University of Luxembourg
Faculty of Sciences, Technology and Communication

Geometric Methods in Mathematical Physics

Authors

Norbert PONCIN

in cooperation with

Edite DE OLIVEIRA
Luca NOTARNICOLA
Massimo NOTARNICOLA

This text is based on a lecture course taught by Norbert Poncin at the University of Luxembourg to second year students of the Master in Secondary Education in Mathematics. The first version of the text was written by the students Edite De Oliveira, Luca Notarnicola, and Massimo Notarnicola.

These notes have been written for students who first attended the lectures and they are certainly not a substitute for the lectures.

Contents

Introduction

Two important physical theories appeared during the first half of the last century: General Relativity and Quantum Mechanics. More recent theories are the standard model and superstring theory. The standard model explains all the elementary particles from which we know that they exist, as well as three of the four fundamental interactions, the electromagnetic force, and the weak and the strong nuclear forces. It is however not able to explain the gravity, which is the main actor in General Relativity. This is not surprising, since General Relativity and Quantum Mechanics are notoriously incompatible. Superstring theory is currently often considered as the best candidate for a unified theory incorporating General Relativity *and* Quantum Theory.

Relativity teaches us that the Universe should be thought of as a 4-dimensional Lorentzian manifold, whereas superstring theory claims that it should be a Calabi-Yau manifold, i.e., a kind of 4-dimensional base manifold together with 7-dimensional fibers made of interlaced circles. The first issue of this course is the study of the concept of **fiber bundle** and of **vector bundle**. General Relativity says roughly that 'matter tells space how to curve and that space curvature tells matter how to move'. We thus take also an interest in **curvature** and **torsion** of fiber and vector bundles. Whereas curvature and torsion of a curve in $\mathbb{R}^3$ can easily be defined at the beginning of the Bachelor and depend only on the curve itself, in the case of vector bundles, torsion is defined only for the tangent bundle and both, curvature and torsion, are defined rather abstractly and do not only depend on the bundle but on the choice of a new concept that can appear in three different forms.

One of the key objectives of the course is the investigation of this new notion. If we think of a plane and then of a sphere, we understand easily that **curvature** is related to **parallel transport**. Remembering the Lie derivative, we realize that parallel transport is tightly linked to a concept of **'covariant' derivative**, or, still, to the choice of a **'connection'**. To simplify, we just identify these three notions in this introduction. Their definitions, properties and relationships will be a major issue in this text. The problem is actually that in most situations there is no *canonical* parallel transport or covariant derivative or connection, but there are many of them. In view of what has been said above, the curvature and the torsion (if defined) depend not only on the chosen vector bundle but on the considered parallel transport, covariant derivative or connection, so that we have to study the *curvature and torsion, not of a bundle, but of a covariant derivative.* In the case of a standard curve of $\mathbb{R}^3$ and of the tangent bundle $T\mathbb{R}^3$, we do actually have a *natural or privileged* connection and its abstract curvature and torsion, which we will define, reduce to the above-mentioned elementary or concrete curvature and torsion of the curve.

As for **applications** of the preceding notions and techniques, the simplest one consists in the interpretation of the electromagnetic potential as a connection on a vector bundle over the Universe. This new standpoint will allow to view Maxwell's equations – which govern electromagnetism – in a new light. Indeed, Maxwell's equations can be written using the electromagnetic tensor and the latter then appears a the curvature of the covariant derivative that represents the potential. This viewpoint deepens our understanding of Maxwell's equations, since two of the four equations are now just natural consequences of our general theory of connections. In other words, in the new approach these two equations become completely natural.

To make this very rough introduction a 'closed circle', let us mention that electromagnetism is a simple example of a so-called **gauge theory**. Gauge theories, which are tightly related to connections, are of crucial importance in Theoretical Physics: for instance, the standard model mentioned at the beginning of the introduction is a quantized Yang-Mills gauge theory.

1 Vector bundles and fiber bundles

1.1 First examples

In this section, we choose an intuitive and informal approach to vector and fiber bundles, looking at examples that we are already familiar with.

Example 1 (Möbius strip)**.** Let us start with the Möbius strip E. It is obtained by taking a rectangle, rotating one extremity of the rectangle 180%, and gluing the two extremities together. Since the rectangle can be viewed as an amalgamation of vertical line segments or intervals $]-1,1[$ for instance, we can imagine the Möbius strip E as an amalgamation of intervals $]-1,1[\subset\mathbb{R}$ or manifolds over the unit circle S^1. It is natural to refer to the manifolds $]-1,1[$ as fibers and to call the amalgamation E a **fiber bundle** over the manifold S^1.

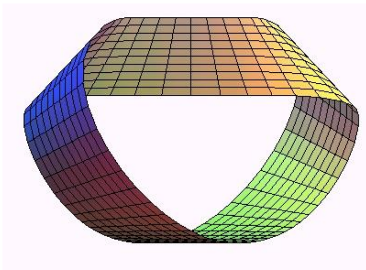

Figure 1: Möbius strip

If we glue the extremities of the initial rectangle without rotating one of them, we get the cylinder $C = S^1\times]-1,1[$. The cylinder C is of course also an amalgamation of manifolds $]-1,1[$ over S^1 and thus a fiber bundle over S^1. Since this fiber bundle is a *product manifold*, we say that it is a *trivial bundle*. Hence the question whether the Möbius strip E is also trivial, i.e., is also a product manifold, or is at least diffeomorphic to a product manifold. One might be tempted to claim that E is diffeomorphic to C or to the initial rectangle. However, this is clearly wrong, since the transformation that leads from E to C or to the rectangle is not smooth because it involves 'cutting' the Möbius strip. The situation changes when looking at a 'local piece' of E, at a slice of the Möbius strip. The latter can obviously be deformed smoothly into a rectangle $R = I\times]-1,1[$, where I denotes an open interval of $\mathbb{R}$, and the inverse of this transformation is of course smooth as well: a local piece of the fiber bundle E is diffeomorphic to a product manifold R, i.e., E is locally trivial. This is the second important aspect of **fiber bundles**: they are locally trivial, locally diffeomorphic to product manifolds. On the other hand, as mentioned above, E is not globally trivial, i.e., the Möbius strip is a non-trivial fiber bundle.

Example 2 (Tangent bundle of a manifold)**.** Let M be an n-dimensional smooth manifold. At any point $m \in M$ we can consider its tangent space T_mM, which is an n-dimensional real vector space. Recall now that a vector field of M is a field of tangent vectors, i.e., we are given a vector $X_m \in T_mM$ at any point $m \in M$. In other words, we are in the presence of a map

$X : M \ni m \mapsto X_m \in T_m M$. Since the target of X must be independent of m, it is natural to consider the disjoint union

$$TM = \bigsqcup_{m \in M} T_m M \ .$$

Since the disjoint union TM is an amalgamation of vector spaces (n-dimensional real vector spaces are n-dimensional smooth manifolds with one-chart-atlases) over M, it seems to be a fiber bundle whose fibers are vector spaces, i.e., it seems to be a **vector bundle**. This idea is corroborated by the fact that in Differential Geometry we called TM the tangent bundle of the manifold M.

To achieve final confirmation that TM is a bundle, we must still show that TM is locally trivial. Actually we proved this already in Differential Geometry, when building an atlas for TM thus establishing that TM is a manifold. Indeed, the proper mental picture of TM imagines M as a horizontal line, TM as a rectangle over this line, and $T_m M$ as the vertical line segment over $m \in M$. We denote by $\pi : TM \to M$ the projection map that associates to any vector $X \in TM$ the corresponding base point $m \in M$. The charts of TM can now be easily obtained from the charts of M. Let

$$\varphi : M \supset U \ni m \mapsto \varphi(m) = (x^1, \ldots, x^n) \in \varphi(U) \subset \mathbb{R}^n \tag{1}$$

be a chart of M. Recall that, for any $m \in U$, this chart provides a basis of the n-dimensional vector space $T_m M$ and that this basis is given by the derivations $(\partial_{x^i}|_m)$ at m, i.e., by the tangent vectors $(\partial_{x^i}|_m)$ at m. Any vector $X \in T_m M$, with $m \in U$, has thus coordinates

$$\phi(X) := (X^1, \ldots, X^n) \in \mathbb{R}^n \tag{2}$$

in this basis. The map

$$\mathbf{\Phi} : \pi^{-1}(U) \supset T_m M \ni X \mapsto (\varphi(m), \phi(X)) \in \varphi(U) \times \mathbb{R}^n \subset \mathbb{R}^{2n} \tag{3}$$

is obviously a chart of TM. The tuple $(\varphi(m), \phi(X))$ are the coordinates of X in this chart $\mathbf{\Phi}$. More precisely, we refer to $\varphi(m) = (x^1, \ldots, x^n)$ as the **base coordinates** of X and to $\phi(X) = (X^1, \ldots, X^n)$ as the **fiber coordinates** of X. Finally, the composition

$$\Phi : \pi^{-1}(U) \xrightarrow{\mathbf{\Phi}} \varphi(U) \times \mathbb{R}^n \xrightarrow{\varphi^{-1} \times \mathrm{id}} U \times \mathbb{R}^n \ , X \stackrel{\Phi}{\longmapsto} (\pi(X), \phi(X)) \tag{4}$$

provides a diffeomorphism, since the resulting map is the composite of two diffeomorphisms: the 'local piece' $\pi^{-1}(U) \subset TM$ is diffeomorphic to the product manifold $U \times \mathbb{R}^n \subset \mathbb{R}^{2n}$.

Let us first emphasize that above and in the following, we use four different letters phi, namely φ, ϕ, Φ, and $\mathbf{\Phi}$. Note also that, for any fixed $m \in U$, we have $T_m M \subset \pi^{-1}(U)$, so that we can restrict Φ to $T_m M$. This restriction, which we denote by Φ_m, reads

$$\Phi_m : T_m M \ni X \mapsto (\pi(X), \phi(X)) = (m, \phi(X)) \simeq \phi(X) \in \mathbb{R}^n \ ,$$

i.e., it assigns to any vector $X \in T_m M$ its (fiber) coordinates $(X^1, \ldots, X^n)$ in the basis $(\partial_{x^i}|_m)$, and it is therefore a vector space isomorphism.

The preceding examples allowed to realize that a fiber bundle (resp., a vector bundle) is an amalgamation of fibers that are finite-dimensional smooth manifolds (resp., finite-dimensional real vector spaces), which is locally trivial, i.e., which is locally diffeomorphic to a product manifold (resp., to a product manifold whose second factor is a finite-dimensional real vector space). In addition, in the case of a vector bundle, the restriction of this diffeomorphism to a fiber is a vector space isomorphism.

1.2 Vector bundles

Remark 1. All manifolds considered in this text are finite-dimensional smooth manifolds. Also all other concepts (e.g., functions, tensor fields) are systematically assumed to be smooth.

In view of the insight that we gained in the preceding subsection, it is natural to define a vector bundle as follows:

Definition 1 (Vector bundle)**.** A manifold E is called a *vector bundle* of rank $r \in \mathbb{N} \setminus \{0\}$ over a base manifold M, with projection or foot map $\pi : E \to M$, if and only if

- the map π is smooth and surjective,
- the fibers $E_m := \pi^{-1}\{m\}$, $m \in M$, are real vector space of dimension r,
- for any $m \in M$, there exists an open neighborhood $U \subset M$ of m and a diffeomorphism – called *local trivialization* –

$$\Phi : \pi^{-1}(U) \ni s \longmapsto (\pi(s), \phi(s)) \in U \times \mathbb{R}^r ,$$

  such that, for all $n \in U$, the restriction

$$\Phi_n : E_n \ni s \longmapsto \phi(s) \in \mathbb{R}^r$$

  is a vector space isomorphism.

Remark 2.

