CHAPTER 4

Tangent spaces

So far, the fact that we dealt with general manifolds M and not just embedded submanifolds of Euclidean spaces did not pose any serious problem: we could make sense of everything right away, by moving (via charts) to opens inside Euclidean spaces.

However, when trying to make sense of "tangent vectors", the situation is quite different. Indeed, already intuitively, when trying to draw a "tangent space" to a manifold M, we are tempted to look at M as part of a bigger (Euclidean) space and draw planes that are "tangent" (in the intuitive sense) to M. The fact that *the tangent spaces of M are intrinsic to M and not to the way it sits inside a bigger ambient space* is remarkable and may seem counterintuitive at first. And it shows that the notion of tangent vector is indeed of a geometric, intrinsic nature. Given our preconception (due to our intuition), the actual general definition may look a bit abstract. For that reason, we describe several different (but equivalent) approaches to tangent spaces. Each one of them starts with one intuitive perception of tangent vectors on $\mathbb{R}^m$, and then proceeds by concentrating on the main properties (to be taken as axioms) of the intuitive perception.

However, before proceeding, it is useful to clearly state what we expect:

- **tangent spaces**: for M-manifold, $p \in M$, have a vector space T_pM. And, of course, for embedded submanifolds $M \subset \mathbb{R}^{\tilde{m}}$, these tangent spaces should be (canonically) isomorphic to the previously defined tangent spaces T_pM.
- **differentials**: for smooths map $F : M \to N$ (between manifolds), $p \in M$, have an induced linear map
$$(dF)_p : T_pM \to T_{F(p)}N,$$
called the differential of F at p. For the differential we expect, like inside $\mathbb{R}^m$:
  - $F = \mathrm{Id}_M : M \to M$ is the identity map, $(dF)_p$ should be the identity map of T_pM.
  - **The chain rule**: for $M \overset{F}{\to} N \overset{G}{\to} P$ smooth maps between manifold and $p \in M$,
$$(d\, G \circ F)_p = (dG)_{F(p)} \circ (dF)_p.$$
- **nothing new in** $\mathbb{R}^m$: Of course, for embedded submanifolds $M \subset \mathbb{R}^{\tilde{m}}$, these tangent spaces should be (canonically) isomorphic to the previously defined tangent spaces T_pM and dF should become the usual differential.

We also expect that embeddings $N \hookrightarrow M$ induce injective maps $T_pN \hookrightarrow T_pM$, but this will turn out to be the case once one imposes the milder condition that, for any open $U \subset M$, the resulting maps $T_pU \to T_pM$ are isomorphisms.

Exercise 4.1. Show that, whatever construction of the tangent spaces T_pM we give so that it satisfies the previous properties, for any diffeomorphism $F : M \to N$ and any $p \in M$, the differential of F at p,
$$(dF)_p : T_pM \to T_{F(p)}N,$$
is a linear isomorphism. Deduce that for any m-dimensional manifold M, T_pM is an m-dimensional vector space.

However, even before these properties, one should keep in mind that the way we should think intuitively about tangent spaces is:

T_pM is made of speeds (at $t = 0$) of curves γ in M passing through p at $t = 0$.

This is clear and precise in the case when $M \subset \mathbb{R}^{\tilde{m}}$, where the embedding was used to make sense of "speeds". For conciseness, we introduce the notation

$$\mathbf{Curves}_p(M) := \{\gamma : (-\epsilon, \epsilon) \to M \textbf{ smooth, with } \epsilon > 0, \gamma(0) = p\}$$

(valid for all manifold M and $p \in M$). Hence, in general, we still want that

$$T_pM = \{\text{"}\frac{d\gamma}{dt}(0)\text{"} : \gamma \in \mathrm{Curves}_p(M)\}, \tag{0.1}$$

the only problem being that "$\frac{d\gamma}{dt}(0)$" does not make sense yet.

One way to proceed is by "brute force", starting from the following remark: although "the speed of γ" (at $t = 0$) is not defined yet, we can make sense right away of the property that two such curves $\gamma_1, \gamma_2 \in \mathrm{Curves}_p(M)$ have the same speed at $t = 0$: if their representations γ_1^χ and γ_2^χ with respect to a/any chart χ around p have this property:

$$\frac{d\gamma_1^\chi}{dt}(0) = \frac{d\gamma_2^\chi}{dt}(0).$$

(recall that, for $\chi : U \to \Omega \subset \mathbb{R}^m$, $\gamma_i^\chi = \chi \circ \gamma_i$ is a curve in $\Omega \subset \mathbb{R}^m$). When this happens, we write

$$\gamma_1 \sim_p \gamma_2.$$

Exercise 4.2. Explain the "a/any" in the discussion above and show that $\sim_p$ is an equivalence relation.

The "brute force" approach to tangent spaces would then be to define:

- T_pM as the quotient of $\mathrm{Curves}_p(M)$ modulo the equivalence relation $\sim_p$.
- the speed $\frac{d\gamma}{dt}(0)$ as the $\sim_p$ equivalence class of γ.

Then equality (0.1) that corresponds to our intuition is forced in tautologically. Although this does work, and it is sometimes taken as definition in some text-books/lecture notes, we find it too abstract for such an intuitive concept. Instead, we will discuss the tangent space via charts, as well as via derivations. Nevertheless, one should always keep in mind the intuition via speeds of paths (and make it precise in whatever model we use).

1. Tangent vectors via charts

The brief philosophy of this approach is:

while a chart χ for M allows one to represent points p by coordinates, it will allow us to represent tangent vectors $v \in T_pM$ by vectors in $\mathbb{R}^m$

Or, concentrating on the tangent vectors rather than on the charts:

a tangent vector to M at p can be seen through each chart, as a vector in the Euclidean space $\mathbb{R}^m$

The actual definition can be discovered as the outcome of "wishful thinking": we want to extend the definition of T_pM from the case of embedded submanifolds of Euclidean spaces to arbitrary manifolds M, so that the properties that we have seen inside Euclidean spaces continue to hold (such as the chain rule- see the general comments itemized above). We see that, for a general manifold M and $p \in M$, choosing a chart (U, χ) for M around p and viewing χ as a smooth map from (an open inside) M to (an open inside) $\mathbb{R}^m$, the wishful thinking tells us that we expect an induced map (even an isomorphism)

$$(d\chi)_p : T_pM \to \mathbb{R}^m.$$

Therefore, an element $v \in T_pM$ can be represented w.r.t. to the chart χ by a vector in the standard Euclidean space:

$$v^\chi := (d\chi)_p(v) \in \mathbb{R}^m.$$

What happens if we change the chart χ by another chart χ' around p? Then $v^{\chi'}$ should be changed in a way that is dictated by the change of coordinates

$$c = c_{\chi,\chi'} = \chi' \circ \chi^{-1}.$$

Indeed, since $\chi' = c \circ \chi$ and we want the chain rule still to hold, we would have:

$$v^{\chi'} = (d\chi')_p(v) = (d\,c)_{\chi(p)}\left((d\chi)_p(v)\right),$$

i.e.

$$v^{\chi'} = (d\,c)_{\chi(p)}(v^\chi). \tag{1.1}$$

Therefore, whatever the meaning of T_pM is, we expect that the elements $v \in T_pM$ can be represented w.r.t local charts χ (around p) by vectors $v^\chi \in \mathbb{R}^m$ which, when we change the chart, transform according to (1.1). Well, if that is what we want from T_pM, let us just define it that way!

Definition 4.1. *Given an m-dimensional manifold M and $p \in M$, a* ***tangent vector of*** *M* ***at*** *p is any function*

$$v : \{\textit{charts of } M \textit{ around } p\} \to \mathbb{R}^m, \quad \chi \mapsto v^\chi,$$

with the property that, for any two charts χ and χ' one has

$$v^{\chi'} = (d\,c)_{\chi(p)}(v^\chi), \tag{1.2}$$

where $c = c_{\chi,\chi'}$ is the change of coordinates from χ to χ'.

We denote by T_pM the vector space of all such tangent vectors of M at p (a vector space using the vector space structure on $\mathbb{R}^m$, i.e. $(v+w)^\chi := v^\chi + w^\chi$, etc).

Let us explain, right away, that this model for the tangent space serves the original purpose: we can make sense of

$$\frac{d\gamma}{dt}(0) \in T_pM$$

for $\gamma \in \mathrm{Curves}_p(M)$. The actual definition should be clear: the representation of this vector with respect to a chart χ should be the speed (at $t = 0$) of the representation γ^χ of γ with respect to χ. In more detail: if $\chi : U \to \Omega$ is a chart of M around p, consider the resulting curve $\gamma^\chi = \chi \circ \gamma$ in $\Omega \subset \mathbb{R}^m$, and take its derivatives at 0- which is a vector in $\mathbb{R}^m$. This gives an operation

$$\{\text{charts of } M \text{ around } p\} \to \mathbb{R}^m, \quad \chi \mapsto \frac{d\gamma^\chi}{dt}(0),$$

which is clearly linear and satisfies the Leibniz identity; therefore we obtain a tangent vector denoted

$$\frac{d\gamma}{dt}(0) \in T_pM.$$

In a short formula,

$$\frac{d\gamma}{dt}(0)^\chi := \frac{d\gamma^\chi}{dt}(0).$$

Of course, there is nothing special for $t = 0$, and in the same way one talks about

$$\frac{d\gamma}{dt}(t) \in T_{\gamma(t)}M.$$

for all t in the domain of γ.

Staring for a minute at the definition of tangent vectors $v \in T_pM$, including (1.2), we see that once we know v^{χ_0} for one single chart around p, we completely know v. More precisely:

Lemma 4.2. *If χ_0 is a chart of M around p, then*

$$T_pM \to \mathbb{R}^m, \quad v \mapsto v^{\chi_0}$$

is an isomorphism of vector spaces.

Proof. As we have already remarked, given $v_0 \in \mathbb{R}^m$ then the $v \in T_pM$ such that $v^{\chi_0} = v_0$ is unique: it must be given by the formula

$$v^{\chi} = (d\, c_{\chi_0,\chi})_{\chi_0(p)}(v_0),$$

for any other chart χ around p. Strictly speaking we still prove (1.2) but that follows from the chain rule and the remark that $c_{\chi_0,\chi'} = c_{\chi,\chi'} \circ c_{\chi_0,\chi}$. □

Definition 4.3. *For a chart χ of M around p, we denote by*

$$\left(\frac{\partial}{\partial \chi_1}\right)_p, \ldots, \left(\frac{\partial}{\partial \chi_m}\right)_p \in T_pM$$

called ***the canonical basis of*** T_pM ***with respect to*** χ*, the basis that corresponds to the canonical one $e_1, \ldots, e_m$ of $\mathbb{R}^m$ via the isomorphism from the previous lemma.*

In other words,

$$\left(\frac{\partial}{\partial \chi_i}\right)_p \in T_pM$$

is the unique tangent vector at p with

$$\left(\frac{\partial}{\partial \chi_i}\right)_p^{\chi} = e_i.$$

Exercise 4.3. If χ and χ' are two charts of M around p, $c = c_{\chi,\chi'}$, show that

$$\left(\frac{\partial}{\partial \chi_i}\right)_p = \sum_j \frac{\partial c_j}{\partial x_i}(\chi(p)) \left(\frac{\partial}{\partial \chi'_j}\right)_p. \tag{1.3}$$

Finally, let us show show that, as we wanted, smooth maps induce differentials at the level of tangent spaces.

Lemma 4.4. *Given a smooth map $f : M \to N$ between two manifolds, and given $p \in M$, there exists a unique linear map*

$$(df)_p : T_pM \to T_{f(p)}N, \quad v \mapsto (df)_p(v)$$

with the property that, for any chart χ of M around p and χ' of N around $f(p)$, one has:

$$((df)_p(v))^{\chi'} = (d\, f_{\chi,\chi'})_{\chi(p)}(v^{\chi})$$

for all $v \in T_pM$.

Proof. Fixing v and χ, the required condition forces the definition of $(d\,f)_p(v)$; to see that one ends up with a tangent vector to N at $f(p)$ (i.e. the condition (1.2) is satisfied in this context) one uses again the chain rule. One still has to check that the resulting map does not depend on the choice of χ- but that follows again from the chain rule. □

DEFINITION 4.5. *For a smooth map $f : M \to N$ and $p \in M$, the resulting linear map $(df)_p : T_pM \to T_{f(p)}N$ is called **the differential of f at the point** p.*

It should be clear (or an easy exercise) that the chain rule continues to hold.

EXERCISE 4.4. Show that $(df)_p$ is uniquely characterized also by the condition that, for any $\gamma \in \mathrm{Curves}_p(M)$, one has:

$$(df)_p(\frac{d\gamma}{dt}(0)) = \frac{df \circ \gamma}{dt}(0).$$

EXERCISE 4.5. For any curve $\gamma \in \mathrm{Curves}_p(M)$ defined on the interval $I = (-\epsilon, \epsilon)$, viewed as a smooth map $\gamma : I \to M$ between manifolds, show that

$$(d\gamma)_0\left(\left(\frac{\partial}{\partial t}\right)_0\right) = \frac{d\gamma}{dt}(0) \quad \text{(an equality of vectors in } \in T_pM).$$

2. Tangent vectors as directional derivatives

Here is another approach to the tangent spaces; the brief philosophy of this approach is:

tangent vectors at p allow us to take directional derivatives (at p) of smooth functions

For this we return to the differential

$$(df)_p(v)$$

of a smooth function $f : \Omega \to \mathbb{R}$ ($\Omega \subset \mathbb{R}^m$ open) at a point $p \in \Omega$, applied to a (tangent) vector $v \in \mathbb{R}^m$, that we reinterpreted it as a v-derivative at p

$$(df)_p(v) = \frac{\partial f}{\partial v}(p) = \partial_v(f)(p).$$

Hence, if we want to understand what the (tangent) vector v really does (at p), one may say that it defines a function

$$\partial_v = \frac{\partial}{\partial v} : \mathcal{C}^\infty(\Omega) \to \mathbb{R};$$

What are the main properties of this function? Well, it is clearly linear and, moreover, it acts on the product of functions according to the Leibniz rule at p:

$$\partial_v(fg) = f(p)\partial_v(g) + g(p)\partial_v(g)$$

for all $f, g \in \mathcal{C}^\infty(\Omega)$. Note that here we are making use of the algebra structure on $C^\infty(\Omega)$: we make reference to the vector space structure (linearity) as well as to the multiplication (to make sense of the Leibniz identity at p).

