Differential and Riemannian Geometry

Olivier Biquard

Sorbonne Université, Institut de Mathématiques de Jussieu-Paris Rive Gauche
6th September 2020

Contents

Introduction

These are notes for an introductory course in differential geometry. The aim is to give some basis on several topics: manifolds, vector fields, connections, curvature, Riemannian geometry, Einstein equation. As an illustration we finish by the calculation of the Schwarzschild metric—the simplest model in general relativity for the gravitational field of a star like our sun or our earth—, and as first application, we explain the deviation of light rays by the sun.

The notes are not intended as a self-contained reference: sometimes the proofs are omitted, short or left to the reader as exercises. The reader should complete these notes by referring to an excellent textbook like [GHL04]. On Einstein metrics at the end of the notes, a standard reference is [Bes87]. A few exercises are proposed in the text, some other ones at the end of the notes.

Chapter I

Submanifolds, manifolds

This chapter is a concise introduction to smooth manifolds, an excellent reference with all the required details is [Lee13].

1 Submanifolds of $\mathbb{R}^N$

1.a Submanifold charts

A submanifold of $\mathbb{R}^N$ of dimension n is a subset of $\mathbb{R}^N$ which is locally diffeomorphic to $\mathbb{R}^n \times \{0\} \subset \mathbb{R}^N$. More formally:

Definition 1.1. A ***submanifold*** of $\mathbb{R}^N$ of dimension n is a subset $M \subset \mathbb{R}^N$ such that, for each $x \in M$, there exists an open neighborhood U of x in $\mathbb{R}^N$, an open set $V \subset \mathbb{R}^N$, and a diffeomorphism $\phi : U \longrightarrow V$ such that

$$\phi(U \cap M) = V \cap (\mathbb{R}^n \times \{0\}).$$

We call such a map ϕ a ***chart*** of M.

Very often, we shall abbreviate 'a n-dimensional submanifold M of $\mathbb{R}^N$' in 'a submanifold M^n of $\mathbb{R}^N$'.

One simple example is the 2-sphere

$$S^2 = \{x^2 + y^2 + z^2 = 1\}$$

which is a 2-dimensional submanifold of $\mathbb{R}^3$. Indeed, consider the open set $U = \{z > 0, x^2 + y^2 < 1\}$ of $\mathbb{R}^3$, then the map $\phi : U \longrightarrow \mathbb{R}^3$ defined by

$$\phi(x, y, z) = (x, y, z - \sqrt{1 - x^2 - y^2})$$

is a diffeomorphism onto an open set of $\mathbb{R}^3$, and takes the sphere into $\mathbb{R}^2 \times \{0\}$. By permuting the variables $\pm x$, $\pm y$ and $\pm z$, one can cover the sphere with similar open sets and charts.

Similarly, the n-dimensional sphere

$$S^n = \{(x^0)^2 + \cdots + (x^n)^2 = 1\}$$

is a submanifold of $\mathbb{R}^{n+1}$.

Exercise. If $f : \mathrm{U} \subset \mathbb{R}^n \longrightarrow \mathbb{R}^{\mathrm{N}-n}$ is a smooth map defined on an open set U in $\mathbb{R}^n$, then the graph $\mathrm{M} = \{(x, f(x)), x \in \mathrm{U}\}$ is a n-dimensional submanifold of $\mathbb{R}^{\mathrm{N}}$.

For example, the n-dimensional hyperbolic space

$$\mathrm{H}^n = \{(x^0, \dots, x^n) \in \mathbb{R}^{n+1}, x^0 > 0, (x^0)^2 - (x^1)^2 - \cdots (x^n)^2 = 1\} \tag{I.1}$$

is a submanifold of $\mathbb{R}^{n+1}$.

1.b Tangent vectors

We now define what it means for a vector of $\mathbb{R}^{\mathrm{N}}$ to be tangent at x to a submanifold passing through x.

Definition 1.2. If M is a submanifold of $\mathbb{R}^{\mathrm{N}}$ and $x \in \mathrm{M}$, then a vector $\mathrm{X} \in \mathbb{R}^{\mathrm{N}}$ is a ***tangent vector*** to M at x if there exists a C^1 curve $c :]-\epsilon, \epsilon[\longrightarrow \mathrm{M} \subset \mathbb{R}^{\mathrm{N}}$, such that $c(0) = x$ and $c'(0) = \mathrm{X}$.

The space of all tangent vectors to M at x is called the ***tangent space*** of M at x and is noted $\mathrm{T}_x\mathrm{M}$.

Example 1.3. 1° If M^n is an affine subspace of $\mathbb{R}^{\mathrm{N}}$, so $\mathrm{M} = x_0 + \mathrm{V}$ where V is a vector subspace of $\mathbb{R}^{\mathrm{N}}$, then for all $x \in \mathrm{M}$, one has $\mathrm{T}_x\mathrm{M} = \mathrm{V}$.

2° Suppose $f : \mathrm{U} \subset \mathbb{R}^{\mathrm{N}} \longrightarrow \mathrm{V} \subset \mathbb{R}^{\mathrm{N}}$ is a diffeomorphism between two open sets U and V. If $c(t) \in \mathrm{M} \cap \mathrm{U}$ with $c(0) = x$, then $f(c(t)) \in f(\mathrm{M}) \cap \mathrm{V}$ and $\frac{d}{dt} f(c(t))|_{t=0} = d_x f(c'(0))$. It follows that $\mathrm{X} \in \mathrm{T}_x\mathrm{M}$ if and only if $d_x f(\mathrm{X}) \in \mathrm{T}_{f(x)}(f(\mathrm{M}))$. So we obtain an isomorphism

$$\mathrm{T}_x\mathrm{M} \xrightarrow{d_x f} \mathrm{T}_{f(x)}(f(\mathrm{M})).$$

Let us now use these two examples together: near a point $x \in \mathrm{M}$ take a chart $\phi : \mathrm{U} \ni x \longrightarrow \mathrm{V} \subset \mathbb{R}^{\mathrm{N}}$, with $\phi(\mathrm{M} \cap \mathrm{U}) = (\mathbb{R}^n \times \{0\}) \cap \mathrm{V}$, then it follows that

$$\mathrm{T}_x\mathrm{M} = (d_x\phi)^{-1}(\mathbb{R}^n \times \{0\}).$$

This proves that $\mathrm{T}_x\mathrm{M}$ is always a n-dimensional vector subspace of $\mathbb{R}^{\mathrm{N}}$.

1.c Submersions

Recall that a map from an open set of $\mathbb{R}^{\mathrm{N}}$ to $\mathbb{R}^{\mathrm{N}-n}$ is called a ***submersion*** if its differential is surjective at any point.

Theorem 1.4. *If a map $f : \mathrm{U} \subset \mathbb{R}^{\mathrm{N}} \longrightarrow \mathbb{R}^{\mathrm{N}-n}$ is a submersion, then for any $b \in \mathbb{R}^{\mathrm{N}-n}$, the set $f^{-1}(b)$ (if non empty) is a n-dimensional submanifold of $\mathbb{R}^{\mathrm{N}}$, and its tangent space at a point $x \in f^{-1}(b)$ is*

$$\mathrm{T}_x f^{-1}(b) = \ker(d_x f).$$

Proof. Let $x \in f^{-1}(b)$. Choose a supplementary subspace $\mathrm{F} \simeq \mathbb{R}^{\mathrm{N}-n}$ of $\ker(d_x f)$ in $\mathbb{R}^{\mathrm{N}}$, such that $\mathbb{R}^{\mathrm{N}} = \ker(d_x f) \oplus \mathrm{F}$. Then the map $d_x f : \mathrm{F} \longrightarrow \mathbb{R}^{\mathrm{N}-n}$ is an isomorphism. Consider the map $\phi : \mathrm{U} \longrightarrow \mathbb{R}^n \times \mathbb{R}^{\mathrm{N}-n}$:

$$\phi(x + (u, v)) = (u, f(x + (u, v)) - b), \quad u \oplus v \in \ker(d_x f) \oplus \mathrm{F} = \mathbb{R}^{\mathrm{N}}.$$

Its differential,

$$d_x\phi(\dot u, \dot v) = (\dot u, d_x f(\dot v)),$$

is an isomorphism, so by the inverse function theorem, the map ϕ is a diffeomorphism on a small neighborhood V of x. Clearly, $f^{-1}(b) \cap V = \phi^{-1}(\mathbb{R}^n \times \{0\} \cap \phi(V))$, so ϕ is a chart for $f^{-1}(b)$ near x, and

$$T_x(f^{-1}(b)) = (d_x\phi)^{-1}(\mathbb{R}^n \times \{0\}) = \ker(d_x f).$$

□

Remark 1.5. Since ϕ is a local diffeomorphism, one can write near x the map f as $f = g\circ\phi + b$, where $g : \mathbb{R}^N \longrightarrow \mathbb{R}^{N-n}$ is given by

$$g(x^1, \dots, x^N) = (x^{n+1}, \dots, x^N). \tag{I.2}$$

So the meaning of the theorem is that, up to a diffeomorphism, any submersion has the local form (I.2).

Example 1.6. 1° The curve $y^2 = x^3 - x$ is a smooth curve (that is a 1-dimensional submanifold of $\mathbb{R}^2$. Indeed, consider $f(x, y) = y^2 - x^3 + x$, then $d_{(x,y)}f = (-3x^2 + 1, 2y)$ which vanishes only at the points $(\pm\frac{1}{\sqrt{3}}, 0)$. Since these two points are not in $f^{-1}(0)$, the result follows from the theorem applied to the map f on the open set $U = \mathbb{R}^2 \setminus \{(\pm\frac{1}{\sqrt{3}}, 0)\}$.

2° The sphere $S^n = \{(x^0)^2 + \cdots + (x^n)^2 = 1\}$ and the hyperbolic space $H^n = \{x^0 > 0, (x^0)^2 - (x^1)^2 - \cdots (x^n)^2 = 1\}$ are submanifolds of $\mathbb{R}^{n+1}$.

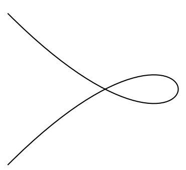
3° (Exercise) The group $O(n)$ is a submanifold of $\mathbb{R}^{n^2}$ (the space of $n \times n$ matrices). Apply the theorem to the map $f(A) = AA^t - 1$ from matrices to symmetric matrices. To prove that f is a submersion at each point $x \in O(n)$, use the invariance $f(Ax) = f(A)$ to reduce to the case $x = 1$.

1.d Immersions

We now pass to immersions: recall that a map f from an open set $U \subset \mathbb{R}^n$ to $\mathbb{R}^N$ is called an ***immersion*** if $d_x f$ is injective at each $x \in U$.

Example 1.7. 1° The map $\mathbb{R}^n \longrightarrow \mathbb{R}^{n+1}$ given by $(x^1, \dots, x^n) \mapsto (1 + (x^1)^2 + \cdots + (x^n)^2, x^1, \dots, x^n)$ is an immersion and a bijection from $\mathbb{R}^n$ to its image $H^n \subset \mathbb{R}^{n+1}$, the hyperbolic space.

2° The two figures below represent immersions $\mathbb{R} \longrightarrow \mathbb{R}^2$ whose image is not a submanifold: the first is not injective, it has a double point; the second one is injective but not proper.



Definition 1.8. We say that a map $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^N$ is an ***embedding***, if f is an immersion and f is an homeomorphism from U to $f(U)$.

Lemma 1.9. *1° A proper injective immersion is an embedding.*
2° If $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^N$ is an embedding, then $f(U)$ is a n-dimensional submanifold of $\mathbb{R}^N$.

Proof. See for example the book [Lee13]. □

2 Manifolds

2.a Atlas and manifold

Definition 2.1. Let M be a topological space. A C^∞ ***atlas*** on M is the data of

1. an open covering $(U_i)_{i\in I}$ of M,
2. homeomorphisms $\phi_i : U_i \longrightarrow V_i \subset \mathbb{R}^n$ onto open sets of $\mathbb{R}^n$,

such that for any i and j, the composite

$$\phi_i \circ \phi_j^{-1} : \phi_j(U_i \cap U_j) \longrightarrow \phi_i(U_i \cap U_j)$$

is a C^∞ diffeomorphism.

The maps $\phi_i \circ \phi_j^{-1}$ are called the ***transition functions*** of the atlas.

Example 2.2. If M^n is a submanifold of $\mathbb{R}^N$, then for any point $x \in M$, we have a chart $\phi_x : U \subset \mathbb{R}^N \longrightarrow \mathbb{R}^n \times \mathbb{R}^{N-n}$ sending $U \cap M$ to $\mathbb{R}^n \times \{0\}$. Denote by π the projection $\mathbb{R}^n \times \mathbb{R}^{N-n} \longrightarrow \mathbb{R}^n$, then the collection $(\pi \circ \phi_x)_{x\in M}$ is an atlas for M.

Two atlas are ***equivalent atlases*** if their union is an atlas. Concretely this means that if ϕ_i and ψ_j are the charts of the first and second atlas, then the compositions $\phi_i \circ \psi_j^{-1}$ are C^∞ on the open sets where they are defined.

Definition 2.3. A ***manifold*** M is the data of:

1. a topological space M which is Hausdorff and countable at infinity (that is a countable union of compact sets);
2. an equivalence class of C^∞ atlas on M.

The integer n appearing in the definition of an atlas is the ***dimension*** of M. It is constant on each connected component of M. (One usually considers manifolds with all the connected components sharing the same dimension).

Example 2.2 shows that a submanifold of $\mathbb{R}^N$ is a manifold, there is only the topological assumption of being countable at infinity to prove (exercise): prove that a submanifold $M \subset \mathbb{R}^N$ is locally closed, that is the intersection of a closed set ($\bar{M}$) and of an open set of $\mathbb{R}^N$ (the union of the images of the submanifold charts); prove that an open set of $\mathbb{R}^N$ is countable at infinity, and conclude.

Remark 2.4. One can define also the notion of a C^k atlas and a C^k manifold by asking that the transition functions $\phi_i \circ \phi_j^{-1}$ be only in C^k. The fact that the dimension is locally constant remains obvious for $k > 0$, but is a more difficult topological result for $k = 0$ (in that case, one says that M is a ***topological manifold***).

Remark 2.5. If the dimension is even, $n = 2m$, then the charts take values in $\mathbb{C}^m = \mathbb{R}^{2m}$. If the transition functions $\phi_i \phi_j^{-1}$ are holomorphic maps, then the manifold is a ***complex manifold***.

Remark 2.6. One can define an abstract atlas on a set M in the same way we defined an atlas, but we drop the assumption that the transitions are homeomorphisms, since we have no topology. Actually an abstract atlas on M induces a unique topology on M: it is defined by saying that $U \subset M$ is open if the image $\phi_i(U \cap U_i)$ by each chart ϕ_i is an open set of $\mathbb{R}^n$. The uniqueness implies that the topology defined by an atlas on a topological space M coincides with the given topology of M.

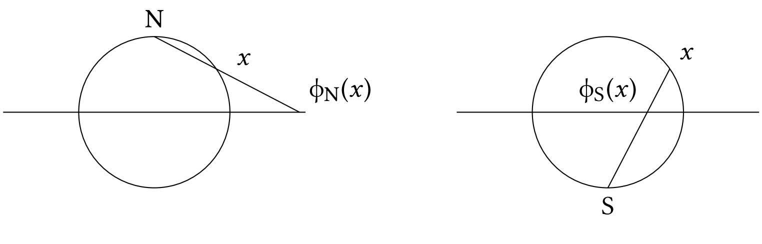
The topological assumptions on M in the definition 2.3 can be checked directly on the atlas:

- M is Hausdorff if for each charts ϕ_i, ϕ_j the set $\{(x, \phi_i \circ \phi_j^{-1}(x)), x \in \phi_j(U_i \cap U_j)\}$ (the graph of the transition) is closed in $\phi_j(U_j) \times \phi_i(U_i)$;
- M is countable at infinity if the atlas contains a countable sub-atlas.

Therefore we can define a manifold structure on a set M just by giving an abstract atlas, and checking the two above conditions to get a good topology on M.

2.b The sphere

It is a good point to stop after this rather abstract definition, and to consider what it means on the concrete example of the sphere S^n. In coordinates $(x^0, \dots, x^n)$, the north pole N and the south pole S are the points $(\pm 1, 0, \dots, 0)$. We now define two charts with values in $\mathbb{R}^n$, considered as the hyperplane $\{x^0 = 0\}$ in $\mathbb{R}^{n+1}$. For $x \in S^n \setminus \{N\}$ define the stereographic projection $\phi_N(x)$ from the north pole to be the point of $\mathbb{R}^n$ where the line passing through N and x meets $\mathbb{R}^n$; the stereographic projection ϕ_S from the south pole is defined similarly:



In formulas:

$$\phi_N(x^0, x^1, \dots, x^n) = \frac{(x^1, \dots, x^n)}{1 - x^0}, \quad \phi_S(x^0, x^1, \dots, x^n) = \frac{(x^1, \dots, x^n)}{1 + x^0}.$$

The transition function is the inversion

$$\phi_N\phi_S^{-1}(x^1,\dots,x^n) = \frac{(x^1,\dots,x^n)}{(x^1)^2+\cdots+(x^n)^2}.$$

This is now an 'abstract' description of the sphere, meaning that it does not rely on seeing it as a submanifold of $\mathbb{R}^{n+1}$.

2.c The projective space

Our next example will be defined directly as an abstract manifold. It is the ***real projective space*** $\mathbb{RP}^n$ of all real lines in $\mathbb{R}^{n+1}$. This can be identified with the quotient $S^n/(\mathbb{Z}/2)$ of the sphere by the antipodal map.

A nonzero vector $(x^0,\dots,x^n)\in\mathbb{R}^n$ generates a line in $\mathbb{R}^n$, that is a point of $\mathbb{RP}^n$ which is denoted $[x^0:\cdots:x^n]$. Therefore, if λ is any non vanishing number, one has

$$[x^0:\cdots:x^n] = [\lambda x^0:\cdots:\lambda x^n].$$

The $[x^0:\cdots:x^n]$ are the ***homogeneous coordinates*** on $\mathbb{RP}^n$. We turn $\mathbb{RP}^n$ into a manifold by giving an explicit atlas, and by checking that the transition functions are smooth: let $U_i\subset\mathbb{RP}^n$ the open set given by $U_i=\{x_i\neq 0\}$. On U_i we have the chart $\phi_i : U_i \longrightarrow \mathbb{R}^n$ given by

$$\phi_i([x^0:\cdots:x^n]) = \Big(\frac{x^0}{x^i},\dots,\widehat{\frac{x^i}{x^i}},\dots,\frac{x^n}{x^i}\Big),$$

where the hat means that this term is omitted. The transition function $\phi_i\phi_j^{-1} : \phi_j(U_i\cap U_j)\longrightarrow\phi_i(U_i\cap U_j)$ is given by

$$\phi_i\phi_j^{-1}(x^1,\dots,x^n) = \Big(\frac{x^1}{x^i},\dots,\frac{x^{j-1}}{x^i},\frac{1}{x^i},\frac{x^{j+1}}{x^j},\dots,\widehat{\frac{x^i}{x^i}},\dots,\frac{x^n}{x^j}\Big).$$

The reader can check that the topological conditions listed in remark 2.6 are satisfied.

The $\mathbb{RP}^n$ for different n's are related in the following way. The chart open set $U_n=\{x^n=1\}$ is diffeomorphic to $\mathbb{R}^n$ by ϕ_n. The complement

$$\mathbb{RP}^n\setminus U_n = \{[x^0:\cdots:x^{n-1}:0]\}$$

identifies naturally with $\mathbb{RP}^{n-1}$. In this way one obtains the $\mathbb{RP}^n$ inductively: starting from $\mathbb{RP}^0$ which is reduced to a point,

- $\mathbb{RP}^1=\mathbb{R}\cup\{\text{pt.}\}$ is a circle;
- $\mathbb{RP}^2=\mathbb{R}^2\cup\mathbb{RP}^1$ is the union of the plane and the line at infinity;
- more generally, $\mathbb{RP}^n=\mathbb{R}^n\cup\mathbb{RP}^{n-1}$.

Finally, observe that all we have done has a meaning if we decide that the x^i are complex coordinates rather than real coordinates. In this way, one obtains the structure of a complex manifold on the ***complex projective space*** $\mathbb{CP}^m$ of complex lines in $\mathbb{C}^{m+1}$. In particular, one obtains that

$$\mathbb{CP}^1=\mathbb{C}\cup\{\text{pt.}\}$$

which will turn to be diffeomorphic to a 2-sphere.

2.d Submanifolds

The notion of submanifold of $\mathbb{R}^N$ studied in section 1 extends to a notion of submanifold of a manifold. Quick definition: $X^n \subset M^N$ is a submanifold if for each chart ϕ defined on an open set $U \subset M$, then $\phi(X \cap U)$ is a submanifold of $\phi(U) \subset \mathbb{R}^N$. This means that up to a diffeomorphism of $\mathbb{R}^N$, one has $\phi(X \cap U) = \phi(U) \cap (\mathbb{R}^n \times \{0\})$. So a more formal definition is:

Definition 2.7. A set $X^n \subset M^N$ is a ***submanifold*** of M^N if near each point of M, there is a chart $\phi : U \subset M \longrightarrow V \subset \mathbb{R}^N$ such that $\phi(X \cap U) = (\mathbb{R}^n \times \{0\}) \cap V \subset \mathbb{R}^n \times \mathbb{R}^{N-n}$.

A submanifold X inherits a manifold structure, for which the charts are the restriction of the submanifold charts to X. In particular the submanifolds of $\mathbb{R}^N$ are manifolds.

2.e Smooth maps

A chart on a manifold M^n is a local map $\phi = (x^1, \dots, x^n)$ to $\mathbb{R}^n$. The $(x^1, \dots, x^n)$ are local functions on M called local coordinates. A map $f : M^n \longrightarrow \mathbb{R}^p$ is locally an application of n variables

$$f(x^1, \dots, x^n) = (f_1(x^1, \dots, x^n), \dots, f_p(x^1, \dots, x^n)),$$

and we declare it to be smooth if each f_i is a smooth function of the variables $(x^1, \dots, x^n)$. If the target is a manifold N^p, we have also coordinates on N^p, and we have to replace the (f_i) by local coordinates on N as well. This leads to the following definition.

Definition 2.8. A ***smooth map between two manifolds*** M^n and N^p (or C^∞ map) is a continuous map $f : M^n \longrightarrow N^p$ such that for any charts $\phi : U \subset M \longrightarrow \mathbb{R}^n$ and $\psi : V \subset N \longrightarrow \mathbb{R}^p$, the map

$$\psi f \phi^{-1} : \phi(U \cap f^{-1}(V)) \subset \mathbb{R}^n \longrightarrow \psi(V) \subset \mathbb{R}^p$$

is C^∞.

As we have just seen, this definition means that in the charts, the coordinates of $f(x)$ are smooth functions of the coordinates of x. Of course the definition does not depend on the choice of charts, because the transition between two charts is always a C^∞ diffeomorphism.

If $f : M^n \longrightarrow N^p$ is smooth and a bijection such that f^{-1} is also smooth, we say that f is a C^∞ ***diffeomorphism***. Of course this implies $n = p$. In that case we say that M and N are diffeomorphic.

Example 2.9. 1° A map $f :]a, b[\longrightarrow M^n$ is smooth if for any chart $\phi : U \subset M \longrightarrow \mathbb{R}^n$ the composite $\phi \circ f$ is smooth. Write $\phi = (x^1, \dots, x^n)$ (the (x^i) are called ***local coordinates*** on M), the map f can be locally written $f = (f^1, \dots, f^n)$ where $f^i = x^i \circ f$. Then f is smooth means that each f^i is C^∞ as a real function of one real variable.

2° A function $f : M^n \longrightarrow \mathbb{R}$ is smooth if for any chart $\phi = (x^1, \dots, x^n)$ as above, the function $f \circ \phi^{-1}$ on $\mathbb{R}^n$ is smooth, that is f is smooth as a function of $(x^1, \dots, x^n)$. One often identifies U with its image in $\mathbb{R}^n$ and then one can write directly $f(x^1, \dots, x^n)$.

Exercise. 1° Prove that the following maps are smooth:

- the quotient by the antipodal map $S^n \longrightarrow \mathbb{R}P^n$;
- the map $S^3 \longrightarrow \mathbb{C}P^1$ taking a vector $x \in S^3$ to the complex line that it generates in $\mathbb{C}^2$.

2° Prove that S^2 and $\mathbb{C}P^1$ are diffeomorphic.

2.f Submersions and immersions

From the definition, a smooth map $f : M^n \longrightarrow N^p$ between two manifolds is just locally a smooth map from $\mathbb{R}^n$ to $\mathbb{R}^p$. It is easy to see that the notions of submersion and immersion do not depend on the choice of the charts, so it makes sense to speak about f being a submersion or an immersion. Then the results of section 1 on submanifolds of $\mathbb{R}^N$ extend to abstract manifolds, in particular we have the notion of embedding of a manifold into another manifold from definition 1.8, and theorem 1.4 and lemma 1.9 can be applied in this more general setting.