- Some authors denote a vector bundle by the pair (E, M) or the triple (E, M, π) to remind the base manifold M and the projection π. Alternatively, we often write $\pi : E \to M$ for a vector bundle E with base M and projection π.
- Although manifolds (of dimension n) can, when considered globally, have complicated shapes, they are locally very simple, since they are locally diffeomorphic to open subsets of $\mathbb{R}^n$. A similar remark holds for vector bundles. They can be globally complicated, but the are locally trivial, i.e., they are locally diffeomorphic to simple product manifolds. The local trivialization or diffeomorphism $\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^r$ identifies the *local piece* $\pi^{-1}(U)$ of the considered vector bundle with the trivial vector bundle $U \times \mathbb{R}^r$. Moreover, the restrictions $\Phi_n : E_n \to \mathbb{R}^r$ of Φ are vector space isomorphisms and thus identify the fibers E_n with $\mathbb{R}^r$, which is therefore referred to as the *typical fiber* of the vector bundle. Eventually, if we consider, in addition to the local trivialization

$$\Phi : \pi^{-1}(U) \to U \times \mathbb{R}^r \quad (5)$$

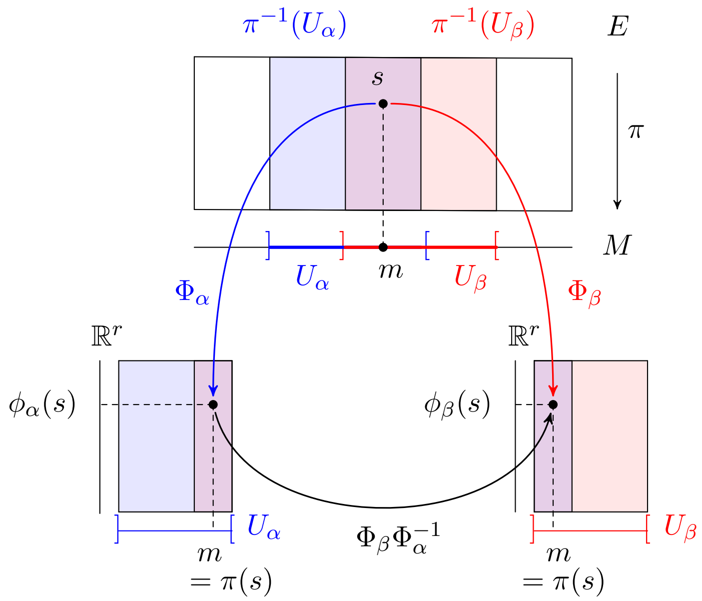

Figure 2: Vector bundle

of the bundle, a local coordinate system φ in the open subset U of M, the composite $(\varphi \times \mathrm{id}) \circ \Phi$ sends any $s \in E_m \subset \pi^{-1}(U)$ to

$$(\varphi(m), \phi(s)) = (x^1, \ldots, x^n, s^1, \ldots, s^r) , \tag{6}$$

i.e., it associates to s its *base and fiber coordinates.*

Let us give some examples of vector bundles.

Example 3.

- For a given manifold M, the tangent bundle TM (see Example 2), the cotangent bundle T^*M, the p-times contravariant and q-times covariant tensor bundle $\otimes^p_q TM$ $(p, q \in \mathbb{N})$, and the anti-symmetric p-covariant tensor bundle $\wedge^p T^*M$ $(p \in \{0, \ldots, \dim M\})$ are all vector bundles. They have been extensively studied in Differential Geometry.

- In view of Equations (1) and (3), the tangent rank n vector bundle $T\mathbb{R}^n$ is not only locally but even globally trivial. Indeed, the manifold $\mathbb{R}^n$ admits the global chart $(U, \varphi) = (\mathbb{R}^n, \mathrm{id})$, so that the induced $T\mathbb{R}^n$-chart $(\pi^{-1}(U), \mathbf{\Phi}) = (T\mathbb{R}^n, \mathbf{\Phi})$ is a diffeomorphism from $T\mathbb{R}^n$ to $\mathbb{R}^n \times \mathbb{R}^n$. The same result holds, more generally, for any n-dimensional real vector space V: $TV \simeq V \times \mathbb{R}^n$.

1.3 Fiber bundles

As suggested above, vector bundles are special fiber bundles. Indeed, the main difference is the nature of the fibers, which are vector spaces in the case of vector bundles and manifolds in the case of fiber bundles. However, (finite-dimensional real) vector spaces are particularly simple manifolds.

Definition 2 (Fiber bundle)**.** A manifold E is called a *fiber bundle* over a base manifold M, with projection $\pi : E \to M$, if and only if

- the map π is smooth and surjective,
- for any $m \in M$, there exists a manifold N, an open neighborhood $U \subset M$ of m, and a diffeomorphism – called *local trivialization* –

$$\Phi : \pi^{-1}(U) \ni s \longmapsto (\pi(s), \phi(s)) \in U \times N \ .$$

It can be shown that, if the requirements of this definition are met, the fibers $E_m = \pi^{-1}(m)$ are automatically manifolds. Moreover, the restriction Φ_m of Φ to a fiber E_m, $m \in U$, is also a diffeomorphism $\Phi_m : E_m \to N$. If the base manifold is connected, all the fibers are thus diffeomorphic to a unique and same manifold N, which is then called the typical fiber of E.

1.4 Transition maps and cocycle condition

1.4.1 Construction of a vector bundle

Recall that a manifold structure on a set M is given by an atlas $(U_\alpha, \varphi_\alpha)_\alpha$ of M, i.e., essentially, by a cover of M by charts or coordinate systems $(U_\alpha, \varphi_\alpha)$ $(\cup_\alpha U_\alpha = M)$, such that the coordinate transformations $\psi_{\beta\alpha} := \varphi_\beta \varphi_\alpha^{-1}$ are smooth and satisfy the condition $\psi_{\gamma\beta}\psi_{\beta\alpha} = \psi_{\gamma\alpha}$ (which is often referred to as the *cocycle condition*). Such an interpretation as a cover by coordinate systems, together with coordinate transformations of a specific type, which satisfy the cocycle condition, is not only possible for manifolds, but for many geometric structures, in particular for vector bundles.

Indeed, a rank r vector bundle $\pi : E \to M$ is covered by local trivializations

$$\Phi_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^r \ , \tag{7}$$

i.e., by fiber coordinate systems (see Equations (5) and (6)). A fiber coordinate transformation is here a map (see Figure: Vector bundle)

$$\Psi_{\beta\alpha} := \Phi_\beta \Phi_\alpha^{-1} : (U_\alpha \cup U_\beta) \times \mathbb{R}^r \to \pi^{-1}(U_\alpha \cap U_\beta) \to (U_\alpha \cap U_\beta) \times \mathbb{R}^r \ ,$$

which can be viewed as a family

$$\Psi_{\beta\alpha}(m) = \Phi_{\beta,m}\Phi_{\alpha,m}^{-1} : \mathbb{R}^r \to E_m \to \mathbb{R}^r \quad (m \in U_\alpha \cap U_\beta) \ ,$$

of vector space automorphisms, which send the fiber coordinates in the trivialization Φ_α to the fiber coordinates in the trivialization Φ_β. Such a family of automorphisms of $\mathbb{R}^r$ can be identified with a family of invertible $r \times r$ matrices, say

$$\eta_{\beta\alpha}(m) \in \mathrm{GL}(r, \mathbb{R}) \quad (m \in U_\alpha \cap U_\beta) \ . \tag{8}$$

Hence, if we denote the fiber coordinates of $s \in E_m$ in Φ_α by $S_\alpha = (s^1_\alpha, \ldots, s^r_\alpha)$ and those in Φ_β by $S_\beta = (s^1_\beta, \ldots, s^r_\beta)$, the fiber coordinate transformation $\Psi_{\beta\alpha}(m)$ in a vector bundle is of the type

$$S_\beta = \eta_{\beta\alpha}(m) S_\alpha \, . \tag{9}$$

If we choose as well coordinates φ_α in U_α and φ_β in U_β (see Figure: Vector bundle), the points $m \in U_\alpha \cap U_\beta$ have base coordinates $x_\alpha = (x^1_\alpha, \ldots, x^n_\alpha)$ in φ_α and $x_\beta = (x^1_\beta, \ldots, x^n_\beta)$ in φ_β. The base coordinate transformation is a coordinate transformation in a manifold, hence it is of the type

$$x_\beta = \psi_{\beta\alpha}(x_\alpha) \, , \tag{10}$$

where $\psi_{\beta\alpha} = \varphi_\beta \varphi_\alpha^{-1}$ is a smooth map and even a diffeomorphism. A full (base and fiber) coordinate transformation in a vector bundle is thus of the type (see Equations (10), (9), (8))

$$\begin{aligned} x' &= \psi(x) \quad (\psi \in \mathrm{Diff}) \, , \\ S' &= \eta(x) S \quad (\eta(x) \in \mathrm{GL(r, \mathbb{R})}) \, . \end{aligned} \tag{11}$$

In principle we are interested here only in the fiber coordinate transformation, i.e., in the second equation in (11). The fiber coordinate transformation or *transition map* $\Psi_{\beta\alpha}$, or, still, the matrix η or $\eta_{\beta\alpha}$, encode the information how to glue the local pieces $U_\alpha \times \mathbb{R}^r$ and $U_\beta \times \mathbb{R}^r$ (see Figure: Vector bundle). Moreover, if we first glue $U_\alpha \times \mathbb{R}^r$ with $U_\beta \times \mathbb{R}^r$ and then glue $U_\beta \times \mathbb{R}^r$ with $U_\gamma \times \mathbb{R}^r$, for some γ, we get the same result as when gluing directly $U_\alpha \times \mathbb{R}^r$ with $U_\gamma \times \mathbb{R}^r$. Indeed, since $\Psi_{\beta\alpha} = \Phi_\beta \Phi_\alpha^{-1}$, we have

$$\Psi_{\gamma\beta} \Psi_{\beta\alpha} = \Psi_{\gamma\alpha} \, . \tag{12}$$

Hence, the cocycle condition is satisfied. In matrix notation, it reads

$$\eta_{\gamma\beta} \, \eta_{\beta\alpha} = \eta_{\gamma\alpha} \, . \tag{13}$$

Finally, a vector bundle over M leads to an open cover of M by fiber coordinate system, with fiber coordinate transformations of the type

$$S_\beta = \eta_{\beta\alpha}(m) S_\alpha, \quad \eta_{\beta\alpha}(m) \in \mathrm{GL}(r, \mathbb{R}), \quad m \in U_\alpha \cap U_\beta \, ,$$

which satisfy the cocycle condition. Conversely,

Proposition 1. *Consider an open cover $(U_\alpha)_\alpha$ of a manifold M by trivial pieces or fiber coordinate systems $U_\alpha \times \mathbb{R}^r$. Assume that transition maps*

$$\Psi_{\beta\alpha} : (U_\alpha \cap U_\beta) \times \mathbb{R}^r \ni (m, S) \mapsto (m, S') \in (U_\alpha \cap U_\beta) \times \mathbb{R}^r \, ,$$

of the type $S' = \eta(m) S$, with $\eta \in C^\infty(U_\alpha \cap U_\beta, \mathrm{GL}(r, \mathbb{R}))$, are given and satisfy the cocycle condition. If we glue the trivial pieces $U_\alpha \times \mathbb{R}^r$ as encoded in the $\Psi_{\beta\alpha}$, we get a rank r vector bundle over M that is locally diffeomorphic to the $U_\alpha \times \mathbb{R}^r$.

In view of what has been said above, this proposition rather natural.

1.4.2 Construction of a fiber bundle

A similar result holds for fiber bundles.

Proposition 2. *Take an open cover $(U_\alpha)_\alpha$ of a manifold M and assume that with any U_α is associated a manifold N_α. If the differential structures of the trivial pieces $U_\alpha \times N_\alpha$ and $U_\beta \times N_\beta$ coincide over $U_\alpha \cap U_\beta$, i.e., if there are transition diffeomorphisms*

$$\Psi_{\beta\alpha} : (U_\alpha \cap U_\beta) \times N_\alpha \to (U_\alpha \cap U_\beta) \times N_\beta$$

that satisfy the cocycle condition, and if we glue the trivial pieces $U_\alpha \times V_\alpha$ as encoded in the $\Psi_{\beta\alpha}$, we get a fiber bundle over M that is locally diffeomorphic to the $U_\alpha \times N_\alpha$.

1.4.3 Applications

Since the tangent bundle of an n-dimensional manifold *is* a vector bundle of rank n, the cocycle condition *is* satisfied in this case. We now check the cocycle condition by direct computation.