The main point of this discussion is that, conversely, any linear map

$$\partial : \mathcal{C}^\infty(\Omega) \to \mathbb{R}$$

which is linear and satisfies the Leibniz identity at p is of type ∂_v for a unique vector $v \in \mathbb{R}^m$ (this will follow from the discussion below). This gives another perspective/possible approach to tangent spaces.

DEFINITION 4.6. *Given an m-dimensional manifold M and $p \in M$, a **derivation of M at** p is any linear map*

$$\partial : \mathcal{C}^\infty(M) \to \mathbb{R}$$

that is linear and satisfies the Leibniz identity at p.

We denote by $T_p^{\boldsymbol{deriv}}M$ the vector space of all such derivations M at p (a vector space using the usual addition and multiplication by scalars of linear maps).

EXAMPLE 4.7. When $M = \Omega$ is an open in $\mathbb{R}^m$ (endowed with the canonical smooth structure), $p \in \Omega$, then the operation of taking the usual partial derivatives at p,

$$\left(\frac{\partial}{\partial x_i}\right)_p : \mathcal{C}^\infty(U) \to \mathbb{R}, \quad f \mapsto \frac{\partial f}{\partial x_i}(p),$$

are derivations at p- hence they can be interpreted as vectors

$$\left(\frac{\partial}{\partial x_1}\right)_p, \ldots, \left(\frac{\partial}{\partial x_m}\right)_p \in T_p^{\text{deriv}}\Omega.$$

As it will follow from the discussion below, they form a basis of $T_p^{\text{deriv}}\Omega$.

Similarly, for any manyfold M, a chart $\chi : U \to \Omega \subset \mathbb{R}^m$ of M around p induces **partial derivatives at p w.r.t. the chart** χ,

$$\frac{\partial f}{\partial \chi_i}(p) := \frac{\partial f_\chi}{\partial x_i}(\chi(p)),$$

where $f_\chi = f \circ \chi^{-1} : \Omega \to \mathbb{R}$. And then vectors

$$\left(\frac{\partial}{\partial \chi_1}\right)_p, \ldots, \left(\frac{\partial}{\partial \chi_m}\right)_p \in T_p^{\text{deriv}}M$$

which form a basis of $T_p^{\text{deriv}}M$ (but this is still to be proven).

Let us point out, right away, that also this model for tangent spaces serves the original purpose: we can make sense of speeds

$$\dot{\gamma}(0) \in T_p^{\text{deriv}}M.$$

Here we use the notation $\dot{\gamma}(0)$ instead of $\frac{d\gamma}{dt}(0)$ to (temporarily) avoid overlaps with the previous section. The actual definition should be clear: just consider the variation of functions along γ (at $t = 0$):

$$\partial_{\dot{\gamma}(0)} : \mathcal{C}^\infty(M) \to \mathbb{R}, \quad \partial_{\dot{\gamma}(0)}(f) := \left.\frac{d}{dt}\right|_{t=0} f(\gamma(t)).$$

Of course, one can use the derivative at any t on the domain of γ, giving rise to

$$\partial_{\dot{\gamma}(t)} \in T_{\gamma(t)}^{\text{deriv}}M.$$

There is also a very natural interaction (which turns out to be an isomorphism) with the tangent space from the previous section. To see this, fix $v \in T_pM$; we try to define $\partial_v(f)$ for $f \in C^\infty(M)$. Once a chart $\chi : U \to \Omega \subset \mathbb{R}^m$ of M around p is chosen, we can use the representation of f with respect to χ, $f_\chi : \Omega \to \mathbb{R}$, as well as the one of v, $v^\chi \in \mathbb{R}^m$. Using again the chain rule, we see that

$$\partial_v(f) := \frac{\partial f_\chi}{\partial v^\chi}(\chi(p)) = (df_\chi)_{\chi(p)}(v^\chi)$$

actually does not depend on the choice of χ. Therefore we obtain

$$\partial_v : \mathcal{C}^\infty(M) \to \mathbb{R};$$

it should be clear that it is a derivation of M at p, hence

$$\partial_v \in T_p^{\text{deriv}}M.$$

For speeds of curves, when $v = \frac{d\gamma}{dt}(0)$ as defined in the previous section, one obtains $\partial_{\dot{\gamma}(0)}$ that we have just discussed. Similarly the vectors $\left(\frac{\partial}{\partial \chi}\right)_p$ from the previous section (Definition 4.3) give rise to the similar vectors defined in this section (Example 4.7). In general, we obtain a linear map

$$I_p : T_pM \to T_p^{\text{deriv}}M, \quad v \mapsto \partial_v.$$

THEOREM 4.8. *The map $I_p : T_pM \to T_p^{deriv}M$ is a linear isomorphism.*

At this point we could just prove this theorem and then transfer all the properties/constructions that we know for T_pM to $T_p^{\text{deriv}}M$, via I. It is true that the proof of the theorem is not completely trivial (well, it is not very difficult either, but requires some preparation). However, more importantly, it is instructive to look at $T_p^{\text{deriv}}M$ independently and check that it does satisfy the properties that we want. That will reveal the advantages of $T_p^{\text{deriv}}M$ as well as the further nature of tangent spaces.

Regarding the nature of $T_p^{\text{deriv}}M$, note that it is completely algebraic. See e.g. the following exercise.

EXERCISE 4.6. Let A be an algebra (with unit 1_A). Recall that a character on A is any map

$$\chi : A \to \mathbb{R}$$

which is a map of algebras, i.e. it is linear, multiplicative ($\chi(ab) = \chi(a)\chi(b)$ for all $a, b \in A$) and $\chi(1_A) = 1$. We denote by $X(A)$ the set of all characters on A.

For instance, when $A = \mathcal{C}^\infty(M)$ is the algebra of smooth functions on a manifold M, then any $p \in M$ gives rise to the character

$$\chi_p : \mathcal{C}^\infty(M) \to \mathbb{R}, \quad \chi_p(f) = f(p).$$

(and, even more is true: any character arises in this way!).

Motivated by this example, for any algebra A and any character χ on A we introduce the notion of χ-derivation on A, by which we mean a linear map $\chi : A \to \mathbb{R}$ satisfying the derivation identity:

$$\partial(ab) = \chi(a)\partial(b) + \partial(a)\chi(b)$$

for all $a, b \in A$. We denote by $\text{Der}_\chi(A)$ the (vector) space of all such derivations. Intuitively, this may be interpreted as "the tangent space of $X(A)$ at χ".

Assume now that M is a manifold, and we want to apply the previous discussion to two algebras: $A = \mathcal{C}(M)$ and $A^\infty = \mathcal{C}^\infty(M)$. Let us mention that the construction $M \ni p \mapsto \chi_p$ defines a bijection between M and both $X(A)$ as well as $X(A^\infty)$ (we are not asking you to prove this here- for details please see "Inleiding Topologie"). Any way, the point is that characters do not see the difference between A and A^∞. However, derivations (i.e. tangent spaces) do. More precisely, while

$$\text{Der}_{\chi_p}(\mathcal{C}^\infty(M)) = T_p^{\text{deriv}}M,$$

we are asking you to prove that $\text{Der}_{\chi_p}(\mathcal{C}(M)) = 0$.

(Hint: the proof is at least ten times shorter than the bla-bla (but otherwise interesting) of the exercise)

EXERCISE 4.7. When $A = \mathbb{R}[X_1, \ldots, X_m]$ is the algebra of polynomials in m variables show that any character is the evaluation χ_x at some point $x \in \mathbb{R}^m$, and then show that $\text{Der}_{\chi_x}(A)$ is a finite dimensional vector space and exhibit a basis.

Then consider another A: the algebra of polynomial functions on the circle S^1. And, again, consider the characters χ_p for $p \in S^1$ and compute $\text{Der}_{\chi_p}(A)$.

We now return to our tangent spaces $T_p^{\text{deriv}}M$ and the properties that we expect from them:

EXERCISE 4.8. Show that any smooth map $F : M \to N$ between two manifolds induces, at any $p \in M$, a linear map

$$(dF)_p : T_p^{\text{deriv}} M \to T_{F(p)}^{\text{deriv}} N, \quad (dF)_p(\partial)(f) = \partial(f \circ F).$$

Then state and prove the chain rule and deduce that any local diffeomorphism induces an isomorphism at the level of the (deriv-)tangent spaces.

We now discuss another aspect that $T_p^{\text{deriv}} M$ reveals: the local nature of derivations.

LEMMA 4.9. *Any derivation of M at p, $\partial \in T_p^{deriv} M$, has the following local property: for $f_1, f_2 \in \mathcal{C}^\infty(M)$ one has:*

$$f_1 = f_2 \text{ in a neighborhood of } p \Longrightarrow \partial(f_1) = \partial(f_2).$$

PROOF. Let $f = f_1 - f_2$. We know that $f = 0$ in a neighborhood U of p and we want to prove that $\partial(f) = 0$. We may assume that U is chosen so that it is diffeomorphic to a ball $B(0, \epsilon) \subset \mathbb{R}^m$, by a diffeomorphism that takes p to 0. As we have already noticed, one can find a smooth function on $\mathbb{R}^m$ that is supported inside $B(0, \epsilon)$ and is non-zero at 0 (e.g. take $x \mapsto g(\frac{1}{\epsilon^2}||x||^2)$, where $g : R \to \mathbb{R}$ is the function from Exercise 2.2). Moving from $B(0, \epsilon)$ to U, and extending by zero outside U, we find a function $\eta \in C^\infty(M)$ which is supported inside U and $\eta(p) \neq 0$. Then $\eta f = 0$, and applying the Leibniz identity for ηf we find $\eta(p)\partial(f) = 0$ hence, since $\eta(p) \neq 0$, one must have $\partial(f) = 0$. □

REMARK 4.10. The previous lemma can be packed into a more conceptual conclusion: any derivations ∂ at p descends to the quotient space

$$\mathcal{C}_p^\infty(M) := \mathcal{C}^\infty(M)/I_p = C^\infty(M)/\sim_p .$$

where I_p is the space of functions that vanish around p and $\sim_p$ is the associated equivalence relation:

$$f_1 \sim_p f_2 \Longleftrightarrow f_1 = f_2 \text{ in a neighborhood of } p.$$

For $f \in \mathcal{C}^\infty(M)$, its equivalence class is denoted

$$\text{germ}_p(f) \in \mathcal{C}_p^\infty(M)$$

and is called **the germ of** f **at** p. The space of germs at p, i.e. our quotient $\mathcal{C}_p^\infty(M)$, is still an algebra (since I_p is an ideal of $\mathcal{C}^\infty(M)$); explicitly, the operations are induced from $\mathcal{C}^\infty(M)$:

$$\text{germ}_p(f + g) = \text{germ}_p(f) + \text{germ}_p(g), \quad \textit{etc.}$$

The locality property from the previous lemma can now be interpreted as an isomorphism

$$T_p^{\text{deriv}} M \cong \text{Der}_p(\mathcal{C}_p^\infty(M))$$

between the tangent space and the space of all derivations of $\mathcal{C}_p^\infty(M)$ at p (with respect to the evaluation at p).

Assume that $U \subset M$ is an open containing p. Then any derivation $\partial \in T_p^{\text{deriv}} U$ gives a derivation $\tilde{\partial} \in T_p^{\text{deriv}} M$ simply by

$$\tilde{\partial}(f) := \partial(f|_U).$$

Of course, the resulting map

$$T_p^{\text{deriv}} U \to T_p^{\text{deriv}} M, \quad \partial \mapsto \tilde{\partial}$$

is just the map induced (cf. Exercise 4.8 above) by the inclusion $U \hookrightarrow M$.

LEMMA 4.11. *Given $p \in U \subset M$ with U open, the canonical map $T_p^{deriv}U \to T_p^{deriv}M$ is a linear isomorphism.*

PROOF. For injectivity, let $\partial : \mathcal{C}^\infty(U) \to \mathbb{R}$ be a derivation at p such that $\partial(\tilde{f}|U) = 0$ for all $\tilde{f} \in \mathcal{C}^\infty(M)$; we must show that $\partial = 0$. This follows immediately from the following:

- $\partial(f)$ depends only on f in an arbitrarily small neighborhood of p. This is the content of the previous lemma.
- for any $f \in \mathcal{C}^\infty(U)$, there exists $\tilde{f} \in \mathcal{C}^\infty(M)$ s.t. $f = \tilde{f}$ in some neighborhood of p.

To prove the last item, we take $\tilde{f} := \eta \cdot f$ where $\eta \in \mathcal{C}^\infty(M)$ is supported inside U (so that $\tilde{f}$ is defined and smooth on the entire M) and $\eta = 1$ in some (smaller) neighborhood of p. The existence of η is a local problem- therefore it suffice to build a smooth function η on $\mathbb{R}^m$ that is supported in the ball $B(0,1)$ and is 1 in a neighborhood of 0 (say on $B(0, \frac{1}{3})$); for that one takes again η of type $\eta(x) = g(||x||^2)$ where $g : \mathbb{R} \to \mathbb{R}$ is any smooth function that is 1 when $|t| < \frac{1}{3}$ and is 0 when $t \geq \frac{1}{2}$.