Remark 2.10. Any manifold M^n can be embedded in $\mathbb{R}^N$ for N large enough, so is diffeomorphic to a submanifold of $\mathbb{R}^N$. A theorem of Whitney gives a bound on N in terms of n: the 'easy' version is $N = 2n + 1$ (see [Lee13]), but Whitney actually proved that it is possible to take $N = 2n$.

3 Tangent vectors

3.a Paths and tangent vectors

We now turn to the notion of a tangent vector at a point x in a manifold M. Recall the case of submanifolds of $\mathbb{R}^N$ (definition 1.2): a tangent vector at x to a submanifold $M \subset \mathbb{R}^N$ is the derivative $c'(0)$ of a path $c :]-\epsilon, \epsilon[\longrightarrow M$ such that $c(0) = x$. Of course two paths c_1 and c_2 define the same tangent vector if $c_1'(0) = c_2'(0)$. It turns out that this point of view leads to a good definition for an abstract manifold:

Definition 3.1. Let M be a manifold and x a point of M.

1° We say that two paths $c_1, c_2 :]-\epsilon, \epsilon[\longrightarrow M$ such that $c_1(0) = c_2(0) = x$ are ***equivalent paths*** if for any local chart ϕ at x, one has

$$(\phi \circ c_1)'(0) = (\phi \circ c_2)'(0).$$

2° A ***tangent vector*** at x to M is an equivalence class of paths for this relation.

3° The set of all tangent vectors at x to M is called the ***tangent space*** of M at x and noted $T_x M$.

Observe that in the first part of the definition, it is equivalent to ask the equality of the derivatives for one chart or for all charts.

If we have a smooth map between two manifolds, $f : M^n \longrightarrow N^p$, then to a path c at $x \in M$ we can associate the path $f(c)$ at $f(x)$. It is easy to check that if c_1 and c_2 are equivalent, then so are $f(c_1)$ and $f(c_2)$. It follows that we obtain a well defined map

$$d_x f : T_x M \longrightarrow T_{f(x)} N. \tag{I.3}$$

If f is a diffeomorphism, then it is easy to check that $(d_x f)^{-1} = d_x(f^{-1})$.

Apply this to a local chart ϕ at x: the map ϕ is a diffeomorphism $U \subset M \longrightarrow V \subset \mathbb{R}^n$, so we obtain an isomorphism

$$d_x\phi : T_xM \xrightarrow{\sim} \mathbb{R}^n, \quad c \longmapsto (\phi \circ c)'(0).$$

We would like to deduce that T_xM is a vector space, since it is identified to $\mathbb{R}^n$. Again, this will be true if it does not depend on the chart ϕ. This is a good place to check this kind of statement, that we are using repeatedly: if we have another chart ψ, so we have a transition function $\psi\phi^{-1}$, then the following diagram is commutative (we can assume that the chart is centered: $\phi(x) = 0$):

$$\begin{array}{ccc} & T_xM & \\ {}^{d_x\phi}\swarrow & & \searrow^{d_x\psi} \\ \mathbb{R}^n & \xrightarrow{d_0(\psi\phi^{-1})} & \mathbb{R}^n \end{array}$$

So the two different identifications of T_xM with $\mathbb{R}^n$ differ by the linear isomorphism $d_0(\psi\phi^{-1})$, which preserves the vector space structure. So the vector space structures induced on T_xM from $d_x\phi$ and $d_x\psi$ coincide.

3.b The tangent bundle

We now turn to the problem of constructing the *manifold* of all tangent vectors at all points of a manifold M. First consider the case of a submanifold $M^n \subset \mathbb{R}^N$. Then we can consider

$$TM = \{(x, X) \in M \times \mathbb{R}^N, X \in T_xM\} \subset \mathbb{R}^N \times \mathbb{R}^N.$$

Then (exercise):

1. TM is a submanifold of $\mathbb{R}^N \times \mathbb{R}^N$: if $\phi : U \subset \mathbb{R}^N \longrightarrow \mathbb{R}^N$ is a local submanifold chart for M, then

$$U \times \mathbb{R}^N \longrightarrow \mathbb{R}^N \times \mathbb{R}^N, \quad (x, X) \mapsto (\phi(x), d_x\phi(X))$$

   is a submanifold chart for TM;

2. the projection $(x, X) \longrightarrow x$ gives a map $\pi : TM \longrightarrow M$, such that $\pi^{-1}(x) = T_xM$, that is the fibers are the tangent spaces of M.

Observe that in particular, for an open set $U \subset \mathbb{R}^n$, we simply have $TU = U \times \mathbb{R}^n$: the tangent vectors to U at a point identify to $\mathbb{R}^n$.

Now pass to an abstract manifold M: there is a way to do the same construction, but the result TM will be a manifold instead of a submanifold of $\mathbb{R}^N \times \mathbb{R}^N$. Let us describe it now:

- as a set, $TM = \coprod_{x\in M} T_xM = \{(x, X), x \in M, X \in T_xM\}$; there is a projection $\pi : TM \longrightarrow M$ given by $\pi(x, X) = x$;
- the manifold structure is described by the following charts: if $\phi : U \subset M \longrightarrow \mathbb{R}^n$ is a chart for M, then a chart $d\phi : \pi^{-1}(U) \subset TM \longrightarrow \mathbb{R}^{2n}$ for TM is given by

$$d\phi(x, X) = (\phi(x), d_x\phi(X)).$$

We now write the transition functions for this atlas. In particular this will prove that TM has indeed a manifold structure. Suppose we have two charts ϕ_1 and ϕ_2 of M defined on open sets U_i, then we have the charts $d\phi_i$ of TM defined on $\pi^{-1}(U_i)$. On the intersection $U_{12} = U_1 \cap U_2$, the two charts are related by the following commutative diagram:

$$\begin{array}{ccc} & \pi^{-1}(U_{12}) & \\ {}^{d\phi_1}\swarrow & & \searrow^{d\phi_2} \\ \phi_1(U_{12}) \times \mathbb{R}^n \subset \mathbb{R}^n \times \mathbb{R}^n & \xrightarrow{d(\phi_2\phi_1^{-1})} & \phi_2(U_{12}) \times \mathbb{R}^n \subset \mathbb{R}^n \times \mathbb{R}^n \end{array} \tag{I.4}$$

where of course $d(\phi_2\phi_1^{-1})$ is the differential of the transition function $\phi_2\phi_1^{-1}$, that is $d(\phi_2\phi_1^{-1})(x, X) = (\phi_2\phi_1^{-1}(x), d_x(\phi_2\phi_1^{-1})(X))$.

Therefore we have an atlas on TM, and using remark 2.6 one can check that it defines a Hausdorff, countable at infinity, topology on TM. Therefore TM is a manifold of dimension $2n$.

Now come back to a smooth map $f : M^n \longrightarrow N^p$, then the collection of the maps $d_xf : T_xM \longrightarrow T_xN$ gives a smooth ***tangent map*** $df : TM \longrightarrow TN$ preserving the vector bundle structure, that is one has the commutative diagram

$$\begin{array}{ccc} TM & \xrightarrow{df} & TN \\ \downarrow & & \downarrow \\ M & \xrightarrow{f} & N \end{array}$$

where for each x the induced map $d_xf : T_xM \longrightarrow T_{f(x)}N$ is a linear map. Let us give concrete formulas (and this will prove that df is indeed a smooth map between manifolds): in local coordinates (x^i) on M and (y^j) on N (recall that this means that the local map $M \longrightarrow \mathbb{R}^n$ given by the (x^i) is a chart), we can write

$$f = (f^1(x^1, \dots, x^n), \dots, f^p(x^1, \dots, x^n)),$$

and then df is calculated as

$$df(x^1, \dots, x^n, X^1, \dots, X^n) = (f^1, \dots, f^p, X^i\frac{\partial f^1}{\partial x^i}, \dots, X^i\frac{\partial f^p}{\partial x^i}). \tag{I.5}$$

In this formula we used the **implicit summation convention**: if we find in a formula the same i as an index and as an exponent, then one must understand that the result is just the sum on all possible i's, so in the above formula $X^i\frac{\partial f^j}{\partial x^i}$ means $\sum_{i=1}^n X^i\frac{\partial f^j}{\partial x^i}$.

So we see that df is nothing but the abstract version (for manifolds) of the differential of a map $\mathbb{R}^n \longrightarrow \mathbb{R}^p$. In particular one has the composition formula

$$d(g \circ f) = dg \circ df.$$

3.c Vector bundles

The transition functions (I.4) of TM are of a special kind: for each $x \in \phi_1(U_{12})$, the restriction of $d(\phi_2\phi_1^{-1})$ to $\{x\} \times \mathbb{R}^n$ is a linear isomorphism (actually this is a way to see that each fiber $\pi^{-1}(x)$ has a vector space structure). This a special example of the more general notion of a ***vector bundle*** over M: a rank p vector bundle E over M is a manifold E with a smooth map $\pi : E \longrightarrow M$, and charts of the type (I.4), called ***local trivializations***:

$$\begin{array}{ccc} \psi_i : \pi^{-1}(U_i) & \longrightarrow & U_i \times \mathbb{R}^p \\ {\scriptstyle\pi} \searrow & & \swarrow \\ & U_i & \end{array}$$

where the diagram is commutative, such that the transition functions

$$\psi_i\psi_j^{-1} : \pi^{-1}(U_{ij}) \times \mathbb{R}^p \longrightarrow \pi^{-1}(U_{ij}) \times \mathbb{R}^p$$

have the form

$$\psi_i\psi_j^{-1}(x, \xi) = (x, u_{ij}(x)(\xi)), \quad u_{ij} : U_{ij} \longrightarrow GL_p\mathbb{R} \text{ smooth.}$$

(Note that here we have not exactly charts as in equation (I.4) because we have not applied a local chart ϕ_i to the x part of (x, ξ); to have true manifold charts for E we should compose ψ_i with $(\phi_i, 1_{\mathbb{R}^p})$.)

From the local trivializations of a vector bundle $\pi : E \longrightarrow M$ and the transitions functions, we see immediately that the fibers $E_x := \pi^{-1}(x)$ have a vector space structure.

The vector bundle structure is completely characterized by the set of $GL_p\mathbb{R}$ valued transition functions (u_{ij}) on U_{ij}. They satisfy $u_{ji} = u_{ij}^{-1}$ and the cocycle identity

$$u_{ij}u_{jk}u_{kj} = 1 \text{ on } U_i \cap U_j \cap U_k.$$

Conversely, such a set of transition functions relative to some covering (U_i) of M defines a vector bundle on M, defined by (exercise)

$$E = (\coprod_i U_i \times \mathbb{R}^p)/(j, x, \xi) \sim (i, x, u_{ij}(x)\xi) \text{ if } x \in U_{ij}.$$

Also remark that one can define similarly complex rank p vector bundles, where the transition functions take their values in the group $GL_p\mathbb{C}$ of complex linear isomorphisms.

Finally, a ***smooth section*** of a vector bundle E is a C^∞ map $s : M \longrightarrow E$ such that $\pi \circ s = 1_M$. This just means $s(x) \in E_x$ for each x. In a local trivialization $E|_{U_i} \simeq U_i \times \mathbb{R}^p$, such a section is given by p coordinates $s_1(x), \dots, s_p(x)$ which are smooth functions. The set of smooth sections (resp. compactly supported smooth sections) of E over M will be denoted $\Gamma(M, E)$ (resp. $C_c^\infty(M, E)$). If there is no ambiguity we may abbreviate into $\Gamma(E)$ and $C_c^\infty(E)$.

If E is the trivial vector bundle $M \times \mathbb{R}$, then sections of E are just smooth functions on M. The algebra of smooth functions on M is noted $C^\infty(M)$.

If $s, t \in \Gamma(M, E)$ and $f \in C^\infty(M)$, then $s + t$ (defined by $(s + t)(x) = s(x) + t(x)$) and fs (defined by $(fs)(x) = f(x)s(x)$) are still smooth sections of E (exercise), that is $s + t$ and $fs \in \Gamma(M, E)$. Therefore $\Gamma(M, E)$ is a $C^\infty(M)$-module.

3.d The cotangent bundle

An important example of vector bundle is the ***cotangent bundle*** of a manifold M^n. This is a vector bundle, denoted T^*M, whose fiber at $x \in M$ is the dual T_x^*M of the tangent space T_xM. If we want to define T^*M by transition functions, we use a covering of M by open sets (U_i) on which TM is trivialized, with transition functions (u_{ij}); then it is easy to see that the transition functions for T^*M are $({}^t u_{ij}^{-1})$.

If $f : M \longrightarrow \mathbb{R}$ is a smooth function, its differential d_xf is a linear form on T_xM, so $d_xf \in T_x^*M$. It follows that df can be interpreted as a section of T^*M. Sections of T^*M are called 1-forms. If (x^i) are local coordinates, then a local basis of T^*M is given by the differential dx^i of the coordinates. Then, for any function f the formula (I.5) can be written

$$df = \frac{\partial f}{\partial x^i} dx^i,$$

which actually proves that df is a smooth section of T^*M.

This construction of the cotangent bundle is just a special case of the following general fact. If we have one or several bundles, any algebraic operation on the underlying vector spaces can be done fiberwise to give rise to a new vector bundle. For example, if E and F are vector bundles, then $E \oplus F$, $E \otimes F$ and Hom(E, F) are vector bundles, whose fibers at $x \in M$ are $E_x \oplus F_x$, $E_x \otimes F_x$ and $\mathrm{Hom}(E_x, F_x)$. The proof is left as an exercise.

4 Vector fields and bracket

4.a Vector fields and derivations

If $f : M \longrightarrow \mathbb{R}$ is a smooth function, then at each point $x \in M$ we have a differential $d_xf : T_xM \longrightarrow \mathbb{R}$. If $X \in T_xM$ is a tangent vector at x, then we can consider the map

$$D_X : f \longrightarrow d_xf(X). \tag{I.6}$$

It satisfies the Leibniz rule

$$D_X(fg) = (D_Xf)g(x) + f(x)(D_Xg).$$

Now suppose that we have a tangent vector $X(x) \in T_xM$ for each $x \in M$, depending smoothly on x, that is X is a section of TM (smooth sections of TM are called ***vector fields***). The map (I.6) considered for all x together gives a map, called ***Lie derivative***

$$\mathscr{L}_X : C^\infty(M) \longrightarrow C^\infty(M), \quad f \longmapsto df(X), \tag{I.7}$$

satisfying the Leibniz rule

$$\mathscr{L}_X(fg) = (\mathscr{L}_Xf)g + f\mathscr{L}_Xg.$$

So the Lie derivative is a derivation of the algebra $C^\infty(M)$.

Let us see what this means in local coordinates (x^i) on M. Then the tangent bundle is locally identified with $\mathbb{R}^n$, and we consider the basis vector fields

$$e_1 = (1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1).$$

Any local vector field X can be written as $X = X^i e_i$ (recall the implicit summation convention). If f is a function on M, we can consider locally f as a function of the local coordinates, $f(x^1, \dots, x^n)$. Then we check easily that

$$\mathscr{L}_{e_i} f = \frac{\partial f}{\partial x^i}. \tag{I.8}$$

This is why the vector field e_i is usually identified with the corresponding derivation, $e_i = \frac{\partial}{\partial x^i}$. In the sequel we will use the notation $\frac{\partial}{\partial x^i}$ instead of e_i.

We now generalize this identification. If we have a commutative $\mathbb{R}$-algebra A, then a ***derivation*** of A is a $\mathbb{R}$-linear map $D : A \longrightarrow A$ satisfying the Leibniz formula $D(ab) = (Da)b + a(Db)$. The set of all derivations of A is noted D(A).

The main result in this section says that a vector field is the same thing as a derivation of the algebra $C^\infty(M)$:

Theorem 4.1. *The map* $X \mapsto \mathscr{L}_X$ *is an isomorphism* $\Gamma(TM) \longrightarrow D(M)$.

Proof. Let D be a derivation of $C^\infty(M)$. We want to find a vector field X such that $D = \mathscr{L}_X$.

First step: D is a local operator, that is if U is open, then

$$f|_U = 0 \Longrightarrow (Df)|_U = 0. \tag{I.9}$$

This implies that if f and g coincide on U, then $D(f - g)|_U = 0$ so Df and Dg coincide on U: in particular $Df(x)$ depends only on the values of f on an arbitrary small neighborhood of x. Now prove (I.9): choose a function χ with compact support in U, then

$$D(\chi f) = \chi Df + (D\chi)f.$$

If $f|_U = 0$, then $\chi f = 0$ so $\chi Df = -(D\chi)f$ vanishes on U.

Note that we can now define Df for f defined on any open set $U \subset M$. Indeed suppose $x \in V \subset\subset U$, we can choose χ with compact support in U such that $\chi|_V = 1$, then $\chi f \in C^\infty(M)$ and we can define $(Df)(x) := D(\chi f)(x)$. This does not depend on the choices.

Second step: $D(1) = 0$. This is clear, since $D(1^2) = 1D(1) + D(1)1 = 2D(1)$.

Third step: local. Take local coordinates (x^i) and write

$$f(x^1, \dots, x^n) = f(0) + x^i g_i(x^1, \dots, x^n)$$

for smooth functions g_i. Because of the first step, we can apply D to these local functions: using Leibniz identity we get

$$\begin{aligned}(Df)(0, \dots, 0) &= \sum_1^n (Dx^i)(0, \dots, 0) g_i(0, \dots, 0)\\ &= \sum_1^n (Dx^i)(0, \dots, 0) \frac{\partial f}{\partial x^i}(0, \dots, 0)\end{aligned}$$

So we must have $X(0) = (Dx^i)(0, \dots, 0)\frac{\partial}{\partial x^i}$. Doing this at each point x of the chart gives

$$X(x) = (Dx^i)(x)\frac{\partial}{\partial x^i}. \tag{I.10}$$

This proves the uniqueness of X, and the local existence of X.

For the existence, we define X on each coordinate open set by (I.10), then by the uniqueness statement the various vector fields defined on the open sets must coincide on the intersections, so they glue together to define the expected vector field. □

The theorem enables to define easily the bracket of two vector fields:

Definition 4.2. If X and Y are two vector fields on M, then their ***bracket*** $[X, Y]$ is the vector field corresponding to the derivation $[\mathscr{L}_X, \mathscr{L}_Y] = \mathscr{L}_X\mathscr{L}_Y - \mathscr{L}_Y\mathscr{L}_X$.

This rather abstract definition corresponds to a simple calculation: taking local coordinates (x^i), we write $X = X^i \frac{\partial}{\partial x^i}$ and $Y = Y^i \frac{\partial}{\partial x^i}$, then

$$\begin{aligned}\mathscr{L}_X\mathscr{L}_Y f - \mathscr{L}_Y\mathscr{L}_X f &= X^j \frac{\partial}{\partial x^j}\big(Y^i \frac{\partial f}{\partial x^i}\big) - Y^j \frac{\partial}{\partial x^j}\big(X^i \frac{\partial f}{\partial x^i}\big)\\ &= \big(X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j}\big)\frac{\partial f}{\partial x^i}\end{aligned}$$

Therefore

$$[X, Y] = \big(X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j}\big)\frac{\partial}{\partial x^i}. \tag{I.11}$$

Lemma 4.3. *1°* $[X, fY] = f[X, Y] + (\mathscr{L}_X f)Y$.

2° Jacobi identity: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$.

3° If N *is a submanifold of* M *and the restrictions of* X *and* Y *to* N *lie inside* $TN \subset TM|_N$, *then* $[X, Y]|_N$ *is tangent to* N *and equals* $[X|_N, Y|_N]$.

Proof. The proof is easy given the explicit formula (I.11) for the bracket, and is left to the reader. The Jacobi identity is best seen as the consequence of the obvious corresponding algebraic identity

$$[\mathscr{L}_X, [\mathscr{L}_Y, \mathscr{L}_Z]] + \cdots = 0.$$

□

4.b First order ordinary differential equations

If we have a smooth curve $c : I \longrightarrow M$ in M, defined on an interval $I \subset \mathbb{R}$, then for each t we can consider the derivative $\dot{c}(t) = d_t c(\frac{\partial}{\partial t}) \in T_{c(t)}M$. Let X be a vector field on the manifold M^n, we look for solutions $c : I \longrightarrow M$ of the equation

$$\dot{c} = X(c). \tag{I.12}$$

Note that for each $t \in I$, both $\dot{c}(t)$ and $X(c(t))$ belong to $T_{c(t)}M$, so the equation makes sense.

For example, if $M = \mathbb{R}^2$, then a vector field is $X = f(x, y)\frac{\partial}{\partial x} + g(x, y)\frac{\partial}{\partial y}$, the curve is $c(t) = (x(t), y(t))$ and the equation (I.12) is the system

$$\begin{cases}\dot{x} &= f(x, y)\\ \dot{y} &= g(x, y).\end{cases}$$

More generally, in local coordinates (x^i), the equation (I.12) becomes

$$\dot{x}^i = X^i(x^1, \dots, x^n).$$

So if we give the initial condition $c(t_0) \in \mathrm{M}$, near t_0 the path $c(t)$ will remain in the open set of coordinates and the equation (I.12) translates into a first order system of ordinary differential equations. Usual results then say that the equation has a unique solution in a small interval containing t_0.

It follows that if we give the initial condition $c(0) = x \in \mathrm{M}$, there is a unique solution defined on a maximal interval $\mathrm{I} \ni x$. We shall denote this solution $c_x(t)$.

Definition 4.4. The vector field X on M is a ***complete vector field*** if for any initial condition x, the solution c_x is defined on $\mathbb{R}$.

Lemma 4.5. *If a vector field* X *has compact support, then* X *is complete.*

Proof. The only way a solution can exist only on a bounded interval is that $c(t)$ gets out of any compact of M. But this is impossible since $\mathrm{X} = 0$ outside a large compact set K so the solutions starting from outside K are constant. □

Now change the perspective: we consider t as fixed and we vary the initial condition x, and define $\phi_t(x) = c_x(t)$. So ϕ_t consists in following the solution of $\dot{c} = \mathrm{X}(c)$ from the initial condition x during a time t. It is the ***flow*** of X at time t. The following result is then a direct consequence of the uniqueness of solutions of the equation:

Lemma 4.6. *Where it is defined, we have* $\phi_t \circ \phi_{t'} = \phi_{t+t'}$. *In particular* $\phi_t \circ \phi_{-t} = 1_{\mathrm{M}}$, *so* ϕ_t *is a local diffeomorphism where it is defined.*

In particular:

Corollary 4.7. *If* X *is complete on* M, *then* $(\phi_t)_{t \in \mathbb{R}}$ *is a 1-parameter group of diffeomorphisms of* M.

Example 4.8. 1° Check that the radial vector field $\mathrm{X} = x^i \frac{\partial}{\partial x^i}$ generates an homothety ϕ_t of ratio e^t.

2° Check that the vector field $\mathrm{X} = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}$ is a vector field on $\mathrm{S}^2 \subset \mathbb{R}^3$ which generates a rotation of angle t around the z axis.

4.c Geometric interpretation of the bracket

If $\phi : \mathrm{M} \longrightarrow \mathrm{N}$ is a diffeomorphism and X a vector field on N, then we can define the ***pull-back*** of X by ϕ, denoted by $\phi^*\mathrm{X}$, as the vector field on M given at each point $x \in \mathrm{M}$ by

$$(\phi^*\mathrm{X})_x = (d_x\phi)^{-1}\mathrm{X}_{\phi(x)}.$$

Lemma 4.9. *1° One has* $\phi^*[\mathrm{X},\mathrm{Y}] = [\phi^*\mathrm{X}, \phi^*\mathrm{Y}]$.

2° If ϕ *and* ψ *are diffeomorphisms, then* $(\psi\phi)^*\mathrm{X} = \phi^*\psi^*\mathrm{X}$.

3° If (ϕ_t) *is the flow of diffeomorphisms generated by the vector field* X *on* M, *and* Y *is another vector field, then*

$$\left.\frac{d}{dt}\right|_{t=0} \phi_t^*\mathrm{Y} = [\mathrm{X},\mathrm{Y}].$$

Any differential geometric object ξ on M can be pulled-back by the flow (ϕ_t) generated by a vector field X, and it is a general definition that the ***Lie derivative*** of ξ with respect to X is $\mathscr{L}_{\mathrm{X}}\xi = \frac{d}{dt}|_{t=0}\phi_t^*\xi$. Up to now we have seen two examples:

- for a function f, one has $(\phi_t^* f)(x) = f(\phi_t(x))$ then $(\mathscr{L}_X f)(x) = \frac{d}{dt}|_{t=0} f(\phi_t(x)) = d_x f(X(x))$ which coincides with definition (I.7);
- for a vector field Y, the above lemma means that $\mathscr{L}_X Y = [X, Y]$.

