$ such that $p \in U$ – such a chart must exist from condition (*i*). We can then define the function $f \circ \phi^{-1} : \mathbb{R}^n \to \mathbb{R}$, see Fig. 3. The function f is said to be $\mathcal{C}^\infty$ (or **smooth**) if these functions are all smooth – because of property (*ii*), the choice of ϕ does not matter in the definition.
I will denote by $\mathcal{F}$ the set of all smooth functions from $\mathcal{M} \to \mathbb{R}$ (or, more generally, form a subset of $\mathcal{M}$ to $\mathbb{R}$):

$$\boxed{\mathcal{F} \equiv \{f : \mathcal{M} \to \mathbb{R} \text{ such that } f \text{ is } \mathcal{C}^\infty\}}.$$

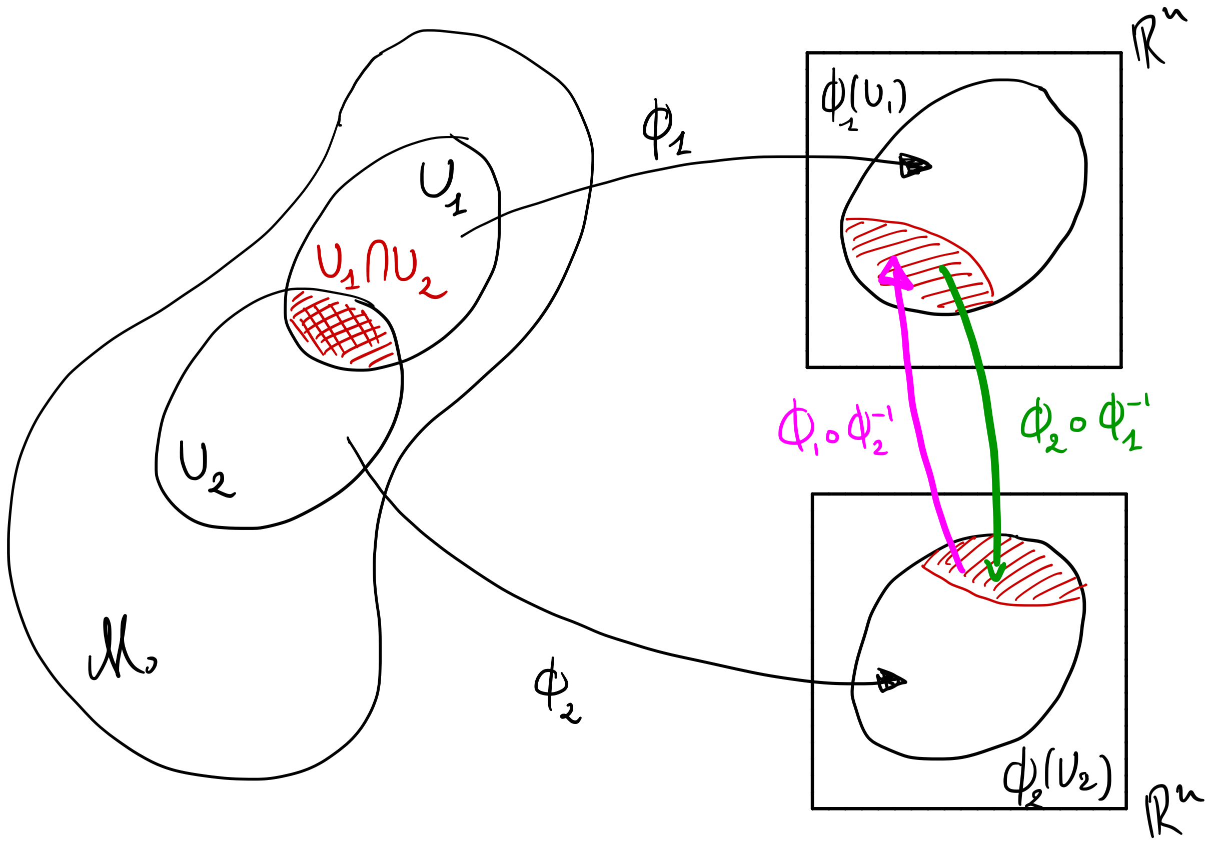

FIG. 1. Domains of various maps used to defined compatiblity of overlapping charts.