For the surjectivity, we start with $\tilde{\partial} : \mathcal{C}^\infty(M) \to \mathbb{R}$ and we want to build ∂. By now the definition should be clear: or $f \in \mathcal{C}^\infty(U)$, choose any $\tilde{f} \in \mathcal{C}^\infty(M)$ that coincides with f near p (possible by the first part) and set $\partial(f) := \tilde{\partial}(\tilde{f})$. The previous lemma shows that the definition of $\partial(f)$ does not depend on the choice of $\tilde{f}$. And this can used also to check that ∂ is still a derivation at p. E.g., for the Leibniz identity, choosing extensions $\tilde{f}$ or f and $\tilde{g}$ or g, then use the extension $\tilde{f}\tilde{g}$ or fg and use the Leibniz identity for $\tilde{\partial}(\tilde{f}\tilde{g})$. □

We are now ready to prove the main theorem of this section,

PROOF OF THEOREM 4.8. Fix a chart $\chi : U \to \mathbb{R}^m$ around p that takes p to $0 \in \mathbb{R}^m$. Since the induced basis of T_pM (of Definition 4.3) are sent by I_p into the similar vectors $\left(\frac{\partial}{\partial \chi}\right)_p$ of this section (see Example 4.3), it suffices to show that the last ones form a basis of $T_p^{\text{deriv}}M$. Or, by the previous lemma, of $T_p^{\text{deriv}}U$. But the chart χ takes $T_p^{\text{deriv}}U$ isomorphically into $T_0^{\text{deriv}}\Omega$ (this follows using e.g. the chain rule- see Exercise 4.8) and, o course, takes our vectors into $\left(\frac{\partial}{\partial x_i}\right)_0 \in T_0^{\text{deriv}}\Omega$. Using again the previous lemma, it suffices to show that

$$\left(\frac{\partial}{\partial x_i}\right)_0 : \mathcal{C}^\infty(\mathbb{R}^m) \to \mathbb{R}$$

form a basis of $T_p^{\text{deriv}}\mathbb{R}^m$. The linear independence is immediate: if $\sum_i \lambda_i \left(\frac{\partial}{\partial x_i}\right)_0 = 0$, applying this to x_i (i.e. to the coordinate function $x \mapsto x_i$) we find $\lambda_i = 0$.

We still have to check that any derivation at 0,

$$\partial : \mathcal{C}^\infty(\mathbb{R}^m) \to \mathbb{R},$$

is a linear combination of the standard partial derivatives at 0:

$$\partial = \sum_i \lambda_i \left(\frac{\partial}{\partial x_i}\right)_0 \quad \text{with } \lambda_i \in \mathbb{R}.$$

We will show that $\lambda_i := \partial(x_i) \in \mathbb{R}$ do the job. To show this, we use the fact that any $f \in \mathcal{C}^\infty(\mathbb{R}^m)$ can be written as

$$f(x) = f(0) + \sum_i x_i g_i(x), \quad \text{with } g_i \in \mathcal{C}^\infty(\mathbb{R}^m) \tag{2.1}$$

(see below). Using the Leibniz identity (which also implies that ∂ is 0 on constant functions) we obtain

$$\partial(f) = \sum_i \lambda_i g_i(0),$$

while $\frac{\partial f}{\partial x_i}(0) = g_i(0)$. Hence we are left with proving that any f can be written in the form (2.1). For this we use the identity

$$h(1) - h(0) = \int_0^1 \frac{dh}{dt}(t)dt$$

for all smooth functions $h : [0,1] \to \mathbb{R}$; of course, we want to choose h such that $h(1) = f(x)$ and $h(0) = f(0)$. Choose then $h(t) = f(tx)$ and we obtain the desired identity (2.1) with

$$g_i = \int_0^1 \frac{\partial f}{\partial x_i}(tx)dt.$$

□

3. Tangent spaces: conclusions

Let us put together the main conclusions on tangent spaces.

3.1. Via charts, or as derivations: In the previous two sections we have seen (for any m-dimensional manifold M, $p \in M$) two ways of looking at the tangent space T_pM:

T1: *its elements $v \in T_pM$ can be represented w.r.t. any chart χ around p by vectors $v^\chi \in M$; the passing from one chart to the other was given by*

$$v^{\chi'} = (d\,c)_{\chi(p)}(v^\chi), \quad \textit{where } c = c_{\chi,\chi'} = \chi' \circ \chi^{-1}.$$

T2: *its elements $v \in T_pM$ can be interpreted as derivations on $\mathcal{C}^\infty(M)$ at p.*

The equivalence between the two is given by the isomorphism $I_p : T_pM \to T_p^{\text{deriv}}M$ of Theorem 4.8. From now on we will identify the two (via I_p) and use only the notation T_pM. The upshot is that, when describing a vector $v \in T_pM$ we have the choice to describe it w.r.t. a chart, or as a derivation at p.

3.2. Speeds: Most importantly, any curve $\gamma \in \text{Curves}_p(M)$ has a speed

$$\frac{d\gamma}{dt}(0) \in T_pM$$

Explicitly:

- w.r.t. a chart χ, it is $\frac{d\gamma^\chi}{dt}(0) \in \mathbb{R}^m$.
- as a derivation at p it sends a smooth function f to $\frac{df\circ\gamma}{dt}(0)$.

(and, with the same descriptions, one has $\frac{d\gamma}{dt}(t) \in T_{\gamma(t)}M$ for any t in the domain of γ).

And this gives the best way to think (and even to work with) tangent spaces:

T0: $T_pM = \{\frac{d\gamma}{dt}(0) : \mathit{Curves}_p(M)\}$, *where we know that two such curves have the same speed at $t = 0$ if and only if that happens for their representations w.r.t. a/any chart.*

3.3. Basis with respect to a chart: The tangent space is an m-dimensional vector space- hence isomorphic to $\mathbb{R}^m$. Each chart χ around p gives rise to a basis of T_pM (hence to a specific isomorphism with $\mathbb{R}^m$):

$$\left(\frac{\partial}{\partial \chi_1}\right)_p, \ldots, \left(\frac{\partial}{\partial \chi_m}\right)_p \in T_pM,$$

called the canonical basis w.r.t. χ:

T1: w.r.t. χ, $\left(\frac{\partial}{\partial \chi_i}\right)_p$ corresponds to $e_i \in \mathbb{R}^m$.

T2: as derivations, they are given by the partial derivatives w.r.t. χ:

$$\frac{\partial f}{\partial \chi_i}(p) := \frac{\partial f_\chi}{\partial x_i}(\chi(p)).$$

T0: as speeds, $\left(\frac{\partial}{\partial \chi_i}\right)_p$ are induced by the paths

$$t \mapsto \chi^{-1}(\chi(p) + te_i).$$

See also Definition 4.3 and Example 4.7. The way that these vectors change when we change the chart is easiest read off from the point of view of [T1], which gives the formula (1.3). Or, using the partial derivatives given by the point of view of [T2], we have the more compact formula:

$$\left(\frac{\partial}{\partial \chi_i}\right)_p = \sum_{j=1}^{m} \frac{\partial c_j}{\partial \chi_i}(\chi(p)) \left(\frac{\partial}{\partial \chi'_j}\right)_p = \sum_{j=1}^{m} \frac{\partial \chi'_j}{\partial \chi_i}(p) \left(\frac{\partial}{\partial \chi'_j}\right)_p \quad \text{(where } c = \chi' \circ \chi^{-1}\text{)}.$$

Or, if one denotes by x and y the charts χ and χ', respectively, and c by y but interpreted as a function of x, one come across the more compact (but a bit sloppy) formula:

$$\frac{\partial}{\partial x_i} = \sum_j \frac{\partial y_j}{\partial x_i} \frac{\partial}{\partial y_j}.$$

3.4. Differentials: While, intuitively, the tangent T_pM is "the linearization of M near p", we can also linearize smooth maps. And this was one of the main properties that we wanted from tangent spaces (and the property that allows one to use them in order to compare different manifolds). More precisely, one can talk about the differential

$$(dF)_p : T_pM \to T_{F(p)}N \quad \text{for any smooth map } f : M \to N.$$

Again, this can described from the three points of view:

T0: It sends the speed (at $t = 0$) of $\gamma \in \text{Curves}_p(M)$ to the one of $F \circ \gamma \in \text{Curves}_{F(p)}(N)$.

T1: For $v \in T_pM$ represented w.r.t. χ by v^χ, $(dF)_p(v)$ is represented w.r.t. a/any chart χ' around $f(p)$ by the image of v^χ by $(dF_{\chi,\chi'})_{\chi(p)}$.

T2: $(dF)_p(v)$, as a derivation, acts on a function $g \in \mathcal{C}^\infty(M)$ by acting with v on $g \circ F$.

Or, if we want to write down $(dF)_p$ w.r.t. bases induced by charts χ around p and χ' around $F(p)$:

$$\left(\frac{\partial}{\partial \chi_i}\right)_p \xmapsto{(dF)_p} \sum_{j=1}^{m} \frac{dF^j_{\chi,\chi'}}{dx_i}(\chi(p)) \left(\frac{\partial}{\partial \chi'_j}\right)_{F(p)},$$

where we use the representation $F_{\chi,\chi'} = \chi' \circ F \circ \chi^{-1}$ of F w.r.t. the two bases. Or, in other words, the matrix corresponding to $(dF)_p$ w.r.t. the two bases is precisely the matrix of the partial derivatives of $F_{\chi,\chi'}$ at $\chi(p)$.

Finally, there is the following characterization of immersions/submersions which is often taken as a definition (we did not do that because we did not want to wait until the tangent spaces were defined. However, this should certainly be the way to think about immersions/submersions from now on!). For embedded submanifolds of Euclidean spaces, this was already pointed out in the previous chapter (Exercise 3.14 there). The following exercise (the general case) should be rather immediate, but it is instructive to do it using the various properties we mentioned above (e.g. without mentioning any formula!).

EXERCISE 4.9. Show that a smooth map $f : M \to N$ between two manifolds is an immersion/submersion at a point $p \in M$ if and only if its differential at p, $(df)_p : T_pM \to T_{f(p)}N$, is injective/surjective.

3.5. Submanifolds: Let us restrict for simplicity to embedded submanifolds. Given such a submanifold $N \subset M$ of a manifold M, the inclusion $i : N \hookrightarrow M$ is an immersion, therefore, for each $p \in N$,

$$(di)_p : T_pN \to T_{i(p)}M$$

is an injection. In the same way that we do not write $i(p)$ but p, this injection will be viewed as an inclusion

$$T_pN \subset T_{i(p)}M.$$

From the point of view of curves this is completely clear: since any curve in N is also one in M, the speed $\frac{d\gamma}{dt}(0) \in T_pN$ is now identified with $\frac{d\gamma}{dt}(0) \in T_pM$ (we do not even have a notation to distinguish the two!). Also, as derivations, if $v \in T_pN$, it acts on $f \in \mathcal{C}^\infty(M)$ by acting with v on $f|_N$.

In this way tangent spaces give us (linear) information on the way that N sits inside M (near p).

Finally, here is the full version of the regular value theorem proven in Chapter 3 (Theorem 3.21 there), this time with exra-information on the tangent spaces.

THEOREM 4.12 (the regular value theorem with tangent spaces). *If $q \in N$ is a regular value of a smooth map*

$$F : M \to N,$$

then the fiber above q, $F^{-1}(q)$, is an embedded submanifold of M of dimension

$$dim(F^{-1}(q)) = dim(M) - dim(N)$$

and the tangent spaces $T_p\left(F^{-1}(q)\right)$, as subspaces of T_pM, coincide with the kernel of the differential $(dF)_p : T_pM \to T_{F(p)}N$:

$$T_p\left(F^{-1}(q)\right) = Ker\,(dF)_p \quad ((for\ all\ p \in F^{-1}(q)).$$

PROOF. Nice exercise (hint: just look back at our reminder on Analysis). □

3.6. Back to Euclidean spaces and their embedded submanifolds: Back to Analysis, i.e. for a manifold of type $\mathbb{R}^k$ (or an open inside it), the general/abstract theory of this chapter tells us that the tangent spaces $T_p\mathbb{R}^k$ come with a canonical basis

$$\left(\frac{\partial}{\partial x_1}\right)_p, \ldots, \left(\frac{\partial}{\partial x_m}\right)_p. \tag{3.1}$$

(the smooth structure on $\mathbb{R}^k$ can be given by one, global, chart- the identity).

As derivations at p, these correspond to the standard partial derivatives. With respect to the canonical chart, they are represented by the canonical basis $e_1, \ldots, e_k$ of $\mathbb{R}^k$. As speeds, they correspond to the paths $t \mapsto p + te_i$. In any case, we have an isomorphism that will be treated from now on as an identification:

$$T_p\mathbb{R}^k = \mathbb{R}^k,$$

but we will still keep the notation (3.1) for the canonical basis. This is in order to indicate that we are thinking about/using tangent vectors.

With the identification above, whenever we have a smooth map $f : M \to \mathbb{R}^k$, we talk about its differentials as maps $(df)_p : T_pM \to \mathbb{R}^k$.

EXERCISE 4.10. For $f \in \mathcal{C}^\infty(M)$ and $v \in T_pM$, check that the action of v (as a derivation) on f is precisely

$$\partial_v(f) = (df)_p(v).$$

EXERCISE 4.11. Check that, via this identification, the speeds $\frac{d\gamma}{dt}(t)$ discussed in this chapter are identified with the standard ones from Analysis.

As in the previous exercise for curves, we have a similar problem for embedded submanifolds:

EXERCISE 4.12. For an embedded submanifold $M \subset \mathbb{R}^k$, $p \in M$, we now have two tangent spaces,

$$T_pM \subset \mathbb{R}^k,$$

namely the one from this chapter $T_pM \subset T_p\mathbb{R}^k$ (see subsection 3.5) combined with the identification $T_p\mathbb{R}^k = \mathbb{R}^k$, and the one from Analysis using derivatives of curves. Show that the two coincide. Try to find different arguments; e.g. one using the previous exercise, one using the regular values theorem, etc.

By now, it should also be clear that for smooth maps between embedded submanifolds of Euclidean spaces, the differential discussed in this chapter becomes (identified with) the one from Analysis (e.g. as recalled in Chapter 2, Exercise 2.10).