Proof. The proof of 1° and 2° is left to the reader. For 3° we use local coordinates (x^i), then we can write $X = (X^i)$, $Y = (Y^i)$ and $\phi_t(x) = (\phi_t^i(x))$. We denote the Jacobian matrix

$$J(\phi_t) = (\frac{\partial \phi_t^j}{\partial x^i})_{ij}$$

to write

$$(\phi_t^* Y)_x = J(\phi_t)_x^{-1} Y^i(\phi_t(x)) \frac{\partial}{\partial x^i}.$$

Differentiating at $t = 0$ and using $\phi_t' = X(\phi_t)$, so $\frac{d\phi_t^j}{dt}(x) = X^j(\phi_t(x))$:

$$\begin{aligned} \frac{d}{dt}\Big|_{t=0} (\phi_t^* Y)_x &= \big(d_x Y^i(X_x) - \frac{d}{dt}\Big|_{t=0} J(\phi_t)_x Y^i(x)\big) \frac{\partial}{\partial x^i} \\ &= \big(\frac{\partial Y^i}{\partial x^j}(x) X^j(x) - \frac{\partial X^j}{\partial x^i}(x) Y^i(x)\big) \frac{\partial}{\partial x^i} \end{aligned}$$

which is exactly $[X, Y]_x$. □

We are now ready for the geometric interpretation of the bracket:

Theorem 4.10. *The flows generated by two vector fields* X *and* Y *commute if and only if* $[X, Y] = 0$.

The typical example of two vector fields with $[X, Y] = 0$ is $X = \frac{\partial}{\partial x^i}$ and $Y = \frac{\partial}{\partial x^j}$. The generated flows are translating by t along the x^i variable (or the x^j variable), so they clearly commute. Somehow this is a very general example, see section 5 on Frobenius theorem.

Proof. Denote by (ϕ_t) and (ψ_u) the flows generated by X and Y. Then I claim that

1. $\frac{d}{du}(\phi_t^{-1} \psi_u \phi_t) = \phi_t^* Y$; in particular for $X = Y$ one obtains $\phi_t^* X = X$;
2. $\frac{d}{dt} \phi_t^* Y = \phi_t^* [X, Y]$.

The lemma follows immediately from the two formulas, whose proof is left to the reader. Indeed, if the flows commute then the first formula implies $Y = \phi_t^* Y$, and then from the second $[X, Y] = 0$. Conversely, if $[X, Y] = 0$, then from the second formula $\phi_t^* Y$ is constant, $\phi_t^* Y \equiv Y$ and then the first equation says that for any t the flow generated by Y is $(\phi_t \psi_u \phi_t^{-1})$. But this flow is (ψ_u) whence $\phi_t \psi_u \phi_t^{-1} = \psi_u$. □

5 Frobenius theorem

Definition 5.1. A p-dimensional ***distribution*** in a manifold M^n is the data at each point $x \in M$ of a p-dimensional subspace $D_p \subset T_x M$ depending smoothly on x.

The smooth dependence means that near each point x, one can find p smooth vector fields which generate the distribution at all nearby points.

For example, a non vanishing vector field defines a 1-dimensional distribution on a manifold (the direction generated by the vector). In this example, the distribution appears as the tangent bundle to the trajectories of the vector field, that is the solutions of $\dot{c} = X(c)$; one says that these trajectories are ***integral curves*** for the distribution. It is natural to ask for a higher dimensional analogue of this phenomenon : for instance, does a 2-dimensional distribution induce some surfaces ? In general, for a p-dimensional distribution, the convenient replacement for the curves c will be immersions from a p-dimensional manifold to M.

Definition 5.2. 1° An ***integral submanifold*** of D (or ***leaf*** of D) is an immersion $i : X^p \longrightarrow M$ such that at each $x \in X$ one has $T_{i(x)} i(M) = D_{i(x)}$.

2° A distribution D is ***integrable*** if every point of M belongs to an integral submanifold of D.

3° A distribution D is ***involutive*** if for any vector fields X and Y lying in D, one has $[X, Y] \in D$.

From the 3° in lemma 4.3, it follows that an integrable distribution must be involutive. The aim of this section is to prove Frobenius theorem, which states the converse.

An involutive distribution has always the following local form:

Lemma 5.3. *If* D *is a p-dimensional involutive distribution, then around any point x there are local coordinates (x^i) such that* D *is generated by the vector fields* $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^p}$.

This means that locally the leaves of the distribution are exactly the submanifolds $\mathbb{R}^p \times \{y\} \subset \mathbb{R}^p \times \mathbb{R}^{n-p}$.

Proof. The first step consists in producing vector fields $X_1, \dots, X_p \in D$ near x such that $[X_i, X_j] = 0$. Choose local coordinates $(x^1, \dots, x^n)$ near x. Up to composing by a (linear) diffeomorphism of $\mathbb{R}^n$, we can suppose that just at the point 0:

$$D_0 = \langle \frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^p} \rangle.$$

Then there is a unique basis of D consisting of vectors

$$X_i = \frac{\partial}{\partial x^i} + \sum_{j=p+1}^{n} f_i^j \frac{\partial}{\partial x^j}, \quad f_i^j(0) = 0.$$

Then we can calculate

$$[X_i, X_j] = \sum_{k=p+1}^{n} (\mathscr{L}_{X_i} f_j^k - \mathscr{L}_{X_j} f_i^k) \frac{\partial}{\partial x^k}.$$

But since D is involutive $[X_i, X_j] \in D$. From the form of the basis (X_i) of D we wee that this implies $[X_i, X_j] = 0$.

The second step in the proof then consists in considering the flows $\phi^1, \dots, \phi^p$ generated by the vector fields $X_1, \dots, X_p$. Let us choose a local submanifold $Y^{n-p} \subset X$ which is transverse to D at x (that is $T_xM = D_x \oplus T_xY$), for example $Y = \{0\} \times \mathbb{R}^{n-p}$ in the local coordinates above. Consider the map

$$f: \begin{array}{ccc} \mathbb{R}^p \times Y & \longrightarrow & M \\ (x^1, \dots, x^p, y) & \longmapsto & \phi^1_{x^1} \cdots \phi^p_{x^p}(y) \end{array}.$$

The differential at $(0, y)$ of this map is

$$(x^1, \dots, x^p, \xi) \longmapsto x^1 X_1 + \cdots + x^p X_p + \xi$$

which is an isomorphism $\mathbb{R}^p \times T_xY \longrightarrow T_xM$, so f is a local diffeomorphism. Since the X_i commute, the $\phi^i_{x^i}$ commute and it follows easily that

$$f^* X_i = \frac{\partial}{\partial x^i}.$$

The wished coordinates on M are therefore obtained by applying f^{-1} and taking coordinates on Y. ☐

Frobenius theorem is now an immediate corollary:

Theorem 5.4 (Frobenius). *A distribution on a manifold is integrable if and only if it is involutive.* ☐

Several problems can be expressed in terms of the integrability of a distribution, and are solved by Frobenius theorem. Here is an example: we explain how the problem of finding a function with given differential can be expressed in these terms. Of course the result is a well-known basic fact, but it will serve for us as a very simple illustration of the use of the theorem.

So suppose we have a 1-form α on a manifold M^n (that is, a section of the cotangent bundle T^*M), and we want to understand conditions on α in order to find a function f such that $df = \alpha$. We consider the manifold $X^{n+1} = M \times \mathbb{R}$, with the distribution

$$D_{(x,t)} = \{(\xi, \alpha_x(\xi)), \xi \in T_xM\}.$$

What is an integral submanifold of D ? a submanifold $Y^n \subset M^n \times \mathbb{R}$ tangent to D at each point; as D never intersects the $\mathbb{R}$ part of $T_xX = T_xM \oplus \mathbb{R}$, such Y is locally the graph of a function $f : M \longrightarrow \mathbb{R}$. Then $T_{(x,f(x))}Y = \{(\xi, d_xf(\xi)), \xi \in T_xM\}$ so Y is a leaf of D if and only if $df = \alpha$.

So we see that the problem of finding locally f such that $df = \alpha$ is equivalent to finding an integral submanifold of D. Now by Frobenius theorem, this is possible if and only if D is involutive. Let us write down the condition in local coordinates (x^i) on M: then $\alpha = \alpha_i dx^i$, the distribution D is generated by the vector fields $X_i = \frac{\partial}{\partial x^i} + \alpha_i \frac{\partial}{\partial t}$, and

$$[X_i, X_j] = \Big(\frac{\partial \alpha_j}{\partial x^i} - \frac{\partial \alpha_i}{\partial x^j}\Big) \frac{\partial}{\partial t}.$$

This belongs to D only if it vanishes, and we recover in this way the fact that α is locally a df if and only if its first derivatives are symmetric.

Exercise. Let x, y, z denote the standard coordinates on $\mathbb{R}^3$. We consider the distribution $D = \ker(dz - ydx) = \langle \frac{\partial}{\partial y}, \frac{\partial}{\partial x} + y\frac{\partial}{\partial z} \rangle$. 1° Check that D is not integrable. 2° Compute the flows ϕ_t (resp. ψ_u) of $\frac{\partial}{\partial y}$ (resp. $\frac{\partial}{\partial x} + y\frac{\partial}{\partial z}$) and then the commutator $\psi_{-u}\phi_{-t}\psi_u\phi_t$. 3° Deduce that any two points in $\mathbb{R}^3$ can be connected by a (piecewise smooth) path that remains tangent to D. Compare this phenomenon with what happens on a foliation (namely, an integrable distribution).

6 Differential forms

Here we give only a brief summary, since we will not use much differential forms in this course (except 1 and 2-forms). The details should be studied in a book.

6.a Linear algebra

If E is a n dimensional vector space, then we define $\Lambda^k E^*$ as the space of alternate k-linear forms on E. if $\alpha \in \Lambda^k E^*$, the integer k is the degree of the form α, and is often denoted by $|\alpha|$. Sometimes one considers all k-forms together: $\Lambda^\bullet E^* = \oplus \Lambda^k E^*$, where $\Lambda^0 E^* = \mathbb{R}$. Also observe that $\Lambda^1 E^* = E^*$.

Concretely, if $(e_i)_{i=1,\dots,n}$ is a basis of E, and (e^i) denotes the dual basis of E^*, then a basis of $\Lambda^k E^*$ consists of $(e^{i_1} \wedge \cdots \wedge e^{i_k})_{i_1 < \cdots < i_k}$, where the exterior product $\alpha_1 \wedge \cdots \wedge \alpha_k$ of k one forms is defined by

$$\alpha_1 \wedge \cdots \wedge \alpha_k(x_1, \dots, x_k) = \sum_{\sigma \in \mathfrak{S}_k} \varepsilon(\sigma)\alpha_1(x_{\sigma(1)}) \cdots \alpha_k(x_{\sigma(k)}).$$

In particular, the dimension of $\Lambda^k E^*$ is $\binom{n}{k}$, and $\Lambda^k E^* = 0$ for $k > n$.

The exterior product extends to all forms to define an associative product sending $\Lambda^k \otimes \Lambda^l \longrightarrow \Lambda^{k+l}$, and satisfying the commutation

$$\beta \wedge \alpha = (-1)^{|\alpha||\beta|}\alpha \wedge \beta.$$

Finally, if $u : E \longrightarrow F$ is a linear map, then it induces $\Lambda^k u^t : \Lambda^k F^* \longrightarrow \Lambda^k E^*$ by

$$\big((\Lambda^k u^t)\omega\big)(x_1, \dots, x_k) = \omega(u(x_1), \dots, u(x_p)).$$

It is easy to check that $\Lambda^k u^t$ preserves the algebra structure:

$$\Lambda^k u^t(\alpha \wedge \beta) = \Lambda^k u^t(\alpha) \wedge \Lambda^k u^t(\beta).$$

6.b Differential forms on a manifold

In local coordinates, we have a basis (dx^i) of 1-forms, and a k-***differential form*** ω is a linear combination

$$\sum_{i_1 < \cdots < i_k} \omega_{i_1 \cdots i_k}(x) dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

More intrinsically, there is a vector bundle $\Lambda^k T^*M$ over M, whose fiber $\Lambda^k T_x^*M$ at a point x is the space of k-forms on the tangent space T_xM (see section 3.d for 1-forms).

Then a k-differential form is a section of this vector bundle, and the space of k-forms on M is denoted $\Omega^k M := \Gamma(M, \Lambda^k T^*M)$. In particular 0-forms are just functions on M, $\Omega^0 M = C^\infty(M)$. We also consider all forms together, $\Omega^\bullet M = \oplus_k \Omega^k M$.

Exercise. Check that the form $4\frac{dx\wedge dy}{(1+x^2+y^2)^2}$ defined on $S^2 \setminus \{N\}$ in the coordinates (x, y) given by stereographic projection extends to a global 2-form on S^2. (As we will see later, this is the volume form of the sphere, and its integral gives the volume of the sphere, that is 4π).

If $f : M \longrightarrow N$ is a smooth map, and α is a k-form on N, then one can define the ***pull-back*** of α by f on M, defined at the point $x \in M$ by

$$(f^*\alpha)_x = \Lambda^k (d_x f)^t \alpha.$$

The pull-back satisfies $f^*(\alpha \wedge \beta) = f^*\alpha \wedge f^*\beta$.
Finally, a k-form ω on M defines an alternate $C^\infty(M)$-linear form on the space $\Gamma(TM)$ of vector fields on M, by

$$(X_1, \dots, X_p) \longrightarrow \omega(X_1, \dots, X_p).$$

Conversely:

Lemma 6.1. *Any* $C^\infty(M)$*-linear alternate* k*-form* α *on* $\Gamma(TM)$ *is induced by some smooth* k*-differential form.*

One says that the form α is ***tensorial***, that is it comes from a section of a tensor bundle (a bundle of the type $\otimes^a TM \otimes \otimes^b T^*M$).

Proof. One first prove that such a $C^\infty(M)$-linear form L is local, as in the proof of theorem 4.1. Then one is reduced to consider only local vector fields, and one can use local coordinates (x^i): if $X_j = X_j^i \frac{\partial}{\partial x^i}$, then by $C^\infty(M)$-linearity

$$L(X_1, \dots, X_k) = \sum_{(i_1,\dots,i_k)} X_1^{i_1} \cdots X_k^{i_k} L(\frac{\partial}{\partial x^{i_1}}, \dots, \frac{\partial}{\partial x^{i_k}})$$

which is induced by the k-differential form

$$\omega = \sum_{i_1<\dots<i_k} L(\frac{\partial}{\partial x^{i_1}}, \dots, \frac{\partial}{\partial x^{i_k}}) dx^{i_1} \wedge \dots \wedge dx^{i_k}.$$

□

6.c Exterior differential

An ***odd derivation*** of the exterior algebra $\Omega^\bullet M$ is a map $D : \Omega^\bullet M \longrightarrow \Omega^\bullet M$ satisfying the modified Leibniz identity:

$$D(\alpha \wedge \beta) = (D\alpha) \wedge \beta + (-1)^{|\alpha|} \alpha \wedge (D\beta).$$

It has degree d if it sends $\Omega^k(M)$ to $\Omega^{k+d}(M)$.

Lemma and Definition 6.2. *The differential of functions, $d : C^\infty(M) \longrightarrow \Omega^1 M$ extends uniquely into an odd derivation of $\Omega^\bullet M$ which satisfies $d(df) = 0$ for any function f. This extension is called the* **exterior derivative**, *and it satisfies:*

1. *$d \circ d = 0$;*
2. *for any smooth map $f : M \longrightarrow N$ and differential form ω on N one has $f^* d\omega = d(f^*\omega)$.*

We let the precise proof to the reader, but it can be deduced from the explicit local formulas that we shall now derive, by following the ideas in the proof of theorem 4.1. Begin by a 1-form $\omega = \omega_i dx^i$, then applying the Leibniz formula and $d(dx^i) = 0$, we obtain

$$d\omega = d\omega_i \wedge dx^i = \frac{\partial \omega_i}{\partial x^j} dx^j \wedge dx^i = \sum_{i<j} \Big(\frac{\partial \omega_j}{\partial x^i} - \frac{\partial \omega_i}{\partial x^j} \Big) dx^i \wedge dx^j.$$

Similarly, for a k-form $\omega = \omega_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}$, we obtain

$$d\omega = d\omega_{i_1\cdots i_k} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} = \frac{\partial \omega_{i_1 \cdots i_k}}{\partial x^j} dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

It is important to be able to calculate $d\omega$ from the point of view of linear forms on vector fields:

Lemma 6.3 (Maurer-Cartan formula). *For a differential k-form, one has the formula*

$$d\alpha(X_0, \dots, X_k) = \sum_{i=0}^{k} (-1)^i \mathscr{L}_{X_i}\big(\alpha(X_0, \dots, \widehat{X_i}, \dots, X_k)\big) + \sum_{0 \leqslant i<j \leqslant k} (-1)^{i+j} \alpha([X_i, X_j], X_0, \dots, \widehat{X_i}, \dots, \widehat{X_j}, \dots, X_k).$$

In particular, for a 1-form one has

$$d\alpha(X, Y) = \mathscr{L}_X\big(\alpha(Y)\big) - \mathscr{L}_Y\big(\alpha(X)\big) - \alpha([X, Y]). \tag{I.13}$$

Proof. One checks that the RHS of the formula is $C^\infty(M)$-linear in $X_0, X_1, \dots, X_k$, and is alternate, so it actually defines a $(k+1)$-differential form. To determine it, it suffices to take the X_j among a local basis of vector fields $(\frac{\partial}{\partial x^i})$. Then the calculation becomes very simple because all the brackets vanish. □

Exercise. Prove the following reformulation of the involutivity condition in Frobenius theorem. Suppose a rank k distribution D on M^n is given locally as $D = \ker \alpha_1 \cap \cdots \cap \ker \alpha_{n-k}$, where the α_i are 1-forms which are linearly independant at each point. Let $I \subset \Omega^\bullet M$ be the ideal generated by the α_i's. Then D is involutive if and only if I is stable by the exterior derivative, $d(I) \subset I$. In particular, deduce that if $D = \ker \alpha$ is a rank 2 distribution on M^3, then D is integrable if and only if $\alpha \wedge d\alpha = 0$.

A vector field X on a manifold M generates a flow (ϕ_t) of diffeomorphisms of M. If ω is a k-differential form on M, then the ***Lie derivative*** of ω with respect to X is

$$\mathscr{L}_X \omega = \frac{d}{dt} \phi_t^* \omega \Big|_{t=0}, \tag{I.14}$$

according to the general principle explained in section 4.c. From the definitions, one obtains immediately the following properties:

- $\mathscr{L}_X(\alpha \wedge \beta) = (\mathscr{L}_X\alpha) \wedge \beta + \alpha \wedge (\mathscr{L}_X\beta)$;
- $\mathscr{L}_X d\alpha = d\mathscr{L}_X\alpha$.

Define the ***interior product*** by X, denoted ι_X, as the map $\iota_X : \Omega^{k+1}M \longrightarrow \Omega^k M$, defined by $(\iota_X\alpha)(X_1, \dots, X_k) = \alpha(X, X_1, \dots, X_k)$. It is an odd derivation of the algebra $\Omega^\bullet M$.

Lemma 6.4 (Cartan's formula). *One has* $\mathscr{L}_X = d\circ\iota_X + \iota_X\circ d$.

Proof. One checks that $d\circ\iota_X + \iota_X\circ d$ and $\mathscr{L}_X$ are derivations of the algebra $\Omega^\bullet M$, therefore they coincide if they coincide on functions and 1-forms. On functions they are both equal to $f \longrightarrow df(X)$. If α is a 1-form and Y is a vector field then one calculates $\mathscr{L}_X\alpha$ via

$$\mathscr{L}_X(\alpha(Y)) = (\mathscr{L}_X\alpha)(Y) + \alpha(\mathscr{L}_X Y) = (\mathscr{L}_X\alpha)(Y) + \alpha([X, Y]).$$

Then from formula (I.13) one obtains $\mathscr{L}_X\alpha = d(\iota_X\alpha) + \iota_X d\alpha$. □

6.d De Rham cohomology

We just mention the definition and a few facts without proof about this important invariant: the k-th group of De Rham cohomology is defined by

$$H^k(M) = \{\alpha \in \Omega^k M, d\alpha = 0\} / \{d\beta, \beta \in \Omega^{k-1}M\}.$$

One can also define a compactly supported version $H^k_c(M)$, by requiring that α and β have compact support.

Of course, it is clear that $H^0(M)$ consists of locally constant functions on M, so

$$H^0(M) = \mathbb{R}^{\#\text{connected components of } M}.$$

Locally, if $d\alpha = 0$ then there exists β such that $d\beta = \alpha$, so the cohomology does not depend on local properties of M. It turns out that $H^k(M)$ is a topological invariant of M (it depends on the class of M modulo homeomorphisms, and even modulo homotopy equivalences). If M is compact, then $H^k(M)$ is finite dimensional and its dimension $b_k(M) = \dim H^k(M)$ is called the k-th Betti number of M.

For example, for the sphere S^n, the cohomology vanishes in every degree, except in degrees 0 and n, and $H^0(S^n) = H^n(S^n) = \mathbb{R}$. For the complex projective space $\mathbb{C}P^n$, the cohomology vanishes in odd degrees, and in even degree $2k$ for $k = 0, \dots, n$ one has $H^{2k}(\mathbb{C}P^n) = \mathbb{R}$.

6.e Orientation

Remark that $\Lambda^n(\mathbb{R}^n)^* = \mathbb{R}$: every alternate n-form is proportional to $dx^1 \wedge \cdots \wedge dx^n$.

On a manifold, $dx^1 \wedge \cdots \wedge dx^n$ is well defined in local coordinates, but of course does not extend in general to the whole manifold. If we change coordinates, $(x^i) = \phi(y^j)$, then

$$\begin{aligned} dx^1 \wedge \cdots \wedge dx^n &= \Big(\frac{\partial x^1}{\partial y^i}\Big) \wedge \cdots \wedge \Big(\frac{\partial x^n}{\partial y^i} dy^i\Big) \\ &= \det\Big(\frac{\partial x^i}{\partial y^j}\Big) dy^1 \wedge \cdots \wedge dy^n \\ &= J(\phi) dy^1 \wedge \cdots \wedge dy^n, \end{aligned}$$

where $J(\phi)$ is the determinant of the Jacobian matrix of ϕ. More generally, if $\omega = f dx^1 \wedge \cdots \wedge dx^n$ then

$$\phi^*\omega = (\phi^* f) J(\phi) dy^1 \wedge \cdots \wedge dy^n. \tag{I.15}$$

Definition 6.5. A manifold M^n is ***orientable*** if there exists an atlas such that all the transitions ϕ have $J(\phi) > 0$. An ***orientation*** is the choice of a maximal such atlas.

Lemma 6.6. *Suppose M^n is connected. Then M^n is orientable if and only if $\Lambda^n T^*M \setminus \{\text{zero section}\}$ has two connected components. An orientation of M is the same as the choice of one component.*

Proof. Elements of $\Lambda^n T^*M$ are locally represented by $f dx^1 \wedge \cdots \wedge dx^n$. Locally we have the two components $f > 0$ and $f < 0$ of $\Lambda^n T^*M \setminus \{0\}$. From equation (I.15) it is clear that these two components do not depend on the chart if we take an atlas with positive $J(\phi)$. So we have proved that if M is orientable then $\Lambda^n T^*M \setminus \{0\}$ has two components, and the orientation of M selects one component. The converse is left to the reader. ☐

The component of $\Lambda^n T^*M \setminus \{0\}$ selected by the orientation is called positive.

Definition 6.7. If M^n is oriented, a ***volume form*** on M is a differential n-form which is positive at every point. The manifold always carries such volume form.

To construct the volume form, one defines $\omega_i = dx^1 \wedge \cdots \wedge dx^n$ in local coordinates, for a covering of M by coordinate charts (U_i). To pass from these local forms to a global form, we use an important tool (see [Lee13]), a ***partition of unity*** subordinate to the (U_i), that is a collection of functions $\chi_i : M \longrightarrow \mathbb{R}$ such that:

1. $\chi_i \geqslant 0$,
2. the support of χ_i is included in U_i,
3. the supports of the χ_i are locally finite (each point has a neighbourhood intersecting only a finite number of supports),
4. $\sum \chi_i = 1$ (by the previous item, each point has a neighbourhood where this sum is finite).

We can now define the desired volume form by

$$\omega = \sum \chi_i \omega_i.$$

At each point all the terms are nonnegative, and at least one is positive since $\sum \chi_i = 1$, so ω is a positive form.

Conversely, note that the existence of a non vanishing n-form proves immediately that M is orientable.

Example 6.8. 1° The sphere S^n is oriented, with volume form $\iota_{\vec{n}}(dx^1 \wedge \cdots \wedge dx^{n+1})$, where $\vec{n} = x^i \frac{\partial}{\partial x^i}$ is the outward normal vector to S^n. This means that at each point, a direct basis of S^n is given by $(e_1, \dots, e_n)$ so that $(\vec{n}, e_1, \dots, e_n)$ is a direct basis of $\mathbb{R}^{n+1}$.