Example 4. As announced we establish the cocycle condition (13) for the tangent bundle $E = TM$ of a base manifold M of dimension n. Since the fiber coordinates are read in the basis $(\partial_{x^i}|_m)$ of T_mM, $m \in U_\alpha$, induced by a base coordinate system $(U_\alpha, (x^1, \ldots, x^n))$ of M, and since the cocycle equation involves three fiber coordinate systems, we choose two additional base coordinate systems $(U_\beta, (y^1, \ldots, y^n))$ and $(U_\gamma, (z^1, \ldots, z^n))$, and consider the induced bases $(\partial_{y^i}|_m)$ and $(\partial_{z^i}|_m)$ of the tangent spaces T_mM at the points $m \in U_\beta$ and $m \in U_\gamma$, respectively. According to Equation (4), a vector $X \in \pi^{-1}(U_\alpha \cap U_\beta \cap U_\gamma)$ has fiber coordinates (X^i), (Y^i), and (Z^i) in all three bases (see Equation (2)):

$$X = X^i \partial_{x^i} = Y^j \partial_{y^j} = Z^k \partial_{z^k} \ ,$$

where we omitted the summation symbol and the subscript m. In view of the chain rule

$$X = Y^i \partial_{y^i} = X^j \partial_{x^j} = X^j \partial_{x^j} y^i \partial_{y^i} \ ,$$

so that

$$Y^i = \partial_{x^j} y^i X^j \ .$$

In other words, we have

$$\eta_{YX}(x) = (\partial_{x^j} y^i)_{ij} \ ,$$

where the RHS is the Jacobian matrix of the base coordinate transformation $x \rightleftarrows y$. Hence, $\eta_{YX}(x) \in \mathrm{GL}(n, \mathbb{R})$ and $\eta_{YX}(x)$ depends smoothly on x (see Proposition 1). Similarly,

$$\eta_{ZY} = (\partial_{y^j} z^i)_{ij} \quad \text{and} \quad \eta_{ZX} = (\partial_{x^j} z^i)_{ij} \ ,$$

so that

$$(\eta_{ZY}\ \eta_{YX})_{ij} = (\eta_{ZY})_{ik}\ (\eta_{YX})_{kj} = \partial_{y^k} z^i \partial_{x^j} y^k = \partial_{x^j} z^i = (\eta_{ZX})_{ij} \ ,$$

due to the chain rule. Hence, the cocycle condition is satisfied.

In the second application, we glue the Möbius strip from local pieces, using Proposition 2.

Example 5. Let $M = S^1$ be the unit circle and let

$$\mathrm{TP}_1 = U_1 \times]-1,1[=]0,2\pi[\times]-1,1[$$

and

$$\mathrm{TP}_2 = U_2 \times]-1,1[=]-\pi,\pi[\times]-1,1[$$

be a cover of S^1 by trivial pieces and more precisely by cylinders cut at $0 \simeq 2\pi$ and $-\pi \simeq \pi$, respectively. The gluing map is defined on the intersection

$$U_1 \cap U_2 =]0,2\pi[\cap]-\pi,\pi[=]0,\pi[\cup]\pi,2\pi[\simeq]0,\pi[\cup]-\pi,0[\ .$$

Since we wish to get the Möbius strip after the gluing process, we glue the parts $]0,\pi[\times]-1,1[$ of TP_1 and TP_2 and we glue their parts $]\pi,2\pi[\times]-1,1[$ and $]-\pi,0[\times]-1,1[$ after a 180% rotation. This leads to the gluing map Ψ_{21}, which is defined on the two connected components of $(U_1 \cap U_2)\times]-1,1[$ by

$$\Psi_{21}|_{]0,\pi[} :]0,\pi[\times]-1,1[\ni (x,s) \mapsto (x',s') = (x,s) \in]0,\pi[\times]-1,1[$$

and

$$\Psi_{21}|_{]\pi,2\pi[} :]\pi,2\pi[\times]-1,1[\ni (x,s) \mapsto (x',s') = (x,-s) \in]-\pi,0[\times]-1,1[\ .$$

It is clear that Ψ_{21} is a diffeomorphism. Further, no problem with the cocycle condition arises, so that the gluing process leads to a fiber bundle and more precisely to the Möbius strip.

If we replace the interval $]-1,1[$ above by the whole real line $\mathbb{R}$, the trivial pieces or fiber coordinate systems are $U_1 \times \mathbb{R}$ and $U_2 \times \mathbb{R}$, and we should thus get a vector bundle. Indeed, the fiber coordinate transformation is given by $s' = 1 \cdot s$ on the first connected component of $U_1 \cap U_2$ and by $s' = (-1) \cdot s$ on the second. Hence, it is given by

$$\eta(x) = \begin{cases} 1 \in \mathrm{GL}(1,\mathbb{R}), \ \forall x \in]0,\pi[\\ -1 \in \mathrm{GL}(1,\mathbb{R}), \ \forall x \in]\pi,2\pi[\end{cases} .$$

In view of Proposition 1, the gluing process leads now to a vector bundle and more precisely to a variant of the Möbius strip.

2 Sections and local frames

We start recalling the notion of section of a vector bundle. Let (E, M, π) be a vector bundle and let $U \subset M$ be an open subset.

Definition 3. A *section* of E over U is a smooth map $s : U \ni m \mapsto s_m \in E_m \subset E$. The set of all sections of E over U is denoted by $\Gamma(U, E)$. If $U = M$, we write $\Gamma(E)$ instead of $\Gamma(M, E)$.

The set $\Gamma(U, E)$ carries two obvious algebraic structures. For $s, s' \in \Gamma(U, E)$, $\lambda \in \mathbb{R}$, and $f \in C^\infty(U)$, we set:

$$s + s' : M \ni m \mapsto (s + s')_m := s_m + s'_m \in E_m \subset E \ ,$$

$$\lambda s : M \ni m \mapsto (\lambda s)_m := \lambda s_m \in E_m \subset E \ ,$$

$$fs : M \ni m \mapsto (fs)_m := f(m)s_m \in E_m \subset E \ .$$

Remark 3. The sets $\Gamma(U, E)$, U open in M, are equipped with a real vector space structure (first two operations) and with a $C^\infty(U)$-module structure (first and third operations).

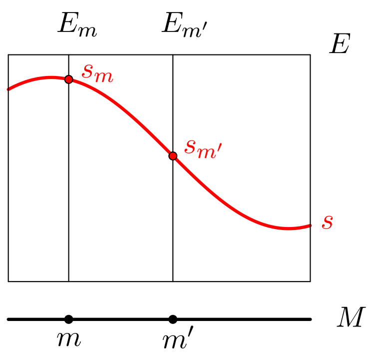

Figure 3: Section of a vector bundle

Consider a trivialization $\Phi : \pi^{-1}(U) \ni s \mapsto (\pi(s), \phi(s)) \in U \times \mathbb{R}^r$ of E over U. It is now easy to define a basis of the vector space E_m, $m \in U$. Indeed, denote by $(e_1, \ldots, e_r)$ the canonical basis of $\mathbb{R}^r$. Since the restriction of Φ to E_m, $m \in U$, is a vector space isomorphism $\Phi_m : E_m \to \mathbb{R}^r$, the vectors

$$\Phi_m^{-1}(e_i) =: \sigma_{i,m} \in E_m \ , \quad i \in \{1, \ldots, r\} \ ,$$

form a basis $(\sigma_{1,m}, \ldots, \sigma_{r,m})$ of E_m. As m is an arbitrary point of U, we get maps

$$\sigma_i : U \ni m \mapsto \sigma_{i,m} \in E_m \subset E \ ,$$

which are local sections $\sigma_i \in \Gamma(U, E)$.

Definition 4. A *local frame* is a family of sections $(\sigma_i)_i \in \Gamma(U, E)$ such that for all $m \in U$, the collection $(\sigma_{i,m})_i$ is a basis of E_m.

A local frame is thus made of local sections that define at any point a basis of the fiber.

Example 6. Consider the tangent bundle TM of a manifold M of dimension n. Let $\varphi : U \ni m \mapsto (x^1, \ldots, x^n) \in \mathbb{R}^n$ be a chart of M. Then the partial derivatives $(\partial_{x^i}|_m)$ provide a basis of T_mM, $m \in U$. Remember from the Differential Geometry course that these derivatives satisfy

$$\partial_{x^i}|_m \simeq (T_m\varphi)^{-1}(e_i) ,$$

where $T_m\varphi$ is the vector space isomorphism $T_m\varphi : T_mM \to \mathbb{R}^n$. Therefore, the partial derivatives $\partial_{x^i} \in \Gamma(U, TM)$ form a local frame of TM over U.

Above we observed that a local trivialization allows to construct a local frame. The converse is true as well. Let $(\sigma_i)_i$ be a local frame over U and try to build a trivialization over U, i.e., essentially, to assign fiber coordinates to any $s \in E_m$, $m \in U$. Since $(\sigma_{i,m})_i$ is a basis of E_m,

$$s = \sum_{i=1}^{r} s^i \sigma_{i,m} ,$$

with $s^1, \ldots, s^r \in \mathbb{R}$. We can now define

$$\Phi : \pi^{-1}(U) \supset E_m \ni s \mapsto (m, s^1, \ldots, s^r) \in U \times \mathbb{R}^r ,$$

which turns out to be a trivialization of E over U.

Remark 4. Given a local frame, one can construct a local trivialization and vice-versa: we sometimes identify local frames and local trivializations.

Let $s \in \Gamma(E)$ be a global section and $(\sigma_i)_i$ be a local frame over U. For all $m \in U$, we can decompose $s_m \in E_m$ in the basis $(\sigma_{i,m})_i$:

$$s_m = \sum_{i=1}^{r} s^i(m)\sigma_{i,m} . \tag{14}$$

We know from the Differential Geometry course, that smoothness of the considered section s (see Definition 3) implies that the functions $s^i : U \ni m \mapsto s^i(m) \in \mathbb{R}$ are smooth, i.e., that $s^i \in C^\infty(U)$. Equation (14) can thus be written

$$s|_U = \sum_{i=1}^{r} s^i\sigma_i, \quad s \in \Gamma(E), s^i \in C^\infty(U), \sigma_i \in \Gamma(U, E) ,$$

which is the *local form of a section.*

The preceding observations lead us to consider the map

$$\Gamma(U, E) \ni s \mapsto (s^1, \ldots, s^r) \in C^\infty(U, \mathbb{R}^r) ,$$

which is an isomorphism of $\mathbb{R}$-vector spaces and $C^\infty(U)$-modules:

Remark 5. We can identify local sections over a trivialization domain U with smooth functions on U valued in the typical fiber $\mathbb{R}^r$:

$$\Gamma(U, E) \simeq C^\infty(U, \mathbb{R}^r) . \tag{15}$$

For instance, $\Gamma(M \times \mathbb{R}^r) = \Gamma(M, M \times \mathbb{R}^r) \simeq C^\infty(M, \mathbb{R}^r)$ and, in particular, $\Gamma(M \times \mathbb{R}) \simeq C^\infty(M \times \mathbb{R}) = C^\infty(M)$.

3 Operations on vector bundles

Operations on vector spaces, e.g., duals, direct sums, tensor products..., give rise to new vector spaces. Similar operations can be defined for vector bundles and they lead to new vector bundles. Let E and E' be vector bundles of ranks r and r' over a same base manifold M. As the fibers E_m and fiber coordinate maps $\Phi_m : E_m \to \mathbb{R}^r$ are the main ingredients of the vector bundle E, we confine ourselves to specifying these data for the new dual, direct sum and tensor product vector bundles.

Dual of a vector bundle.

The dual vector bundle E^* of E is made of the fibers

$$(E^*)_m := (E_m)^*$$

and admits the fiber coordinate maps $(\Phi^*)_m : (E_m)^* \to (\mathbb{R}^r)^* \simeq \mathbb{R}^r$ defined by

$$(\Phi^*)_m := (\Phi^t_m)^{-1} ,$$

where Φ^t_m denotes the transpose of the linear map Φ_m. Since no confusion can arise, we write E^*_m instead of $(E^*)_m = (E_m)^*$. Note that E^* is, just as E, a vector bundle of rank r.

Direct sum of vector bundles.

The direct sum $E \oplus E'$ of E and E' is given by the fibers

$$(E \oplus E')_m := E_m \oplus E'_m$$

and the fiber coordinate maps $(\Phi^{\oplus})_m : E_m \oplus E'_m \to \mathbb{R}^r \oplus \mathbb{R}^{r'} \simeq \mathbb{R}^{r+r'}$ defined by

$$(\Phi^{\oplus})_m := \Phi_m \oplus \Phi'_m .$$

We observe that $E \oplus E'$ is a vector bundle of rank $r + r'$.

Tensor product of vector bundles.