4. Vector fields

We now discuss vector fields. The brief philosophy is:

While for embedded submanifolds $M \subset \mathbb{R}^k$ each single tangent space $T_pM \subset \mathbb{R}^k$ was interesting (e.g. because it reflects the position of M inside $\mathbb{R}^k$ near p), for a general manifold M this is less so. Instead, it is the entire family $\{T_pM\}_{p\in M}$ that is interesting. I.e. not one single tangent vector, but a family $\{v_p\}_{p\in M}$ of vectors (of course, "varying smothly" with respect to p). I.e. vector fields.

Looking back at the previous sections of this chapter, we see that we insisted in making sense of T_pM as vector spaces independently of the way that M may sit in some larger Euclidean space. Now, just as a vector space, T_pM is rather boring: it is isomorphic to $\mathbb{R}^m$. And the same is true for any finite dimensional vector space V; however, one should keep in mind that, for a bare m-dimensional vector space V, realizing an explicit isomorphism with $\mathbb{R}^m$ amounts to extra choices (namely a basis of V). In particular, for the tangent spaces T_pM, we obtained identifications with $\mathbb{R}^m$ once we fixed a chart χ (and these identifications work for p in the domain of χ only!). Actually this is a "problem" (or better: "interesting phenomena") not only for general Ms, but even for embedded submanifolds $M \subset \mathbb{R}^k$: while isomorphisms

$$\mathbb{R}^m \cong T_pM \subset \mathbb{R}^k$$

can be chosen for each p, often cannot be chosen so that they depend smoothly on $p \in M$. The "hairy ball theorem" implies that this is the case already for $M = S^2 \subset \mathbb{R}^3$:

EXERCISE 4.13. Assume that for each $p \in S^2$ one can find a linear isomorphism

$$\Phi_p : \mathbb{R}^2 \to T_pS^2 \subset \mathbb{R}^3$$

so that the Φ_ps vary smoothly with respect to p, in the sense that the map

$$\Phi : S^2 \times \mathbb{R}^2 \to \mathbb{R}^3, \quad (p, v) \mapsto \Phi_p(v)$$

is smooth. Show that there exists a no-where vanishing vector field on S^2, i.e. a smooth map

$$X : S^2 \to \mathbb{R}^3,$$

nowhere vanishing, such that $X(p) \in T_pS^2$ for all $p \in S^2$. Try now to construct such an X! Then state the "hairy ball theorem".

More generally, for an embedded submanifold $M \subset \mathbb{R}^k$, a vector field on M is a function

$$M \ni p \mapsto X(p) \in T_pM \subset \mathbb{R}^k$$

which, when interpreted as a function with values in $\mathbb{R}^k$, is smooth. What if an embedding into some Euclidean space is not fixed (or if we use a different one)? First of all, we can still look at maps

$$X : M \ni p \mapsto X_p \in T_pM$$

(interpreted also as families $X = \{X_p\}_{p\in M}$). These will be called **set-theoretical vector fields** on M. To make sense of their smoothness, we use charts $\chi : U \to \Omega \subset \mathbb{R}^m$ of M. Recall that any such chart induces a basis $\left(\frac{\partial}{\partial\chi_i}\right)_p$ of T_pM hence any set-theoretical vector field X can be written as

$$X_p = \sum_{i=1}^{m} X^i_\chi(\chi(p)) \left(\frac{\partial}{\partial\chi_i}\right)_p \tag{4.1}$$

for all $p \in M$, where each X^i_χ is a function

$$X^i_\chi : \Omega \to \mathbb{R}.$$

These will be called **the coordinates (functions) of X w.r.t. the chart χ.**

Definition 4.13. *A vector field on a manifold is any set-theoretical vector field X with the property that its coordinate functions X^i_χ w.r.t any chart χ are smooth.*
We denote by $\mathfrak{X}(M)$ the set of all vector fields on M.

Note the algebraic structure on $\mathfrak{X}(M)$ that is present right away:

- it is a vector space, with the addition an multiplication of scalars defined pointwise:
$$(X + Y)_p := X_p + Y_p, \quad (\lambda \cdot X)_p := \lambda \cdot X_p.$$
- it is a $\mathcal{C}^\infty(M)$-module, i.e. there is an operation of multiplication of vector fields by smooth functions,
$$\mathcal{C}^\infty(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M), \quad (f, X) \mapsto f \cdot X.$$
Again, this is defined pointwise by the obvious formula:
$$(f \cdot X)_p := f(p) \cdot X_p.$$

Next, we look for alternative ways to characterize the smoothness of vector fields; as a bonus, these will reveal more of the structure that vector fields carry.

First of all, while tangent vectors could be interpreted as derivations (at given points $p \in M$), for any $X \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$ one has $\partial_{X_p}(f) \in \mathbb{R}$ for each $p \in M$, therefore a function

$$L_X(f) : M \to \mathbb{R}, \quad p \mapsto \partial_{X_p}(f) = (df)_p(X_p).$$

This is called **the Lie derivative of f along the vector field X**. Of course, we expect that $L_X(f)$ is again smooth.

PROPOSITION 4.14. *A set-theoretical vector field X on M is smooth if and only if*

$$L_X(f) \in C^{\infty}(M) \text{ for all } f \in C^{\infty}(M).$$

PROOF. Assume first that X is smooth in sense of the definition. For $f \in C^{\infty}(M)$, we can check the smoothness of $L_X(f)$ locally: for an arbirary chart χ we have to make sure that $L_X(f) \circ \chi^{-1}$ is smooth. But applying the definitions, we see that this expression is precisely

$$\sum_i X^i_{\chi} \frac{\partial f_{\chi}}{\partial x_i}$$

and then smoothness follows.

For the converse, by the type of arguments we have already seen, it suffices to show that for any point $p \in M$ there exists a chart around p such that the coefficients X_i^{χ} are smooth. For an arbitrary chart $\chi : U \to \Omega \subset \mathbb{R}^m$ around p, we choose $f_i \in \mathcal{C}^{\infty}(M)$ such that, in a smaller neighborhood $U_0 \subset U$ of p, f_i coincides with χ_i. By hypothesis, $L_X(f_i)$ is smooth; but over U_0, this function is precisely X_i^{χ}. Hence taking $\chi_0 := \chi|_{U_0}$, the coefficients $X_i^{\chi_0}$ □

The interpretation of tangent vectors at p as derivations at p can be pushed further to a similar characterization of vector fields. The algebraic objects are **derivations of** $\mathcal{C}^{\infty}(M)$, by which we mean linear maps

$$L : \mathcal{C}^{\infty}(M) \to \mathcal{C}^{\infty}(M)$$

satisfying the derivation low (Leibniz identity):

$$L(fg) = L(f)g + fL(g) \quad \text{for all } f, g \in \mathcal{C}^{\infty}(M).$$

We denote by $\mathrm{Der}(\mathcal{C}^{\infty}(M))$ the (vector) space of such derivations (note: again, this is a purely algebraic construction, that applies to any algebra A).

THEOREM 4.15. *For any $X \in \mathfrak{X}(M)$, the Lie derivative along X,*

$$L_X : \mathcal{C}^{\infty}(M) \to \mathcal{C}^{\infty}(M), \quad f \mapsto L_X(f)$$

is a derivation. Moreover,

$$I : \mathfrak{X}(M) \to Der(\mathcal{C}^{\infty}(M)), \quad X \mapsto L_X$$

is an isomorphism of vector spaces.

PROOF. The fact that L_X is a derivation follows right away from the similar property of the ∂_{X_p}s. The injectivity of I follows from the fact that, if a tangent vector $v_p \in T_pM$ is non-zero, there exists a smooth function f on M such that $\partial_{v_p}(f) \neq 0$ (why?).

For the surjectivity of I, let L be a derivation. Then, for each $p \in M$,

$$X_p : \mathcal{C}^{\infty}(M) \to \mathbb{R}, \quad f \mapsto L(f)(p)$$

is a derivation at p, hence defines a tangent vector $X_p \in T_pM$. Since$L_X(f) = L(f)$ is smooth for any smooth f, the previous proposition implies that X is smooth. □

And here is one clear advantage of the point of view of derivations: the presence of yet another structure on $\mathfrak{X}(M)$, namely a bracket operation

$$\mathfrak{X}(M) \times \mathfrak{X}(M) \to\in \mathfrak{X}(M).$$

More precisely, given two derivations on $\mathcal{C}^{\infty}(M)$, necessarily of type L_X and L_Y, their composition

$$L_X \circ L_Y : \mathcal{C}^{\infty}(M) \to \mathcal{C}^{\infty}(M)$$

is not a derivation anymore. Indeed, what we have is:

$$(L_X \circ L_Y)(fg) = (L_X \circ L_Y)(f)g + f(L_X \circ L_Y)(g) + L_X(f)L_Y(g) + L_Y(f)L_X(g),$$

i.e. a "defect term" $L_X(f)L_Y(g) + L_Y(f)L_X(g)$ which spoils the Leibniz identity. However, the other composition $L_Y \circ L_X$ comes with the same "defect term"; therefore, their difference

$$[L_X, L_Y] := L_X \circ L_Y - L_Y \circ L_X : \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$$

is again a derivation.

DEFINITION 4.16. *Given $X, Y \in \mathfrak{X}(M)$, we denote by $[X, Y] \in \mathfrak{X}(M)$, called* ***the Lie bracket of** X **and** Y, the (unique) vector field on M with the property that*

$$L_{[X,Y]} = L_X \circ L_Y - L_Y \circ L_X.$$

The following exercise should convince you of the advantage of the point of view of derivations.

EXERCISE 4.14. Compute the coordinate functions of $[X, Y]$ with respect to a chart χ, in terms of the coordinates functions of X and Y defined by (4.1).

EXERCISE 4.15. In this exercise we point out the main properties of the Lie bracket operation $[\cdot, \cdot]$ on $\mathfrak{X}(M)$.

First of all, the way it interacts with the other structure present on on $\mathfrak{X}(M)$- show that:

- w.r.t. the vector space structure: it is a bilinear and skew-symmetric operation
$$[\cdot, \cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \in \mathfrak{X}(M).$$
- w.r.t. the $\mathcal{C}^\infty(M)$-module structure: it satisfies the derivation rule:
$$[X, f \cdot Y] = f \cdot [X, Y] + L_X(f) \cdot Y,$$
for all $X, Y \in XX(M)$, $f \in \mathcal{C}^\infty(M)$.

And then a property that $[\cdot, \cdot]$ has on its own- **the Jacobi identity**:

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0 \quad \text{for all } X, Y, Z \in \mathfrak{X}(M).$$

EXERCISE 4.16. Here we point out the main properties of the operation $[\cdot, \cdot]$ on $\mathfrak{X}(M)$.

First of all, to see the way it interacts with the other structures present on $\mathfrak{X}(M)$, show that:

- w.r.t. the vector space structure: it is a bilinear and skew-symmetric operation
$$[\cdot, \cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M).$$
- w.r.t. the $\mathcal{C}^\infty(M)$-module structure: it satisfies the derivation rule:
$$[X, f \cdot Y] = f \cdot [X, Y] + L_X(f) \cdot Y,$$
for all $X, Y \in XX(M)$, $f \in \mathcal{C}^\infty(M)$.

Then show the following property that $[\cdot, \cdot]$ has on its own:

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$$

for all $X, Y, Z \in \mathfrak{X}(M)$ (this is known as **the Jacobi identity**).

Back to the the very definition of vector fields and of $\mathfrak{X}(M)$, here is another way to characterize smoothness. The idea is to put all the spaces T_pM (disjointly) in one big space

$$TM := \sqcup_{p \in M} T_pM = \{(p, v_p) : p \in M, v_p \in T_pM\}$$

and make this into a manifold, so that the smoothness of a set-theoretical vector field X is the same as the smoothness of X viewed as a map between manifolds:

$$X : M \to TM, \quad p \mapsto X_p.$$

To put a smooth structure on TM, i.e. to exhibit charts, we need to understand how to parametrize the points of TM by coordinates. Of course, we should start from a chart

$$\chi : U \to \Omega \subset \mathbb{R}^m$$

for M. The point is that such a chart also allows us to parametrize tangent vectors $v_p \in T_pM$, for $p \in U$, by coordinates $\hat{\chi}_i(v_p)$ w.r.t. to the induced basis $\left(\frac{\partial}{\partial \chi_i}\right)_p$ of T_pM: just decompose v_p as

$$v_p = \sum_i \lambda_i \left(\frac{\partial}{\partial \chi_i}\right)_p \quad \text{and set } \hat{\chi}_i(v_p) := \lambda_i. \tag{4.2}$$

Therefore, on the subset $TU \subset TM$, we have a natural "chart":

$$\tilde{\chi} : TU \to \Omega \times \mathbb{R}^m \subset \mathbb{R}^{2m},$$
$$\tilde{\chi}(p, v_p) = (\chi_1(p), \ldots, \chi_m(p), \hat{\chi}_1(v_p), \ldots, \hat{\chi}_m(v_p)).$$

Proposition 4.17. *TM can be made into a manifold in a unique way so that, for any chart (U, χ) of M, $(TU, \tilde{\chi})$ is a chart of TM.*

Moreover, this is also the unique smooth structure with the property that a set-theoretical tangent vector X on M is smooth if and only if it is smooth as a map $X : M \to TM$.

Proof. We first compute the change of coordinates between charts of type $(TU, \tilde{\chi})$. We start with two charts (U, χ) and (U', χ') for M, with change of coordinates $c = \chi' \circ \chi^{-1}$. We want to compute $\tilde{c} := \tilde{\chi}' \circ \tilde{\chi}^{-1}$. Hence we try to write $\tilde{\chi}'(p)$ as a function depending on $\tilde{\chi}(p)$. While the first m coordinates are $c(\chi(p))$, we need to understand the last m. For that we use:

$$\left(\frac{\partial}{\partial \chi_i}\right)_p = \sum_j \frac{\partial c_j}{\partial x_i}(\chi(p)) \left(\frac{\partial}{\partial \chi'_i}\right)_p.$$

(see (1.3)). TRhen, for a tangent vector $p \in M$, replacing this in (4.2), we find that

$$\hat{\chi'}_j(v_p) = \sum_i \hat{\chi}_i(v_p) \frac{\partial c_j}{\partial x_i}(\chi(p)).$$

Denoting by $(x, \hat{x})$ the coordinates in $\mathbb{R}^{2m} = \mathbb{R}^m \times \mathbb{R}^m$, we find that

$$\tilde{c}(x, \hat{x}) = (c(x), \sum_i \hat{x}_i \frac{\partial c}{\partial x_i}(x).$$

Hence the changes of coordinates is smooth.