2° The projective space is orientable if n is odd. Indeed consider the map $\pi : S^n \longrightarrow \mathbb{R}P^n$. This is a 2:1 local diffeomorphism, given by quotient by the antipodal map a. If ω is a volume form on $\mathbb{R}P^n$, then $\pi^*\omega$ is a nowhere vanishing n-form on S^n (since π is a local diffeomorphism), satisfying $a^*\pi^*\omega = \pi^*\omega$ since $\pi \circ a = \pi$. This implies that a preserves the orientation of S^n. Now remark that $a^*\vec{n} = \vec{n}$, so

$$a^*(\iota_{\vec{n}}(dx^1 \wedge \cdots \wedge dx^{n+1})) = (-1)^{n+1}\iota_{\vec{n}}(dx^1 \wedge \cdots \wedge dx^{n+1}),$$

so a preserves the orientation of S^n if and only n is odd. So if $\mathbb{R}P^n$ is orientable then n is odd. Conversely if n is odd, then the standard volume form of S^n is invariant under a, so descends to a well-defined volume form on $\mathbb{R}P^n$.

3° (Exercise) Let $X^{2n} = T^*M^n$ denote the cotangent bundle of a manifold M^n. We denote by $p : X \longrightarrow M$ the natural projection. For any point $a = (x, \alpha) \in X$ (namely $x \in M$ and $\alpha \in T^*_x M$) and $\xi \in T_a X$, we set $\lambda_a(\xi) = \alpha(d_a p(\xi))$. Check that this defines a (canonical !) one-form λ on X, known as the Liouville form. The choice of some local coordinates x^i on M gives rise to local coordinates (x^i, p_i) on X, where $\alpha = p_i dx^i$. What is λ in these coordinates ? Prove that the closed two-form $\omega = d\lambda$ is non-degenerate, that is $\xi \longmapsto \iota_\xi \omega$ is a linear isomorphism between $T_a X$ and $T^*_a X$ for every $a \in X$ (such 2-form is called a ***symplectic form***). As a consequence, prove that $\omega^n = \omega \wedge \cdots \wedge \omega$ (n times) is a canonical volume form on the cotangent bundle.

6.f Integration of forms

Suppose that M^n is an oriented manifold. We are now going to define $\int_M \omega$ for any compactly supported n-form ω on M. First suppose that ω has his support contained in coordinate chart. Then $\omega = f dx^1 \wedge \cdots \wedge dx^n$ where f has compact support, and we can define

$$\int_M \omega = \int f(x) dx^1 \cdots dx^n.$$

Suppose we have other coordinates (y^j) such that $(x^i) = \phi(y^j)$, then one has the well-known formula for the change of variables:

$$\int f(x) dx^1 \cdots dx^n = \int f(y) |J(\phi)| dy^1 \cdots dy^n.$$

In view of formula (I.15), if we have $J(\phi) > 0$ (which is the case we have chosen coordinates compatible with the orientation), then our definition of $\int_M \omega$ does not depend on the choice of coordinates.

The definition of $\int_M \omega$ is then extended to any ω by a partition of unity (χ_i) subordinate to a covering of M by coordinate charts: $\omega = \sum \chi_i \omega_i$ and $\int_M \omega = \sum \int_M (\chi_i \omega_i)$.

Theorem 6.9 (Stokes). *If M^n is an oriented manifold with boundary, and ω is a compactly supported n-form, and $i : \partial M \hookrightarrow M$ the inclusion of its boundary. Then*

$$\int_M d\omega = \int_{\partial M} i^* \omega.$$

We have not encountered before the notion of ***manifold with boundary***, so a few words are needed here. The model example is $\mathbb{R}_- \times \mathbb{R}^{n-1}$, with boundary $\{0\} \times \mathbb{R}^{n-1}$. The general definition is modeled on this example: a manifold with boundary M^n has the same definition as a ***closed manifold*** (closed means without boundary), except that some coordinate charts have values in $\mathbb{R}_- \times \mathbb{R}^{n-1}$, and the corresponding transition functions preserve globally $\{0\} \times \mathbb{R}^{n-1}$. Then the set $x^1 = 0$ in the corresponding charts define a submanifold ∂M of M which is the ***boundary*** of M. A basic example is the ball of $\mathbb{R}^n$, whose boundary is the unit sphere S^{n-1}.

If M is oriented, then ∂M inherits an orientation. In the model $\mathbb{R}_- \times \mathbb{R}^{n-1}$, we decide that $\{0\} \times \mathbb{R}^{n-1}$ is oriented by the basis $(e_2, \dots, e_n)$, if $(e_1, \dots, e_n)$ is an oriented basis of $\mathbb{R}^n$. More intrinsically maybe, at a point $x \in \partial M$, we decide that a basis $(e_2, \dots, e_n)$ of $T_x \partial M$ is direct if $(\vec{n}, e_2, \dots, e_n)$ is a direct basis of $T_x M$, where $\vec{n}$ is a vector pointing outward M. See example 6.8.

Now the statement of the theorem is well-defined. This is the hardest part, since the proof is very simple:

Proof. Using a partition of unity, it is sufficient to check the case where the support of ω is contained in a coordinate chart (x^i), where $M = \{x^1 \leqslant 0\}$. Then $\omega = \omega_i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n$, so

$$d\omega = (-1)^{i-1} \frac{\partial \omega_i}{\partial x^i} dx^1 \wedge \cdots \wedge dx^n,$$

and

$$\int_M d\omega = \int_{x^1 \leqslant 0} (-1)^{i-1} \frac{\partial \omega_i}{\partial x^i} dx^1 \cdots dx^n.$$

Because ω has compact support, by integration by parts all the terms for $i > 1$ give 0, and for $i = 1$ we get

$$\int_M d\omega = \int_{x^1 = 0} \omega_1 dx^2 \cdots dx^n = \int_{\partial M} i^* \omega$$

since $i^*\omega = \omega_1(0, x^2, \dots, x^n) dx^2 \wedge \cdots \wedge dx^n$. □

The Stokes theorem contains all special cases of integration by parts known in small dimension, as the Green-Riemann formula in dimension 2 or the Ostrogradsky formula in dimension 3.

If M^n has no boundary, an important consequence of the theorem is that $\int_M d\omega = 0$ for any compactly supported $(n-1)$-form, so $\omega \mapsto \int_M \omega$ is actually well defined on $H^n_c(M)$. This completely determines $H^n_c(M)$, as stated in the following theorem, that we will not prove.

Theorem 6.10. *If* M^n *is a connected oriented closed manifold then the map* $\omega \mapsto \int_M \omega$ *induces an isomorphism* $H^n_c(M) \approx \mathbb{R}$.

If M is compact then $H^n_c(M) = H^n(M)$ so the theorem gives the calculation of $H^n(M)$.

Chapter II

Riemannian metric, connection, geodesics

7 Riemannian metrics

7.a Definition and examples

Definition 7.1. Let M be a manifold. A ***Riemannian metric*** on M is the data for each point $x \in M$ of a positive definite quadratic form g_x on T_xM, depending smoothly on the point x.

In another words, a Riemannian metric is a measure of the length of tangent vectors. In local coordinates (x^i), it is given by a positive definite matrix $(g_{ij}(x)) = \big(g(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j})\big)$, with the $g_{ij}(x)$ being smooth functions, and one writes simply

$$g = g_{ij} dx^i dx^j.$$

If we have other coordinates (y^j), then it is easy to see that

$$g = g_{ij} \frac{\partial x^i}{\partial y^k} \frac{\partial x^j}{\partial y^l} dy^k dy^l.$$

Example 7.2. 1° The flat metric $g = (dx^1)^2 + \cdots + (dx^n)^2$ on $\mathbb{R}^n$. At each point the tangent space identifies to $\mathbb{R}^n$ and the metric is the standard metric of $\mathbb{R}^n$.

2° The metric $g = (dx^1)^2 - (dx^2)^2 - \cdots - (dx^n)^2$ on $\mathbb{R}^{1,n-1}$. This is the same as the previous example, except that it is not a Riemannian metric since it is not positive definite. This metric is Lorentzian (only one positive direction). In general indefinite (but non degenerate) metrics are called ***pseudo-Riemannian metrics***.

3° The flat metric of $\mathbb{R}^2$ in polar coordinates (r, θ) writes $g = dr^2 + r^2 d\theta^2$. This is clear because the basis $(\frac{\partial}{\partial r}, \frac{1}{r}\frac{\partial}{d\theta})$ is orthonormal.

More generally, the flat metric of $\mathbb{R}^n \setminus \{0\} =]0, +\infty[\times S^{n-1}$ can be written

$$g = dr^2 + r^2 g_{S^{n-1}}.$$

4° The sphere $S^n \subset \mathbb{R}^{n+1}$ inherits a metric g from $\mathbb{R}^{n+1}$. In the coordinates given by the stereographic projection (section 2.b) one calculates

$$g = 4\frac{\sum (dx^i)^2}{(1+r^2)^2}. \tag{II.1}$$

The proof is by writing from the coordinates (x^i) of the stereographic projection the coordinates of the corresponding point in $\mathbb{R}^{n+1}$: this is the point

$$\frac{1}{1+r^2}(r^2-1, 2x^1, \dots, 2x^n),$$

and therefore

$$g = d\Big(\frac{r^2-1}{1+r^2}\Big)^2 + d\Big(\frac{2x^1}{1+r^2}\Big)^2 + \cdots + d\Big(\frac{2x^n}{1+r^2}\Big)^2.$$

Developing this expression simplifies to the formula above.

5° Any submanifold of a Riemannian manifold inherits a Riemannian metric by restricting the metric of the manifold to the tangent bundle of the submanifold.

6° For the hyperbolic space $H^n \subset \mathbb{R}^{1,n}$ (see (I.1)), there is also a stereographic projection $H^n \longrightarrow \mathbb{R}^n$ from the point $(-1, 0, \dots, 0)$. This projection is a global diffeomorphism on the unit ball $\{r < 1\}$, and one obtains (exercise)

$$g = 4\frac{\sum (dx^i)^2}{(1-r^2)^2}. \tag{II.2}$$

7° A torus $T = \mathbb{R}^n/\Lambda$, where Λ is a lattice of $\mathbb{R}^n$. The projection map $\mathbb{R}^n \longrightarrow T$ is a local diffeomorphism, and the action of Λ is by translations (which preserve the metric), so the metric of $\mathbb{R}^n$ induces a metric on T.

8° A surface of revolution in $\mathbb{R}^3$, say around the z axis. We take polar coordinates (r, θ) in the xy plane. The surface is given by an equation of the type $r = f(z)$, but it is more convenient to parameterize it in a different way: the intersection with the xz plane is a curve, which we parameterize by the length u. Then the metric of the surface is

$$g = du^2 + r(u)^2 d\theta^2.$$

7.b Volume form

Suppose (M^n, g) is an oriented Riemannian manifold. Then at each point x there is a privileged n-form, namely a positive form of norm 1 (if (e_i) is an orthonormal basis of E, the $(e^{i_1} \wedge \cdots \wedge e^{i_k})$ give the orthonormal basis of a canonical metric on $\Lambda^\bullet E^*$). This form is called the ***volume form*** of g, and is given in local coordinates by the formula

$$\mathrm{vol}^g = \sqrt{\det(g_{ij})}\, dx^1 \wedge \cdots \wedge dx^n.$$

The volume of M (which can be infinite if M is non compact) is then

$$V = \int_M \mathrm{vol}^g.$$

This is the most basic Riemannian invariant. For example, one can calculate

$$V(S^{2n}) = (4\pi)^n \frac{(n-1)!}{(2n-1)!}, \quad V(S^{2n+1}) = 2\frac{\pi^{n+1}}{n!}.$$

7.c Isometries

Definition 7.3. A diffeomorphism $\phi : (M, g) \longrightarrow (N, h)$ is an ***isometry*** if $\phi^* h = g$.

The definition means $h_{\phi(x)}(d_x\phi(X), d_x\phi(Y)) = g_x(X, Y)$ for all $X, Y \in T_xM$, or equivalently $d_x\phi$ is a linear isometry between T_xM and $T_{\phi(x)}N$.

Theorem 7.4. *The group of isometries of a Riemannian manifold is a Lie group.*

We do not prove this theorem, see [Kob95].

Example 7.5. 1° The antipodal map $x \longrightarrow -x$ on S^n is an isometry. As a consequence, since $\mathbb{R}P^n$ is the quotient of S^n by this isometry, the metric of S^n induces a metric on $\mathbb{R}P^n$.

2° The isometries of $\mathbb{R}^n$ consists of orthogonal transformations and translations: $\mathrm{Isom}(\mathbb{R}^n) = \mathbb{R}^n \ltimes O(n)$.

3° $\mathrm{Isom}(S^n) = O(n+1)$, and $\mathrm{Isom}(H^n) = O_o(1, n)$, where the index means that we take the subgroup preserving the nap $\{x^0 > 0\}$ of $\{(x^0)^2 - (x^1)^2 - \cdots (x^n)^2 = 1\}$. If we write SO instead of O in these examples, we obtain the orientation-preserving isometries. These two spaces are homogeneous spaces, that is the isometry group acts transitively. Therefore they are quotient of the isometry group by the isotropy group of a point:

$$S^n = O(n+1)/_{O(n)}, \quad H^n = O_0(1, n)/_{O(n)}.$$

For $M = \mathbb{R}^n, S^n$ or H^n, we have written a group which is clearly a group of isometries, but we have not proved that there is no other isometry. Nevertheless it is easy to see that these groups have a stronger property than being just homogeneous: actually, for any points x et y and any isometry $u : T_xM \longrightarrow T_yM$, there exists an element ϕ of the group such that $\phi(x) = y$ and $T_x\phi = u$. (This is because the stabilizer of a point is each time $O(n)$). We will see later in corollary 10.4 that for a complete connected Riemannian manifold, there is at most one isometry with given (x, y, ϕ), so this proves that there is no possible other isometry.

8 Connections

8.a Connections and Christoffel symbols

Here we address the following problem: find a way to take derivatives of sections of bundles. Indeed, if we consider the section of a bundle in a local trivialization, we can calculate a derivative, but taking another trivialization will result in another derivative. What we need is a *covariant derivative.*

More precisely, what do we need ? suppose E is a vector bundle over M, and s is a section of E. Choose a tangent vector $X \in T_xM$, we wish to define a derivative of s along X at x, denoted $\nabla_X s$. This should depend only on the value of X at x and be linear in $X \in T_xM$, so at the point x the object $(\nabla s)_x$ should belong to

$$\mathrm{Hom}(T_xM, E_x) = T_x^*M \otimes E_x.$$

If we now take a vector field $X \in \Gamma(M, TM)$, then this means that the covariant derivative ∇s should be a section of the bundle $T^*M \otimes E$, that is the bundle of 1-forms with values in E. We denote by $\Omega^1(M, E)$ the space of sections of $T^*M \otimes E$, and more generally $\Omega^k(M, E) = \Gamma(M, \Lambda^k T^*M \otimes E)$.

Definition 8.1. A ***connection***, or ***covariant derivative***, on a real (resp. complex) vector bundle E over M is a $\mathbb{R}$ (resp. $\mathbb{C}$)-linear operator

$$\nabla : \Gamma(M, E) \longrightarrow \Omega^1(M, E),$$

satisfying the following Leibniz rule: if $f \in C^\infty(M)$ and $s \in \Gamma(M, E)$, then

$$\nabla(fs) = df \otimes s + f\nabla s.$$

As we have already seen in other contexts, the Leibniz rule implies immediately that ∇ is a local operator: if U is an open set, $(\nabla s)|_U$ depends only on $s|_U$. Another way to say the same thing is to say that ∇ induces as well an operator $\Gamma(U, E) \longrightarrow \Omega^1(U, E)$. Therefore we can take U to be a coordinate chart on which we have a trivialization of E and write down explicit formulas. Suppose (x^i) are local coordinates and $(e_1, \dots, e_r)$ is a local basis of sections of E. Define the ***Christoffel symbols*** Γ^b_{ia} by

$$\nabla e_a = \Gamma^b_{ia} dx^i \otimes e_b.$$

A general section of E writes $s = s^a e_a$ and applying Leibniz rule:

$$\begin{aligned} \nabla s &= ds^a \otimes e_a + s^a \nabla e_a \\ &= \Big(\frac{\partial s^a}{\partial x^i} + \Gamma^a_{ib} s^b \Big) dx^i \otimes e_a \end{aligned}$$

or, equivalently,

$$\nabla_{\frac{\partial}{\partial x^i}} s = \Big(\frac{\partial s^a}{\partial x^i} + \Gamma^a_{ib} s^b \Big) e_a.$$

Therefore we shall write (in this trivialization)

$$\nabla = d + \Gamma_i dx^i,$$

where $\Gamma_i = (\Gamma^b_{ia})$ is a matrix (an endomorphism of E). This tells us that the connection ∇ is locally given by a 1-form with values in End E. In a more synthetic way, considering s as a column vector, the formula above means

$$\nabla s = ds + dx^i \otimes \Gamma_i s.$$

The 1-form $\Gamma = dx^i \otimes \Gamma_i$ (with values in End E) is called the ***connection 1-form***.

Let us see what is happening by a change of trivialization. If we have a new basis (f_b) of E, such that $e_a = u^b_a f_b$, then a section $s = s^b f_b$ has coordinates $u^{-1}s$ in the basis (e_a) and therefore in this basis ∇s writes $d(u^{-1}s) + dx^i \Gamma_i u^{-1} s$. Coming back to the basis (f_b), we obtain

$$\nabla s = u\Big(d(u^{-1}s) + dx^i \Gamma_i u^{-1} s\Big) = ds + \Big(-du u^{-1} + dx^i u \Gamma_i u^{-1} \Big) s.$$

In particular, we see that the matrices (Γ'_i) in the basis (f_b) can be expressed as

$$\Gamma'_i = -\frac{\partial u}{\partial x^i} u^{-1} + u\Gamma_i u^{-1}. \tag{II.3}$$

This formula is important, it shows that the Γ's are not tensorial objects (they do not give a section of $\Omega^1 \otimes \mathrm{End}\,E$), since the law when we change the basis involves derivatives of the transition u. Still, we see from the formula that the difference between two connections is tensorial: if we have two connections ${}^1\nabla$ and ${}^2\nabla$, then by a change of trivialization the equality (II.3) gives

$${}^1\Gamma'_i - {}^2\Gamma'_i = u({}^1\Gamma_i - {}^2\Gamma_i)u^{-1},$$

which means now that ${}^1\nabla - {}^2\nabla$ is tensorial: ${}^1\nabla - {}^2\nabla \in \Omega^1(M, \mathrm{End}\,E)$.

This can also be seen directly: using the Leibniz formula for both connections, we obtain immediately that

$$({}^1\nabla - {}^2\nabla)(fs) = f({}^1\nabla - {}^2\nabla)s,$$

that is the difference is $C^\infty(M)$-linear, implying that it is tensorial. Conversely, if ∇ is a connection and $a \in \Omega^1(M, \mathrm{End}\,E)$, it is easy to check that $\nabla + a$ is again a connection, so we have proved:

Lemma 8.2. *The space of connections on a given bundle* E *is an affine space with direction* $\Omega^1(M, \mathrm{End}\,E)$. □

8.b Examples of connections

1° The tangent bundle $T\mathbb{R}^n$ of $\mathbb{R}^n$ with the trivial connection $\nabla = d$. This means $\nabla_{\frac{\partial}{\partial x^i}} X^j \frac{\partial}{\partial x^j} = \frac{\partial X^j}{\partial x^i}\frac{\partial}{\partial x^j}$.

2° Here we introduce the bundle $\mathcal{O}(-1)$ over $\mathbb{C}P^1$. This is the 'tautological' bundle whose fiber over a point $x \in \mathbb{C}P^1$ is the complex line $x \subset \mathbb{C}^2$. In homogeneous coordinates $[z^1 : z^2]$ on $\mathbb{C}P^1$ we can write two sections: $s_1 = (1, \frac{z^2}{z^1})$ and $s_2 = (\frac{z^1}{z^2}, 1)$, defined on the open sets $U_1 = \{z^1 \neq 0\}$ and $U_2 = \{z^2 \neq 0\}$. One has $s_1 = \frac{z^2}{z^1} s_2$ on U_{12} so the transition function for the bundle $\mathcal{O}(-1)$ is $u = \frac{z^2}{z^1}$.

Now we define a connection on $\mathcal{O}(-1)$ in the following way: locally we can consider a section as a map $s : \mathbb{C}P^1 \longrightarrow \mathbb{C}^2$ such that $s(x) \in x$, and we define

$$\nabla_X s = \pi_x(d_x s(X)), \tag{II.4}$$

where π_x is the orthogonal projection on x. If we take a coordinate z on U_1 by considering the point $[1 : z]$, then we can write $s_1(z) = (1, z)$ and therefore

$$\nabla_X s_1 = \pi_{(1,z)}(0, X) = \frac{X\bar{z}}{1 + |z|^2} s_1.$$

Similarly, with the same coordinate z,

$$\nabla_X s_2 = -\frac{X}{z(1 + |z|^2)} s_2.$$

So in the two charts we have the Christoffel symbols $\Gamma = \frac{\bar{z}dz}{1+|z|^2}$ and $\Gamma' = -\frac{dz}{z(1+|z|^2)}$. The difference $\Gamma' - \Gamma = -\frac{dz}{z}$ indeed coincides with $-duu^{-1}$ since $u = z$. (The connection is well defined on the whole $\mathbb{C}P^1$ by the formula (II.4); nevertheless, to define it completely in trivializations, it remains to check that the formula for ∇s_2 in U_{12} extends to the whole U_2 by taking the coordinate $z' = \frac{1}{z}$).

3° If $M \hookrightarrow \mathbb{R}^N$ is an immersed submanifold, then much as in the previous example one can define a connection on TM: consider at each point the tangent space T_xM as a subspace of $\mathbb{R}^n$ and denote π_{T_xM} the orthogonal projection $\mathbb{R}^n \longrightarrow T_xM$, then one defines

$$\nabla^M_X s = \pi_{T_xM}(\nabla^{\mathbb{R}^n}_X s), \quad X \in T_xM. \tag{II.5}$$

It is easy to check that it is indeed a connection on TM.

4° *Induced connections:* a connection on a vector bundle E induces a connection on E^*, by the rule, for $s \in \Gamma(E)$, $\sigma \in \Gamma(E^*)$ and $X \in \Gamma(TM)$:

$$\mathscr{L}_X\langle\sigma, s\rangle = \langle\nabla^{E^*}_X \sigma, s\rangle + \langle\sigma, \nabla^E_X s\rangle. \tag{II.6}$$

If (e_a) is a local basis of sections of E, then the dual basis (e^a) is a local basis for E^*, and the duality bracket writes, for $s = s^a e_a$ and $\sigma = \sigma_b e^b$,

$$\langle\sigma, s\rangle = \sigma_a s^a.$$

The equation (II.6) then gives immediately

$$\nabla_{\frac{\partial}{\partial x^i}} \sigma = \Big(\frac{\partial\sigma_a}{\partial x^i} - \Gamma^b_{ia}\Big)e^a = \frac{\partial\sigma}{dx^i} - \Gamma^t_i\sigma.$$

Therefore the connection 1-form for E^* is $-\Gamma^t$.

5° Suppose we have connections ∇^E and ∇^F on the vector bundles E and F. Then there is a naturally induced connection on $G = \mathrm{Hom}(E, F) = E^* \otimes F$, defined similarly: one requests that if $s \in \Gamma(E)$ and $u \in \mathrm{Hom}(E, F)$, then

$$\nabla^F(u(s)) = (\nabla^G u)(s) + u(\nabla^E s). \tag{II.7}$$

From this it follows quickly that

$$\nabla^G_{\frac{\partial}{\partial x^i}} u = \frac{\partial u}{\partial x^i} + \Gamma^F_i \circ u - u \circ \Gamma^E_i.$$

(Remark that for $F = \mathbb{R}$ we recover the previous case $G = E^*$).

More generally, by asking that the Leibniz rule like in (II.7) is true for algebraic operations, one easily extends a connection on E to all associated bundles (tensor products, exterior products).

8.c Metric connections

A ***metric*** on a vector bundle E is the smooth data of a definite positive quadratic form g_x in each fiber E_x (Hermitian form if E is complex). An example we have already seen is a Riemannian metric on the tangent bundle. Another example is the bundle $\mathscr{O}(-1)$

in section 8.b: each fiber is naturally a complex line of $\mathbb{C}^2$ and so inherits a Hermitian metric from that of $\mathbb{C}^2$.