The tensor product $E \otimes E'$ of E and E' has the fibers

$$(E \otimes E')_m := E_m \otimes E'_m$$

and the fiber coordinate maps $(\Phi^{\otimes})_m : E_m \otimes E'_m \to \mathbb{R}^r \otimes \mathbb{R}^{r'} \simeq \mathbb{R}^{rr'}$ defined by

$$(\Phi^{\otimes})_m := \Phi_m \otimes \Phi'_m .$$

Hence, the product $E \otimes E'$ is a vector bundle of rank rr'.

Generalizing the preceding construction, we get the vector bundle

$$\otimes^p_q E := E \otimes \underset{(p)}{\ldots} \otimes E \otimes E^* \otimes \underset{(q)}{\ldots} \otimes E^* .$$

4 Characterization of tensor fields

In this Section, we prove a basic result that we will use throughout the present text.

Recall that for any vector space V and any non-negative integers $p, q \in \mathbb{N}$, we define the vector space

$$\otimes_q^p V = V \otimes \underset{(p)}{\ldots} \otimes V \otimes V^* \otimes \underset{(q)}{\ldots} \otimes V^*$$

of (p, q)-*tensors*, or p times contravariant and q times covariant tensors, as the vector space of $\mathbb{R}$-*valued* $\mathbb{R}$-*multilinear maps defined on the dual spaces*

$$V^* \times \underset{(p)}{\ldots} \times V^* \times V \times \underset{(q)}{\ldots} \times V \ :$$

$$\otimes_q^p V = \mathcal{L}_{\mathbb{R}-(p+q)-\mathrm{lin}}(V^* \times \ldots \times V^* \times V \times \ldots \times V, \mathbb{R}) = \mathcal{L}_{\mathbb{R}-(p+q)-\mathrm{lin}}((V^*)^{\times p} \times V^{\times q}, \mathbb{R}) \ . \quad (16)$$

The following theorem gives a similar characterization of tensor fields, i.e., of sections $\Gamma(\otimes_q^p E)$ of the tensor bundle $\otimes_q^p E$.

Theorem 1. *Let (E, M, π) be a vector bundle and let $p, q \in \mathbb{N}$. Then,*

$$\Gamma(\otimes_q^p E) \simeq \mathcal{L}_{C^\infty(M)-(p+q)-\mathrm{lin}}(\Gamma(E^*) \times \underset{(p)}{\ldots} \times \Gamma(E^*) \times \Gamma(E) \times \underset{(q)}{\ldots} \times \Gamma(E), C^\infty(M)) \ , \quad (17)$$

i.e., tensor fields can be viewed as function-valued function-multilinear maps defined on sections of the dual bundles.

Note that, with respect to the characterization (16) of tensors, the preceding characterization (17) of tensor fields (tensors depending on the point where they are 'measured') just replaces reals by functions (reals depending on the point where they are 'measured') and 'vectors' by 'vector' fields (sections).

Proof. We just check that a tensor field can be interpreted as multilinear map and vice versa.

Let $\mathcal{T} \in \Gamma(\otimes_q^p E)$. We have to define a $C^\infty(M)$-multilinear map

$$\begin{array}{rccc} \mathcal{T}: & \Gamma(E^*) \times \underset{(p)}{\ldots} \times \Gamma(E^*) \times \Gamma(E) \times \underset{(q)}{\ldots} \times \Gamma(E) & \longrightarrow & C^\infty(M) \\ & (t^1, \ldots, t^p, s_1, \ldots, s_q) & \mapsto & \mathcal{T}(t^1, \ldots, t^p, s_1, \ldots, s_q) \ . \end{array} \quad (18)$$

Therefore, we must construct $\mathcal{T}(t^1, \ldots, t^p, s_1, \ldots, s_q)(m) \in \mathbb{R}$, for all $m \in M$.

Since $\mathcal{T}$ is a tensor field $\mathcal{T} \in \Gamma(\otimes_q^p E)$, it is a tensor $\mathcal{T}_m \in (\otimes_q^p E)_m = \otimes_q^p E_m$ that depends on $m \in M$. It follows now from (16) that

$$\mathcal{T}_m \in \otimes_q^p E_m = \mathcal{L}_{\mathbb{R}-(p+q)-\mathrm{lin}}((E_m^*)^{\times p} \times (E_m)^{\times q}, \mathbb{R}) \ .$$

It is thus natural to set

$$\mathcal{T}(t^1, \ldots, t^p, s_1, \ldots, s_q)(m) := \mathcal{T}_m(t_m^1, \ldots, t_m^p, s_{1,m}, \ldots, s_{q,m}) \in \mathbb{R} \ .$$

It is easily seen that the map (18) is function-multilinear.

Conversely, let $\mathcal{T}$ be a function-multilinear map in the RHS of (17). To show that $\mathcal{T}$ is a section in $\Gamma(\otimes^p_q E)$, we have to build a tensor $\mathcal{T}_m \in \otimes^p_q E_m$, $m \in M$, i.e., again by (16), we have to construct a multilinear map $\mathcal{T}_m \in \mathcal{L}_{\mathbb{R}-(p+q)-\mathrm{lin}}((E^*_m)^{\times p} \times (E_m)^{\times q}, \mathbb{R})$, i.e., a multilinear map

$$\begin{array}{rcl} \mathcal{T}_m : \quad E^*_m \times \underset{(p)}{\ldots} \times E^*_m \times E_m \times \underset{(q)}{\ldots} \times E_m & \longrightarrow & \mathbb{R} \\ (\tau^1, \ldots, \tau^p, \sigma_1, \ldots, \sigma_q) & \mapsto & \mathcal{T}_m(\tau^1, \ldots, \tau^p, \sigma_1, \ldots, \sigma_q) \; . \end{array}$$

Now, for all $\tau^i \in E^*_m$ and all $\sigma_j \in E_m$, one can choose sections $t^i \in \Gamma(E^*)$ and $s_j \in \Gamma(E)$, which pass at m through τ^i and σ_j, respectively: $t^i_m = \tau^i$ and $s_{j,m} = \sigma_j$. Hence, it is natural to set

$$\mathcal{T}_m(\tau^1, \ldots, \tau^p, \sigma_1, \ldots, \sigma_q) := \mathcal{T}(t^1, \ldots, t^p, s_1, \ldots, s_q)(m) \in \mathbb{R} \; . \tag{19}$$

However, the map $\mathcal{T}_m$ that we thus defined is not necessarily well-defined, since another choice of the sections t^i and s_j could lead to a different image.

It now suffices to show that $\mathcal{T}_m$ is actually well-defined. There exists a rigorous proof, but we prefer here a more intuitive approach – which is more instructive. It is known that almost all operators that appear in Differential Geometry are *local operators*, i.e., operators $\mathcal{T}$ such that the value $\mathcal{T}(t^1, \ldots, t^p, s_1, \ldots, s_q)(m)$ at m of the image only depends on restrictions of the arguments t^i, s_j to a neighborhood of m. The prototypical local operators are the *differential operators*. For the operators ∂_{x^i} for instance, the value $\partial_{x^i} f|_m$ at m of the image $\partial_{x^i} f$ only depend on the restriction of f to a neighborhood of m. An important result, which is referred to as *Peetre's Theorem*[1] states roughly that the converse holds as well: any local operator is (locally) a differential operator. Hence, if we assume that the multilinear map or multilinear operator $\mathcal{T}$ is, as most operators, a local one, it follows from Peetre's result that $\mathcal{T}$ is a differential operator. Consider for simplicity that $\mathcal{T}$ has only one argument, say $f \in \Gamma(U \times \mathbb{R}) \simeq C^\infty(U)$ (U coordinate domain of M). Then,

$$\mathcal{T}(f) = \sum_{|\alpha| \leq k} g_\alpha \, \partial^\alpha_x f \; ,$$

with $g_\alpha \in C^\infty(U)$ and $k \in \mathbb{N}$. Since $\mathcal{T}$ belongs to the RHS of (17), it is function-multilinear, i.e., for any $h \in C^\infty(U)$, we have $\mathcal{T}(hf) = h\mathcal{T}(f) = \mathcal{T}(fh) = f\mathcal{T}(h)$. Taking in particular $h = 1$, we obtain

$$\mathcal{T}(f) = \mathcal{T}(f \cdot 1) = f \sum_{|\alpha| \leq k} g_\alpha \, \partial^\alpha_x 1 = f \cdot g_0 \; ,$$

or, still,

$$\mathcal{T}(f)(m) = f(m) g_0(m) \; .$$

We thus see that $\mathcal{T}(f)(m)$, not only depends only on the restriction of f to a neighborhood of m, but even depends only on the value of f *at* the point m. In our situation,

[1] Jaak Peetre (1935–) is a Swedish mathematician.

this means that $\mathcal{T}(t^1, \ldots, t^p, s_1, \ldots, s_q)(m)$ (see Equation (19)) only depends on the values of $t^1, \ldots, t^p, s_1 \ldots, s_q$ at m, i.e., on

$$t^1_m = \tau^1, \ldots, t^p_m = \tau^p, s_{1,m} = \sigma_1, \ldots, s_{q,m} = \sigma_q \ .$$

Therefore, the definition (19) does not depend on the choice of t^i and s_j, so that the map $\mathcal{T}_m$ is well-defined. □

Here are two applications of Theorem 1.

Example 7. 1. Any section $s \in \Gamma(E) = \Gamma(\otimes^1_0 E)$ can be viewed as a map

$$s \in \mathcal{L}_{C^\infty(M)-\mathrm{lin}}(\Gamma(E^*), C^\infty(M)) \ .$$

2. Let us interpret the map $\Delta \in \mathcal{L}_{C^\infty(M)-\mathrm{lin}}(\Gamma(TM), \Gamma(E))$ as a tensor field. In view of Point 1, we can view Δ as a map

$$\Delta : \Gamma(TM) \xrightarrow{C^\infty(M)-\mathrm{lin}} \left(\Gamma(E^*) \xrightarrow{C^\infty(M)-\mathrm{lin}} C^\infty(M)\right) \ ,$$

or, equivalently, as a map

$$\Delta \in \mathcal{L}_{C^\infty(M)-\mathrm{bilin}}(\Gamma(TM) \times \Gamma(E^*), C^\infty(M)) \ .$$

It follows now from Theorem 1 that $\Delta \in \Gamma(T^*M \otimes E)$:

$$\mathcal{L}_{C^\infty(M)-\mathrm{lin}}(\Gamma(TM), \Gamma(E)) \simeq \Gamma(T^*M \otimes E) \ .$$

In the following, we will write $\mathcal{L}_{C^\infty(M)}$ instead of $\mathcal{L}_{C^\infty(M)-\mathrm{lin}}$, $\mathcal{L}_{C^\infty(M)-\mathrm{bilin}}$ ···

5 Covariant derivative on a vector bundle

In this section, we introduce the notion of *covariant derivative* $\nabla_X s$ of a section $s \in \Gamma(E)$ of a vector bundle $\pi : E \to M$ in the direction of a vector field $X \in \Gamma(TM)$.

5.1 Motivation

Take $M = \mathbb{R}^3$ and $\pi : E = \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3 = M$, and consider the temperature $\tau_m \in \mathbb{R}$ at a point $m \in \mathbb{R}^3$. We view τ as a function $\tau \in C^\infty(\mathbb{R}^3) = C^\infty(M) \simeq \Gamma(E)$. Let us compute the derivative $\nabla_X \tau$ of τ in the direction of a vector field $X \in \Gamma(TM) = \Gamma(\mathbb{R}^3 \times \mathbb{R}^3) \simeq C^\infty(\mathbb{R}^3, \mathbb{R}^3)$. Since a derivative does not change the nature of the object that it derives, the derivative $\nabla_X \tau$ is, just as τ, a section $\nabla_X \tau \in \Gamma(E) \simeq C^\infty(M) = C^\infty(\mathbb{R}^3)$. To determine the derivative $(\nabla_X \tau)_m \in \mathbb{R}$ at a point $m \in \mathbb{R}^3$ in the direction of X, we follow the standard idea of a derivative, and measure the temperature τ_m at m, then the temperature $\tau_{m'}$ at a point m' close to m and in the direction of X_m, and compute finally the limit for $m' \to m$ of the relative difference of these values. It is clear that the result depends only on X_m and not on the whole vector field X or its values around m. As understood in Section 4, this means that the derivative $\nabla_X \tau$ is function-linear in X. The latter property is a natural and fundamental requirement for the directional or covariant derivative $\nabla_X s$ of a section s in the direction of a vector field X: for any $f \in C^\infty(M)$, we must have

$$\nabla_{fX} s = f \nabla_X s \ .$$

We now examine the concept of covariant derivative in the case of simple vector bundles $\pi : E \to M$ or simple section spaces $\Gamma(E)$.