However, strictly speaking, to show that TM is a manifold, there are still a few things to be done. First of all, we should have made it into a topological space. But, as we discussed at the end of Section 4 from Chapter 3, the topology can actually be recovered from an atlas. Let us give the details, independently of our earlier discussion. Given

$$\text{a chart } \chi : U \to \Omega \text{ for } M, \quad \tilde{\Omega} - \text{open in } \Omega \times \mathbb{R}^m \tag{4.3}$$

we would like

$$\tilde{\chi}^{-1}(\tilde{\Omega}) \subset TM$$

to be open in M. We denote by $\mathcal{B}$ the collection of subsets of TM of this type. We will declare a subset $D \subset TM$ to be open if for each $(p, v_p) \in D$ there exists $B \in \mathcal{B}$ such that

$$(p, v_p) \in B \subset D.$$

Of course, what is happening here is that $\mathcal{B}$ is a topology basis, and we consider the associated topology (hence one can also say that a subset of TM is open iff it is an union of members of $\mathcal{B}$).

It should be clear that the resulting topology is Hausdorff. For 2nd countability, note that, in (4.3), we can restrict ourselves to χ belonging to an atlas inducing the smooth structure on M and, for each χ, to $\tilde{\Omega}$ belonging to a basis of the Euclidean topology on $\Omega \times \mathbb{R}^m$. As long as the domains of the charts from $\mathcal{A}$ for a basis for the topology of M, the resulting $\mathcal{B}$ would still be a topology basis for the topology we defined on TM. Since M and the $\Omega \times \mathbb{R}^m$s are 2nd countable, we are able to produce a countable basis for the topology on TM.

The fact that a set-theoretical vector field is smooth if and only if it is smooth as a map $X : M \to TM$ should be clear (one just has to check it locally, and then it is really the definition of the smoothness of X). The uniqueness is left as an exercise for the interested student. □

Note the object that results from this discussion: TM. It is now a manifold that relates to M by an obvious "projection map"

$$\pi : TM \to M, \quad (p, v_p) \mapsto p.$$

And the fibers $\pi^{-1}(p)$ are vector spaces (precisely the tangent spaces of M). This is precisely what vector bundles are- but we will return to them later on.

4.1. Some interesting questions/phenomena with vector fields. (... for the curious students) Here are some very simple questions about vector fields whose answer is non-trivial question but very exciting. Let us start from the "hairy ball theorem" which can be seen as an interesting property of the two-sphere S^2: it does not admit a nowhere vanishing vector field. The first question in our list is now obvious:

Question: *Which manifolds admit a non-where vanishing vector field?*

It is worth looking at this question even for the spheres $S^m \subset \mathbb{R}^{m+1}$. Then the question becomes pretty elementary: does there exist a smooth function

$$X : S^m \to \mathbb{R}^{m+1} \setminus \{0\}$$

with the property that

$$x_0 \cdot X_0(x) + x_1 \cdot X_1(X) + \ldots + X_m \cdot X_m(x) = 0 \quad \forall\, x = (x_0, x_1, \ldots, x_m) \in S^{m+1}?$$

For S^1 the answer is clearly yes (think on the picture!) and, as we have already mentioned, for S^2 the answer is no. What about S^3? Well, using quaternions as in Example 3.34 we see that the answer is yes. More precisely, viewing $S^3 \subset \mathbb{H} \cong \mathbb{R}^4$, the multiplication by i, j and k provide three nowhere vanishing (and even linearly independent!) vector fields on S^3:

$$X^i(x) = i \cdot x, \quad X^j(x) = j \cdot x, \quad X^k(x) = k \cdot x.$$

(why are these vector fields on spheres?).

It turns out that, among the spheres, it is precisely the odd dimensional ones that do admit such vector fields.

For general manifolds the answer to the previous question is closely related to the topology of M- namely to the Euler number $\chi(M)$ of M. We will briefly discuss the Euler number later on when we will introduce DeRham cohomology and we will see that $\chi(S^m) = 1 + (-1)^m$. For now, let us mention that this number is, in principle, easy to compute using a triangulation of M, by the Euler formula

$$\chi(M) = \#\,\text{vertices} - \#\,\text{edges} + \#\,\text{faces} - \ldots.$$

(using this you should convince yourself that, indeed, $\chi(S^m) = 1 + (-1)^m$). The answer to the previous question is: a compact manifold M admits a no-where vanishing vector field if and only $\chi(M) = 0$!

What about more than one no-where vanishing vector fields? Of course, the question is interesting only if we look for linearly independent ones- and then the number of such vector fields will be bounded from above by the dimension of the manifold. E.g., on S^3, there are at most three such; and the vector field X^i, X^j and X^k described above show that the upper bound can be achieved. More generally:

Question: *An m-dimensional manifold M is said to be parallelizable if it admits m vector fields which, at each point, are linearly independent.*

For instance, we already know that S^1 and S^3 are parallelizable. One possible explanation for this is the fact that both S^1 as well as S^3 are Lie groups (and they are the only spheres that can be made into Lie groups- see Example 3.34 from Chapter 3). More precisely:

EXERCISE 4.17. *Show that all Lie groups are parallelizable. Here is how to do it. Let G be a k-dimensional Lie group. Consider its tangent space at the identity, $\mathfrak{g} := T_eG$ (a k-dimensional vector space). For $v \in \mathfrak{g}$, we define the vector field $\vec{v}$ on G by*

$$\vec{v}_g := (dL_g)_e(v)$$

where $L_g : G \to G$ is given by $L_g(h) = gh$ and $(dL_g)_e : \mathfrak{g} \to T_gG$ is its differential.

Show that if $\{v_1, \ldots, v_k\}$ is a basis of the vector space $\mathfrak{g}$, then $\vec{v}_1, \ldots, \vec{v}_k$ are k-linearly independent vector fields on G.

While the only spheres that can be made into Lie groups are S^0, S^1 and S^3, are there any other spheres that are parallelizable? Can one reproduce the argument that we used for S^3 so that it applies to other spheres S^m? We see that we need some sort of "product operation" on $\mathbb{R}^{m+1}$, so that we can define vector fields X^l by

$$X^l(x) = e_l \cdot x, \quad l \in \{1, \ldots, m\}$$

where $\{e_0, e_1, \ldots, e_m\}$ is the canonical basis. We end up again with the question mentioned in Example 3.34 from Chapter 3, of whether $\mathbb{R}^{m+1}$ can be made into a normed division algebra. And, as we mentioned there, the only possibilities were $\mathbb{R}, \mathbb{R}^2$ and R^4 for associative products, and also $\mathbb{R}^8$(octonions) in full generality. However, while the failure of associativity was a problem in making the corresponding sphere into a Lie group, it is not difficult to see that it does not pose a problem in producing the desired vector fields. Therefore: yes, the arguments that used $\mathbb{H}$ works also for octonions, hence also S^7 is parallelizable.

And, as you may expect by now, it turns out that S^0, S^1, S^3 and S^7 are the only spheres that are parallelizable!

So far we have looked at the extreme cases: one single nowhere vanishing vector field, or the maximal number of linearly independent vector fields. Of course, the most general question is:

Question: *Given a manifold M, what is the maximal number of linearly independent vector fields that one can find on M?*

Already for the case of the spheres S^m, this deceivingly simple question turns out to be highly non-trivial. The solution (found half way in the 20th century) requires again the machinery of Algebraic Topology. But here is the answer: write $m+1 = 2^{4a+r}m'$ with m' odd, $a \geq 0$ integer, $r \in \{0, 1, 2, 3\}$. Then the maximal number of linearly independent vector fields on S^m is

$$r(m) = 8a + 2^r - 1.$$

Note that $r(m) = 0$ if m is even- and that corresponds to the fact that there are no nowhere vanishing vector fields on even dimensional spheres. The parallelizability of S^m is equivalent to $r(m) = m$, i.e. $8a + 2^r = 2^{4a+r}m'$, which is easily seen to have the only solutions

$$m = 1, \quad a = 0, \quad r \in \{0, 1, 2, 3\},$$

giving $m = 0, 2, 4, 8$. Hence again the spheres S^0, S^1, S^3 and S^7.

4.2. Push-forwards of vector fields. We now discuss how vector fields can be pushed forward from one manifold to another. The best scenario is when we start with a diffeomorphism $F : M \to N$ between two manifolds; then induces a push-forward operation:

$$F_* : \mathfrak{X}(M) \to \mathfrak{X}(N), \quad X \mapsto F_*(X)$$

where, for $X \in \mathfrak{X}(M)$, $F_*(X) \in \mathfrak{X}(N)$ is defined by describing how it acts on functions $f \in \mathcal{C}^\infty(N)$: one first moves f to N via F, i.e. consider $f \circ F \in \mathcal{C}^\infty(M)$, then one acts by X, i.e. consider $L_X(f \circ F) \in \mathcal{C}^\infty(M)$, then one moves back to N to obtain $L_X(f \circ F) \circ F^{-1} \in \mathcal{C}^\infty(N)$. One checks right away that this operation on X is a derivation, hence there exists a unique vector field $F_*(X) \in \mathfrak{X}(N)$ such that

$$L_{F_*(X)}(f) = L_X(f \circ F) \circ F^{-1} \quad \text{for all } f \in \mathcal{C}^\infty(N).$$

A more explicit description of $F_*(X)$ (that gives $F_*(X)_q \in T_qN$ for each $q \in N$) is obtained using the formula $L_Y(g)(p) = (dg)_p(Y_p)$ for vector fields Y and smooth functions g.

Exercise 4.18. Deduce that, for all $q \in N$,

$$F_*(X)_q = (dF)_p(X_p) \quad \text{where } p = F^{-1}(q). \tag{4.4}$$

Also show that, if $G : N \to P$ is another diffeomorphism, then $(G \circ F)_* = G_* \circ F_*$.

Proposition 4.18. *The push forward operation preserves the Lie bracket of vector fields: for any $X_1, X_2 \in \mathfrak{X}(M)$, one has*

$$F_*([X_1, X_2]) = [F_*(X_1), F_*)(X_2)].$$

Proof. We start with the definition of $[F_*(X_1), F_*)(X_2)]$:

$$L_{[F_*(X_1),F_*)(X_2)]}(f) = L_{F_*(X_1)}\left(L_{F_*)(X_2)}(f)\right) - L_{F_*(X_2)}\left(L_{F_*)(X_1)}(f)\right)$$

(for all $f \in \mathcal{C}^\infty(N)$, in which we plug in the definition of $F_*(X_i)$ to obtain

$$L_{F_*(X_1)}\left(L_{X_2}(f \circ F) \circ F^{-1}\right) - \text{ the similar one} = L_{X_1}\left(L_{X_2}(f \circ F)\right) \circ F - \text{ the similar one };$$

using again the formula defining $[X_1, X_2]$, and then the one for F_*, the last expression is

$$L_{[X_1,X_2]}(f \circ F) \circ F^{-1} = L_{F_*([X_1,X_2])}(f).$$

□

Exercise 4.19. By similar arguments defined also a pull-back operation $F^* : \mathfrak{X}(N) \to \mathfrak{X}(M)$. And show that this makes sense not only for diffeomorphisms $F : M \to N$, but for all local diffeomorphisms. Also: is it true that $(G \circ F)^* = G^* \circ F^*$?

The construction of push-forwards cannot be extended for arbitrary smooth maps $F : M \to N$. Instead, one can talk about a vector field $X \in XX(M)$ being "related via F" to a vector field $Y \in \mathfrak{X}(N)$:

Definition 4.19. *Given a smooth map $F : M \to N$, we say that $X \in \mathfrak{X}(M)$ is F-**projectable** to $Y \in \mathfrak{X}(N)$ if*

$$(dF)_p(X_p) = Y_{F(p)} \quad \text{for all } p \in M.$$

Note that, when F is a diffeomorphism, this corresponds precisely to $Y = F_*(X)$. However, in general, there may be different vector fields on M that are F-projectable to the same vector field on N (think e.g. of the case when N is a point) or there may be vector fields on M that are not projectable to any vector field on N. However, Proposition 4.18 still has a version that applies to this general setting:

Exercise 4.20. *For any smooth map $F : M \to N$, if $X_i \in \mathfrak{X}(M)$ are F-projectable to $Y_i \in \mathfrak{X}(N)$ for $i \in \{1, 2\}$, then $[X_1, X_2]$ is F-projectable to $[Y_1, Y_2]$. (Hint: proceed locally and remember the formula that you obtained when you solved Exercise 4.14).*

5. The Lie algebra of a Lie group

We now discuss the infinitesimal counterpart of Lie groups.

Definition 4.20. *A Lie algebra is a vector space $\mathfrak{a}$ endowed with an operation*

$$[\cdot, \cdot] : \mathfrak{a} \times \mathfrak{a} \to \mathfrak{a}$$

which is bi-linear, antisymmetric and which satisfies the Jacobi identity:

$$[[u, v], w] + [[v, w], u] + [[w, u], v = 0 \quad \text{for all } u, v, w \in \mathfrak{a}.$$

Example 4.21 (commutators of matrices). The space $\mathcal{M}_n(\mathbb{R})$ of $n \times n$ matrices, together with the commutator of matrices,

$$[A, B] := AB - BA$$

is a Lie algebra. More generally, for any vector space V, the space $\mathrm{Lin}(V, V)$ of linear maps from V can be made into a Lie algebra by the same formula (just that AB becomes the composition of A and B).

Example 4.22 (vector fields). As we have already pointed out, for any manifold M, the space of vector fields $\mathfrak{X}(M)$ endowed with the Lie bracket of vector fields becomes a (infinite dimensional) Lie algebra.