If the bundle E has a metric g, we say that a connection ∇ on E is a ***metric connection*** (or ***unitary connection***) if for any sections s, t of E and any vector field X:

$$\mathscr{L}_X\big(g(s,t)\big) = g(\nabla_X s, t) + g(s, \nabla_X t).$$

What does it mean on the Christoffel symbols ? suppose that (e_a) is a local *orthonormal* basis of E, then for all a, b we must have $g(\nabla_X e_a, e_b) + g(e_a, \nabla_X e_b) = 0$, whence

$$\begin{cases} \Gamma_{ia}^b = -\Gamma_{ib}^a & \text{if E is real,} \\ \Gamma_{ia}^b = -\overline{\Gamma_{ib}^a} & \text{if E is complex.} \end{cases}$$

This condition characterizes the metric connections. It means that the matrices Γ_i take values in antisymmetric or anti-Hermitian endomorphisms of E. So for example it is obvious that the flat connection on $T\mathbb{R}^n$ is a metric connection. We shall denote the bundle of antisymmetric endomorphisms $\mathfrak{so}(E)$, and the bundle of anti-Hermitian endomorphisms $\mathfrak{u}(E)$. So we have proved the following version of lemma 8.2:

Lemma 8.3. *The space of metric connections of* (E, g) *is an affine space with direction* $\Gamma(\Omega^1 \otimes \mathfrak{so}(E))$ *in the real case,* $\Gamma(\Omega^1 \otimes \mathfrak{u}(E))$ *in the complex case.*

Exercises. 1° Check the connection we defined on $\mathscr{O}(-1)$ is a metric connection.

2° The connection induced on TM by an immersion $M \hookrightarrow \mathbb{R}^N$ (see example 3° in 8.b) is a metric connection for the metric induced from the embedding.

3° If we have a metric on E, we can identify E^* with E using the metric. Therefore we have two connections on E^*: the connection of $E \simeq E^*$ and the connection as the dual of E. Prove that these two connections coincide.

8.d Parallel transport

If we have a trivial vector bundle $E = M \times \mathbb{R}^k$ or $M \times \mathbb{C}^k$, then all fibers of the bundle are identified with a fixed vector space $\mathbb{R}^k$ or $\mathbb{C}^k$. But for a general vector bundle E over M, there is no canonical way to identify the fibers of E_x, say with E_{x_0} for x close to x_0. We will see that a connection provides exactly the tool for such an identification.

Lemma 8.4. *Suppose that* (E, ∇) *is a vector bundle with connection over* M. *If we have a path* $c(t) \in M$*, and a section* $s(t) \in E_{c(t)}$ *of* E *over c, then* $\nabla_{\dot c(t)} s(t)$ *depends only on* $s(t)$.

Proof. In a local trivialization over a coordinate chart, let $c(t) = (x^i(t))$ and s has values in $\mathbb{R}^k$, then we have the formula

$$\nabla_{\dot c} s = \dot x^i \Big(\frac{\partial s}{\partial x^i} + \Gamma_i s\Big) = \dot s + \Gamma_{\dot c} s. \tag{II.8}$$

□

The same formula shows that the equation

$$\nabla_{\dot c} s = 0 \tag{II.9}$$

is a first order linear ordinary differential equation on s. Therefore given some initial condition $s(0)$ one can construct a unique solution of (II.9) along $c(t)$. This leads to the following definition:

Definition 8.5. Let (E, ∇) be a bundle with connection over M. If $(c(t))_{t\in[a,b]}$ is a path in M, then the ***parallel transport*** along c is the application $\mathrm{E}_{c(a)} \longrightarrow \mathrm{E}_{c(b)}$, $s(a) \mapsto s(b)$ obtained by solving the equation (II.9) along c.

The parallel transport $\mathrm{E}_{c(a)} \longrightarrow \mathrm{E}_{c(b)}$ is always a linear isomorphism, since the inverse is obtained by parallel transport along c in the reverse direction. If ∇ is a metric connection, then the parallel transport is an isometry (this can be seen abstractly using the definition of a metric connection, or directly from equation (II.8) written in an orthonormal trivialization).

9 Riemannian connection, geodesics

9.a The Levi-Civita connection

If we have a connection defined on the tangent bundle TM of M, then there is an interesting invariant:

Lemma and Definition 9.1. *If* ∇ *is a connection defined on the tangent bundle* TM *of* M*, then*

$$\mathrm{T}_{\mathrm{X,Y}} := \nabla_{\mathrm{X}}\mathrm{Y} - \nabla_{\mathrm{Y}}\mathrm{X} - [\mathrm{X}, \mathrm{Y}]$$

is tensorial. Therefore it defines a 2-form on M *with values in* TM*, called the* **torsion** *of* M.

The proof of the lemma is left to the reader, who will also check the following local formula for the torsion:

$$\mathrm{T} = (\Gamma_{ij}^k - \Gamma_{ji}^k). \tag{II.10}$$

In particular, a ***torsion-free connection*** (that is, a connection with zero torsion), satisfies the symmetry $\Gamma_{ij}^k = \Gamma_{ji}^k$ of its Christoffel symbols.

Theorem and Definition 9.2. *If* (M, g) *is a Riemannian or pseudo-Riemannian manifold, then* TM *admits a unique torsion-free metric connection, called the* **Levi-Civita connection** *of* M.

Proof. First the uniqueness: if we have two such connections, then by lemma 8.3 the difference between these connections is a 1-form $a = (a_{ij}^k)$ with values into antisymmetric endomorphisms of TM. The torsion free condition gives the symmetry condition $a_{ij}^k = a_{ji}^k$. Now instead of choosing a local coordinate trivialization $(\frac{\partial}{\partial x^i})$ of TM, let us choose an orthonormal trivialization (e_i) of TM, and write a in this trivialization. Then the two conditions write:

$$a_{ij}^k = -a_{ik}^j, \quad a_{ij}^k = a_{ji}^k.$$

This immediately implies $a = 0$, so the two connections are equal.

Now let us consider the existence. For a submanifold of $\mathbb{R}^N$ with the induced metric, it is easy to check that the connection defined in example 3° of section 8.b, has vanishing torsion, so this is the connection we want to construct. For an abstract manifold, one has the formula

$$2\langle \nabla_X Y, Z\rangle = \mathscr{L}_X\langle Y, Z\rangle + \mathscr{L}_Y\langle Z, X\rangle - \mathscr{L}_Z\langle X, Y\rangle \\ + \langle [X, Y], Z\rangle - \langle [X, Z], Y\rangle - \langle [Y, Z], X\rangle. \quad \text{(II.11)}$$

The reader will check that this indeed defines a metric torsion-free connection. □

The formula (II.11) gives immediately an expression in local coordinates. We have

$$2g_{kl}\Gamma^l_{ij} = 2\Big\langle \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{dx^j}, \frac{\partial}{\partial x^k}\Big\rangle = \frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}$$

so that

$$\Gamma^l_{ij} = \frac{1}{2} g^{kl}\Big(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\Big). \quad \text{(II.12)}$$

9.b Geodesics

We will now see how the Levi-Civita connection gives us the correct tool to find the equation satisfied by the curves which minimize the distance between two points x and y. Suppose that $c : [a, b] \longrightarrow M$ is a path, then its ***length*** is

$$L(c) = \int_a^b \sqrt{g(\dot c(t), \dot c(t))}dt.$$

When there is no ambiguity, we will write more simply $L(c) = \int_a^b |\dot c| dt$. This is independent of the parameterization of c, that is $L(c\circ\phi) = L(c)$ for any diffeomorphism ϕ from an interval of $\mathbb{R}$ to $[a, b]$. In particular it is easy to change the parameterization so that c is parameterized by arc length: $|\dot c| = \text{cst}$. We want to analyze the paths realizing the minimum distance from x to y (which we call ***minimizing paths***), and for this we will find the critical points of L. We consider a family of paths $c_s : [a, b] \longrightarrow M$ depending on $s \in] - \epsilon, \epsilon[$, and we wish to calculate $\frac{d}{ds}L(c_s)$ at $s = 0$. We turn c into a map $[0, 1]\times] - \epsilon, \epsilon[\longrightarrow M$.

We note $X = Tc(\frac{\partial}{\partial t})$ the tangent vectors to the curves, and $N = Tc(\frac{\partial}{\partial s})$ the vector tangent to the deformation. Hence X and N are vector fields defined along c, in the same sense used in lemma 8.4. In particular, the covariant derivatives of X and N are well defined along X and N. The reader will check that $[X, N]$ still makes sense along c, and that we have the relations

$$[X, N] = Tc\Big(\Big[\frac{\partial}{\partial t}, \frac{\partial}{\partial s}\Big]\Big) = 0 \quad \text{(II.13)}$$

$$\nabla_X N - \nabla_N X - [X, N] = T_{X,N}. \quad \text{(II.14)}$$

Then one can do the following calculation:

$$\frac{d}{ds}L(c_s) = \frac{d}{ds}\int_a^b |X| dt \\ = \int_a^b \frac{g(X, \nabla_N X)}{|X|} dt$$

because ∇ is torsion-free and using (II.13) and (II.14):

$$\begin{aligned} &= \int_a^b \frac{g(\mathrm{X}, \nabla_{\mathrm{X}}\mathrm{N})}{|\mathrm{X}|} dt \\ &= \int_a^b \frac{1}{|\mathrm{X}|}\Big(\frac{d}{dt} g(\mathrm{X},\mathrm{N}) - g(\nabla_{\mathrm{X}}\mathrm{X},\mathrm{N})\Big). \end{aligned}$$

Now up to re-parameterizing c_0 by arc length, that is $|\mathrm{X}|_{s=0} = \mathrm{cst}$, we obtain the formula for the variation of length:

$$\frac{d}{ds}\Big|_{s=0} \mathrm{L}(c_s) = \frac{1}{|\mathrm{X}|}\Big(- \int_a^b g(\nabla_{\mathrm{X}}\mathrm{X},\mathrm{N}) dt + g(\mathrm{X},\mathrm{N})|_{t=1} - g(\mathrm{X},\mathrm{N})|_{t=0}\Big). \tag{II.15}$$

Now c_0 being a critical point of L among paths from x to y means that for any deformation c_s of c with the same endpoints, the derivative of $\mathrm{L}(c_s)$ at $s = 0$ vanishes. This implies that (II.15) must vanish for any normal vector field N such that $\mathrm{N}(a) = 0$ and $\mathrm{N}(b) = 0$. It follows that if c is parametrized by arc length, then c is a critical point of L if and only if $\nabla_{\mathrm{X}}\mathrm{X} = 0$, that is $\nabla_{\dot c}\dot c = 0$.

Definition 9.3. A path $c : [a, b] \longrightarrow \mathrm{M}$ is called a ***geodesic*** if $\nabla_{\dot c}\dot c = 0$.

Remark that the definition implies $\frac{d}{dt}|\dot c|^2 = 2\langle \dot c, \nabla_{\dot c}\dot c\rangle = 0$, so a geodesic is always parametrized by arc length. We then summarize the above calculation in the following:

Lemma 9.4. *A path $c : [a, b] \longrightarrow \mathrm{M}$ parametrized by arc length is a critical point of the length among paths from $c(a)$ to $c(b)$ if and only if it is a geodesic.*

Exercise. Prove that the critical points of the ***energy***

$$\mathrm{E}(c) = \int_a^b |\dot c|^2 dt \tag{II.16}$$

among paths from $c(a)$ to $c(b)$ are exactly the geodesics.

Let us write now the geodesic equation in local coordinates (x^i): if $c(t) = (x^i(t))$, then $\dot c = \dot x^i \frac{\partial}{\partial x^i}$ and

$$\begin{aligned} \nabla_{\dot c}\dot c &= \dot x^j \Big(\frac{\partial \dot x^i}{\partial x^j} + \Gamma^i_{jk}\dot x^k\Big)\frac{\partial}{\partial x^i} \\ &= \big(\ddot x^i + \Gamma^i_{jk}\dot x^j \dot x^k\big)\frac{\partial}{\partial x^i} \end{aligned} \tag{II.17}$$

This is a nonlinear second order differential equation on $(x^i(t))$. It has a unique solution on some maximal interval as soon as $c(0)$ and $\dot c(0)$ are given, that is the initial position and the initial speed.

Example 9.5. 1° On $\mathbb{R}^n$, the equation reads $\ddot x^i = 0$, so the solutions are lines in $\mathbb{R}^n$.

2° On $\mathrm{S}^n \subset \mathbb{R}^{n+1}$ the Levi-Civita connection is the projection of the Levi-Civita connection of $\mathbb{R}^{n+1}$ (see the proof of theorem 9.2). Then check that the solutions are the great circles (draw a picture).

3° On $\mathrm{H}^n \subset \mathbb{R}^{1,n}$, this is similar, the geodesics are the intersections of H^n with the hyperplanes of $\mathbb{R}^{1,n}$.

4° On a torus $\mathrm{M}^n = \mathbb{R}^n/\mathbb{Z}^n$, the projection $\pi : \mathbb{R}^n \longrightarrow \mathrm{M}^n$ is a local isometry, so it sends a geodesic of $\mathbb{R}^n$ to a geodesic of M^n. Therefore the geodesics are the projections of straight lines in $\mathbb{R}^n$.

9.c Killing fields

We will now see that symmetries of a Riemannian manifold enable to calculate more easily the geodesics. Let us begin by introducing the infinitesimal version of an isometry.

Lemma and Definition 9.6. *A vector field* X *on a Riemannian manifold generates a flow of isometries if and only if for any vector fields* Y *and* Z *one has*

$$\langle \nabla_Y X, Z\rangle + \langle \nabla_Z X, Y\rangle = 0.$$

Such a vector field is called a **Killing field***.*

Proof. The vector X generates a flow (ϕ_t) of diffeomorphisms, and $\frac{d}{dt}\phi_t^* g = \frac{d}{d\epsilon}|_{\epsilon=0}\phi_{t+\epsilon}^* g = \frac{d}{d\epsilon}|_{\epsilon=0}\phi_t^*\phi_\epsilon^* g = \phi_t^* \mathscr{L}_X g$, so ϕ_t is a flow of isometries ($\phi_t^* g = g$) if and only if $\mathscr{L}_X g = 0$. Now using the properties of the Levi-Civita connection:

$$\begin{aligned}(\mathscr{L}_X g)(Y,Z) &= \mathscr{L}_X(g(Y,Z)) - g(\mathscr{L}_X Y, Z) - g(Y, \mathscr{L}_X Z)\\ &= g(\nabla_X Y - [X,Y], Z) + g(Y, \nabla_X Z - [X,Z])\\ &= g(\nabla_Y X, Z) + g(Y, \nabla_Z X).\end{aligned}$$

□

Remark 9.7. It follows from the definition, or from a direct calculation, that the bracket of two Killing vector fields is again a Killing vector field, so the space of Killing vector fields is an algebra for the bracket. This algebra turns out to be the Lie algebra of the isometry group.

Lemma 9.8. *If* X *is a Killing vector field and* c *a geodesic, then* $\langle \dot c, X\rangle$ *is constant along* c*.*

Proof. One has $\mathscr{L}_{\dot c}\langle \dot c, X\rangle = \langle \dot c, \nabla_{\dot c} X\rangle = 0$ (the first equality by the geodesic equation, the second by the Killing condition). □

The quantity $\langle \dot c, X\rangle$ is preserved along a geodesic, it is a ***first integral*** of the geodesic equation. This is useful for finding the solutions of the geodesic equation when the metric has symmetries, and we shall now give an example.

Example 9.9. Suppose we have a surface of revolution, with metric $g = du^2 + r(u)^2 d\theta^2$ (see example 7.2). The rotation vector $X = \frac{\partial}{\partial\theta}$ generates the flow of rotations of the surface, and is therefore a Killing field. Then our first integral says immediately that along a geodesic c, the quantity $r^2\dot\theta$ is a constant, say C. On the other hand, if we suppose c parametrized by arc length, then $\dot u^2 + r^2\dot\theta^2 = 1$. Therefore we obtain the system

$$\dot\theta = \frac{C}{r^2}, \quad \dot u = \sqrt{1 - \frac{C^2}{r^2}}. \tag{II.18}$$

The geodesic equation is now reduced to a system of first order differential equations, which is completely integrable (one can solve it). Two special kinds of solutions are interesting:

- $C = 0$, then $u(t) = t$ and $\theta = \text{cst}$: these are the meridians;
- $u(t) = \text{cst} = u_0$, then $C = u_0$ and $\dot\theta = \frac{1}{r(u_0)}$: there are horizontal circles, but they are geodesics if and only if $\frac{dr}{du}|_{u=u_0} = 0$. Question: why ? and why do we find solutions which are not geodesics ?

Exercise. On the 2-sphere S^2 we consider the metric of revolution

$$g = \frac{(1+f(z))^2}{1-z^2}dz^2 + (1-z^2)d\theta^2.$$

Show that if f is an odd function ($f(-z) = -f(z)$), then all geodesics of g are circles (Zoll, 1903).

10 Exponential map

10.a Exponential map and injectivity radius

Let (M^n, g) be a Riemannian manifold. Let $x \in M$, $X \in T_xM$, and γ be the geodesic such that $\gamma(0) = x$ and $\dot\gamma(0) = X$. Then we define

$$\exp_x(X) = \gamma(1). \tag{II.19}$$

So $\exp_x$ is a map from some subset of T_xM to M. Remark that if X is tangent to the geodesic $\gamma(t)$ and $\lambda \in \mathbb{R}$, then λX is tangent to the geodesic $\gamma(\lambda t)$, and it follows that

$$\exp_x(\lambda X) = \gamma(\lambda). \tag{II.20}$$

Since the geodesic $\gamma(t)$ exists at least for small t, we see that $\exp_x(\lambda X)$ is well defined for small enough λ. Varying X, it follows that $\exp_x$ is defined on some open neighborhood of the origin in T_xM.

Taking the derivative of (II.20) with respect to λ at $\lambda = 0$, we see that $d_0 \exp_x(X) = X$, and therefore

$$d_0 \exp_x = 1_{T_xM}. \tag{II.21}$$

It follows that $\exp_x : T_xM \longrightarrow M$ is a local diffeomorphism on a neighborhood of the origin. Now restrict $\exp_x$ on an open set $U \subset T_xM$ on which it is a diffeomorphism onto an open set $V \subset M$, and consider

$$\exp_x^{-1} : V \longrightarrow U \subset T_xM \simeq \mathbb{R}^n. \tag{II.22}$$

This gives a canonical local coordinate chart for M (given the Riemannian metric g), and the coordinates obtained in this way are called ***normal coordinates***.

An important notion in Riemannian geometry is the ***injectivity radius***: the injectivity radius at x is the supremum of all $r > 0$ such that $\exp_x$ is a diffeomorphism on a ball of radius r, and the injectivity radius of M is the infimum for all $x \in M$ of the injectivity radius at the point x.

10.b Normal coordinates

Let us write the metric in normal coordinates: $g = g_{ij}dx^i dx^j$. Since $d_0 \exp_x$ is the identity, it follows that

$$g_{ij}(0) = \delta_{ij}. \tag{II.23}$$

In these coordinates, the straight rays from the origin are geodesics:

$$\nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial r} = 0. \tag{II.24}$$

From this equation used at the origin, it follows that $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}(0) = 0$ and therefore all the Christoffel symbols vanish at the origin:

$$\Gamma_{ij}^k(0) = 0. \tag{II.25}$$

Finally,

$$\frac{\partial g_{ij}}{\partial x^k} = \left\langle \nabla_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle + \left\langle \frac{\partial}{\partial x^i}, \nabla_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^j} \right\rangle$$

which vanishes at the origin, so it follows that

$$g_{ij} = \delta_{ij} + \mathrm{O}(r^2). \tag{II.26}$$

This means that in normal coordinates, the metric is approximated up to second order by the Euclidean metric $\sum(dx^i)^2$. As we shall see later, it is not possible in general to obtain a better approximation, because the second derivatives of the coefficients g_{ij} can be interpreted as curvatures of the metric.

Example 10.1. 1° In $\mathbb{R}^n$, one has $\exp_x(\mathrm{X}) = x + \mathrm{X}$ since the geodesics are the straight lines. The injectivity radius is $+\infty$.

2° In S^n, in the stereographic projection from the north pole, the geodesics issued from the south pole become straight lines, but the velocity in the coordinates is not constant, see formula (II.1). To obtain the normal coordinates, it is therefore sufficient to re-parameterize each ray by arc length: this gives the change of coordinates $\rho = 2\arctan r$ (so $\rho < \pi$), and the formula

$$g = d\rho^2 + \sin^2(\rho) g_{s^{n-1}}. \tag{II.27}$$

The injectivity radius is π.

3° Similarly prove that the hyperbolic metric can be written in normal coordinates as

$$g = d\rho^2 + \sinh^2(\rho) g_{\mathrm{S}^{n-1}}. \tag{II.28}$$

The injectivity radius is $+\infty$.

Lemma 10.2 (Gauss). *If* $\gamma(t)$ *is a ray issued from* $\gamma(0) = x$, *and* $\mathrm{Y} \in \mathrm{T}_x\mathrm{M}$ *is orthogonal to* $\dot\gamma(0)$, *then* $d_{t\dot\gamma(0)} \exp_x(\mathrm{Y}) \perp \dot\gamma(t) = d_{t\dot\gamma(0)} \exp_x(\dot\gamma(0))$.

This means that despite the fact that the exponential map is not an isometry, the orthogonality with the rays is preserved.

Proof. We use the formula (II.15) for the variation of the length. Note $X = \dot\gamma(0)$, and consider a family of geodesics $(\gamma_s(t))_{t\in[0,T]}$ issued from x, on a fixed interval $[0,T]$, such that $\gamma_0 = \gamma$, $|\dot\gamma_s(0)| = |X|$ and $\frac{d}{ds}\dot\gamma_s|_{s=0} = Y$ (this is possible since $Y \perp X$). Then $L(\gamma_s) = T|X|$ is constant, so the formula for the variation of the length gives us

$$0 = \frac{\langle X, N\rangle}{|X|}, \quad \text{with } N = \frac{d\gamma_s(T)}{ds}\Big|_{s=0}.$$

But N is exactly $d_{T\dot\gamma(0)} \exp_x(Y)$ and we get the lemma. □

This lemma has several important consequences. First it tells us that in normal coordinates, the rays from the origin are orthogonal to the concentric spheres, which imply that

$$g = dr^2 + g_r, \quad g_r = r^2 g_{S^{n-1}} + O(r^4), \tag{II.29}$$

with g_r a family of metrics on the sphere S^{n-1}. For the Euclidean metric $g_r = r^2 g_{S^{n-1}}$, for the sphere and the hyperbolic space see the formulas (II.27) and (II.28).

The second consequence, which follows immediately from (II.29), is that on a ball $B \subset M$ on which exp is a diffeomorphism, then for any $x \in B$ the shortest path from 0 to x is the geodesic from 0 to x, and it is unique. This expresses the fact that *the geodesics are locally minimizing.*

The third consequence is that *small balls are convex*: for any x, for r small enough, any two points of the ball $B(x,r)$ are joined by a geodesic which is the unique shortest path between these two points. The proof is left as an exercise.

Fourth and final consequence is that *any minimizing path between two points is a smooth geodesic.* This follows immediately from the fact geodesics are locally minimizing.

10.c Hopf-Rinow theorem

On a Riemannian manifold (M, g) we have a natural distance

$$d(p,q) = \inf L(\text{path from } p \text{ to } q).$$

We say that (M, g) is a ***complete Riemannian manifold*** if the metric space (M, d) is complete. In general a Riemannian manifold is not complete: for example $\mathbb{R}^n \setminus \{0\}$ is not complete, since a ray going towards the origin must stop after finite time.

Theorem 10.3 (Hopf-Rinow). *Let* (M, g) *be a connected Riemannian manifold. Then the following are equivalent:*

1. (M, g) *is complete;*
2. *for any* $x \in M$, $\exp_x$ *is defined on* T_xM*;*
3. *there exists* $x \in M$*, such that* $\exp_x$ *is defined on* T_xM.

If (M, g) *is complete, then any two points of* M *can be joined by a minimizing geodesic.*

Proof. 1. $\Longrightarrow$ 2. If one has a geodesic c defined on a maximal interval $[0, T[$ and T is finite, then because it is parametrized by arc length and M is complete, it follows that $c(t)$ has a limit at $t = T$. The derivative $\dot{c}$ also converges because $|\dot{c}| = 1$ and the sphere is compact (this is not true in pseudo-Riemannian geometry). Since the geodesic equation is a second order ODE, it follows that one can extend a bit c beyond T. So $T = +\infty$.

3. $\Longrightarrow$ (every point of M can be joined to x by a minimizing geodesic). So fix $y \in M$, we shall construct a geodesic from x to y of length $d(x, y)$. Let $S_\delta(x) \subset M$ the sphere of (small) radius $\delta > 0$ around x, then there exists $z_1 \in S_\delta(x)$ such that $d(x, y) = d(x, z_1) + d(z_1, y)$. Let c be the geodesic ray from x passing through z_1, we shall prove that $c(d(x, y)) = y$. Let

$$I = \{t \geqslant 0, d(x, c(t)) + d(c(t), y) = d(x, y)\}.$$

It is clear that $\delta \in I$ and I is closed. Let $T \in I$, and suppose that $T < d(x, y)$. Again for small $\epsilon > 0$ there exists $z_2 \in S_\epsilon(c(T))$ such that

$$d(c(T), y) = d(c(T), z_2) + d(z_2, y).$$

Then it follows that

$$\begin{aligned} d(x, z_2) &\geqslant d(x, y) - d(z_2, y) = d(x, c(T)) + d(c(T), y) - d(z_2, y) \\ &\geqslant d(x, c(T)) + d(c(T), z_2) \end{aligned}$$

therefore $d(x, z_2) = d(x, c(T)) + d(c(T), z_2)$. This implies that the path c from x to $c(T)$ followed by the geodesic from $c(T)$ to z_2 is a minimizing path, and therefore is a smooth geodesic: so z_2 is on the geodesic c, meaning $z_2 = c(T + \epsilon)$ so I is open in $[0, d(x, y)]$. This finally proves that $d(x, y) \in I$.