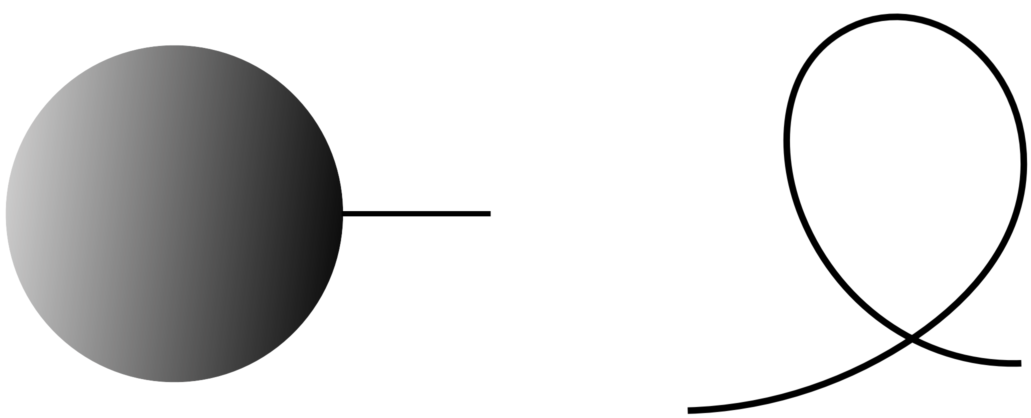

FIG. 2. Example of sets which are *not* manifolds: a line attached to a sphere, or a self-intersecting line.

TANGENT VECTORS

When we think of $\mathbb{R}^n$, we often blur the line between points and vectors: after all, vectors are differences between two points, so we can identify a point and a vector attached to the origin of coordinates.

On a general manifold $\mathcal{M}$, there is no such notion of addition or subtraction of points. Intuitively, we can easily visualize that vectors live in a space **tangent** to $\mathcal{M}$, at any given point $p \in \mathcal{M}$. In $\mathbb{R}^n$, the tangent spaces at each point are unambiguously mapped onto one another, and onto the original space $\mathbb{R}^n$ itself. On a general curved surface, tangent spaces at each point are not parallel, and there is **no obvious mapping between tangent spaces at distant points** – consider for instance the tangent plane at the pole of a sphere, and that at one point on the equator.

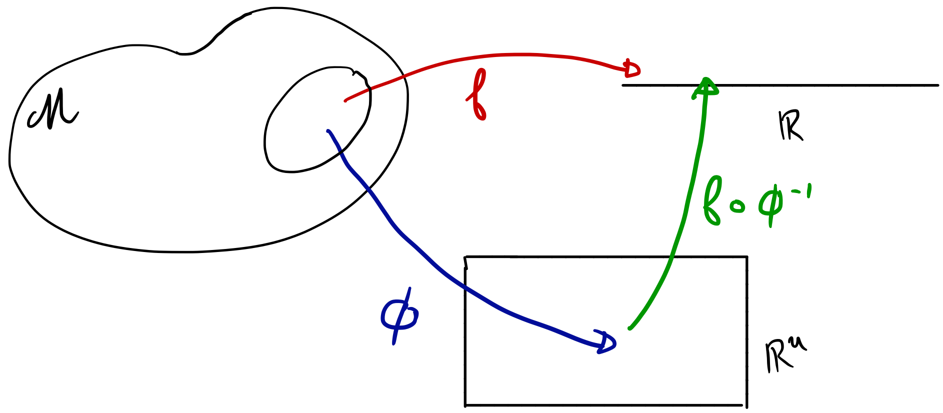

FIG. 3. Domains of different maps used to defined a smooth function on $\mathcal{M}$.

Tangent vectors convey the notion of **infinitesimal displacements along a given direction**, at a given point p. For instance, you are probably familiar with the unit vectors $\hat{e}_r, \hat{e}_\theta, \hat{e}_\varphi$ in spherical polar coordinates: they indicate the direction towards which a point $p(r, \theta, \varphi)$ would move upon an infinitesimal increase of a given coordinate. This infinitesimal business means that vectors have something to do with derivatives, and indeed, their formal definition should remind you of derivatives.

Definition: a tangent vector $\overline{V}$ at a point $p \in \mathcal{M}$ is a an *operator* $\overline{V} : \mathcal{F} \to \mathbb{R}$ which
(i) is **linear**: for any two smooth functions f, g, and any two real numbers α, β, $\overline{V}(\alpha f + \beta g) = \alpha \overline{V}(f) + \beta \overline{V}(g)$,
(ii) satisfies **Leibniz' rule**: $\overline{V}(f \times g) = f(p) \times \overline{V}(g) + g(p) \times \overline{V}(f)$.
I will denote by **$\mathcal{V}_p$ the set of all tangent vectors at a point $p \in \mathcal{M}$**. This forms a **vector space** in the mathematical sense of the term.

In the **homework**, you will show that any vector $\overline{V}$ acting on a constant function gives zero: $\overline{V}(\text{constant}) = 0$.