Exercise 4.21. What is the relationship between the previous two examples?

We now explain how a Lie group G gives rise to a finite dimensional Lie algebra $\mathfrak{g}$. As a linear space, it is defined as the tangent space of G at the identity element $e \in G$:

$$\mathfrak{g} := T_eG.$$

For the Lie bracket, we will related $\mathfrak{g}$ to vector fields on G, and then use the Lie bracket of vector fields. Hence, as a first step, we note that any $v \in \mathfrak{g}$ gives rise to a vector field

$$\vec{v} \in \mathfrak{X}(G).$$

defined as follows. For $g \in G$, to define $\vec{v}_g \in T_gG$, we use the left translation by g:

$$L_g : G \to G, \quad L_g(h) = gh.$$

This is smooth (and even a diffeomorphism!) and make use of its differential at e, $(dL_g)_e : \mathfrak{g} \to T_gG$. Define then

$$\vec{v}_g := (dL_g)_e(v) \in T_gG.$$

EXERCISE 4.22. Show that, indeed, $\vec{v}$ is smooth.

PROPOSITION 4.23. *With the same notations as above (a Lie group G and $\mathfrak{g} := T_eG$), for any $u, v \in \mathfrak{g}$, the Lie bracket $[\vec{u}, \vec{v}] \in \mathfrak{X}(G)$ is again of type $\vec{w}$ for some $w \in \mathfrak{g}$. Denoting the resulting w by $[v, w] \in \mathfrak{g}$, we obtain an operation*

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

that makes $\mathfrak{g}$ into a Lie algebra.

PROOF. Note that the map

$$\mathrm{j} : \mathfrak{g} \to \mathfrak{X}(G), \quad v \mapsto \vec{v} \tag{5.1}$$

is injective. That is simply because $\vec{v}_e = v$. This implies the uniqueness of w; and it also indicates how to define it:

$$w := [\vec{u}, \vec{v}]_e.$$

We still have to show that $[\vec{u}, \vec{v}] = \vec{w}$. For that we note that

$$(L_a)_*(\vec{v}) = \vec{v}$$

for all $a \in G$ and all $v \in \mathfrak{g}$, where we use the push-forward as defined in subsection 4.2. Indeed, using the formula (4.4) for the push-forward and then the definition of $\vec{v}$:

$$(L_a)_*(\vec{v})_g = (dL_a)_{a^{-1}g}\left((dL_{a^{-1}g})_e(v)\right)$$

and then, using the chain rule and $L_a \circ L_{a^{-1}g} = L_g$, we end up with $\vec{v}_g$.

Since $(L_a)_*(\vec{v}) = \vec{v}$ and similarly for u, Proposition 4.18 implies that $[\vec{u}, \vec{v}] = (L_a)_*([\vec{u}, \vec{v}])$. Evaluating this at $a \in G$ and using again formula formula (4.4) (where $F = L_a, q = a$ so that $p = L_a^{-1}(a) = e$):

$$[\vec{u}, \vec{u}]_a = (dL_a)_e([\vec{u}, \vec{u}]_e) = (dL_a)_e(w_e) = \vec{w}_a \ .$$

To fact that the resulting operation $[\cdot, \cdot]$ on $\mathfrak{g}$ is a bilinear and skew-symmetric is immediate. For the Jacobi identity, since the map j from (5.1) is injective and satisfies $j([v, w]) = [j(v), j(w)]$, it suffices to remember that the Lie bracket of vector fields does satisfy the Jacobi identity. □

DEFINITION 4.24. ***The Lie algebra of the Lie group** G is defined as $\mathfrak{g} = T_eG$ endowed with the Lie algebra structure from the previous proposition.*

We now compute some examples.

EXAMPLE 4.25. *We start with the general linear group $GL_n(\mathbb{R})$, and we claim that its Lie algebra is just the algebra $\mathcal{M}_n(\mathbb{R})$ of $n \times n$ matrices endowed with the commutator bracket.*

PROOF. As we already mentioned, $\mathcal{M}_n(\mathbb{R})$ is seen as an Euclidean space; we denote coordinates functions by

$$x^i_j : \mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$$

(sending a matrix to the element on the position (i, j)). We also remarked already that, while $GL_n = GL_n(\mathbb{R})$ sits openly inside $\mathcal{M}_n(\mathbb{R})$, its tangent space at the identity matrix I (and similarly at any point) is canonically identified with $\mathcal{M}_n(\mathbb{R})$: any $X \in \mathcal{M}_n(\mathbb{R})$ is identified with the speed at $t = 0$ of $t \mapsto (I + tX)$. After left translating, we find that the corresponding vector field $\vec{X} \in \mathfrak{X}(GL_n)$ is given, at an arbitrary point $A \in GL_n$, by

$$\vec{X}_A = \frac{d}{dt}\bigg|_{|t=0} A \cdot (I + tX) \in T_AGL_n. \tag{5.2}$$

Let $X, Y \in \mathcal{M}_n(\mathbb{R})$ and let $[X, Y] \in \mathcal{M}_n(\mathbb{R})$ be the Lie algebra bracket- hence defined by $\overrightarrow{[X, Y]} = [\vec{X}, \vec{Y}]$, where the last bracket is the Lie bracket of vector fields on GL_n. To compute it, we use the defining equation for the Lie bracket of two vector fields:

$$L_{[\vec{X},\vec{Y}]}(F) = L_{\vec{X}}L_{\vec{Y}}(F) - L_{\vec{Y}}L_{\vec{X}}(F) \tag{5.3}$$

for all smooth functions

$$F : GL_n \to \mathbb{R}.$$

From the previous description of $\vec{X}$ we deduce that

$$L_{\vec{X}}(F)(A) = \frac{d}{dt}\bigg|_{t=0} F(A \cdot (I + tX)) = \sum_{i,j,k} \frac{\partial F}{\partial x^i_j}(A)\, A^i_k X^k_j.$$

In particular, applied to a coordinate function $F = x^i_j$, we find

$$L_{\vec{X}}(x^i_j)(A) = \sum_k A^i_k X^k_j$$

or, equivalently, $L_{\vec{X}}(x^i_j) = \sum_k x^i_k X^k_j$. Therefore

$$L_{\vec{Y}}(L_{\vec{X}}(x^i_j)) = \sum_k L_{\vec{Y}}(x^i_k) X^k_j = \sum_{k,l} x^i_l Y^l_k X^k_j = \sum_l x^i_l (Y \cdot X)^l_j,$$

which is precisely $L_{\overrightarrow{Y \cdot X}}(x^i_j)$. Since this holds for all coordinate functions, we have

$$L_{\vec{Y}} \circ L_{\vec{X}} = L_{\overrightarrow{Y \cdot X}};$$

we deduce that $[\vec{X}, \vec{Y}] = \overrightarrow{Y \cdot X} - \overrightarrow{X \cdot Y}$ and then that $[X, Y]$ is $X \cdot Y - Y \cdot X$, i.e. the usual commutator of matrices. □

EXAMPLE 4.26. In section 11 of Chapter 3 we have looked at several classical Lie groups $G \subset GL_n$ and we have computed their tangent spaces at the identity matrix I, as subspaces of $T_I GL_n = \mathcal{M}_n(\mathbb{R})$:

- for $O(n)$ and $SO(n)$ we obtained
$$o(n) = \{X \in \mathcal{M}_n(\mathbb{R}) : X + X^T = 0\}.$$
- for $SL_n(\mathbb{R})$ we obtained
$$sl_n(\mathbb{R}) = \{A \in \mathcal{M}_n(\mathbb{R}) : Tr(X) = 0\}.$$
- $U(n)$ we obtained
$$u(n) = \{X \in \mathcal{M}_n(\mathbb{C}) : X + X^* = 0\}.$$
- for $SU(n)$ we obtained
$$su(n) = \{X \in \mathcal{M}_n(\mathbb{C}) : X + X^* = 0, Tr(X) = 0\}.$$

We claim that the resulting Lie brackets on these subspaces of $\mathcal{M}_n(\mathbb{R})$ are still given by the commutators of matrices (just for fun, you should check that, indeed, all these subspaces are closed under taking commutators).

Of course, all these examples can be treated at once: given $G \subset GL_n$ which is a Lie subgroup (i.e. a subgroup and an embedded sub manifold), the resulting Lie algebra structure on $\mathfrak{g} = T_I G \subset T_I GL_n = \mathcal{M}_n(\mathbb{R})$ is just the commutator bracket.

Proof. Note that for $X \in \mathfrak{g}$ we have two induced vector fields: one on G, denoted as above by $\overrightarrow{X} \in \mathfrak{X}(G)$ and then, since $\mathfrak{g} \subset \mathcal{M}_n(\mathbb{R})$ we will have a similar one on GL_n; to distinguish the two we will denote the second one by $\overset{\rightarrowtail}{X} \in \mathfrak{X}(GL_n)$. Note that, denoting by $i : G \hookrightarrow GL_n$ the inclusion, we are in the situation of Definition 4.19: $\overrightarrow{X}$ is i-projectable to $\overset{\rightarrowtail}{X}$. Hence, by Exercise 4.20, $[\overrightarrow{X}, \overrightarrow{Y}]$ is i-projectable to $[\overset{\rightarrowtail}{X}, \overset{\rightarrowtail}{Y}]$ (for all $X, Y \in \mathfrak{g}$). By the previous example, the last expression is $\overset{\rightarrowtail}{[X, Y]}_{\text{comm}}$ where, for clarity, we denote by $[X, Y]_{\text{comm}}$ the commutator bracket of the matrices X and Y. Hence, for the Lie bracket of X and Y as elements of the Lie algebra $\mathfrak{g}$ of G we find

$$[X, Y] = [\overset{\rightarrowtail}{X}, \overset{\rightarrowtail}{Y}]_I = \left(\overset{\rightarrowtail}{[X, Y]}_{\text{comm}}\right)_I = [X, Y]_{\text{comm}}.$$

□

Remark 4.27 (for the extra-curious students). The Lie algebra $\mathfrak{g}$ of a Lie group G contains (almost) all the information about G! This may sound surprising at first, since $\mathfrak{g}$ is only "the linear approximation" of G and only around the identity $e \in G$. The fact that it is enough to consider only the identity e can be explained by the fact that, using the group structure (and the resulting left translations) one can move around from e to any other element in g. The fact that "the linear approximation" contains (almost) all the information is more subtle and the real content of this is to be found on the Lie bracket structure of $\mathfrak{g}$.

And here is what happens precisely:

- As we have seen, any Lie group has an associated (finite dimensional) Lie algebra.
- Any finite dimensional Lie algebra $\mathfrak{g}$ comes from a Lie group.
- For any finite dimensional Lie algebra $\mathfrak{g}$ there exists and is unique a Lie group $G(\mathfrak{g})$ whose Lie algebra is $\mathfrak{g}$ and which is 1-connected (i.e. both connected as well as simply-connected). The Lie group $G(\mathfrak{g})$ can be constructed explicitly out of $\mathfrak{g}$. And any other connected Lie group which has $\mathfrak{g}$ as Lie algebra is a quotient of $G(\mathfrak{g})$ by a discrete subgroup.
- Explanation for the 1-connectedness condition: starting with a Lie group G, the connected component of the identity element, is itself a Lie group with the same Lie algebra as G. And so is the universal cover of G. I.e. one can always replace G by a Lie group that is 1-connected without changing its Lie algebra.
- the construction $G \mapsto \mathfrak{g}$ gives an equivalence between (the category of) 1-connected Lie groups and (the category of) finite dimensional Lie algebras. At the level of morphisms even more is true: while a morphism of Lie groups $F : G \to H$ induces a morphism of the corresponding Lie algebras, $f : \mathfrak{g} \to \mathfrak{h}$, if G is 1-connected (but H is arbitrary), then F can be recovered from f (i.e.: any morphism f of Lie algebras comes from an unique morphism F of Lie groups!). That is remarkable since f is just the linearization of F at the identity $f = (dF)_e$.

6. Integral curves (and flows)

The brief philosophy is:

If your world is a manifold M and your are sitting at a point p, then a tangent vector $X_p \in T_pM$ gives you a direction in which you can start walking. A vector field however gives you an entire path to walk on.

In the ideal situation (no friction etc, no opposition from your side), and if you think of a vector field as a wind, its integral curve at p is the trajectory that you will follow, blown by the wind.

DEFINITION 4.28. *Given a vector field $X \in \mathfrak{X}(M)$,* ***an integral curve of*** *X is any curve $\gamma : I \to M$ defined on some open interval $I \subset \mathbb{R}$ ($I = \mathbb{R}$ not excluded) such that*

$$\frac{d\gamma}{dt}(t) = X_{\gamma(t)} \quad \textit{for all } t \in I.$$

We say that γ ***starts at*** *p if $0 \in I$ and $\gamma(0) = p$.*

EXAMPLE 4.29. Consider $M = \mathbb{R}^2$ and the vector field X given by

$$X = \frac{\partial}{\partial x} + \frac{\partial}{\partial y}.$$

Then a curve in $\mathbb{R}^2$, written as $\gamma(t) = (x(t), y(t))$, is an integral curve if and only if

$$\dot{x}(t) = 1, \quad \dot{y}(t) = 1.$$

Hence we find that the integral curves of X are those of type

$$\gamma(t) = (t + a, t + b)$$

with $a, b \in \mathbb{R}$ constants. Prescribing the starting point of γ is the same thing as fixing the constants a and b (and then we have an unique integral curve).

A more interesting vector field is:

$$X = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}.$$

Then the equations to solve are $\dot{x} = y$ and $\dot{y} = -x$. Hence $\ddot{x} = -x$ from which one can deduce (or guess, and then use the uniqueness recalled below) that

$$x(t) = a\cos t + b\sin t, \quad y(t) = b\cos t - a\sin t,$$

for some (arbitrary) constants $a, b \in \mathbb{R}$ (and, again, prescribing the starting point of γ is the same thing as fixing the constants a and b).