5.1.1 Covariant derivative of a function in the direction of a vector field

Let first $\pi : E = M \times \mathbb{R} \to M$, and consider $X \in \Gamma(TM)$ and $f \in \Gamma(E) = C^\infty(M)$. Since the Lie derivative $L_X f$, which has been studied in Differential Geometry, is also interpreted as the derivative of f in the direction of X, it is natural to set

$$\nabla_X f := L_X f = (df)(X) \in C^\infty(M) \ , \tag{20}$$

where d is the de Rham differential. Recall that the differential of a differential 0-form $f \in \Omega^0(M) = C^\infty(M)$ is a differential 1-form $df \in \Omega^1(M) = \Gamma(T^*M)$, so that, in view of Theorem 1,

$$df \in \mathcal{L}_{C^\infty(M)}(\Gamma(TM), C^\infty(M)) \ .$$

It follows that $df(X) \in C^\infty(M)$ as announced, and that the covariant derivative $\nabla_X f$ defined in (20) satisfies

$$\nabla_{gX} f = g \nabla_X f \ ,$$

for any $g \in C^\infty(M)$. Since our main condition for a covariant derivative is thus fulfilled, we can accept Equation (20) as the definition of the covariant derivative of functions.

5.1.2 Covariant derivative of a vector field in the direction of a vector field

After the case $\pi : E = M \times \mathbb{R} = \otimes^0_0 TM \to M$, we study the case $\pi : E = \otimes^1_0 TM = TM \to M$, and consider $X \in \Gamma(TM)$ and $Y \in \Gamma(E) = \Gamma(TM)$, i.e., we consider two vector fields X, Y. Let us see whether the covariant derivative $\nabla_X Y$ can also be defined as the Lie derivative, that is, whether we can set

$$\nabla_X Y := L_X Y = [X, Y] = X \circ Y - Y \circ X \in \Gamma(TM) \ , \tag{21}$$

where, in the RHS, X, Y are viewed as derivations, i.e., $X, Y \in \mathrm{Der}(C^\infty(M)) \simeq \Gamma(TM)$. The question is whether this derivative satisfies our function-linearity requirement with respect to X. When writing X and Y in local coordinates $(x^1, \ldots, x^n)$, we get $X = \sum_i X^i \partial_{x^i}$ and $Y = \sum_j Y^j \partial_{x^j}$, so that Equation (21) reads

$$[X, Y] = \sum_j \sum_i (X^i \partial_{x^i} Y^j - Y^i \partial_{x^i} X^j) \partial_{x^j} \ .$$

Therefore, for $g \in C^\infty(M)$, we have locally

$$\nabla_{gX} Y = [gX, Y] = \sum_j \sum_i (gX^i \partial_{x^i} Y^j - Y^i \partial_{x^i}(gX^j))\partial_{x^j} = g\nabla_X Y - \sum_j \sum_i Y^i X^j (\partial_{x^i} g)\partial_{x^j} \ ,$$

so that the covariant derivative of vector fields defined as the Lie derivative of vector fields is **not** function-linear. Hence, we must **reject** the definition proposed in Equation (21)!

5.1.3 Covariant derivative of an arbitrary section in the direction of a vector field

Let $\pi : E \to M$ be an arbitrary vector bundle, in particular a tensor bundle $\pi : E = \otimes_q^p TM \to M$, and consider $X \in \Gamma(TM)$ and $s \in \Gamma(E)$. Try now to define the covariant derivative $\nabla_X s$ so that it satisfies the function-linearity condition. Since we cannot use the Lie derivative (it works only for $(p,q) = (0,0)$, but, for instance, not for $(p,q) = (1,0)$), we try to mimic the usual definition of a derivative or here of a directional derivative, i.e., the one used already above in the case of the temperature. In other words, to define $\nabla_X s \in \Gamma(E)$ or $(\nabla_X s)_m \in E_m$, we consider the value s_m of s at m and the value $s_{m'}$ of s at a point m' close to m in the direction of X_m, and compute the limit of the relative difference of these values. However, since $s_m \in E_m$ and $s_{m'} \in E_{m'}$, these vectors belong to different vector spaces and their difference does not make sense. Therefore, we must transport $s_{m'}$ into the vector space E_m by means of some transportation rule – we call it a **parallel transport**. Differently stated, we need some rule that connects the fibers E_m and $E_{m'}$, i.e., we need a **connection** on E. But which parallel transport or connection should we choose? In Differential Geometry, we encountered the same problem for $E = \otimes_q^p TM$, and the searched transport was implemented by the maximal integral curves $\varphi_t^X(m)$ of X. Indeed, the map $\varphi_t^X : M \to M$ is (at least locally) a diffeomorphism with inverse φ_{-t}^X, so that, if $m' = \varphi_t^X(m)$ with $t \simeq 0$, the derivative $T_{m'}\varphi_{-t}^X : T_{m'}M \to T_m M$ is a vector space isomorphism, which can be extended to an isomorphism $T_{m'}^\otimes \varphi_{-t}^X : \otimes_q^p T_{m'}M \to \otimes_q^p T_m M$, or, still, $T_{m'}^\otimes \varphi_{-t}^X : E_{m'} \to E_m$. The point is that if we use the transportation rule $T_{m'}^\otimes \varphi_{-t}^X$, the value $(\nabla_X s)_m$ depends, as immediately seen, on the values of the vector field X in a neighborhood of m, instead of only depending on the value of X at m. Another way to understand this is to observe that with the transportation rule $T_{m'}^\otimes \varphi_{-t}^X$ we do get exactly the Lie derivative of the considered tensor field s, i.e., we obtain $(\nabla_X s)_m = (L_X s)_m$, what is not acceptable since the Lie derivative is not function-linear.

The preceding discussion shows that to be able to define a 'covariant derivative' ∇, we need a convenient rule of 'parallel transport' or a 'connection' between the fibers. Actually, the notions of 'covariant derivative', 'parallel transport', and 'connection' are three different concepts, but they are in some sense equivalent and are thus sometimes even used as synonyms. However, which appropriate parallel transport rule should we choose? It turns out that there are many such rules, but that in general there is no privileged one, i.e., no canonical rule of parallel transport. This means that there exist in general many covariant derivatives or connections, but that none is more natural than the others. Therefore, we will define a covariant derivative by listing all the properties it should have, then we will show that it is always possible to define many such derivatives.

5.2 Definition and existence

Let $\pi : E \to M$ be a vector bundle over M.

Definition 5. A *covariant derivative* (or connection) on E is an $\mathbb{R}$-bilinear map

$$\nabla : \Gamma(TM) \times \Gamma(E) \ni (X, s) \mapsto \nabla_X s \in \Gamma(E) ,$$

such that, for any $X \in \Gamma(TM), s \in \Gamma(E)$, and $f \in C^\infty(M)$,

$$\nabla_{fX} s = f \nabla_X s \tag{22}$$

and

$$\nabla_X (fs) = (L_X f)s + f \nabla_X s . \tag{23}$$

Since for functions $\nabla_X f = L_X f = (df)(X)$, the condition (23) is nothing but the Leibniz rule.

There exists an equivalent formulation of this definition. Indeed, as explained in Example 7, the map

$$\nabla : \Gamma(E) \xrightarrow{\mathbb{R}-\text{lin}} \Gamma(TM) \xrightarrow{C^\infty(M)-\text{lin}} \Gamma(E)$$

can be viewed as a map

$$\nabla : \Gamma(E) \xrightarrow{\mathbb{R}-\text{lin}} \Gamma(T^*M \otimes E) .$$

Hence the following reformulation of Definition 5:

Definition 6. A *covariant derivative* (or connection) on E is an $\mathbb{R}$-linear map

$$\nabla : \Gamma(E) \ni s \mapsto \nabla s \in \Gamma(T^*M \otimes E) ,$$

such that, for any $s \in \Gamma(E)$ and $f \in C^\infty(M)$,

$$\nabla(fs) = df \otimes s + f \nabla s . \tag{24}$$

Remark 6.

1. In Definition 6, we consider ∇ as a map of the variable $s \in \Gamma(E)$ only, the variable $X \in \Gamma(TM)$ being encrypted in the target space

   $$\Gamma(T^*M \otimes E) = \mathcal{L}_{C^\infty(M)}(\Gamma(TM), \Gamma(E)) .$$

   This target space encodes also the condition (22). Moreover, we use the notation $df \otimes s$, since, when considering $\nabla_X s$, we get $(df \otimes s)(X)$, which is, according to the standard definition of the tensor product of maps, given by $(df)(X)s = (L_X f)s$.

2. Equations (22) and (23) mean that ∇ is a differential operator of order 0 with respect to the variable X and a differential operator of order 1 with respect to the variable s. Hence, the operator ∇ is a local operator and it can thus be restricted to any open subset $U \subset M$. This means that, starting from

   $$\nabla : \Gamma(TM) \times \Gamma(E) \to \Gamma(E) ,$$

we can, via common differential geometric methods, define

$$\nabla^U : \Gamma(U, TM) \times \Gamma(U, E) \to \Gamma(U, E) ,$$

in a way such that

$$(\nabla_X s)|_U = \nabla^U_{X|_U} s|_U .$$

3. Let $E = \otimes^p_q TM$ be the vector bundle of p-times contravariant and q-times covariant tensors. Then

$$\nabla : \Gamma(\otimes^p_q TM) \ni s \mapsto \nabla s \in \Gamma(T^*M \otimes (\otimes^p_q TM)) = \Gamma(\otimes^p_{q+1} TM) ,$$

i.e., ∇ increases the covariant degree of a tensor field by 1. This observation motivates the name of covariant derivative.

Example 8. Let $\pi : E \to M$ be a trivial vector bundle of rank r over a base manifold M, so that $E \simeq M \times \mathbb{R}^r$. Due to Equation (15), we have

$$\Gamma(E) \simeq C^\infty(M, \mathbb{R}^r) = (C^\infty(M))^{\times r} .$$

This shows that in the special case of a trivial bundle, the sections that ∇ derives are tuples of smooth functions. Therefore, in contrast with ordinary vector bundles, there exists, on a trivial bundle, a canonical covariant derivative, namely, as mentioned above, the derivative $\nabla = d$, where d is the de Rham differential.

Proposition 3. *A trivial vector bundle admits a privileged covariant derivative – the de Rham differential. This canonical derivative $\nabla = d$ is referred to as the trivial covariant derivative on the considered trivial bundle.*

To construct a covariant derivative on an arbitrary vector bundle, we need the next

Lemma 1. *Let $(\nabla^i)_i$ (resp., $(f_i)_i$) be a family of covariant derivatives on a vector bundle $\pi : E \to M$ (resp., a family of smooth functions on M). The linear combination $\sum_i f_i \nabla^i$ is again a covariant derivative on E if and only if it is an affine combination of covariant derivatives, i.e., if and only if $\sum_i f_i = 1$.*

Proof. Function-linearity of $\nabla := \sum_i f_i \nabla^i$ is obvious. The condition that the combination be affine appears when one checks the Leibniz rule. Indeed,

$$\nabla_X(fs) = \sum_i f_i \nabla^i_X(fs) = \sum_i f_i(L_X f)s + \sum_i f_i f \nabla^i_X s = (L_X f)s \sum_i f_i + f \nabla_X s .$$

This shows that ∇ is a covariant derivative if and only if $\sum_i f_i = 1$. □

A priori the families in the preceding lemma are finite. However, the lemma still holds in the infinite case, provided there are no convergence issues.

The next theorem guarantees the existence of covariant derivatives on any vector bundle.

Theorem 2. *Covariant derivatives do exist on any vector bundle.*

Proof. Let $\pi : E \to M$ be a rank r vector bundle and let $(U_\alpha)_\alpha$ be an open cover of M by local trivialization domains.

Each trivial bundle $\pi^{-1}(U_\alpha) \simeq U_\alpha \times \mathbb{R}^r$ admits the trivial covariant derivative ∇^{U_α} given, for any $X_{U_\alpha} \in \Gamma(U_\alpha, TM)$ and any $s_{U_\alpha} \in \Gamma(U_\alpha, E)$, by

$$\nabla^{U_\alpha}_{X_{U_\alpha}} s_{U_\alpha} := (ds_{U_\alpha})\,(X_{U_\alpha}) \in \Gamma(U_\alpha, E)$$

(see Proposition 3).