A simple (and important) example of a vector at p is as follows: pick a coordinate system at p, i.e. a map $\phi : U \subset \mathcal{M} \to \mathbb{R}^n$. For any point $q \in u$, $\phi(q) = \{x^1, ..., x^n\} \in \mathbb{R}^n$, where the n real numbers $\{x^\mu\}$ are the **coordinates of q**. Then define the following operator on $\mathcal{F}$:

$$\partial_{(\mu)}|_p : \begin{cases} \mathcal{F} & \to \mathbb{R} \\ f & \mapsto \frac{\partial (f \circ \phi^{-1})}{\partial x^\mu}(\phi(p)). \end{cases}$$

This operator is linear, and satisfies Leibniz' rule, hence is a tangent vector at p. From now on I will just write $\partial_{(\mu)}$ rather than $\partial_{(\mu)}|_p$, but it should be implicit that this is defined at a specific point p. Rather than writing all the ϕ's, we will often write, for short,

$$\partial_{(\mu)} f = \frac{\partial f}{\partial x^\mu}.$$

COORDINATE BASES

Let us now show that the n vectors $\partial_{(\mu)}$ form a **basis** of the tangent space $\mathcal{V}_p$, meaning that (i) they are linearly independent and (ii) they span $\mathcal{V}_p$, i.e. any vector $\overline{V}$ can be written as $\overline{V} = V^\mu \partial_{(\mu)}$.

As a preiminary, let's define the n functions $\phi^\mu(q) \equiv x^\mu$, for q in the neighborhood of a point $p \in \mathcal{M}$. These functions are clearly smooth, and by construction, they are such that $\partial_{(\mu)} \phi^\nu = \delta^\nu_\mu$.

Let us start by proving the linear independence of the $\partial_{(\mu)}$'s. Suppose there exist n coefficients $\{\alpha^\mu\}$ such that $\alpha^\mu \partial_{(\mu)} = 0$, meaning that, for any function $f \in \mathcal{F}, \alpha^\mu \partial_{(\mu)} f = 0$. Applying this to any of the f^ν functions, we see that all coefficients α^ν must be zero, which proves linear independence.

Now, given a fucntion $f \in \mathcal{F}$, consider the function $F(x) \equiv f \circ \phi^{-1} : \mathbb{R}^n \to \mathbb{R}$. This function is $\mathcal{C}^\infty$, thus there exists $\mathcal{C}^\infty$ functions $H_\mu(x)$ such that

$$F(x) = \text{constant} + (x^\mu - a^\mu) H_\mu(x).$$

If we wanted, we could explicitly write the full function $H_\mu(x)$ in terms of the derivatives of F, but that's not needed. All we need to know is that $H_\mu(a) = \frac{\partial F}{\partial x^\mu}(a)$.

Let's apply this at $a = \phi(p)$ and $x = \phi(q)$, define $h_\mu(q) \equiv H_\mu(\phi(q))$, and recall that $x^\mu = \phi^\mu(q)$: we get

$$f(q) = \text{constant} + (\phi^\mu(q) - \phi^\mu(p)) \times h_\mu(q),$$

which we can also write as an equality between functions:

$$f = \text{constant} + (\phi^\mu - \phi^\mu(p)) \times h_\mu.$$

Let us now apply a vector $\overline{V} \in \mathcal{V}_p$ to this function, using linearity and Leibniz' rule:

$$\overline{V}(f) = \overline{V}(\text{constant}) + (\phi^\mu(p) - \phi^\mu(p)) \times \overline{V}(h_\mu) + h_\mu(p) \times \overline{V}(f^\mu) = h_\mu(p) \times \overline{V}(\phi^\mu).$$

On the other hand, we have

$$h_\mu(p) = H_\mu(a) = \frac{\partial(f \circ \phi^{-1})}{\partial x^\mu}(\phi(p)) = \partial_{(\mu)}|_p f.$$

Thus we found that

$$\overline{V}(f) = V^\mu \partial_\mu f, \qquad V^\mu \equiv \overline{V}(\phi^\mu) = \overline{V}(x^\mu) \qquad \Rightarrow \overline{V} = V^\mu \partial_{(\mu)}.$$

In other words, we have showed that any tangent vector can be written as a linear combination of the partial derivative operators. **The components V^μ are obtained by applying $\overline{V}$ to the coordinate functions x^μ.**

In summary, we have exhibited a basis of $\mathcal{V}_p$: the n vectors $\{\partial_{(\mu)}\}$. This is called a **coordinate basis** of the tangent space. The fact that these n vectors form a basis implies that **the tangent space is n-dimensional**.