Note that in both examples the integral curves that we were obtaining are defined on the entire $I = \mathbb{R}$. However, this need not always be the case. For instance, for

$$X = -x^2\frac{\partial}{\partial x} - y\frac{\partial}{\partial y},$$

starting with any constants $a, b \in \mathbb{R}$ one has the integral curve given by

$$\gamma(t) = (\frac{a}{at+1},\ be^{-t});$$

but, for $a > 0$, the largest interval containing 0 on which this is defined is $(-\frac{1}{a}, \infty)$. However, still as above, for any point $p = (a, b) \in \mathbb{R}^2$, one finds an integral curve that starts at p.

Of course, in the previous examples, the existence of integral curves starting at a given (arbitrary) point is no accident:

PROPOSITION 4.30 (local existence and uniqueness). *Given a vector field $X \in \mathfrak{X}(M)$:*

1. *for any $p \in M$ there exists an integral curve $\gamma : I \to M$ of X that starts at p.*
2. *any two integral curves of X that coincide at a certain time t_0 must coincide in a neighborhood of t_0.*

Since this is a local statement, we just have to look in a chart. Then, as we shall see, what we end up with is a system of ODEs and the previous proposition becomes the standard local existence/uniqueness result for ODEs. Here are the details. Fix a chart $\chi : U \to \Omega \subset M$ of M around p. Then, as in (4.1), X can be written on U as

$$X_p = \sum_{i=1}^{m} X^i_\chi(\chi(p)) \left(\frac{\partial}{\partial \chi_i}\right)_p$$

with the coordinate functions X^i_χ of X w.r.t. χ being smooth functions on Ω. Also, a curve γ in U becomes, via χ, a curve γ_χ in Ω. Since

$$\gamma(t) = \chi^{-1}(\gamma_\chi(t)),$$

its derivatives becomes

$$\frac{d\gamma}{dt}(t) = \frac{d\gamma^i_\chi}{dt}(t) \left(\frac{\partial}{\partial \chi_i}\right)_{\gamma(t)}.$$

Hence the integral curve condition becomes the system of ODEs:

$$\frac{d\gamma^i_\chi}{dt}(t) = X^i_\chi(\gamma^1_\chi(t), \ldots, \gamma^m_\chi(t)) \quad i \in \{1, \ldots, m\}$$

or, more compactly,

$$\frac{d\gamma_\chi}{dt}(t) = X_\chi(\gamma_\chi(t)),$$

where $X^\chi = (X^1_\chi, \ldots, X^m_\chi) : \Omega \to \mathbb{R}^m$. And here is the standard result on such ODEs:

THEOREM 4.31. *Let $F : \Omega \to \mathbb{R}^m$ be a smooth function. Then, for any $x \in \Omega$, the following ordinary differential equation with initial condition:*

$$\frac{d\gamma}{dt}(t) = F(\gamma(t)), \quad \gamma(0) = x \tag{6.1}$$

has a solution γ defined on an open interval containing $0 \in \mathbb{R}$, and any two such solutions must coincide in a neighborhood of 0.

Furthermore, for any $x_0 \in \Omega$ there exists an open neighborhood Ω_0 of x_0 in Ω, $\epsilon > 0$ and a smooth map

$$\phi : (-\epsilon, \epsilon) \times \Omega_{x_0} \to \Omega,$$

such that, for any $x \in \Omega_0$, $\phi(\cdot, x) : (-\epsilon, \epsilon) \to \Omega$ is a solution of (6.1).

We see that this the first part of the theorem immediately implies (well, it is actually the same as) the previous proposition. For a more global version of the proposition, we have to look at maximal integral curves.

DEFINITION 4.32. *Given a vector field $X \in \mathfrak{X}(M)$, a* ***maximal integral curve*** *of X is any integral curve $\gamma : I \to M$ which admits no extension to a strictly larger interval $\tilde{I}$ and which is still an integral curve of X.*

With this, the local result implies quite easily the following one, with a more global flavour.

COROLLARY 4.33. *Given a vector field $X \in \mathfrak{X}(M)$, for any $p \in M$ there exists an unique maximal integral curve of X,*

$$\gamma_p : I_p \to M,$$

that starts at p.

PROOF. The main remark is that if γ_1 and γ_2 are two integral curves defined on intervals I_1 and I_2, respectively, then the set I of points in $I_1 \cap I_2$ on which the two coincide is closed in $I_1 \cap I_2$ (because it is defined by an equation) and open by the second part of the previous proposition. Hence, since $I_1 \cap I_2$ is connected, if $I \neq \emptyset$ (i.e. the two coincide at some t), then $I = I_1 \cap I_2$ (i.e. the two must coincide on their common domain of definition). Therefore, to obtain I_p and γ_p we can just put together all the integral curves that start at p. □

The best possible scenario is:

DEFINITION 4.34. *We say that $X \in \mathfrak{X}(M)$ is **complete** if all the maximal integral curves are defined on the entire $\mathbb{R}$.*

As we shall see below, all the vector fields that are compactly supported (i.e. which vanish outside a compact subset of M) are complete. In particular, when M is compact, all vector fields on M are complete.

We now put together all the maximal integral curves in one object: the flow of X. First of all, we define the domain of the flow of X as

$$\mathcal{D}(X) := \{(p,t) \in M \times \mathbb{R} : t \in I_p\} \subset M \times \mathbb{R}$$

and then **the flow of** X is defined as the resulting map

$$\phi_X : \mathcal{D}(X) \to M, \quad \phi_X(p,t) := \gamma_p(t).$$

For complete vector fields one has $\mathcal{D}(X) = M \times \mathbb{R}$. What we can say in general is that

$$M \times \{0\} \subset \mathcal{D}(X) \subset M \times \mathbb{R}.$$

and $\mathcal{D}(X)$ is open in $M \times \mathbb{R}$. Actually, applying the second part of the Theorem 4.31, we deduce:

COROLLARY 4.35. *For any $X \in \mathfrak{X}(M)$, the domain $\mathcal{D}(X)$ of its flow is open in $M \times \mathbb{R}$ and the flow ϕ_X is a smooth map.*

One of the main uses of the flow of a vector field X comes from the maps one obtains whenever one fixes a time t; we will use the notation:

$$\phi_X^t(\cdot) := \phi_X(\cdot, t)$$

(whenever the right hand side is defined). The best scenario is when X is complete, when each ϕ^t will be defined on the entire M:

THEOREM 4.36. *If $X \in \mathfrak{X}(M)$ is complete then $\phi_X^t : M \times M$ satisfy:*

$$\phi_X^t \circ \phi_X^s = \phi_X^{t+s}, \quad \phi_X^0 = Id$$

(for all $t, s \in \mathbb{R}$). In particular, each ϕ_X^t is a diffeomorphism and

$$\left(\phi_X^t\right)^{-1} = \phi_X^{-t}.$$

PROOF. Fixing $x \in M$ and $s \in \mathbb{R}$, the two curves

$$t \mapsto \phi_X^{t+s}(p), \quad t \mapsto \phi_X^t(\phi_X^s(p))$$

are both integral curves of X and they coincide at $t = 0$- hence they coincide everywhere. That ϕ_X^0 is the identity is clear; we see that the last part follows by taking $s = -t$ in the first part. □

For general (possibly non-complete) vector fields one has to be a bit more careful.

DEFINITION 4.37. *For $X \in \mathfrak{X}(M)$ and $t \in \mathbb{R}$, define **the flow of X at time** t as the map*

$$\phi_X^t : \mathcal{D}_t(X) \to M, \quad p \mapsto \phi_X^t(p) := \phi_X(p,t) = \gamma_p(t),$$

defined on:

$$\mathcal{D}_t(X) := \{p \in M : t \in I_p\}.$$

THEOREM 4.38. *For any $X \in \mathfrak{X}(M)$:*

(1) *for any $t \in \mathbb{R}$, the domain $\mathcal{D}_t(X)$ of ϕ_X^t is open in M, ϕ_X^t takes values in $\mathcal{D}_{-t}(X)$, and*
$$\phi_X^t : \mathcal{D}_t(X) \to \mathcal{D}_{-t}(X)$$
is a diffeomorphism.

(2) *For $t, s \in \mathbb{R}$ one has*
$$\phi^t \circ \phi^s = \phi^{t+s}$$
in the sense that, for each $p \in M$ on which the left hand side is defined, also $\phi^{t+s}(p)$ is defined and the two expressions coincide.

PROOF. The fact that $\mathcal{D}_t(X)$ is open and ϕ_X^t is smooth follows right away from the previous the standard local result (Theorem 4.31). For the last part we consider again the two curves from the proof of Theorem 4.36; this time we have to be more careful with their domain of definition. Hence we fix $p \in M$, $s \in I(p)$ and we look at the curves:
$$a : -s + I(p) \to M, \quad a(t) = \phi_X^{t+s}(p),$$
$$b : I(\phi_X^s(p)) \to M, \quad b(t) = \phi_X^t(\phi_X^s(p)).$$
Note that both a and b are integral curves of X, and their are both maximal (the second one is maximal from its very definition; for the first one, if $\tilde{a}$ was an extension to $J \supset -s + I(p)$, then $s + J \ni t \mapsto \tilde{a}(t - s)$ would be an integral curve defined on $s + J \supset I(p)$, and starting at p). We then obtain the last part of the theorem together with the equality
$$I(\phi_X^s(p)) = I(p) - s$$
for all s. Since $0 \in I(p)$ we find that $-s \in I(\phi_X^s(p))$, hence
$$\phi_X^s(\mathcal{D}_s) \subset \mathcal{D}_{-s}$$
and, for $p \in \mathcal{D}_s$, $\phi_X^{-s}(\phi_X^s(p)) = p$. We still have to show that, in the previously centered inclusion, equality holds. Using the inclusion for $-s$ instead of s and applying ϕ_X^s we obtain
$$\phi_X^s\left(\phi_X^{-s}(\mathcal{D}_{-s})\right) \subset (\phi_X^s(\mathcal{D}_s);$$
but the first term equals to $\phi_X^0(\mathcal{D}_s) = \mathcal{D}_s$, hence we obtain the reverse inclusion and this finishes the proof. ☐

THEOREM 4.39. *If M is compact then any vector field $X \in \mathfrak{X}(M)$ is complete.*
More generally, for any M, any $X \in \mathfrak{X}(M)$ that is compactly supported (i.e. is zero outside some compact subset of M) is complete.

PROOF. When M is compact, since $\mathcal{D}(X)$ is an open in $M \times \mathbb{R}$ containing $M \times \{0\}$, the tube lemma implies that there exists $\epsilon > 0$ such that:
$$M \times (-\epsilon, \epsilon) \subset \mathcal{D}(X). \tag{6.2}$$
(proof: for any $p \in M$, using that ϕ_X is continuous at $(p, 0)$ and that $\mathcal{D}(X)$ is open, find $\epsilon_p > 0$ and U_p open containing p with $U_p \times (-\epsilon_p, \epsilon_p) \subset \mathcal{D}(X)$. Then $\{U_p : p \in M\}$ is an open cover of M hence, by compactness, M is covered by a finite number of them, say the ones corresponding to $p_1, \ldots, p_k \in M$; set ϵ as the smallest of the epsilons corresponding to the points p_i.)

We claim that the existence of ϵ such that the inclusion (6.2) holds implies (without further using the compactness of M!) that X is complete. Indeed, we would have that $\phi_X^t(p)$ is defined for all $t \in (-\epsilon, \epsilon)$ and all $p \in M$. But then so would be $\phi_X^t(\phi_X^t(p))$, hence $\phi_X^{2t}(p)$, hence (6.2) holds also with 2ϵ instead of ϵ. Repeating the process, we find that $\mathbb{R} \times M \subset \mathcal{D}(X)$, hence X must be complete.

For the last part (M general and X is compactly supported) it is enough to remark that an inclusion of type (6.2) can still be achieved: if X is zero outside the compact K then the tube lemma implies that $(-\epsilon, \epsilon) \times K \subset \mathcal{D}(X)$ for some $\epsilon > 0$, while all the points (t, p) with $p \notin K$ are clearly in $\mathcal{D}(X)$ since X vanishes at p (hence the integral curve of X starting at p is simply $\gamma(t) = p$, defined for all t). □

For latter reference, let us also cast the last part of the argument into:

COROLLARY 4.40. *If $X \in \mathfrak{X}(M)$ has the property that $M \times (-\epsilon, \epsilon) \subset \mathcal{D}(X)$ for some $\epsilon > 0$, then X is complete.*

EXERCISE 4.23. Let $F : M \to N$ be a smooth function. Recall that $X \in \mathfrak{X}(M)$ is said to be F-projectable to $Y \in \mathfrak{X}(N)$ if

$$(dF)_p(X_p) = Y_{F(p)} \quad \text{for all } p \in M.$$

If X is F-projectable to Y show that, for any $t \in \mathbb{R}$ one has

$$F(\mathcal{D}_t(X)) \subset \mathcal{D}_t(Y)$$

and, one $\mathcal{D}_t(X)$,

$$F \circ \phi_X^t = \phi_Y^t \circ F.$$

7. Lie derivatives along vector fields

The general philosophy of taking Lie derivatives L_X along vector fields is rather simple:

for any "type of objects" ξ on manifolds M (functions, vector fields, etc etc), which are natural (in the sense that a diffeomorphism $\phi : M \to N$ allows one pull-back the objects from N to ones on M), the Lie derivative $\mathcal{L}_X(\xi)$ of ξ along X measures the variation of ξ along the flow of X:

$$\mathcal{L}_X(\xi) := \left.\frac{d}{dt}\right|_{t=0} (\phi_X^t)^*(\xi). \tag{7.1}$$

This is a very general principle that is applied over and over in Differential Geometry, depending on the "type of objects" one is interested in. Here we illustrate this principle for functions and vector fields, and compute the outcome.