Consider now a locally finite partition of unity $(\psi_\alpha)_\alpha$ subordinate to the cover $(U_\alpha)_\alpha$, let $X \in \Gamma(TM)$ and $s \in \Gamma(E)$, and set

$$\nabla_X s := \sum_\alpha \psi_\alpha \nabla^{U_\alpha}_{X|_{U_\alpha}} s|_{U_\alpha} \in \Gamma(E)\ . \tag{25}$$

Indeed, in view of the standard argument $\psi_\alpha \nabla^{U_\alpha}_{X|_{U_\alpha}} s|_{U_\alpha} \in \Gamma(E)$, although $\nabla^{U_\alpha}_{X|_{U_\alpha}} s|_{U_\alpha} \in \Gamma(U_\alpha, E)$. Further, since the partition of unity is locally finite, the sum over α does not give rise to convergence problems, so that the RHS of (25) is actually a section in $\Gamma(E)$. By Lemma 1, the operator ∇ defined by (25), is indeed a covariant derivative, since $\sum_\alpha \psi_\alpha = 1$ and the ∇^{U_α} are covariant derivatives. □

The next result discloses the structure of the set $C(E, M)$ of all covariant derivatives on a vector bundle $\pi : E \to M$.

Let $\nabla, \nabla' \in C(E, M)$ and consider their naturally defined difference

$$\nabla' - \nabla : \Gamma(TM) \times \Gamma(E) \ni (X, s) \mapsto (\nabla' - \nabla)_X s := \nabla'_X s - \nabla_X s \in \Gamma(E)\ .$$

For all $X \in \Gamma(TM)$, $f \in C^\infty(M)$, and $s \in \Gamma(E)$, we have

$$(\nabla' - \nabla)_X(fs) = \nabla'_X(fs) - \nabla_X(fs) = (L_X f)s + f\nabla'_X s - (L_X f)s - f\nabla_X s = f(\nabla' - \nabla)_X s\ .$$

In other words, the difference $\nabla' - \nabla$ is not only function-linear in X but also in s. Since $\Gamma(E) = \mathcal{L}_{C^\infty(M)}(\Gamma(E^*), C^\infty(M))$, the difference $\nabla' - \nabla$, viewed as map

$$\nabla' - \nabla : \Gamma(TM) \times \Gamma(E) \times \Gamma(E^*) \to C^\infty(M)\ ,$$

is function-linear in all three variables. Hence,

$$\nabla' - \nabla \in \Gamma(T^*M \otimes E^* \otimes E) = \Gamma(T^*M \otimes \mathrm{End}(E))\ .$$

Indeed, since for a finite-dimensional vector space V, we have $V^* \otimes V = \mathcal{L}(V, V) = \mathrm{End}(V)$, we define the endomorphism bundle $\mathrm{End}(E)$ by $\mathrm{End}(E) := E^* \otimes E$.

Eventually we defined a subtraction

$$-\ : C(E, M) \times C(E, M) \ni (\nabla', \nabla) \mapsto \nabla' - \nabla \in \Gamma(T^*M \otimes \mathrm{End}(E))\ ,$$

i.e., a subtraction on $C(E,M)$ valued in the vector space $\Gamma(T^*M \otimes \mathrm{End}(E))$. Recall now that if a subtraction '$-$' defined on a set A and valued in a vector space V satisfies Weyl's axioms

$$\forall a \in A,\ \forall v \in V,\ \exists!\, a' \in A : a' - a = v\ ,$$

$$\forall a, b, c \in A,\ (c-b) + (b-a) = c - a\ ,$$

the set A is an affine space modelled on the vector space V. It is straightforwardly checked that the difference considered above satisfies these axioms, so that:

Theorem 3. *The set of all covariant derivatives on a vector bundle $\pi : E \to M$ is an affine space modelled on the vector space $\Gamma(T^*M \otimes \mathrm{End}(E))$ of differential 1-forms on M valued in the endomorphism bundle of E.*

5.3 Coordinate form of a covariant derivative – connection 1-form

Let ∇ be a covariant derivative on a rank r vector bundle $\pi : E \to M$. As explained in Item 2 of Remark 6, the derivative ∇ can be localized to any open subset $U \subset M$. Recall also that if this localization

$$\nabla^U : \Gamma(U, TM) \times \Gamma(U, E) \ni (X_U, s_U) \mapsto \nabla^U_{X_U} s_U \in \Gamma(U, E)$$

is computed, not on an arbitrary X_U and s_U but on restrictions $X_U = X|_U$ and $s_U = s|_U$ of globally defined sections $X \in \Gamma(TM)$ and $s \in \Gamma(E)$, we have

$$\nabla^U_{X|_U} s|_U = (\nabla_X s)|_U\ .$$

Let now U be a trivialization domain with trivialization Φ or, equivalently, with local frame $(\sigma_i)_i$. Any section $s_U \in \Gamma(U, E)$ can be decomposed in this frame,

$$s_U = \sum_{i=1}^{r} s^i \sigma_i\ , \tag{26}$$

with $s_i \in C^\infty(U)$. We take now an interest in the coordinates of the section $\nabla^U_{X_U} s_U \in \Gamma(U, E)$, which we denote in the following simply by $\nabla_X s$. In view of (26) and (23), we obtain

$$\nabla_X s = \nabla_X \sum_i s^i \sigma_i = \sum_i (L_X s^i)\sigma_i + \sum_i s^i \nabla_X \sigma_i\ . \tag{27}$$

Of course, the sections $\nabla_X \sigma_i \in \Gamma(U, E)$ can also be decomposed in the considered frame:

$$\nabla_X \sigma_i = \sum_j \mathcal{A}(X)\big|_i^j \sigma_j\ ,$$

where $\mathcal{A}(X)\big|_i^j \in C^\infty(U)$. We thus get maps

$$\mathcal{A}\big|_i^j : \Gamma(U, TM) \ni X \mapsto \mathcal{A}(X)\big|_i^j \in C^\infty(U)\ ,$$

which are $C^\infty(U)$-linear since the $\nabla_X\sigma_i$ are $C^\infty(U)$-linear in X. It follows that

$$\mathcal{A}\big|_i^j \in \Gamma(U, T^*M) = \Omega^1(U) \ ,$$

i.e., that the $\mathcal{A}\big|_i^j$ are differential 1-forms on U. Finally $\mathcal{A}$ is an $r \times r$ matrix with entries in $\Omega^1(U)$. We also say that $\mathcal{A}$ is a differential 1-form on U with values in $r \times r$ matrices and write

$$\mathcal{A} \in \Omega^1(U) \otimes \mathrm{gl}(r, \mathbb{R}) \ .$$

Remember also that

$$\mathcal{A}(X) \in C^\infty(U) \otimes \mathrm{gl}(r, \mathbb{R}) \ .$$

Having the matrix $\mathcal{A}$ at hand, we can rewrite Equation (27) as follows:

$$\begin{aligned} \nabla_X s &= \sum_i (L_X s^i)\sigma_i + \sum_i s^i \sum_j \mathcal{A}(X)\big|_i^j \sigma_j \\ &= \sum_i (L_X s^i)\sigma_i + \sum_i \sum_j \mathcal{A}(X)\big|_j^i s^j \sigma_i \\ &= \sum_i \left(L_X s^i + \sum_j \mathcal{A}(X)\big|_j^i s^j \right) \sigma_i \ . \end{aligned}$$

This means that the coordinates of $\nabla_X s$ in the frame $(\sigma_i)_i$ are

$$L_X s^i + \sum_j \mathcal{A}(X)\big|_j^i s^j \ , \quad i \in \{1, \dots, r\} \ .$$

The isomorphism (15) between sections of E over U and functions on U with values in the typical fiber of E, is valid for a domain U of a local trivialization Φ or local frame $(\sigma_i)_i$. We therefore denote this isomorphism by Φ:

$$\Phi : \Gamma(U, E) \ni s = \sum_i s^i \sigma_i \mapsto s^\Phi = (s^1, \dots, s^r) \in C^\infty(U)^{\times r} \ .$$

When writing now s^Φ (resp., $(\nabla_X s)^\Phi$) for the coordinates of s (resp., $\nabla_X s$) in the considered trivialization Φ, we eventually obtain

$$(\nabla_X s)^\Phi = L_X(s^\Phi) + \mathcal{A}(X) s^\Phi \ . \tag{28}$$

Theorem 4. *Let ∇ be a covariant derivative or connection on a vector bundle $\pi : E \to M$ of rank r. Locally, in a trivialization Φ of E over an open subset U of M, the connection reads*

$$(\nabla_X s)^\Phi = L_X(s^\Phi) + \mathcal{A}(X) s^\Phi$$

($X \in \Gamma(U, TM)$, $s \in \Gamma(U, E)$), where $\mathcal{A}$ is a matrix $\mathcal{A} \in \Omega^1(U) \otimes \mathrm{gl}(r, \mathbb{R})$ with entries in differential 1-forms. This shows that locally in a trivialization a connection ∇ is completely characterized by the 1-form $\mathcal{A}$ with values in matrices. We refer to $\mathcal{A}$ as the connection 1-form *of ∇ in the considered trivialization.*

5.3.1 Transformation rule for connection 1-forms – construction of a connection from connection 1-forms

Recall that a vector v of an n-dimensional real vector space $\mathcal{V}$ has in every basis $(e_i)_i$ of $\mathcal{V}$ coordinates $V = {}^t(v^1, \ldots, v^n)$, and that these coordinates transform according to the rule

$$V = AV' , \tag{29}$$

where V and V' are the coordinates of v in a basis $(e_i)_i$ and a basis $(e'_i)_i$, respectively, and where A is the transition matrix from the basis $(e_i)_i$ to the basis $(e'_i)_i$, i.e., the matrix whose column j is made of the coordinates of e'_j in the basis $(e_i)_i$.

Conversely, if we are given, for any basis of $\mathcal{V}$, an n-tuple V of real numbers and if these tuples satisfy the transformation rule $V = AV'$, then these tuples are the coordinates of a unique vector $v \in \mathcal{V}$, i.e., the tuples define a vector v. We can thus construct a vector from tuples of real numbers that satisfy the transformation rule (29).

Similarly, a connection ∇ on a rank r vector bundle E has in every trivialization Φ of E the coordinate form (28) characterized by its connection 1-form $\mathcal{A}$. Our goal is to find the transformation rule that allows passing from the connection 1-form $\mathcal{A}^\beta$ in a trivialization Φ_β over an open subset U_β to the connection 1 form $\mathcal{A}^\alpha$ in a trivialization Φ_α over U_α.

In view of Theorem 4, we have

$$(\nabla_X s)^{\Phi_\alpha} = L_X(s^{\Phi_\alpha}) + \mathcal{A}^\alpha(X) s^{\Phi_\alpha} \tag{30}$$

and

$$(\nabla_X s)^{\Phi_\beta} = L_X(s^{\Phi_\beta}) + \mathcal{A}^\beta(X) s^{\Phi_\beta} \tag{31}$$

(here X and s are defined on $U_\alpha \cup U_\beta$). Recall the fiber coordinate transformation (9) in a vector bundle:

$$s^{\Phi_\alpha} = \eta_{\alpha\beta}\, s^{\Phi_\beta} \quad \text{and} \quad s^{\Phi_\beta} = \eta_{\beta\alpha}\, s^{\Phi_\alpha} , \tag{32}$$

where $\eta_{\alpha\beta} \in C^\infty(U_\alpha \cap U_\beta, \mathrm{GL}(r, \mathbb{R}))$ is an invertible matrix depending smoothly on $m \in U_\alpha \cap U_\beta$ and whose inverse is $\eta_{\beta\alpha}$ (the preceding equalities hold on $U_\alpha \cap U_\beta$). Applying this transformation rule to the fiber coordinates $(\nabla_X s)^{\Phi_\alpha}$ and $(\nabla_X s)^{\Phi_\beta}$ of the section $\nabla_X s$, we obtain

$$(\nabla_X s)^{\Phi_\alpha} = \eta_{\alpha\beta}\, (\nabla_X s)^{\Phi_\beta} . \tag{33}$$