3. $\Longrightarrow$ 1. If (x_i) is a Cauchy sequence, then by the previous statement one has $x_i = \exp_x(X_i)$ with $|X_i| = d(x, x_i)$. Since (x_i) is bounded, the sequence (X_i) is bounded so some subsequence converges: $X_{i'} \longrightarrow X$ and $x_i \longrightarrow \exp_x(X)$. □

An isometry between Riemannian manifolds M and N sends the metric of M to the metric of N, and therefore preserves all associated objects, in particular the Levi-Civita connection. The image of a geodesic by an isometry is therefore a geodesic: if $f : M \longrightarrow N$ is an isometry, then

$$f(\exp_x^M X) = \exp_{f(x)}^N d_x f(X)$$

where $\exp^M$ exists. In particular, if M is complete, by the Hopf-Rinow theorem $\exp_x^M$ is well defined on T_xM and therefore the LHS with x fixed and $X \in T_xM$ determines f. We deduce:

Corollary 10.4. *Suppose $f_1, f_2 : M \longrightarrow N$ are isometries between connected complete Riemannian manifolds. If for some $x \in M$ one has $f_1(x) = f_2(x)$ and $d_x f_1 = d_x f_2$, then $f_1 = f_2$.* □

Chapter III

Curvature

11 Curvature and integrability

11.a Horizontal distribution

Let $\pi : E \longrightarrow M$ be a vector bundle over M with a connection ∇, and $x \in M$. We have seen in section 8.d that if we have a path $(c(t))_{t\in[0,1]}$ in M and an initial value $s_0 \in E_x$, then c can be lifted to a path s in E such that $\nabla_{\dot c} s = 0$.

Actually there is an infinitesimal version of this process: if $X = \dot c(0) \in T_x M$, then we define the ***horizontal lift*** of X at $s_0 \in E_x$ to be

$$\tilde X = \frac{ds}{dt}\Big|_{t=0}. \tag{III.1}$$

We claim that $\tilde X \in T_{s_0}E$ does not depend of the choice of c, so it depends only of X. One way to see that is by calculating $\tilde X$ in a local trivialization $(e_1, \dots, e_r)$ of E, over a coordinate open set U of M, with coordinates (x^i). Therefore locally

$$E|_U \simeq U \times \mathbb{R}^r$$

with coordinates $(x^i, s^a)_{i=1,\dots,n,\, a=1,\dots,r}$, and the corresponding vector fields $\frac{\partial}{\partial x^i}$ and $\frac{\partial}{\partial s^a}$. Observe that the later ones are tangent to the fibers of E. The connection is written $\nabla = d + \Gamma$, where Γ is a 1-form with values in End E. From equation (II.8) we obtain $\dot s(0) = -\Gamma_X s_0$ and therefore

$$\tilde X = (X, -\Gamma_X s_0). \tag{III.2}$$

For example for $X = \frac{\partial}{\partial x^i}$, noting $s_0 = s^a e_a$, we obtain

$$\tilde X = \Big(\frac{\partial}{\partial x^i}, -\Gamma_i s_0\Big) = \frac{\partial}{\partial x^i} - \Gamma^a_{ib} s^b \frac{\partial}{\partial s^a}.$$

.

Definition 11.1. The ***horizontal distribution*** of (E, ∇) at each $s_0 \in E$ is the vector space of horizontal lifts $\{\tilde X, X \in T_{\pi(s_0)}M\} \subset T_{s_0}E$.

At each point $s \in E$ over $x \in M$ we have the tangent space to the fiber: $E_x \subset T_sE$ (the ***vertical space***), and the horizontal distribution (which will be denoted H_s), so that

$$T_sE = E_x \oplus H_s. \tag{III.3}$$

We see that the connection ∇ enables to choose a canonical supplementary subspace to E_x in each tangent space T_sE. (This choice actually characterizes ∇). From formula (III.2) it follows that at 0 the horizontal distribution coincides with the tangent space to the zero section $\{(y,0), y \in M\}$. From this point of view, the parallel transport is interpreted as transporting s over a path c in M by following the horizontal distribution.

11.b Integrability and curvature

We shall now study the integrability of the horizontal distribution of a connection ∇. Suppose it is integrable, therefore for a small enough open set $V \subset E$ containing a given point $(x,0)$ we can suppose that $V = L \times W$ with $L \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^r$, and the leaves of H are the $L \times \{s\}$. Now consider the application $f : V \longrightarrow M \times W$ defined by

$$f(l,w) = (\pi(l,w), w).$$

Since the leaf $L \times \{0\}$ is just (an open set of) M itself, the differential at $(x,0)$ is an isomorphism. By the inverse function theorem, f is a diffeomorphism from a smaller open set V' to an open set of $M \times W$ that we choose of the form $U \times W$, where $U \subset M$ is a small open set containing x. So we now obtain via f a diffeomorphism

$$V' \simeq U \times W$$

such that

- the leaves of H are still of the form $\{y\} \times W$;
- furthermore in these new coordinates $p(y,w) = y$.

Here W parametrizes the leaves, it can be identified to the open subset $\{x\} \times W$ of the fiber E_x.

Because the horizontal distribution is preserved by the homotheties $s \mapsto \lambda s$ (see equation (III.2)), this decomposition extends to the whole inverse image $\pi^{-1}(U)$:

$$E|_U = U \times E_x = U \times \mathbb{R}^r \tag{III.4}$$

and the horizontal leaves are obtained by fixing a point on the factor $\mathbb{R}^r$, therefore the parallel transport in E is given by the identity of $\mathbb{R}^r$.

Here one must be careful that the identification $E|_U = U \times \mathbb{R}^r$ is a priori only a diffeomorphism, with an identification of $\mathbb{R}^r$ with the fiber at x. But since the parallel transport between two points of U is given by the identity on the factor $\mathbb{R}^r$, and the parallel transport acts by linear isomorphisms, we see that over any y the resulting diffeomorphism $E_y = \{y\} \times \mathbb{R}^r$ is actually a linear isomorphism. This means that (III.4) is a local trivialization of E as a vector bundle. Then the equation (III.2) gives immediately $\Gamma = 0$ everywhere in this trivialization, so ∇ is the trivial connection $\nabla = d$. We have proved:

Lemma 11.2. *The horizontal distribution of* ∇ *is integrable if and only if* E *admits local trivializations in which* ∇ *is trivial.*

Remark that if ∇ is a metric connection, then the local trivializations can be supposed to be orthonormal, since the parallel transport preserves the metric.

In view of Frobenius theorem 5.4, the condition of integrability of the horizontal distribution is equivalent to its involutivity. Let us study the condition in a local trivialization as above. Suppose X and Y are two local vector fields on M, then

$$\begin{aligned}[\tilde{X}, \tilde{Y}] &= \big[X - \Gamma^a_{Xb} s^b \frac{\partial}{\partial s^a}, Y - \Gamma^a_{Yb} s^b \frac{\partial}{\partial s^a}\big] \\ &= [X, Y] + \big(-\mathscr{L}_X \Gamma^a_{Yb} + \mathscr{L}_Y \Gamma^a_{Xb} + \Gamma^a_{Xc}\Gamma^c_{Yb} - \Gamma^c_{Xb}\Gamma^a_{Yc}\big) s^b \frac{\partial}{\partial s^a}\end{aligned}$$

using the differential of the 1-form $\Gamma^a_b : X \mapsto \Gamma^a_{Xb}$:

$$= \widetilde{[X, Y]} - \big(d\Gamma^a_b(X, Y) + \Gamma^a_{Xc}\Gamma^c_{Yb} - \Gamma^c_{Xb}\Gamma^a_{Yc}\big) s^b \frac{\partial}{\partial s^a}.$$

We can rewrite this formula in a more concise way:

$$\begin{aligned}[\tilde{X}, \tilde{Y}](s) &= \widetilde{[X, Y]}(s) - \big((d\Gamma)_{X,Y} s + [\Gamma_X, \Gamma_Y] s\big) \\ &= \widetilde{[X, Y]}(s) - F_{X,Y} s, \end{aligned} \tag{III.5}$$

where we define

$$F_{X,Y} = (d\Gamma)_{X,Y} + [\Gamma_X, \Gamma_Y]. \tag{III.6}$$

This is a 2-form with values in End E, so an element of $\Omega^2(\operatorname{End} E)$. From equation (III.5) it is clear that it does not depend on the trivialization, but this can also be checked directly: if we choose a different trivialization $(e_a) = \phi(f_b)$, then by (II.3) we have the transformation $\Gamma \longrightarrow \phi \circ \Gamma \circ \phi^{-1} - d\phi \circ \phi^{-1}$, and the reader can check that $F \longrightarrow \phi \circ F \circ \phi^{-1}$.

Definition 11.3. The ***curvature*** of the connection ∇ is the 2-form with values in End E defined locally by the formula (III.6).

Recall that if the connection is metric, then Γ is a 1-form with values in $\mathfrak{so}(E)$ or $\mathfrak{u}(E)$ in an orthonormal trivialization, and one can read from formula (III.6) that F is a 2-form with values in the same bundle.

Example 11.4. If we take the line bundle $\mathscr{O}(-1)$ over $\mathbb{C}P^1$, with the connection $\Gamma = \frac{\bar{z} dz}{1+|z|^2}$ constructed in section 8.b, then the brackets in (III.6) vanish since we have only a line bundle, and we get

$$F^{\mathscr{O}(-1)} = -\frac{dz \wedge d\bar{z}}{(1+|z|^2)^2} = \frac{2i dx \wedge dy}{(1+x^2+y^2)^2},$$

which is a 2-form with values in $i\mathbb{R} = \mathfrak{u}_1$. (As we shall see later, the first Chern number of $\mathscr{O}(-1)$ is $\frac{i}{2\pi}\int_{\mathbb{C}P^1} F = -1$).

Definition 11.5. A connection is ***flat*** if its curvature vanishes.

Since the vanishing of the curvature is equivalent to the involutivity of the horizontal distribution, lemma 11.2 implies that ∇ is flat if and only if E has local trivializations in which ∇ is the trivial connection: $\nabla = d$.

In the special case of the Levi-Civita connection, we obtain:

Lemma 11.6. *If* (M^n, g) *is a Riemannian manifold, then its Levi-Civita connection is flat if and only if near any point there exist local coordinates* (x^i) *such that* $g = \sum(dx^i)^2$.

Proof. If the Levi-Civita connection on TM is flat, then near each point we have an orthonormal basis $(X_1, \dots, X_n)$ of parallel vector fields: $\nabla X_i = 0$. In particular, since ∇ is torsion-free, one has $[X_i, X_j] = \nabla_{X_i} X_j - \nabla_{X_j} X_i = 0$. From the proof of lemma 5.3, it follows that there exists local coordinates such that $X_i = \frac{\partial}{\partial x^i}$. □

11.c Second derivatives and curvature

Here we present another definition of the curvature, less geometric but of equal importance. Suppose (E, ∇) is a fiber bundle.

Lemma and Definition 11.7. *If* X *and* Y *are vector fields and* s *is a section of* E, *then*

$$F_{X,Y} s = \nabla_X \nabla_Y s - \nabla_Y \nabla_X s - \nabla_{[X,Y]} s$$

is tensorial in X, Y *and* s, *and defines a 2-form* F *with values in* $\operatorname{End} E$ *called the* **curvature** *of* E.

The proof is left to the reader. In a trivialization, $\nabla = d + \Gamma_i dx^i$, and one can calculate directly

$$F_{ij} = \frac{\partial \Gamma_j}{\partial x^i} - \frac{\partial \Gamma_i}{\partial x^j} + [\Gamma_i, \Gamma_j] \tag{III.7}$$

which is the same result as in equation (III.6), so we defined the same object.

This point of view will be used later in various calculations, in particular to establish the differential Bianchi identity 16.2.

12 Riemannian curvature

12.a Symmetries of the Riemannian curvature

The Levi-Civita connection satisfies additional properties, which make sense only for connections on the tangent bundle. In general we will denote the curvature of the Levi-Civita connection by R, so R is an element of $\Omega^2(M, \mathfrak{o}(TM))$.

Lemma 12.1 (***Bianchi identity***). *The Levi-Civita connection* R *of a Riemannian manifold satisfies*

$$R_{X,Y} Z + R_{Y,Z} X + R_{Z,X} Y = 0.$$

Proof. This quantity is

$$\underline{\nabla_X \nabla_Y Z} - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z + \nabla_Y \nabla_Z X - \nabla_Z \nabla_Y X \underline{- \nabla_{[Y,Z]} X} + \nabla_Z \nabla_X Y \underline{- \nabla_X \nabla_Z Y} - \nabla_{[Z,X]} Y$$

The three underlined terms give $[X, [Y, Z]]$, and gathering the other terms similarly, we get

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

which vanishes by the Jacobi identity. (This proof extends to any torsion-free connection). □

Corollary 12.2. *The Riemannian curvature satisfies*

$$\langle R_{X,Y}Z, T\rangle = \langle R_{Z,T}X, Y\rangle.$$

Therefore R *defines a symmetric endomorphism of* Λ^2TM *called the* **curvature operator**.

The proof is left to the reader, who should play with the Bianchi identity and the fact that $R_{X,Y}$ is an anti-symmetric endomorphism.

This gives symmetries satisfied by the Riemannian curvature. If we take an orthonormal basis of vector fields $(X_1, \dots, X_n)$, we can write down the curvature as $R = (R_{abc}{}^d)$, with ab the indices of the 2-form, and cd that of the anti-symmetric endomorphism. Then we have the following symmetries:

$$R_{abc}{}^d = -R_{bac}{}^d, \quad R_{abd}{}^c = -R_{abc}{}^d, \quad R_{cda}{}^b = R_{abc}{}^d.$$

12.b Sectional curvature

From these symmetries we see for example that in dimension 2, the curvature has only one coefficient $K = R_{122}{}^1$ which is the Gauss curvature of the surface. In higher dimension, one defines analogous 2-dimensional curvatures in the following way:

Definition 12.3. Let (M, g) be a Riemannian manifold, and $P \subset T_xM$ a 2-plane. One defines the ***sectional curvature*** of the plane P as the number

$$K(P) = \frac{g(R_{X,Y}Y, X)}{g(X,X)g(Y,Y) - g(X,Y)^2}$$

for any basis (X, Y) of P.

It is easy to check that the definition does not depend on the basis (X, Y). Also the definition still makes sense in pseudo-Riemannian manifolds, provided that the denominator does not vanish.

One can prove that the data of the sectional curvatures of all 2-planes in T_xM completely determines the curvature tensor at the point x.

Example 12.4. 1° The curvature of the flat $\mathbb{R}^n$ vanishes and therefore all the sectional curvatures vanish.

2° For the sphere S^n, we first observe that the isometry group $SO(n+1)$ is transitive on 2-planes: indeed it is transitive on the points of S^n, and the isotropy group of a point is $SO(n)$ which acts transitively on 2-planes of $\mathbb{R}^n$. Since the curvature and the sectional curvatures are canonically defined from the metric, they are preserved by isometries and it follows that all the sectional curvatures of S^n equal a fixed constant (+1, as we shall see later).

3° Similarly the hyperbolic space H^n has constant sectional curvature.

4° The sectional curvatures of $\lambda^2 g$, where λ is a positive number, are related to that of g by the relation

$$K^{\lambda^2 g} = \frac{1}{\lambda^2}K^g. \tag{III.8}$$

This comes immediately from the fact that g and $\lambda^2 g$ have the same Levi-Civita connection, and therefore the same curvature tensor R. The formula corresponds to the idea that if we make a sphere very big (λ big), then its curvature becomes small, that is it becomes almost flat. Indeed the earth looks locally very flat !

13 Second fundamental form

13.a Covariant derivative and second fundamental form

Suppose that (M^n, g) is an oriented Riemannian manifold, and $N^{n-1} \subset M$ is a submanifold oriented by the normal vector $\vec{n}$. Similarly to the case of submanifolds of $\mathbb{R}^n$, it is easy to check that the Levi-Civita connection of N is

$$\nabla^N = \pi \circ \nabla^M, \tag{III.9}$$

where $\pi : T_xM \longrightarrow T_xN$ is the orthogonal projection. Therefore, for two vector fields X, Y on N, the covariant derivative $\nabla^M_X Y$ decomposes as

$$\nabla^M_X Y = \nabla^N_X Y + \mathbb{I}(X, Y)\vec{n}. \tag{III.10}$$

Actually, in the pseudo-Riemannian case, it will be useful to use the following normalization:

$$\nabla^M_X Y = \nabla^N_X Y + \mathbb{I}(X, Y)\frac{\vec{n}}{\langle \vec{n}, \vec{n}\rangle}. \tag{III.11}$$

Developing the torsion-free condition $\nabla^M_X Y - \nabla^M_Y X = [X, Y]$ with (III.10), we obtain the symmetry condition

$$\mathbb{I}(X, Y) = \mathbb{I}(Y, X). \tag{III.12}$$

A priori the formula (III.10) defines $\mathbb{I}(X, Y)$ as a tensorial object only with respect to X. Then the symmetry (III.12) proves that it is tensorial also with respect to Y (this can also be checked directly). We can now state:

Definition 13.1. The formula $\mathbb{I}(X, Y) = \langle \nabla^M_X Y, \vec{n}\rangle$ defines a symmetric 2-tensor on N, called the ***second fundamental form*** of N.

Directly from the definition, using the properties of ∇^M, one also gets:

$$\mathbb{I}(X, Y) = -\langle \nabla^M_X \vec{n}, Y\rangle. \tag{III.13}$$

This gives another formula for the second fundamental form: $\mathbb{I} = -\nabla^M \vec{n}$.

13.b Curvature and second fundamental form

There is an application $\phi : \mathbb{R} \times N \longrightarrow M$, defined by

$$\phi(r, x) = \exp_x(r\vec{n}).$$

(If M is not complete, then ϕ may be defined only on an open subset of $\mathbb{R} \times N$). This means that from each point $x \in N$ we draw the geodesic from x which is orthogonal to N, and we parameterize it by its arc length r.

The differential of ϕ at a point $(0, x)$ is $d_{(0,x)}\phi(r, X) = r\vec{n} + X$, so it is an isomorphism $\mathbb{R} \times T_xN \longrightarrow T_xM$. It follows that ϕ is a diffeomorphism from a neighborhood of $\{0\} \times N \subset \mathbb{R} \times N$ onto a neighborhood of $N \subset M$ (say, if N is compact, or at least locally near any point of N).

Lemma 13.2. *The geodesics normal to* N *are orthogonal to the hypersurfaces* $\phi(\{r\} \times N)$.

Proof. This is a version of Gauss lemma 10.2, the proof is similar. □

It follows that on the open set where ϕ is a diffeomorphism, one has

$$\phi^* g = dr^2 + g_r, \qquad g_r \text{ metric on N.} \tag{III.14}$$

The normal vector $\vec{n}$ can then be extended to a neighborhood of N as $\phi_* \frac{\partial}{\partial r}$. If X is a vector field on N, one can extend it to a neighborhood of N as being independent of the $\mathbb{R}$ variable in the product $\mathbb{R} \times \mathrm{N}$; equivalently, this is the unique extension so that $[\vec{n}, \mathrm{X}] = 0$. Choose two vector fields X, Y on N and extend them in this way: then from (III.13) one deduces

$$\mathbb{I}(\mathrm{X}, \mathrm{Y}) = -\langle \nabla_{\vec{n}} \mathrm{X}, \mathrm{Y} \rangle + \langle [\vec{n}, \mathrm{X}], \mathrm{Y} \rangle = -\langle \nabla_{\vec{n}} \mathrm{X}, \mathrm{Y} \rangle;$$

by symmetry we get

$$\mathbb{I}(\mathrm{X}, \mathrm{Y}) = -\frac{1}{2}\big(\langle \nabla_{\vec{n}} \mathrm{X}, \mathrm{Y} \rangle + \langle \nabla_{\vec{n}} \mathrm{Y}, \mathrm{X} \rangle\big) = -\frac{1}{2}\mathscr{L}_{\vec{n}} \langle \mathrm{X}, \mathrm{Y} \rangle.$$

This proves the formula:

$$\mathbb{I} = -\frac{1}{2} \frac{\partial g_r}{\partial r}\Big|_{r=0}, \tag{III.15}$$

which gives a concrete formula to calculate $\mathbb{I}$.

Often we will need to consider $\mathbb{I}$ as a symmetric endomorphism of N rather than a quadratic form: therefore we define the ***Weingarten endomorphism*** A by the formula

$$\mathbb{I}(\mathrm{X}, \mathrm{Y}) = g(\mathrm{A}(\mathrm{X}), \mathrm{Y}), \qquad \mathrm{A}(\mathrm{X}) = -\nabla_{\mathrm{X}} \vec{n}, \tag{III.16}$$

so that we have the formula

$$\mathrm{A} = g^{-1} \mathbb{I} = -\frac{1}{2} g^{-1} \frac{\partial g}{\partial r}. \tag{III.17}$$

Lemma 13.3. *If* X *and* Y *are two vectors of* N*, one has the formulas*

$$\mathrm{K}^{\mathrm{M}}(\mathrm{X} \wedge \mathrm{Y}) = \mathrm{K}^{\mathrm{N}}(\mathrm{X} \wedge \mathrm{Y}) - \langle \vec{n}, \vec{n} \rangle \frac{\mathbb{I}(\mathrm{X}, \mathrm{X}) \mathbb{I}(\mathrm{Y}, \mathrm{Y}) - \mathbb{I}(\mathrm{X}, \mathrm{Y})^2}{\langle \mathrm{X}, \mathrm{X} \rangle \langle \mathrm{Y}, \mathrm{Y} \rangle - \langle \mathrm{X}, \mathrm{Y} \rangle^2}; \tag{III.18}$$

$$\mathrm{K}^{\mathrm{M}}(\mathrm{X} \wedge \vec{n}) = \frac{\mathscr{L}_{\vec{n}} \mathbb{I}(\mathrm{X}, \mathrm{X}) + |\mathrm{AX}|^2}{\langle \mathrm{X}, \mathrm{X} \rangle \langle \vec{n}, \vec{n} \rangle}. \tag{III.19}$$

In the second formula, A *is the Weingarten endomorphism of the hypersurfaces* $\{r\} \times \mathrm{N}$.

It is important to note that this lemma is written for a pseudo-Riemannian manifold: that is why we kept $\langle \vec{n}, \vec{n} \rangle$ which may equal -1.

Proof. We only deal with the Riemannian case, and let the reader check the signs in the pseudo-Riemannian case.

Let X and Y be tangent vectors to N at the point x. We can extend X and Y so that $[\mathrm{X}, \mathrm{Y}] = 0$. Let Z, T be vector fields on N, and denote ∇ the connection of M and ∇^{N} that of N. Then

$$\begin{aligned} \langle \nabla_{\mathrm{X}} \nabla_{\mathrm{Y}} \mathrm{Z}, \mathrm{T} \rangle &= \langle \nabla_{\mathrm{X}} (\nabla^{\mathrm{N}}_{\mathrm{Y}} \mathrm{Z} + \mathbb{I}(\mathrm{Y}, \mathrm{Z}) \vec{n}), \mathrm{T} \rangle \\ &= \langle \nabla^{\mathrm{N}}_{\mathrm{X}} \nabla^{\mathrm{N}}_{\mathrm{Y}} \mathrm{Z} + \mathbb{I}(\mathrm{Y}, \mathrm{Z}) \nabla_{\mathrm{X}} \vec{n}, \mathrm{T} \rangle \\ &= \langle \nabla^{\mathrm{N}}_{\mathrm{X}} \nabla^{\mathrm{N}}_{\mathrm{Y}} \mathrm{Z}, \mathrm{T} \rangle - \mathbb{I}(\mathrm{Y}, \mathrm{Z}) \mathbb{I}(\mathrm{X}, \mathrm{T}). \end{aligned}$$

Therefore

$$\langle R_{X,Y}Z, T\rangle = \langle R^N_{X,Y}Z, T\rangle - \mathbb{I}(Y, Z)\mathbb{I}(X, T) + \mathbb{I}(X, Z)\mathbb{I}(Y, T). \tag{III.20}$$

The first formula follows, applying to Z = Y and T = X.