For functions, there is an obvious pull-back operation associated to any smooth function $F : M \to N$ (diffeomorphism or not):

$$F^* : \mathcal{C}^\infty(N) \to \mathcal{C}^\infty(M), \quad F^*(f) = f \circ F.$$

Therefore, when $F = \phi_X^t : M \to M$ is the flow of a complete vector field, the guiding equation (7.1) for $\xi = f \in \mathcal{C}^\infty(M)$ makes sense pointwise as

$$\mathcal{L}_X(f)(p) := \left.\frac{d}{dt}\right|_{t=0} f(\phi_X^t(p)). \tag{7.2}$$

Even more, written in this way, the completeness of X is not even necessary for this formula to make sense (indeed, we only need $\phi_X^t(p)$ to be defined for t near 0, which is always the case).

PROPOSITION 4.41. *For $X \in \mathfrak{X}(M)$ and $f \in \mathcal{C}^\infty(M)$, $\mathcal{L}_X(f)$ defined by (7.2) is precisely $L_X(f) \in \mathcal{C}^\infty(M)$.*

PROOF. The expression in the right hand side of (7.2) is precisely

$$(df)_p \left(\left.\frac{d}{dt}\right|_{t=0} \phi_X^t(p) \right) = (df)_p(X_p) = L_X(f)(p).$$

□

The discussion is similar for vector fields. In this case, a diffeomorphism $F : M \to N$ allows one to pull-back vector fields on N to vector fields on M,

$$F^* : \mathfrak{X}(N) \to \mathfrak{X}(M)$$

where, for $Y \in \mathfrak{X}(N)$, its pull-back $F^*(Y) \in \mathfrak{X}(M)$ is defined by

$$F^*(Y)_p := (dF^{-1})_{F(p)}(Y_{F(p)}).$$

With this, when $F = \phi_X^t : M \to M$ is the flow of a complete vector field, the guiding equation (7.1) for $\xi = Y \in \mathfrak{X}(M)$ makes sense pointwise as

$$\mathcal{L}_X(Y)(p) := \left.\frac{d}{dt}\right|_{t=0} (d\phi_X^{-t})_{\phi_X^t(p)}(Y_{\phi_X^t(p)}). \tag{7.3}$$

And, as for functions, this formula makes sense without any completeness assumption on X.

Proposition 4.42. *For $X \in \mathfrak{X}(M)$ and any other $Y \in \mathfrak{X}(M)$, $\mathcal{L}_X(f)$ defined by (7.3) is precisely the Lie bracket $[X, Y] \in \mathfrak{X}(M)$.*

Proof. Let $Z = \mathcal{L}_X(Y)$. For $f \in \mathcal{C}^\infty(M)$ we compute $L_Z(f)(p) = (df)_p(Z_p)$ and we find

$$\left.\frac{d}{dt}\right|_{t=0} (df)_p \left((d\phi_X^{-t})_{\phi_X^t(p)}(Y_{\phi_X^t(p)})\right) = \left.\frac{d}{dt}\right|_{t=0} d\left(f \circ \phi_X^{-t}\right)_{\phi_X^t(p)} \left(Y_{\phi_X^t(p)}\right).$$

Note that $f \circ \Phi_X^{-t}$ is f at $t = 0$ and has the derivative w.r.t. t at $t = 0$ equal to $-L_X(f)$, hence

$$f \circ \Phi_X^{-t} = f - tL_X(f) + t^2 \cdot ?,$$

where ? is smooth. Continuing the computation started above, we find

$$\left.\frac{d}{dt}\right|_{t=0} (df - td\,(L_X(f)))_{\phi_X^t(p)} \left(Y_{\phi_X^t(p)}\right) = \left.\frac{d}{dt}\right|_{t=0} (df)_{\phi_X^t(p)} \left(Y_{\phi_X^t(p)}\right) - (dL_X(f))\,(Y_p) =$$

$$\left.\frac{d}{dt}\right|_{t=0} L_Y(f)(\phi_X^t(p)) - L_Y(L_X(f))(p) = L_X(L_Y(f))(p) - L_Y(L_X(f))(p),$$

i.e. $L_Z(f)(p) = L_{[X,Y]}(f)(p)$ for all p and f. This implies that $Z = [X, Y]$. □

We obtain the following characterization of the commutation relation $[X, Y] = 0$; for simplicity, we restrict here to the case of complete vector fields.

Corollary 4.43. *For $X, Y \in \mathfrak{X}(M)$ complete, one has:*

$$[X, Y] = 0 \iff \Phi_X^t \circ \Phi_Y^s = \Phi_Y^s \circ \Phi_X^t \quad \forall\, t, s \in \mathbb{R}.$$

Exercise 4.24. Using the previous Proposition and Exercise 4.23 prove again that if $F : M \to N$ is smooth and $X_1, X_2 \in \mathfrak{X}(M)$ are F-projectable to Y_1 and $Y_2 \in \mathfrak{X}(N)$, respectively, then $[X_1, X_2]$ is F-projectable to $[Y_1, Y_2]$.

8. Lie groups again: the exponential map

Proposition 4.44. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. For any $u \in \mathfrak{g}$, the vector field $\vec{u}$ is complete and its flow $\phi_{\vec{u}}^t$ satisfies*

$$\phi_{\vec{u}}^t(ag) = a\phi_{\vec{u}}^t(g).$$

In particular, the flow can be reconstructed from what it does at time $t = 1$, at the identity element of G, i.e. from

$$exp : \mathfrak{g} \to G, \quad exp(u) := \phi_{\vec{u}}^1(e),$$

by the formula

$$\phi^t_{\vec{u}}(a) = a\, exp(t\, u).$$

Moreover, the exponential is a local diffeomorphism around the origin: it sends some open neighborhood of the origin in $\mathfrak{g}$ diffeomorphically into an open neighborhood of the identity matrix in G.

PROOF. For the first part the key remark is that if $\gamma : I \to G$ is an integral curve of $\vec{u}$ defined on some interval I, i.e. if

$$\frac{d\gamma}{dt}(t) = \vec{u}\,(\gamma(t)) \quad \forall\, t \in I$$

then, for any $a \in G$, the left translate $a \cdot \gamma : t \mapsto a \cdot \gamma(t)$ is again an integral curve (of the same $\vec{u}$); this follows by writing $a \cdot \gamma = L_a \circ \gamma$ and using the chain rule. Hence, if γ_g is the maximal integral curve starting at $g \in G$ then $a \cdot \gamma_g$ will be the one starting at $a \cdot g$- and that proves the first identity (for as long as both terms are defined). And we also obtain that $I_g = I_{ag}$ for all $a, g \in G$, hence $I_g = I_e$ for all $g \in G$. Using Corollary 4.40, we obtain that $\vec{u}$ is complete.

We are left with the last part. For that it suffices to show that $(d\exp)_0$ is an isomorphism. But

$$(d\exp)_0 : \mathfrak{g} \to T_eG = \mathfrak{g}$$

sends $u \in \mathfrak{g}$ to

$$\left.\frac{d}{dt}\right|_{t=0} \exp(tu) = \left.\frac{d}{dt}\right|_{t=0} \phi^t_{\vec{u}}(e) = \vec{u}_e = u,$$

i.e. the differential is actually the identity map. □

EXAMPLE 4.45. In Example 4.25 we have seen that the Lie algebra of the general linear group $GL_n(\mathbb{R})$ is just the algebra $\mathcal{M}_n(\mathbb{R})$ of $n \times n$ matrices endowed with the commutator bracket. We claim that the resulting exponential map is just the usual exponential map for matrices:

$$\mathrm{Exp} : \mathcal{M}_n(\mathbb{R}) \to GL_n(\mathbb{R}),\ X \mapsto e^X := \sum_{k=0}^{\infty} \frac{X^k}{k!}$$

(where we work with the conventions that $X^0 = I$ the identity matrix and $0! = 1$). Let us compute $\exp(X)$ for an arbitrary $X \in \mathcal{M}_n(\mathbb{R}) = T_I GL_n$; according to the definition, it is $\gamma(1)$, where γ is the unique path in GL_n satisfying

$$\gamma(0) = I,\ \frac{d\gamma}{dt}(t) = \vec{X}\,(\gamma(t)).$$

The last equality takes place in $T_{\gamma(t)}GL_n$, which is canonically identified with $\mathcal{M}_n(\mathbb{R})$ (as in Example 4.25 , $Y \in \mathcal{M}_n(\mathbb{R})$ is identified with the speed at $\epsilon = 0$ of $\epsilon \mapsto \gamma(t) + \epsilon Y$). By this identification, $\frac{d\gamma}{dt}(t)$ goes to the usual derivative of γ (as a path in the Euclidean space $\mathcal{M}_n(\mathbb{R})$) and $\vec{X}\,(\gamma(t))$ goes to $\gamma(t) \cdot X$. Hence γ is the solution of

$$\gamma(0) = I,\ \gamma'(t) = X \cdot \gamma(t),$$

which is precisely what $t \mapsto e^{tX}$ does. We deduce that $\exp(X) = e^X$.

9. For the curious students: application to closed subgroups of GL_n

As an application of the exponential map we present here, for the curious students, a proof of Theorem 3.31 from Chapter 3; actually, of an improved version of it.

THEOREM 4.46. *Any closed subgroup G of GL_n is an embedded Lie subgroup of GL_n (i.e. an embedded submanifold which, together with the resulting smooth structure, becomes a Lie group). Moreover, it Lie algebra is*

$$\mathfrak{g} := \{X \in \mathcal{M}_n(\mathbb{R}) : e^{tX} \in G \ \ \forall \ \ t \in \mathbb{R}\} \subset \mathcal{M}_n(\mathbb{R})$$

(endowed with the commutator bracket).

PROOF. For the first part we have to prove that G is an embedded submanifold of GL_n; for simplicity, assume we work over $\mathbb{R}$. We will show that $\mathfrak{g}$ is a linear subspace of $\mathcal{M}_n(\mathbb{R})$; then we will choose a complement $\mathfrak{g}'$ of $\mathfrak{g}$ in $\mathcal{M}_n(\mathbb{R})$ and we show that one can find an open neighborhood V of the origin in $\mathfrak{g}$, and a similar one V' in $\mathfrak{g}'$, so that

$$\phi : V \times V' \to GL_n, \phi(X, X') = e^X \cdot e^{X'}$$

is a local diffeomorphism onto an open neighborhood of the identity matrix and so that

$$G \cap \phi(V \times V') = \{\phi(v, 0) : v \in V\}. \tag{9.1}$$

This means that the submanifold condition is verified around the identity matrix and then, using left translations, it will hold at all points of G. Note also that the differential of ϕ at 0 (with ϕ viewed as a map defined on the entire $\mathfrak{g} \times \mathfrak{g}'$) is the identity map:

$$(d\phi)_0 : \mathfrak{g} \times \mathfrak{g}' \to T_I GL_n = \mathcal{M}_n(\mathbb{R}), (X, X') \mapsto \frac{d}{dt}_{t=0} e^{tX} e^{tX'} = X + X'.$$

In particular, ϕ is a indeed a local diffeomorphism around the origin. Hence the main condition that we have to take care of is (9.1), i.e. that if $X \in V$, $X' \in V'$ satisfy $e^X \cdot e^{X'} \in G$, then $X' = 0$. Hence it suffices to show that one can find V' so that

$$X' \in V', e^{X'} \in G \Longrightarrow X' = 0$$

(then just choose any V small enough such that ϕ is a diffeomorphism from $V \times V'$ into an open neighborhood of the identity in GL_n). For that we proceed by contradiction. If such V' would not exist, we would find a sequence $Y_n \to 0$ in $\mathfrak{g}'$ such that $e^{Y_n} \in G$. Since Y_n converges to 0, we find a sequence of integers $k_n \to \infty$ such that $X_n := k_n Y_n$ stay in a closed bounded region of $\mathfrak{g}'$ not containing the origin (e.g. on $\{X \in \mathfrak{g}' : 1 \leq ||X|| \leq 2\}$, for some norm on $\mathfrak{g}'$); then, after eventually passing to a subsequence, we may then assume that $X_n \to X \in \mathfrak{g}'$ non-zero. In particular, X cannot belong to $\mathfrak{g}$. Setting $t_n = \frac{1}{k_n}$, we see that $e^{t_n X_n} = e^{Y_n} \in G$, and then the desired contradiction will follow from:

LEMMA 4.47. *Let $X_n \to X$ be a sequence of elements in $\mathfrak{gl}_n$ and $t_n \to 0$ a sequence of non-zero real numbers. Suppose $e^{t_n X_n} \in G$ for all n: then $e^{tX} \in G$ for all t, i.e., $X \in \mathfrak{g}$.*

For a proof of this Lemma, please see the book by Sternberg on Differential Geometry (Lemma 4.2, Chapter V). This Lemma can also be used to prove the claim we made (and used) at the start of the argument: that $\mathfrak{g}$ is a linear subspace of $\mathcal{M}_n(\mathbb{R})$. The $\mathbb{R}$-linearity follows immediately from the definition. Assume now that $X, Y \in \mathfrak{g}$ and we show that $X + Y \in \mathfrak{g}$. Write now

$$e^{tX} e^{tY} = e^{f(t)}, \quad \text{with} \quad \lim_{t \to 0} \frac{f(t)}{t} = X + Y$$

(exercise: why is this possible?). Taking t_n a sequence of real numbers converging to 0 and $X_n = \frac{f(t_n)}{t_n}$ in the lemma, to conclude that $X + Y \in \mathfrak{g}$.

This closes the proof of the fact that G is a submanifold of GL_n (hence also a Lie group). It should now also be clear that $\mathfrak{g}$ is the Lie algebra of G since $\phi|_{V \times \{0\}}$ is precisely the restoration of the exponential map e to the open neighborhood V of the origin in $\mathfrak{g}$. □

EXERCISE 4.25. In particular, it follows that $\mathfrak{g}$ is closed under the commutator bracket of matrices:

$$X, Y \in \mathfrak{g} \Longrightarrow [X, Y] \in \mathfrak{g}.$$

However, prove this directly, using the previous Lemma, by an argument similar to the one we used to prove that $\mathfrak{g}$ is a linear subspace of gl_n.