It follows from (33), (31), and (32) that

$$\begin{aligned}
(\nabla_X s)^{\Phi_\alpha} &= \eta_{\alpha\beta}\, L_X(s^{\Phi_\beta}) + \eta_{\alpha\beta}\, \mathcal{A}^\beta(X) s^{\Phi_\beta} \\
&= \eta_{\alpha\beta}\, L_X(\eta_{\beta\alpha}\, s^{\Phi_\alpha}) + \eta_{\alpha\beta}\, \mathcal{A}^\beta(X)\, \eta_{\beta\alpha}\, s^{\Phi_\alpha} \\
&= \eta_{\alpha\beta}\, L_X(\eta_{\beta\alpha})\, s^{\Phi_\alpha} + \eta_{\alpha\beta}\, \eta_{\beta\alpha}\, L_X(s^{\Phi_\alpha}) + \eta_{\alpha\beta}\, \mathcal{A}^\beta(X)\, \eta_{\beta\alpha}\, s^{\Phi_\alpha} \\
&= \eta_{\alpha\beta}\, d(\eta_{\beta\alpha})(X)\, s^{\Phi_\alpha} + L_X(s^{\Phi_\alpha}) + \eta_{\alpha\beta}\, \mathcal{A}^\beta(X)\, \eta_{\beta\alpha}\, s^{\Phi_\alpha} ,
\end{aligned}$$

where we used the definition (20) of the Lie derivative of a function since the matrix $\eta_{\beta\alpha}$ is a matrix of functions. When comparing now the latter result with the result (30), we find

$$\mathcal{A}^\alpha(X)\, s^{\Phi_\alpha} = \eta_{\alpha\beta}\, d(\eta_{\beta\alpha})(X)\, s^{\Phi_\alpha} + \eta_{\alpha\beta}\, \mathcal{A}^\beta(X)\, \eta_{\beta\alpha}\, s^{\Phi_\alpha} ,$$

and when omitting the variables X and s^{Φ_α}, we finally obtain the searched transformation rule:

$$\mathcal{A}^\alpha = \eta_{\alpha\beta}\, d(\eta_{\beta\alpha}) + \eta_{\alpha\beta}\, \mathcal{A}^\beta\, \eta_{\beta\alpha}\ . \tag{34}$$

The rule (34) for connections corresponds to the rule (29) for vectors. Just as we can construct a vector from tuples V that satisfy the transformation rule (29), we can construct a connection from connection 1-forms $\mathcal{A}$ that satisfy the transformation rule (34) (note that (32) is actually the geometric variant of (29)).

Theorem 5. *Let $\pi : E \to M$ be a vector bundle of rank r. If ∇ is a connection on E, the connection 1-forms $\mathcal{A}^\alpha$ and $\mathcal{A}^\beta$, which characterize ∇ locally in trivializations (U_α, Φ_α) and (U_β, Φ_β), are related by the transformation rule*

$$\mathcal{A}^\alpha = \eta_{\alpha\beta}\, d(\eta_{\beta\alpha}) + \eta_{\alpha\beta}\, \mathcal{A}^\beta\, \eta_{\beta\alpha}\ , \tag{35}$$

where d is the de Rham differential of M, where $\eta_{\alpha\beta} \in C^\infty(U_\alpha \cap U_\beta, \mathrm{GL}(r,\mathbb{R}))$ is the fiber coordinate transition matrix from coordinates s^{Φ_β} to coordinates s^{Φ_α}, and where $\eta_{\beta\alpha} = \eta_{\alpha\beta}^{-1}$. Conversely, if U_α is a family of trivialization domains that cover M, if $\mathcal{A}^\alpha \in \Omega^1(U_\alpha) \otimes \mathrm{gl}(r,\mathbb{R})$ is a family of connection 1-forms, and if this family satisfies the transformation rule (35), then the local connections defined by the $\mathcal{A}^\alpha$ can be glued so as to define a unique connection ∇ on E.

5.4 Induced covariant derivatives

We will explain that a covariant derivative ∇ on a vector bundle $E \to M$ induces a covariant derivative on each tensor bundle $\otimes^p_q E$.

Induced covariant derivative on the dual bundle $\otimes^0_1 E = E^*$. A covariant derivative ∇^* on E^* is a map

$$\nabla^* : \Gamma(TM) \times \Gamma(E^*) \ni (X, t) \mapsto \nabla^*_X t \in \Gamma(E^*)\ , \tag{36}$$

where

$$\Gamma(E^*) = \mathcal{L}_{C^\infty(M)}(\Gamma(E), C^\infty(M))\ . \tag{37}$$

Therefore, we must define, for any $s \in \Gamma(E)$, a function $(\nabla^*_X t)(s) \in C^\infty(M)$. We set

$$(\nabla^*_X t)(s) := L_X(t(s)) - t(\nabla_X s) \in C^\infty(M)\ . \tag{38}$$

Note that we defined ∇^* in a way that the Lie derivative of the interior product $t(s)$ is given by the Leibniz rule, i.e., in a way that

$$L_X(t(s)) = (\nabla^*_X t)(s) + t(\nabla_X s)\ .$$

It is easily seen that the map $\nabla^*_X t$ defined by (38) is $C^\infty(M)$-linear in s, so that $\nabla^*_X t \in \Gamma(E^*)$ in view of (37). We now have a map ∇^* of the type (36) and must still prove that ∇^* has the two properties of a covariant derivative, i.e., that it is $C^\infty(M)$-linear in X and satisfies the Leibniz rule with respect to ft. We leave these proofs as an exercise to the reader.

Induced covariant derivative on the tensor bundle $\otimes^p_q E$. A covariant derivative $\nabla^\otimes$ on $\otimes^p_q E$ is a map

$$\nabla^\otimes : \Gamma(TM) \times \Gamma(\otimes^p_q E) \ni (X,T) \mapsto \nabla^\otimes_X T \in \Gamma(\otimes^p_q E) \ , \tag{39}$$

where

$$\Gamma(\otimes^p_q E) = \mathcal{L}_{C^\infty(M)}(\Gamma(E^*)^{\times p} \times \Gamma(E)^{\times q}, C^\infty(M)) \ .$$

Let $(t^1, \ldots, t^p, s_1, \ldots, s_q) \in \Gamma(E^*)^{\times p} \times \Gamma(E)^{\times q}$ and define

$$\begin{aligned}(\nabla^\otimes_X T)(t^1, \ldots, t^p, s_1, \ldots, s_q) \ := \ & L_X(T(t^1, \ldots, t^p, s_1, \ldots, s_q)) \\ & - \sum_{i=1}^p T(t^1, \ldots, \nabla^*_X t^i, \ldots, t^p, s_1, \ldots, s_q) \\ & - \sum_{j=1}^q T(t^1, \ldots, t^p, s_1, \ldots, \nabla_X s_j, \ldots, s_q) \ .\end{aligned}$$

This definition of $\nabla^\otimes$ is based, as the above one of ∇^*, on the idea that the Leibniz rule should hold for the Lie derivative of an interior product, here the product $T(t^1, \ldots, t^p, s_1, \ldots, s_q) \in C^\infty(M)$. Further, the same checks as above are necessary here to show that $\nabla^\otimes$ is indeed a covariant derivative on $\otimes^p_q E$.

5.5 Christoffel's symbols

Let $\pi : E \to M$ be a vector bundle and let ∇ be a covariant derivative on E. Locally, in a trivialization (U, Φ) of E, the connection ∇ is completely determined by its local connection 1-form $\mathcal{A}$:

$$(\nabla_X s)^\Phi = L_X s^\Phi + \mathcal{A}(X) s^\Phi \ .$$

Hence, the ith component of $\nabla_X s$ reads

$$(\nabla_X s)^i = L_X s^i + \mathcal{A}(X)\big|^i_k s^k \ . \tag{40}$$

We now further reduce the data needed for the complete local determination of ∇. Therefore, let $\varphi : U \ni m \mapsto x = (x^1, \ldots, x^n) \in \varphi(U)$ be a coordinate system of M in the trivialization domain U of Φ (it suffices to reduce U if necessary). The vector field X (which is defined in U) then reads $X = \sum_j X^j \partial_{x^j}$ and, since

$$(\nabla_X s)^i = \sum_j X^j (\nabla_{\partial_{x^j}} s)^i \ ,$$

it is enough to know the components $(\nabla_{\partial_{x^j}} s)^i$, which are given by

$$(\nabla_{\partial_{x^j}} s)^i = \partial_{x^j} s^i + \mathcal{A}(\partial_{x^j})\big|^i_k s^k \ ,$$

where we identified the derivation $L_{\partial_{x^j}} \in \mathrm{Der}(C^\infty(U))$ with the vector field $\partial_{x^j} \in \Gamma(U, TM)$, as we often do[2]. It follows that:

[2] Actually ∂_{x^j} is a derivation viewed as vector field, i.e., ∂_{x^j} is in fact $L^{-1}_{\partial_{x^j}}$, so that $L_{\partial_{x^j}}$ is in fact $L_{L^{-1}_{\partial_{x^j}}} = \partial_{x^j}$.

Proposition 4. *A covariant derivative ∇ on a vector bundle $E \to M$ is, in the domain U of a trivialization Φ of E and a chart φ of M, completely defined by the functions $\mathcal{A}(\partial_{x^j})\big|_k^i \in C^\infty(U)$. For $E = TM$, these functions are referred to as* Christoffel's symbols *and they are denoted by Γ^i_{jk}:*

$$\Gamma^i_{jk} := \mathcal{A}(\partial_{x^j})\big|_k^i \; . \tag{41}$$

Over U, we then have

$$(\nabla_{\partial_{x^j}} s)^i = \partial_{x^j} s^i + \Gamma^i_{jk} s^k \; . \tag{42}$$

5.6 Christoffel's symbols in Mathematical Physics

5.6.1 Christoffel's symbols as trivial connection

Formula (42) is actually known from elementary Mathematical Physics. To see this, consider the vector bundle $E = T\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2$. Recall that any coordinate system $x = (x^1, x^2)$ of $\mathbb{R}^2$ – the canonical system of cartesian coordinates, the system of polar coordinates ... – induces the frame $\partial_x = (\partial_{x^1}, \partial_{x^2})$ of the tangent bundle $T\mathbb{R}^2$ made of the vector fields

$$\partial_{x^j} = \partial_{x^j}|_x \; .$$

It follows that any section $s \in \Gamma(E)$, i.e., any vector field $s \in \Gamma(T\mathbb{R}^2)$ reads

$$s = \sum_k s^k \partial_{x^k} \; ,$$

where both, the component functions s^k and the fields ∂_{x^k} depend (smoothly) on x. This aspect should be kept in mind! Since the considered bundle $T\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2$ is trivial, we choose on $T\mathbb{R}^2$ the trivial covariant derivative ∇ given by the Lie derivative L and compute as in Equation (42) the component

$$(\nabla_{\partial_{x^j}} s)^i = (L_{\partial_{x^j}} s)^i = (\partial_{x^j} s)^i \; .$$

Without any knowledge of vector bundles and covariant derivatives, one gets in elementary courses

$$\partial_{x^j} s = \partial_{x^j}(s^k \partial_{x^k}) = (\partial_{x^j} s^k)\partial_{x^k} + s^k \partial_{x^j}(\partial_{x^k})$$

and

$$(\partial_{x^j} s)^i = \partial_{x^j} s^i + s^k \partial_{x^j}(\partial_{x^k})\big|^i \; . \tag{43}$$

Since the derivative $\partial_{x^j}(\partial_{x^k})$ of the vector field ∂_{x^k} is again a vector field, one can decompose the latter in the frame of the partial derivatives and one sets

$$\partial_{x^j}(\partial_{x^k}) = \sum_i \Gamma^i_{jk} \partial_{x^i} \; ,$$

so that the coordinate function $\partial_{x^j}(\partial_{x^k})\big|^i$ is

$$\partial_{x^j}(\partial_{x^k})\big|^i = \Gamma^i_{jk} \; . \tag{44}$$

Combining (43) and (44), one obtains that

$$(\partial_{x^j} s)^i = \partial_{x^j} s^i + \Gamma^i_{jk} s^k \ ,$$

refers to $\partial_{x^j} s$ as a covariant derivative, uses the notation $\nabla_{\partial_{x^j}} s$ instead of $\partial_{x^j} s$, and finally writes

$$(\nabla_{\partial_{x^j}} s)^i = \partial_{x^j} s^i + \Gamma^i_{jk} s^k \ ,$$

which is the same equation as (42). The preceding computation is the origin of covariant derivatives. In Equation (42), we recover this origin as a local aspect of our general theory of connections on vector bundles.

5.6.2 Christoffel's symbols in cartesian and polar coordinates

We compute Christoffel's symbols Γ^i_{jk} for the trivial covariant derivative L on the trivial bundle $T\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2$ in cartesian and polar coordinates of $\mathbb{R}^2$ (and the corresponding frames of $T\mathbb{R}^2$).