Now let us prove the second formula. We write the metric as in (III.14), so we get an extension of the normal vector $\vec{n}$ outside N as the velocity vector of the geodesics normal to N, in particular

$$\nabla_{\vec{n}}\vec{n} = 0.$$

We now start with vector fields X and Y on N, which we extend in M by deciding that

$$[\vec{n}, X] = [\vec{n}, Y] = 0.$$

It follows that

$$\nabla_{\vec{n}}Y = \nabla_Y\vec{n} = -A(Y). \tag{III.21}$$

Now:

$$\begin{aligned}\langle R_{\vec{n},X}Y, \vec{n}\rangle &= \langle(\nabla_{\vec{n}}\nabla_X - \nabla_X\nabla_{\vec{n}})Y, \vec{n}\rangle\\ &= \mathscr{L}_{\vec{n}}\langle\nabla_X Y, \vec{n}\rangle + \langle\nabla_X(A(Y)), \vec{n}\rangle\\ &= \mathscr{L}_{\vec{n}}\big(\mathbb{I}(X, Y)\big) + \langle A(Y), A(X)\rangle.\end{aligned}$$

Because $\mathscr{L}_{\vec{n}}X = \mathscr{L}_{\vec{n}}Y = 0$, one has $(\mathscr{L}_{\vec{n}}\mathbb{I})(X, Y) = \mathscr{L}_{\vec{n}}(\mathbb{I}(X, Y))$, and we obtain

$$\langle R_{\vec{n},X}Y, \vec{n}\rangle = (\mathscr{L}_{\vec{n}}\mathbb{I})(X, Y) + \langle A(X), A(Y)\rangle. \tag{III.22}$$

Applying to Y = X we get the formula. □

Remark 13.4. Because of Gauss lemma 13.2, the covariant derivative $\nabla_{\vec{n}}$ preserves $\{0\} \times TN \subset TM = \mathbb{R} \times TN$, and the equation (III.21) gives the following expression of $\nabla_{\vec{n}}$ acting on TN:

$$\nabla_{\vec{n}} = \mathscr{L}_{\vec{n}} - A. \tag{III.23}$$

It follows that, since $\mathbb{I}$ is a section of S^2T^*N,

$$\nabla_{\vec{n}}\mathbb{I}(\cdot, \cdot) = \mathscr{L}_{\vec{n}}\mathbb{I}(\cdot, \cdot) + \mathbb{I}(A\cdot, \cdot) + \mathbb{I}(\cdot, A\cdot).$$

But $\mathbb{I}(A\cdot, \cdot) = \mathbb{I}(\cdot, A\cdot) = \langle A^2\cdot, \cdot\rangle$, so the second formula of the lemma can also be written as

$$K^M(X \wedge \vec{n}) = \frac{\nabla_{\vec{n}}\mathbb{I}(X, X) - |AX|^2}{\langle X, X\rangle\langle\vec{n}, \vec{n}\rangle}. \tag{III.24}$$

Because of equation (III.23), this is the same as

$$K^M(X \wedge \vec{n}) = \frac{\langle(\mathscr{L}_{\vec{n}}A - A^2)X, X\rangle}{\langle X, X\rangle\langle\vec{n}, \vec{n}\rangle}. \tag{III.25}$$

Example 13.5. 1° $g_{\mathbb{R}^{n+1}} = dr^2 + r^2 g_{S^n}$ and $\mathbb{I} = -r g_{S^n}$ so $A = -\frac{1}{r}$; by the first formula $0 = K^{S^n} - 1$, which gives us the the curvature $K^{S^n} = 1$.

2° Similarly, $g_{\mathbb{R}^{1,n}} = dr^2 - r^2 g_{H^n}$; this is the same formula as before, so again $K(-g_{H^n}) = 1$; observing that the sectional curvatures of $-g$ and g are opposite, we obtain $K^{H^n} = -1$.

13.c Surfaces in $\mathbb{R}^3$

If we have a surface $S \subset \mathbb{R}^3$, then the two eigenvalues λ_1 and λ_2 of $\mathbb{I}$ are called the ***principal curvatures*** of S. The first equation gives us the well-known formula for the Gauss curvature:

$$K^S = \lambda_1\lambda_2.$$

The principal curvatures depend on the embedding $S \subset \mathbb{R}^3$ but the product depends only on the intrinsic geometry of S: this is the content of the theorem Egregium of Gauss. Also

$$H = \lambda_1 + \lambda_2$$

is called the ***mean curvature***. Surfaces with $H = 0$ are called ***minimal surfaces***: this is the equation satisfied by the soap bubbles.

Finally, if the surface S is given by an equation $z = f(x, y)$, then the reader will check the following explicit formulas: the metric on S is given by

$$g_{11} = 1 + (\partial_x f)^2, \quad g_{12} = \partial_x f \partial_y f, \quad g_{22} = 1 + (\partial_y f)^2$$

the normal vector is

$$\vec{n} = \frac{(-\partial_x f, -\partial_y f, 1)}{\sqrt{1 + (\partial_x f)^2 + (\partial_y f)^2}},$$

from which one deduces the second fundamental form:

$$\mathbb{I}_{11} = \frac{\partial^2_{xx} f}{\sqrt{1+(\partial_x f)^2+(\partial_y f)^2}}, \quad \mathbb{I}_{12} = \frac{\partial^2_{xy} f}{\sqrt{1+(\partial_x f)^2+(\partial_y f)^2}}, \quad \mathbb{I}_{22} = \frac{\partial^2_{yy} f}{\sqrt{1+(\partial_x f)^2+(\partial_y f)^2}}.$$

It follows that the curvature of S is given by

$$K = \frac{\det(\mathbb{I}_{ij})}{\det(g_{ij})} = \frac{\partial^2_{xx} f \partial^2_{yy} f - (\partial^2_{xy} f)^2}{1 + (\partial_x f)^2 + (\partial_y f)^2}.$$

The geometric meaning is then clear: in particular for $K > 0$ we get a convex surface.

13.d A geometric interpretation of the curvature

The sectional curvatures of a Riemannian metric g can be seen as coefficients is the Taylor development of the metric in normal coordinates. Recall that in normal coordinates,

$$\exp_x^* g = dr^2 + r^2(g_{S^{n-1}} + r^2\gamma + \cdots).$$

Note $\bar{\gamma} = g_{S^{n-1}}^{-1}\gamma$ the corresponding endomorphism on TS^{n-1}, then

$$\mathbb{I} \sim -\frac{1}{2}\frac{\partial g}{\partial r} \sim -r g_{S^{n-1}} - 2r^3\gamma, \quad A \sim -\frac{1}{r}(1 + r^2\bar{\gamma}).$$

Fix a vector $X \in S^{n-1} \subset T_x M$, which we can suppose to be an eigenvector of γ for the eigenvalue $\bar{\gamma}$, then

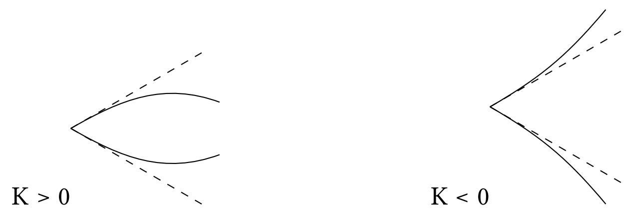
$$\begin{aligned} K(\frac{\partial}{\partial r} \wedge X) &= \frac{\mathscr{L}_{\frac{\partial}{\partial r}} \mathbb{I}(X, X) + g(A^2X, X)}{g(X, X)} \\ &\sim \frac{(-1 - 6r^2\bar{\gamma}) + (1 + 3r^2\bar{\gamma})}{1 + r^2\bar{\gamma}} \\ &\sim -3\bar{\gamma}. \end{aligned}$$

We deduce:

$$g(X,X) \sim r^2 - \frac{1}{3}K(\frac{\partial}{\partial r} \wedge X)r^4 + \cdots \tag{III.26}$$

For the flat $\mathbb{R}^n$, one has $K = 0$ and the geodesics are straight lines. The equation (III.26) means that, at least near the point x, comparing with the straight lines:

- when $K > 0$ the geodesics get closer (think of the sphere: two great circles starting from the same point finally meet again);
- when $K < 0$ the geodesics get far away from each other.



14 Constant curvature metrics

Lemma 14.1. *If* (M^n, g) *has constant sectional curvature, then in normal coordinates* $\exp^* g$ *coincides with the metric of* $\mathbb{R}^n$, S^n *or* H^n *(up to a multiplicative constant).*

Proof. We use normal coordinates around a point x: then $g = dr^2 + g_r$, with g_r a metric on S^{n-1}. The second formula of lemma 13.3, under the form (III.24), gives us

$$\frac{\partial A}{\partial r} - A^2 = k, \tag{III.27}$$

where k is the (constant) sectional curvature. When $r \longrightarrow 0$ we have the asymptotic behavior $A \sim -\frac{1}{r}$, and in particular A is invertible near 0. Consider $B = A^{-1}$, then B extends at 0 with $B(0) = 0$ and satisfies

$$\frac{\partial B}{\partial r} = -1 - kB^2.$$

One can solve this ODE and deduce g_r:

- if $k = 0$, then $B = -r$, $A = -\frac{1}{r}$ and $g_r = r^2$;
- if $k > 0$, then $A = -\frac{\cot(\sqrt{k}r)}{\sqrt{k}}$ and $g_r = \frac{\sin^2(\sqrt{k}r)}{k}$;
- if $k < 0$, then $A = -\frac{\coth(\sqrt{-k}r)}{\sqrt{-k}}$ and $g_r = \frac{\sinh^2(\sqrt{-k}r)}{\sqrt{-k}}$.

□

Corollary 14.2. *A Riemannian manifold with constant curvature is locally isometric to* $\mathbb{R}^n$, S^n *or* H^n *(up to a constant). If moreover the manifold is connected and simply connected, then it is exactly* $\mathbb{R}^n$, S^n *or* H^n.

The first part of the corollary is an immediate consequence of the lemma. The second part is more global and will be proved in the next section.

15 Riemannian curvature and topology

Here we will see how the curvature—more precisely its sign—has influence on the topology of the manifold. This is an important area of research in geometry, and we give only two basic results, the Cartan-Hadamard theorem and the Bonnet-Myers theorem.

15.a The conjugacy radius

In section 10 we have seen the notion of injectivity radius—the supremum of the $r > 0$ such that $\exp_x$ is a diffeomorphism on the ball of radius r. Here we will use another notion, the ***conjugacy radius***, that is the supremum of the $r > 0$ such that $\exp_x$ is a local diffeomorphism on the ball of radius r. This is equivalent to $\exp_x^* g$ being a metric on the ball of radius r, so we can define alternatively the conjugacy radius at x by

$$\rho_{conj}(x) = \inf\{r > 0, \det(\exp_x^* g) \text{ vanishes at some point of } S(r)\}. \quad \text{(III.28)}$$

As in the proof of lemma 14.1, by Gauss lemma we have $\exp_x^* g = dr^2 + g_r$ and on the ball of radius $\rho_{conj}(x)$, one has, for $|X|= 1$,

$$\langle(\nabla_{\frac{\partial}{\partial r}} A - A^2)X, X\rangle = K(\frac{\partial}{\partial r} \wedge X). \quad \text{(III.29)}$$

So for example if $K \leqslant 0$, then $\nabla_{\frac{\partial}{\partial r}} A - A^2 \leqslant 0$ which implies $A \leqslant -\frac{1}{r}$ and $g_r \geqslant r^2 g_{S^{n-1}}$. It follows that $\det(\exp_x^* g)$ can never vanish and we obtain the first part of:

Corollary 15.1. *If* $K \leqslant 0$, *then* $\rho_{conj} = +\infty$. *If* $K \leqslant k$ *with* $k > 0$, *then* $\rho_{conj} \geqslant \frac{\pi}{\sqrt{k}}$.

Proof. It remains to deal with $K \leqslant k$: the same proof gives us $g_r \geqslant \frac{\sin^2(\sqrt{k}r)}{k}$ and therefore $\det(\exp_x^* g)$ cannot vanish for $r < \frac{\pi}{\sqrt{k}}$. □

15.b The Cartan-Hadamard theorem

Theorem 15.2 (Cartan-Hadamard). *If* (M^n, g) *is a complete connected Riemannian manifold with* $K \leqslant 0$, *then* $\exp_x : T_xM \longrightarrow M$ *is a covering. In particular, if* M *is simply connected, then* M *is diffeomorphic to* $\mathbb{R}^n$.

Proof. We have just seen that $\exp_x$ is a local diffeomorphism. It follows that $\exp_x : (T_xM, \exp_x^* g) \longrightarrow (M, g)$ is an isometry. But an isometry is always a covering, because if $r > 0$ is small (smaller than the injectivity radius at x), then $\exp_x^{-1}(B(x, r))$ is the union of balls of radius r. □

For constant curvature metrics, we deduce:

Proof of corollary 14.2. In the case of negative or zero curvature, this is just the Cartan-Hadamard theorem, and the fact that we have an explicit formula for a constant curvature metric in normal coordinates.

In the case of positive curvature, we can suppose that $K = 1$. Then by corollary 15.1, the map $\exp_x$ is a local diffeomorphism on the ball $B(0,\pi) \subset T_xM$, and since $K = 1$,

$$\exp_x^* g = dr^2 + \sin^2(r) ds^2_{S^{n-1}} \quad \text{on } B(0,\pi).$$

Since this is exactly the expression of the metric of S^n in normal coordinates, we deduce an isometric map $f_N : S^n \setminus \{S\} \longrightarrow M$ by the composition

$$S^n \setminus \{S\} \xrightarrow{\exp_N^*} (B(0,\pi), dr^2 + \sin^2(r) ds^2_{S^{n-1}}) \xrightarrow{\exp_x} M.$$

We can do the same thing from another point $y = f(p)$ where $p \in S^n \setminus \{S\}$, and we obtain another isometry $f_p : S^n \setminus \{q\} \longrightarrow M$, where q is the antipodal point to p, such that

$$f_p(p) = y = f_N(p), \quad d_p f_p = d_p f_N.$$

Since a local isometry sends a geodesic to a geodesic, it then follows that f_p and f_N coincide, except maybe on the segment joining q and S. As f_p is defined on $S^n \setminus \{q\}$ and f_S on $S^n \setminus \{N\}$, it follows that they coincide everywhere, and therefore define together an isometric map $f : S^n \longrightarrow M$, which is therefore a covering. If M is simply connected, f must be a global diffeomorphism. □

15.c The Bonnet-Myers theorem

Theorem 15.3 (Bonnet-Myers). *If (M^n, g) is a complete connected Riemannian manifold, satisfying (Bonnet)*

$$K \geqslant k > 0,$$

or the weaker hypothesis (Myers)

$$\mathrm{Ric} \geqslant (n-1)k,$$

then the diameter of M is not bigger than $\frac{\pi}{\sqrt{k}}$. In particular M is compact. Moreover $\pi_1(M)$ is finite.

Here Ric denotes the ***Ricci tensor***, which will be studied in chapter IV. For this theorem, we just need to know the definition:

$$\mathrm{Ric}(X,Y) = \mathrm{Tr}(Z \longmapsto R_{Z,X}Y). \tag{III.30}$$

It is a symmetric 2-tensor. In particular, if $|X| = 1$, we complete X into an orthonormal basis $(X = e_1, \dots, e_n)$, and

$$\mathrm{Ric}(X,X) = \sum_1^n \langle R_{e_i,X}X, e_i\rangle = \sum_2^n K(X \wedge e_i). \tag{III.31}$$

Then it is clear that the first hypothesis of the theorem is stronger than the second one.

Before giving a proof of the theorem, we outline another approach, more in the spirit of what we have just done. Recall that at a point x we have $\exp_x^* g = dr^2 + g_r$, with

$A = -\frac{1}{2}g_r^{-1}\frac{\partial g_r}{\partial r}$ satisfying equation (III.29) where $\exp_x$ is non degenerate. Taking the trace of the equation, we obtain

$$\frac{\partial \operatorname{Tr}(A)}{\partial r} - \frac{\operatorname{Tr}(A)^2}{n-1} \geqslant \operatorname{Ric}(\frac{\partial}{\partial r}, \frac{\partial}{\partial r}) \geqslant (n-1)k > 0,$$

which implies $\operatorname{Tr}(A) \geqslant -(n-1)\sqrt{k}\cot(\sqrt{k}r)$ and $\det(g_r) \leqslant \big(\frac{\sin(\sqrt{k}r)}{\sqrt{k}}\big)^{n-1}$. On each ray from the origin, we see that $\det(g_r)$ must vanish at a radius $r \leqslant R$ with $R = \frac{\pi}{\sqrt{k}}$, that is $d\exp_x$ has a kernel on each ray at most at distance R. One says that x has a ***conjugate point*** on every geodesic from x at distance at most R. But it is known that a geodesic cannot be minimizing after a conjugate point, and it follows that all points of M are at most at distance R from x. The proof of this last fact requires the theory of Jacobi fields that will not be developed in these notes, see for example [GHL04].

Proof. So we now turn to the proof of the theorem. It relies on the second variation of arc length: if $(c_s(t))$ is a family of paths defined on $[a,b]$, with fixed endpoints, and c_0 is a geodesic, then

$$\frac{d^2L(c_s)}{ds}\Big|_{s=0} = \int_a^b \big(|\nabla_{\dot c_0}\tilde N|^2 - \langle R_{\dot c_0,\tilde N}\tilde N, \dot c_0\rangle\big)\,dt \qquad \text{(III.32)}$$

where $N = \frac{\partial c}{\partial s}$ and $\tilde N$ is the projection of N orthogonally to $\dot c_0$. The proof of this formula is similar to that of the first variation of arc length (II.15) and is left to the reader.

Suppose we have two points $x, y \in M$. By the Hopf-Rinow theorem 10.3, there exists a minimizing geodesic c from x to y, of length $L = d(x,y)$. Now choose vectors $E_1,\dots,E_{n-1}$ along c_0 such that $(\dot c, E_1, \dots, E_{n-1})$ is a parallel basis of orthonormal vectors along c. For $i = 1, \dots, n-1$ choose

$$N_i = \sin(\pi\frac{t}{L})E_i.$$

These vectors vanish at the endpoints of $[0,L]$. Fix i and choose a variation (c_s) of c with fixed endpoints, such that $\frac{\partial c}{\partial s}|_{s=0} = N_i$. Since c is a minimizing geodesic, we have $\frac{d^2L(c_s)}{ds^2}|_{s=0} \geqslant 0$, and therefore

$$\int_0^L |\nabla_{\dot c}N_i|^2 - \langle R_{\dot c,N_i}N_i, \dot c\rangle \geqslant 0.$$

But $\nabla_{\dot c}N_i = \frac{\pi}{L}\cos(\frac{\pi t}{L})E_i$ so after summation over i we obtain

$$(n-1)\frac{\pi^2}{L^2}\int_0^L \cos(\tfrac{\pi t}{L})^2 \geqslant \int_0^L \sin(\tfrac{\pi t}{L})^2 \sum_1^{n-1} K(\dot c, E_i) = \int_0^L \sin(\tfrac{\pi t}{L})^2 \operatorname{Ric}(\dot c, \dot c).$$

The hypothesis gives $(n-1)\frac{\pi^2}{L^2} \geqslant (n-1)k$ so $L^2 \leqslant \frac{\pi^2}{k}$.

The diameter begin finite implies immediately that M is compact. Also remark that one can pullback the metric of M on its universal covering $\tilde M$, so $\tilde M$ also satisfies the hypothesis: it is therefore compact, which implies that $\pi_1(M)$ is finite. □

Remark that the formula (III.32) can be rewritten as

$$\frac{d^2 L(c_s)}{ds}\Big|_{s=0} = -\int_a^b \langle \nabla_{\dot c}\nabla_{\dot c}\tilde N + R_{\dot c,\tilde N}\tilde N, \dot c\rangle dt \tag{III.33}$$

so the vector fields satisfying the second order linear ODE $\nabla_{\dot c}\nabla_{\dot c}\tilde N + R_{\dot c,\tilde N}\tilde N = 0$ are exactly the kernel of the Hessian of L. These are the Jacobi fields alluded to above.

16 Chern-Weil construction

We now pass to a different topic: the construction of topological invariants of a vector bundle from the curvature of a connection on the bundle. This first requires to study more properties of the connection and the curvature.

16.a Extension of a connection

Let $E \longrightarrow M$ be a vector bundle with a connection $\nabla : \Gamma(E) \longrightarrow \Omega^1(E)$. In section 6.c we extended the derivative of functions to the exterior differential on forms. In the same way, we extend ∇ uniquely to an exterior differential on E-valued differential forms:

$$d^\nabla : \Omega^k(E) \longrightarrow \Omega^{k+1}(E) \tag{III.34}$$

satisfying the Leibniz identity, for α a differential form and σ an E-valued differential form:

$$d^\nabla(\alpha\wedge\sigma) = d\alpha\wedge\sigma + (-1)^{|\alpha|}\alpha\wedge d^\nabla\sigma. \tag{III.35}$$

This extension can be defined by the local formula $d^\nabla\sigma = d\sigma + a\wedge\sigma$ in a trivialization of E in which $\nabla = d + a$ for an End(E)-valued 1-form a. There is also a formula analogous to lemma 6.3:

$$\begin{aligned}(d^\nabla\sigma)_{X_0,\dots,X_k} = {}& \sum_0^k (-1)^k \nabla_{X_i}(\sigma_{X_0,\dots,\widehat{X_i},\dots,X_k}) \\ &+ \sum_{0\leqslant i<j\leqslant k} (-1)^{i+j}\sigma([X_i,X_j],X_0,\dots,\widehat{X_i},\dots,\widehat{X_j},\dots,X_k).\end{aligned} \tag{III.36}$$

This extension leads to a nice interpretation of the curvature: recall that the exterior differential satisfies $d\circ d = 0$. The curvature is precisely the defect for $d^\nabla\circ d^\nabla$ to vanish:

Lemma 16.1. *1° The curvature* F^∇*, seen as an operator* $\Gamma(E) \longrightarrow \Omega^2(E)$*, is* $F^\nabla = d^\nabla\circ d^\nabla$*.*
2° As an operator $\Omega^k(E) \longrightarrow \Omega^{k+2}(E)$*, one has* $F^\nabla = d^\nabla\circ d^\nabla$*.*

Proof. Let us choose a local trivialization of E, and write the connection $\nabla = d + a = d + a_i dx^i$, where each a_i is End(E)-valued. Then, for a section s of E, we have $d^\nabla s = ds + as$ and

$$\begin{aligned} d^\nabla(d^\nabla s) &= (d+a)(d+a)s \\ &= d(as) + a\wedge ds + a\wedge as \\ &= (da + a\wedge a)s. \end{aligned}$$

This proves the first statement. The proof of the second one is similar. ☐

We deduce the following important identity:

Proposition 16.2 (***differential Bianchi identity***). *The curvature of a connection satisfies the identity*

$$d^{\nabla}F^{\nabla} = 0.$$

Remark that $F^{\nabla} \in \Omega^2(\mathrm{End}(E))$ so d^{∇} is the exterior derivative associated to the connection ∇ on $\mathrm{End}\,E$, and $d^{\nabla}F^{\nabla} \in \Omega^3(\mathrm{End}(E))$.

Proof. In this proof let us distinguish ∇ on E and $\bar{\nabla}$ on $\mathrm{End}\,E$. Recall that, as a linear operator on E, for $u \in \Gamma(\mathrm{End}\,E)$ one has $\bar{\nabla}u = \nabla \circ u - u \circ \nabla$. Then the reader will check that, as operators $\Gamma(E) \longrightarrow \Omega^3(E)$, one has $d^{\bar{\nabla}}F^{\nabla} = d^{\nabla} \circ F^{\nabla} - F^{\nabla} \circ d^{\nabla}$. But since $F^{\nabla} = d^{\nabla} \circ d^{\nabla}$, we obtain

$$d^{\bar{\nabla}}F^{\nabla} = d^{\nabla} \circ d^{\nabla} \circ d^{\nabla} - d^{\nabla} \circ d^{\nabla} \circ d^{\nabla} = 0.$$

□

16.b Chern-Weil homomorphism

We fix a $\mathbb{C}^r$ vector bundle E over M. Let $P : \mathfrak{gl}_r\mathbb{C} \otimes \cdots \otimes \mathfrak{gl}_r\mathbb{C} \longrightarrow \mathbb{C}$ a symmetric k-linear form on $\mathfrak{gl}_r\mathbb{C}$. As is well-known this is equivalent to giving the homogeneous polynomial $P(A, \dots, A)$ of degree k on $\mathfrak{gl}_r\mathbb{C}$. We suppose that P is $GL_r\mathbb{C}$ invariant, that is

$$P(gA_1g^{-1}, \dots, gA_kg^{-1}) = P(A_1, \dots, A_k)$$

for all $g \in GL_r\mathbb{C}$. Differentiating at the identity, this is equivalent to

$$P([X, A_1], A_2, \dots, A_k) + \cdots + P(A_1, \dots, A_{k-1}, [X, A_k]) = 0$$

for all $X \in \mathfrak{gl}_r\mathbb{C}$. A basis of such polynomials is $(\mathrm{Tr}(A^k))_{k=0,\dots,r}$.