Cartesian coordinates. Let $x = (x^1, x^2)$ be the canonical cartesian coordinates of $\mathbb{R}^2$. The corresponding frame $\partial_x = (\partial_{x^1}, \partial_{x^2})$ of $T\mathbb{R}^2$ is, in view of its definition

$$\partial_{x^i}|_m = (T_m\varphi)^{-1}(e_i)$$

in the case of an arbitrary smooth n-dimensional manifold M (where φ is a chart of M around m and $(e_i)_i$ is the standard basis of $\mathbb{R}^n$) and in view of the fact that the chart φ is in our case $M = \mathbb{R}^2$ the global chart $\varphi = \mathrm{id}$, given by

$$\partial_{x^i}|_x = e_i$$

(points of $M = \mathbb{R}^2$ are not denoted by m but by x). Since the basis $(e_1, e_2) = ((1,0),(0,1))$ is constant with respect to x, we obtain

$$\Gamma^i_{jk} = \partial_{x^j}(e_k)\big|^i = 0 \ ,$$

i.e., for the trivial connection of $T\mathbb{R}^2$ all Christoffel symbols vanish in cartesian coordinates.

Polar coordinates. Let (r, θ) be the polar coordinates of $\mathbb{R}^2$, given by $x^1 = r\cos\theta, x^2 = r\sin\theta$ (these are not really global coordinates, but are coordinates in the open subset $U = \mathbb{R}^2 \setminus [O, e_1)$; see Figure: Polar Coordinates). The chain rule provides the decomposition of the corresponding frame $(\partial_r, \partial_\theta)$ of $T\mathbb{R}^2$ in the frame $(\partial_{x^1}, \partial_{x^2}) = (e_1, e_2)$ of the preceding paragraph:

$$\begin{pmatrix} e_r \\ e_\theta \end{pmatrix} := \begin{pmatrix} \partial_r \\ \partial_\theta \end{pmatrix} = \begin{pmatrix} \partial_r x^1 \\ \partial_\theta x^1 \end{pmatrix} \partial_{x^1} + \begin{pmatrix} \partial_r x^2 \\ \partial_\theta x^2 \end{pmatrix} \partial_{x^2} \ ,$$

so that

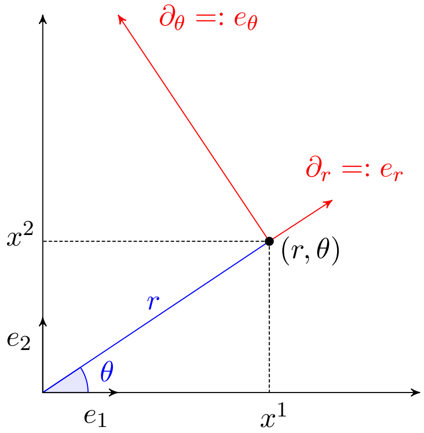

Figure 4: Polar coordinates

$$\begin{pmatrix} e_r \\ e_\theta \end{pmatrix} = \begin{pmatrix} \cos\theta \\ -r\sin\theta \end{pmatrix} e_1 + \begin{pmatrix} \sin\theta \\ r\cos\theta \end{pmatrix} e_2 \,.$$

These decompositions show that $\partial_r = e_r = e_r(\theta)$ and that $\partial_\theta = e_\theta = e_\theta(r,\theta)$ are, in contrast with $\partial_{x^1} = e_1$ and $\partial_{x^2} = e_2$, not constant. It follows that

$$\Gamma^\theta_{r\theta} = \partial_r(e_\theta)\big|^\theta = \partial_r(-r\sin\theta e_1 + r\cos\theta e_2)\big|^\theta = (-\sin\theta e_1 + \cos\theta e_2)\big|^\theta = \frac{1}{r} e_\theta\big|^\theta = \frac{1}{r} \,.$$

The other Christoffel symbols are obtained similarly:

$$\Gamma^\theta_{r\theta} = \Gamma^\theta_{\theta r} = \frac{1}{r},\; \Gamma^r_{\theta\theta} = -r, \quad \text{and} \quad \Gamma^\theta_{rr} = \Gamma^\theta_{\theta\theta} = \Gamma^r_{rr} = \Gamma^r_{r\theta} = \Gamma^r_{\theta r} = 0 \,.$$

5.7 Transformation rule for Christoffel's symbols

Let $\pi : TM \to M$ be the tangent bundle of an n-dimensional manifold M and let ∇ be a covariant derivative on TM. Any coordinate system (U, x) of M implements a frame (U, ∂_x), or, equivalently, a trivialization (U, Φ) of TM, and the connection ∇ is in U completely defined by its Christoffel symbols $\Gamma^i_{jk} = \Gamma^i_{jk}(x) \in C^\infty(U)$.

In Differential Geometry and Mathematical Physics, it's well-known that such packages of functions Γ^i_{jk}, Γ'^a_{bc} ... obtained in different frames ∂_x, $\partial_{x'}$ ... (induced by different coordinate systems x, x' ...) define a (unique) $(1,2)$-tensor field, if they transform into each other via the $(1,2)$-tensor rule

$$\Gamma^i_{jk} = A^i_a A'^b_j A'^c_k \Gamma'^a_{bc} \,, \tag{45}$$

where A is the transition matrix from the frame ∂_x to the frame $\partial_{x'}$, whereas A' is the inverse of A. Column j of the transition matrix A is made of the components of the new frame vector $\partial_{x'^j}$ in the old frame ∂_x. Since $\partial_{x'^j} = \partial_{x'^j}x^i\partial_{x^i}$, we see that $A^i_j = \partial_{x'^j}x^i$, or, still, that $A = \partial_{x'}x$ is the Jacobian matrix of the diffeomorphism or coordinate transformation $x = x(x') \rightleftarrows x' = x'(x)$. Hence, the inverse matrix A' is given by $A' = \partial_x x'$. The $(1,2)$-tensor field transformation condition (45) thus reads

$$\Gamma^i_{jk} = \partial_{x'^a}x^i\partial_{x^j}x'^b\partial_{x^k}x'^c\Gamma'^a_{bc}\ , \tag{46}$$

or, completely precisely,

$$\Gamma^i_{jk}(x) = \partial_{x'^a}x^i|_{x'=x'(x)}\,\partial_{x^j}x'^b\,\partial_{x^k}x'^c\,\Gamma'^a_{bc}(x'(x))\ .^3 \tag{47}$$

However, this transformation condition cannot hold for Christoffel's symbols. Indeed, the computations of Subsection 5.6.2 show that, in the case of the trivial connection of $T\mathbb{R}^2$, Christoffel's symbols Γ'^a_{bc} in cartesian coordinates (and the corresponding frame of partial derivatives or the corresponding trivialization Φ^β) vanish, whereas the three Christoffel symbols Γ^i_{jk} in polar coordinates (and the corresponding frame of partial derivatives or the corresponding trivialization Φ^α) with exactly two θ-indices don't vanish. However, would the condition (46) be satisfied, the annihilation of all the Γ'^a_{bc} would imply the annihilation of all the Γ^i_{jk} – what is not the case. Hence, Christoffel's symbols don't satisfy the condition (46) and are thus not tensorial, i.e., they are not the components of a $(1,2)$-tensor field.

We establish now the correct transformation rule for Christoffel's symbols. By definition,

$$\Gamma^i_{jk} = \mathcal{A}^\alpha(\partial_{x^j})\big|^i_k \quad \text{and} \quad \Gamma'^a_{bc} = \mathcal{A}^\beta(\partial_{x'^b})\big|^a_c\ , \tag{48}$$

where x and x' refer to the considered coordinate systems of M and the superscripts α and β to the corresponding trivializations Φ_α and Φ_β of TM. The transformation rule for Christoffel's symbols is a direct consequence of the transformation rule for connection 1-forms:

$$\mathcal{A}^\alpha = \eta_{\alpha\beta}d(\eta_{\beta\alpha}) + \eta_{\alpha\beta}\mathcal{A}^\beta\eta_{\beta\alpha}\ .$$

Recall that over a trivialization domain U of a rank r vector bundle $E \to M$ that is also a chart domain of the underlying n-dimensional manifold M, the elements $s \in E$ over U are characterized by their base coordinates $x \in \mathbb{R}^n$ and their fiber coordinates $S \in \mathbb{R}^r$, and that the transformation of these coordinates for two such domains U and U' is given by $x = x(x') \rightleftarrows x' = x'(x)$, for the base coordinates, and by

$$S^\alpha = \eta_{\alpha\beta}(x)S^\beta\ , \tag{49}$$

with $\eta_{\alpha\beta}(x) \in \mathrm{GL}(r,\mathbb{R})$, for the fiber coordinates.

In the case $E = TM$ that we consider here, any coordinate chart x on U induces a trivialization Φ_α or frame ∂_x over U, so that an element $s \in TM$ over $U \cap U'$ reads

$$s = S^{\alpha,i}\partial_{x^i} = S^{\beta,j}\partial_{x'^j} = S^{\beta,j}\partial_{x'^j}x^i\partial_{x^i}\ .$$

[3] If x (resp., x') are the coordinates of a system (U,φ) (resp., (U',φ')), the transformation diffeomorphism $\varphi'\varphi^{-1} : x \mapsto x'(x)$ is defined on $\varphi(U \cap U')$ and the condition (47) must thus hold in $\varphi(U \cap U')$.

This means that $S^{\alpha,i} = \partial_{x'^j} x^i S^{\beta,j}$, or, still, that

$$S^\alpha = \partial_{x'} x S^\beta \, ,$$

so that the comparison with (49) shows that, in the case $E = TM$, we have $\eta_{\alpha\beta} = \partial_{x'} x$ and $\eta_{\beta\alpha} = \eta^{-1}_{\alpha\beta} = \partial_x x'$. Therefore,

$$\begin{aligned}
\Gamma^i_{jk} = \mathcal{A}^\alpha(\partial_{x^j})\big|^i_k &= [(\partial_{x'} x) d(\partial_x x')(\partial_{x^j})]\big|^i_k + [(\partial_{x'} x)\mathcal{A}^\beta(\partial_{x^j})\partial_x x']\big|^i_k \\
&= \partial_{x'^a} x^i d(\partial_x x')(\partial_{x^j})\big|^a_k + \partial_{x'^a} x^i \mathcal{A}^\beta(\partial_{x^j})\big|^a_c \partial_{x^k} x'^c \\
&= \partial_{x'^a} x^i \partial_{x^j} \partial_{x^k} x'^a + \partial_{x'^a} x^i \mathcal{A}^\beta(\partial_{x^j})\big|^a_c \partial_{x^k} x'^c \, .
\end{aligned}$$

In view of Equation (48), we write

$$\mathcal{A}^\beta(\partial_{x^j})\big|^a_c = \mathcal{A}^\beta(\partial_{x^j} x'^b \partial_{x'^b})\big|^a_c = \partial_{x^j} x'^b \mathcal{A}^\beta(\partial_{x'^b})\big|^a_c = \partial_{x^j} x'^b \, \Gamma'^a_{bc} \, ,$$

where we used the fact that the connection 1-form is function-linear. We have thus the following proposition, which is similar to Theorem 5:

Proposition 5. *Consider the tangent bundle TM of a manifold M. If ∇ is a connection on TM, Christoffel's symbols of ∇ in two coordinate systems x and x' of M (and in the induced trivializations ∂_x and $\partial_{x'}$ of TM) satisfy the transformation rule*

$$\Gamma^i_{jk} = \partial_{x'^a} x^i \, \partial_{x^j} x'^b \, \partial_{x^k} x'^c \, \Gamma'^a_{bc} + \partial_{x'^a} x^i \, \partial_{x^j} \partial_{x^k} x'^a \, . \qquad (50)$$

Equation (50) *is the tensor rule* (46) *corrected by the second term of the RHS. It proves that Christoffel's symbols are not the components of a tensor field. Conversely, if U_α is a family of coordinate domains that covers M, if $(\Gamma^\alpha)^i_{jk} \in C^\infty(U_\alpha)$ is a family indexed by α of packages of functions indexed by i, j, k, and if this family of packages satisfies the transformation rule* (50)*, then the local connections of TU_α, which are defined by the Γ^α via* (42)*, can be glued so as to define a unique connection ∇ on TM.*

6 Remark

The second part of the lecture notes is not displayed in this reading sample...