If $\alpha_1, \dots, \alpha_k$ are differential forms with values in $\mathrm{End}\,E$, then $P(\alpha_1, \dots, \alpha_k)$ can be defined in any local trivialization (e_i) of E by taking the exterior product on the form part and P on the End(E) part. The result does not depend on the trivialization because in another trivialization $(e_i) = g(f_i)$, the α_i are transformed into $g\alpha_ig^{-1}$, but by invariance

$$P(g\alpha_1g^{-1}, \dots, g\alpha_kg^{-1}) = P(\alpha_1, \dots, \alpha_k).$$

So we get a well-defined form $P(\alpha_1, \dots, \alpha_k)$ of degree $|\alpha_1| + \cdots + |\alpha_k|$. Note also that the symmetry implies

$$P(\alpha_2, \alpha_1, \dots) = (-1)^{|\alpha_1|+|\alpha_2|}P(\alpha_1, \alpha_2, \dots).$$

Lemma 16.3. *Let* ∇ *be a connection on* E *and* α_i *be* End(E) *valued differential forms, then*

$$d\big(P(\alpha_1, \dots, \alpha_k)\big) = P(d^{\nabla}\alpha_1, \alpha_2, \dots, \alpha_k) + (-1)^{|\alpha_1|}P(\alpha_1, d^{\nabla}\alpha_2, \dots, \alpha_k) + \cdots + (-1)^{|\alpha_1|+\cdots+|\alpha_{k-1}|}P(\alpha_1, \dots, \alpha_{k-1}, d^{\nabla}\alpha_k).$$

Proof. In a local trivialization, we obtain immediately

$$d\big(P(\alpha_1, \dots, \alpha_k)\big) = P(d\alpha_1, \alpha_2, \dots, \alpha_k) + (-1)^{|\alpha_1|}P(\alpha_1, d\alpha_2, \dots, \alpha_k) + \cdots + (-1)^{|\alpha_1|+\cdots+|\alpha_{k-1}|}P(\alpha_1, \dots, \alpha_{k-1}, d\alpha_k).$$

If $\nabla = d + a$, the difference between the quantity in the statement of the lemma and the RHS of this equality is

$$P([a \wedge \alpha_1], \alpha_2, \dots, \alpha_k) + (-1)^{|\alpha_1|} P(\alpha_1, [a \wedge \alpha_2], \dots, \alpha_k) \\ + \dots + (-1)^{|\alpha_1| + \dots + |\alpha_{k-1}|} P(\alpha_1, \dots, \alpha_{k-1}, [a \wedge \alpha_k]),$$

but this vanishes by invariance of P (the notation $[a \wedge \alpha]$ means that we are doing exterior product on the form part and bracket on the endomorphism part). □

Lemma 16.4. *If we have a family of connections* $(\nabla_t)_{t \in [0,1]}$ *then* $\frac{d}{dt} F(\nabla_t) = d^{\nabla_t} \frac{d\nabla_t}{dt}$.

Proof. In a local trivialization one has $\nabla_t = d + a_t$ and $F(\nabla_t) = da_t + a_t \wedge a_t$. Therefore

$$\frac{d}{dt} F(\nabla_t) = d\dot{a}_t + a_t \wedge \dot{a}_t + \dot{a}_t \wedge a_t.$$

But for an End(E)-valued 1-form a one has

$$d^{\nabla_t} a = da + [a_t \wedge a] = da + a_t \wedge a + a \wedge a_t.$$

□

Proposition 16.5 (Chern-Weil construction). *Let* E *be a* $\mathbb{C}^r$ *vector bundle over* M, *and* P *an invariant k-form on* $\mathfrak{gl}_r\mathbb{C}$. *Then for any connection* ∇ *on* E *the expression* $P(F^\nabla, \dots, F^\nabla)$ *defines a closed* $(2k)$*-differential form, whose cohomology class does not depend on* ∇.

Proof. By Bianchi differential identity $d^\nabla F^\nabla = 0$ so lemma 16.3 implies $d\big(P(F^\nabla, \dots, F^\nabla)\big) = 0$. So there remains to prove that the cohomology class does not depend on ∇, that is if we have two connections ∇_0 and ∇_1, then $P(F^{\nabla_1}, \dots, F^{\nabla_1}) - P(F^{\nabla_0}, \dots, F^{\nabla_0}) = d\beta$ for some 1-form β. Using the fact that the space of connections is an affine space (lemma 8.2), we can interpolate between ∇_0 and ∇_1 by considering for $t \in [0, 1]$:

$$\nabla_t = \nabla_0 + ta, \quad a = \nabla_1 - \nabla_0.$$

Then by lemma 16.4 one has $\frac{d}{dt} F_t = d^{\nabla_t} a$ and therefore, again using Bianchi identity,

$$\begin{aligned} \frac{d}{dt} P(F_t, \dots, F_t) &= P(d^{\nabla_t} a, F_t, \dots, F_t) + \dots + P(F_t, \cdots, F_t, d^{\nabla_t} a) \\ &= k\, d\big(P(a, F_t, \dots, F_t)\big). \end{aligned}$$

Therefore

$$P(F_1, \dots, F_1) - P(F_0, \dots, F_0) = k\, d\Big(\int_0^1 P(a, F_t, \dots, F_t)\Big). \tag{III.37}$$

□

Remark 16.6. The RHS of (III.37) can be made explicit, since $F_t = F_0 + t d^{\nabla_0} a + t^2 a \wedge a$. This leads to the so-called ***transgression formulas***. For example if $P(A) = \mathrm{Tr}(A^2)$, then

$$\int_0^1 P(a, F_t, \dots, F_t) = \mathrm{Tr}(a \wedge d^{\nabla_0} a + \frac{2}{3} a \wedge a \wedge a). \tag{III.38}$$

This is the beginning of the famous Chern-Simons theory, which, roughly speaking, enables to define invariants of 3-manifolds by integration of this 3-form.

Remark 16.7. The theory is actually much more general: there is a notion of G-connection for any Lie group G, and then one can define $P(F, \dots, F)$ for any G-invariant polynomial on the Lie algebra of G. We will apply later this remark to the group $SO_n \subset GL_n$. As $SO(n)$-connections are represented as $GL(n)$-connections on a vector bundle, preserving a metric, our treatment 16.5 is still valid in this case.

16.c Chern classes

Let $E \longrightarrow M$ be still a $\mathbb{C}^r$ bundle with a connection ∇. We define ***Chern classes*** $c_i \in H^{2i}(M, \mathbb{C})$ using the Chern-Weil homomorphism:

$$\begin{aligned} \det(\mathrm{Id} + \frac{iF}{2\pi}) &= 1 \ + \mathrm{Tr}(\frac{iF}{2\pi}) + \cdots \\ &= c_0 + \quad c_1 \quad + c_2 + \cdots \end{aligned}$$

Actually one can prove that $c_i \in H^{2i}(M, \mathbb{R})$. In topology one refines the definition to obtain classes $c_i \in H^{2i}(M, \mathbb{Z})$.

Example 16.8. 1° The first Chern class c_1 is associated to the polynomial $\mathrm{Tr}(\frac{iF}{2\pi})$. As we have seen in example 11.4, for the bundle $\mathcal{O}(-1)$ on S^2, one obtains $\int \frac{iF}{2\pi} = -1$, so the first Chern class is -1.

2° The second Chern class c_2 is associated to the polynomial

$$\frac{1}{8\pi^2}\mathrm{Tr}(F \wedge F) + \frac{1}{2}\frac{\mathrm{Tr}(F)^2}{4\pi^2} = \frac{1}{8\pi^2}\big(\mathrm{Tr}(F \wedge F) - \mathrm{Tr}(F)^2\big). \tag{III.39}$$

3° One proves easily the identity

$$c(E \oplus F) = c(E)c(F). \tag{III.40}$$

4° One can define invariants of real vector bundles by taking Chern classes of their complexification. For example, for an $\mathbb{R}^3$-bundle V, one defines the first Pontryagin class by

$$p_1(V) = -c_2(V \otimes \mathbb{C}) \in H^4(M). \tag{III.41}$$

16.d Euler form

Before defining the Euler form, we need some algebraic preliminaries.

First, if we have an Euclidean vector space $(\mathbb{R}^n, g)$, then there is an identification

$$\begin{aligned} \Phi : \mathfrak{o}(\mathbb{R}^n) &\xrightarrow{\sim} \Lambda^2\mathbb{R}^n \\ u &\longmapsto \Phi(u)_{X,Y} = \langle u(X), Y\rangle \end{aligned} \tag{III.42}$$

For example the antisymmetric endomorphism $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ corresponds to the form $e^1 \wedge e^2$.

Now the second algebraic preliminary. Suppose that our Euclidean space is even dimensional, $n = 2m$. Then to an antisymmetric endomorphism $A \in \Lambda^2\mathbb{R}^n$ we can associate its ***Pfaffian*** $\mathrm{Pf}(A)$ such that

$$\mathrm{Pf}(A)^2 = \det(A). \tag{III.43}$$

This is defined by the scalar product

$$\mathrm{Pf}(A) := \left\langle \frac{A^m}{m!}, e^1 \wedge \cdots \wedge e^n \right\rangle. \tag{III.44}$$

For example suppose $m = 1$, then $A = \theta e^1 \wedge e^2 = \theta\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and we obtain $\mathrm{Pf}(A) = \theta$ and $\det(A) = \theta^2$.

More generally, every antisymmetric endomorphism A can be put in an orthonormal basis such that

$$A = \theta_1 e^1 \wedge e^2 + \theta_2 e^3 \wedge e^4 + \cdots + \theta_m e^{n-1} \wedge e^n.$$

Then $\mathrm{Pf}(A) = \theta_1\theta_2 \cdots \theta_m$.

After these algebraic preliminaries, we can now state the main result of this section.

Theorem and Definition 16.9. *Let n be even and E be a rank n real oriented vector bundle over M. Then the n-form $\mathrm{Pf}(F^\nabla)$ defined for any metric connection on E is closed, and its cohomology class does not depend on the choice of metric and connection on E.*

The form $\frac{-F^\nabla}{2\pi}$ is called the **Euler form** *of ∇.*

An orientation of E is defined in the same way as the orientation of a manifold, by the choice of a non vanishing section of $\Lambda^n E$. If $E = TM$ an orientation of TM is the same as an orientation of M.

Proof. First fix a metric. Then the form $\mathrm{Pf}(F^\nabla)$ is obtained from the Chern-Weil construction applied to the group $SO(n)$, as Pf is an invariant polynomial on $\mathfrak{so}(n)$. Therefore it is closed and its cohomology class does not depend on the connection.

Now we have to prove that the form does not depend on the choice of metric. If we have two metrics g_0 and g_1 on E, they are related by an automorphism φ of E, by $g_1(e, f) = g_0(\varphi e, \varphi f)$. We can suppose that φ preserves the orientation of E. Then remark that if ∇_0 is a g_0-metric connection, then $\nabla_1 = \varphi^{-1} \circ \nabla_0 \circ \varphi$ is a g_1-metric connection, and $F(\nabla_1) = \varphi^{-1} F(\nabla_0)\varphi$. So $\mathrm{Pf}(F(\nabla_1)) = \mathrm{Pf}((\nabla_0))$. □

Fact 16.10. *Let (M^n, g) be an even dimensional compact Riemannian manifold. Then the Euler form of TM satisfies*

$$\int_M \mathrm{Pf}\left(\frac{-F}{2\pi}\right) = \chi(M),$$

where $\chi(M) := \sum_{i=0}^n (-1)^i b_i(M)$ is the Euler characteristic of M.

We shall not prove this statement, see [BGV04].

Let us calculate the formula if $n = 2$. We know that on a Riemann surface Σ, the Riemannian curvature of a metric is $F_{e_1,e_2} = -K e^1 \wedge e^2$, and therefore $\mathrm{Pf}(\frac{-F}{2\pi}) = \frac{K}{2\pi} e^1 \wedge e^2$. Then we obtain the Gauss-Bonnet formula:

$$\int_\Sigma \frac{K}{2\pi} \mathrm{vol} = \chi(\Sigma) = 2 - 2g(\Sigma). \tag{III.45}$$

In general, if we have a rank n real vector bundle E over an oriented n-dimensional manifold M, then the ***Euler number*** $\int_M \mathrm{Pf}(\frac{-F}{2\pi})$ of E is a topological invariant which equals the number of zeros (with multiplicities and signs) of a section of E. Again see [BGV04] for details.

Chapter IV

Einstein equation

17 Ricci tensor

17.a Ricci and scalar curvature

We defined the Ricci tensor Ric of a pseudo-Riemannian metric in section 15 from the curvature tensor by the formula

$$\mathrm{Ric}(X, Y) = \mathrm{Tr}(Z \longrightarrow R_{Z,X}Y).$$

In an orthonormal basis (e_i) of the tangent bundle, one has

$$\mathrm{Ric}(X, Y) = \sum \epsilon_i \langle R_{e_i,X}Y, e_i \rangle. \tag{IV.1}$$

where $\epsilon_i = \langle e_i, e_i \rangle$ (this is to cover the pseudo-Riemannian case as well). The symmetries of the curvature tensor (corollary 12.2) imply immediately

$$\mathrm{Ric}(X, Y) = \mathrm{Ric}(Y, X), \tag{IV.2}$$

so the Ricci tensor is a symmetric 2-tensor.

Definition 17.1. The ***scalar curvature*** of the metric is the function defined by

$$\mathrm{Scal} = \mathrm{Tr}(g^{-1}\,\mathrm{Ric}) = \sum_1^n \epsilon_i \,\mathrm{Ric}(e_i, e_i).$$

For example, in dimension 2, in an orthonormal basis (e_1, e_2), if $K = R_{122}{}^1$ is the (Gauss) curvature, then one obtains immediately

$$\mathrm{Ric} = Kg, \quad \mathrm{Scal} = 2K.$$

For the sphere S^n one has $\mathrm{Ric} = (n-1)g$ and $\mathrm{Scal} = n(n-1)$.

For the hyperbolic space H^n one has $\mathrm{Ric} = -(n-1)$ and $\mathrm{Scal} = -n(n-1)$.

17.b The Bianchi identity

Proposition 17.2 (***differential Bianchi identity***). *The Riemannian curvature satisfies the identity*

$$(\nabla_X R)_{Y,Z} + (\nabla_Y R)_{Z,X} + (\nabla_Z R)_{X,Y} = 0.$$

Proof. This is just a way of writing the Bianchi identity 16.2, using formula (III.36) with the help of the connection induced on $\Omega^2 \otimes \mathfrak{o}(TM)$. □

Proposition 17.3 (***Bianchi identity***). *One has*

$$\delta \operatorname{Ric} = -\frac{1}{2} d \operatorname{Scal},$$

where the divergence $\delta\phi$ *of a 2-tensor* ϕ *is the 1-form defined by* $(\delta\phi)_X = -\sum_1^n \epsilon_i (\nabla_{e_i}\phi)(e_i, X)$.

Proof. We treat only the Riemannian case. We choose an orthonormal basis (e_i) of TM such that just at the point x one has $\nabla e_i(x) = 0$, and we calculate only at the point x. We can also suppose that $\nabla X(x) = 0$, then we have

$$(d \operatorname{Scal})_X(x) = \mathscr{L}_X \sum_{i,j=1}^n \langle R_{e_i,e_j} e_j, e_i \rangle = \sum_{i,j=1}^n \langle \nabla_X R_{e_i,e_j} e_j, e_i \rangle.$$

Then, using the differential Bianchi identity,

$$\begin{aligned}(\delta \operatorname{Ric})_X(x) &= -\sum_1^n \nabla_{e_i} \operatorname{Ric}(e_i, X) = -\sum_{i,j=1}^n \nabla_{e_i} \langle R_{e_j,X} e_i, e_j \rangle \\ &= \sum_{i,j=1}^n \langle \nabla_{e_j} R_{X,e_i} e_i + \nabla_X R_{e_i,e_j} e_i, e_j \rangle \\ &= -(\delta \operatorname{Ric})_X + (ds)_X.\end{aligned}$$

□

From the definition, if f is a function then $\delta(fg) = -df$, so the Bianchi identity can also be written

$$\delta\Big(\operatorname{Ric} - \frac{\operatorname{Scal}}{2}\Big) = 0. \quad \text{(IV.3)}$$

17.c Einstein equation

The Einstein equation in the vacuum is

$$\operatorname{Ric} - \frac{\operatorname{Scal}}{2} g = 0. \quad \text{(IV.4)}$$

The unknown is the metric g. Of course, in general relativity the metric is a Lorentzian metric on a 4-dimensional manifold. Sometimes one adds a 'cosmological constant' λ and the equation becomes

$$\operatorname{Ric} - \frac{\operatorname{Scal}}{2} g = \lambda g. \quad \text{(IV.5)}$$

It is still discussed among physicists whether the cosmological constant should vanish. In general, the Einstein equation takes the form

$$\mathrm{Ric} - \frac{\mathrm{Scal}}{2} g = \lambda g + \mathrm{T}, \tag{IV.6}$$

where T is the energy-impulsion tensor, which for physical reasons satisfies $\delta \mathrm{T} = 0$. The tensor $\mathrm{Ric} - \frac{\mathrm{Scal}}{2} g$ appearing in the Einstein equation is the only divergence-free combination of Ric and Scal, and this is why it must be the left-hand side of equation (IV.6). Nevertheless, the Einstein equation (IV.5) can be simplified: taking the trace of the equation, we obtain $(1 - \frac{n}{2})\,\mathrm{Scal} = n\lambda$, and therefore if $n \neq 2$,

$$\mathrm{Ric} = \frac{2\lambda}{2-n} g.$$

So in general we will call an ***Einstein metric*** a metric g satisfying the equation

$$\mathrm{Ric}(g) = \lambda g \tag{IV.7}$$

for some real constant λ. Remark that since $\mathrm{Ric}(\mu g) = \frac{1}{\mu}\mathrm{Ric}(g)$ for any $\mu > 0$, up to scaling the metric by a constant, one can have the Einstein constant λ to be $+1$, 0 or -1.

The most basic examples of Einstein metrics are of course $\mathbb{R}^n$, S^n and H^n with Einstein constants 0, $n(n-1)$ and $-n(n-1)$.

Remark that in dimension 2, since $\mathrm{Ric} = \mathrm{K}g$, an Einstein metric is just a constant curvature metric. The same holds in dimension 3 (actually in dimension 3 the Ricci tensor determines the whole Riemannian curvature). Things become very different in higher dimension.

18 Schwarzschild metric

This is the first nontrivial example of a solution of the equations of general relativity. We are looking for a Lorentzian manifold (M^4, g), with a global decomposition $\mathrm{M} = \mathbb{R} \times \mathbb{R}^3$ (the $\mathbb{R}$ is the time direction and the $\mathbb{R}^3$ is the space direction), such that g is independent of the time variable t, and has a spherical symmetry. Concretely, if ρ is the radius on $\mathbb{R}^3$, we are looking for a metric

$$g = \mathrm{F}^2(\rho) dt^2 - d\rho^2 - \mathrm{G}^2(\rho) ds^2_{\mathrm{S}^2} \tag{IV.8}$$

satisfying the equation

$$\mathrm{Ric}^g = 0.$$

One can write down directly the equation, but some qualitative considerations simplify greatly the calculations, so we will first shortly digress on totally geodesic submanifolds.

18.a Totally geodesic submanifolds

Definition 18.1. A ***totally geodesic submanifold*** of (M, g) is a submanifold $\mathrm{N} \subset \mathrm{M}$ if any geodesic of N is also a geodesic of M.

For example a subspace $\mathbb{R}^k \subset \mathbb{R}^n$ is totally geodesic, as is a sub-sphere $S^k \subset S^n$ or a sub-hyperbolic space $H^k \subset H^n$.

Lemma 18.2. *1° A submanifold* N ⊂ M *is totally geodesic if and only if its second fundamental form vanishes.*

2° If N *is the set of fixed points of an isometry, then* N *is totally geodesic.*

3° If $N^{n-1} \subset M^n$ *is a totally geodesic hypersurface, then for any* X, Y *tangent to* N *one has* $R_{X,Y}\vec{n} = 0$ *and* $\mathrm{Ric}(X, \vec{n}) = 0$.

Proof. 1° First one has to define the second fundamental form in the case of an arbitrary submanifold (rather than a codimension 1 submanifold). It is defined for $X, Y \in TN$ by $\nabla^M_X Y = \nabla^N_X Y + \mathbb{I}(X, Y)$. This is now a symmetric 2-tensor with values in $(TN)^\perp$.

Now it is clear that a geodesic c of N is also a geodesic of M if and only if $\mathbb{I}(\dot{c}, \dot{c}) = 0$. Since for any tangent vector X we can find a geodesic c such that $\dot{c} = X$ it follows that N is totally geodesic if and only if $\mathbb{I}(X, X) = 0$ for all X.

2° If $N = \mathrm{Fix}(\tau)$ where τ is an isometry, and $x \in N$, then $T_xM = T_xN \oplus (T_xN)^\perp$ and $d_x\tau = (+1) \oplus u$ in this decomposition, where u is an isometry of $(T_xN)^\perp$ without eigenvalue +1. Since τ is an isometry, one has $\tau^*\mathbb{I} = \mathbb{I}$ so $\mathbb{I}(X, Y)$ is τ-invariant. But there is no τ-invariant in $(T_xN)^\perp$, so $\mathbb{I} = 0$.

3° One has

$$R_{X,Y}\vec{n} = (\nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]})\vec{n} = -\nabla_X A(Y) + \nabla_Y A(X) + A([X, Y]) = 0. \qquad \text{(IV.9)}$$

Therefore $\mathrm{Ric}(X, \vec{n}) = \sum\langle R_{e_i,X}\vec{n}, e_i\rangle = 0$. □

Remark 18.3. In the course of the proof, we proved the Gauss-Codazzi equation: considering A as a 1-form with values in TN, then equation (IV.9) tells us that

$$(d^\nabla A)_{X,Y,Z} = \langle R_{X,Y}Z, \vec{n}\rangle.$$

18.b The equations and the Schwarzschild metric

Now we come back to the Lorentzian metric g given by (IV.8) and the Einstein equation. We use the variables $(t, \rho, x) \in \mathbb{R} \times I \times S^2$, where I is some interval. Then:

1. each slice $\{t\} \times \mathbb{R}^3$ is totally geodesic, as e.g. $\{0\} \times \mathbb{R}^3$ is a fixed point of the isometry $t \longrightarrow -t$; therefore by lemma 18.2, 3°, the vector $\frac{\partial}{\partial t}$ is an eigenvector of Ric;
2. in the same way, $\mathbb{R} \times I \times$ equatorial circle is totally geodesic, therefore every direction of S^2 is an eigenvector for Ric;
3. finally, since Ric is diagonalizable in an orthonormal basis, the last vector $\frac{\partial}{\partial \rho}$ is also an eigenvector of Ric.

So without any calculation, we have found a diagonal basis for Ric. There remains to compute the various sectional curvatures involved. We write $g = -d\rho^2 + g_\rho$ with

$g_\rho = F^2(\rho)dt^2 - G^2(\rho)ds^2_{S^2}$. Then $\mathbb{I} = -\frac{1}{2}\frac{\partial g_\rho}{\partial\rho} = -FF'dt^2 + GG'ds^2$. Therefore, in a basis $(\frac{\partial}{\partial t}, e_1, e_2)$ with (e_1, e_2) an orthonormal basis of TS^2:

$$A = g^{-1}\mathbb{I} = \begin{pmatrix} -F'/F & & \\ & -G'/G & \\ & & -G'/G \end{pmatrix}.$$

Applying the first formula of lemma 13.3, one obtains

$$K(e_1 \wedge e_2) = -\frac{1}{G^2} + \big(\frac{G'}{G^2}\big)^2, \quad K(\frac{\partial}{\partial t} \wedge e_i) = \frac{F'G'}{FG},$$

while using the second formula:

$$K(\frac{\partial}{\partial\rho} \wedge \frac{\partial}{\partial t}) = -\Big\langle(\frac{\partial A}{\partial\rho} - A^2)\frac{1}{F}\frac{\partial}{\partial t}, \frac{1}{F}\frac{\partial}{\partial t}\Big\rangle = \frac{F''}{F},$$

$$K(\frac{\partial}{\partial\rho} \wedge e_i) = \Big\langle(\frac{\partial A}{\partial\rho} - A^2)\frac{e_i}{G}, \frac{e_i}{G}\Big\rangle = \frac{G''}{G}.$$

Finally, using formula (III.31) with signs, that is in an orthonormal basis $\mathrm{Ric}(e_i, e_i) = \langle e_i, e_i\rangle \sum_j K(e_i \wedge e_j)$, we obtain the equalities

$$\mathrm{Ric}(\frac{1}{F}\frac{\partial}{\partial t}, \frac{1}{F}\frac{\partial}{\partial t}) = \frac{F''}{F} + 2\frac{F'G'}{FG} = \frac{(F'G^2)'}{FG^2} \tag{IV.10}$$

$$\mathrm{Ric}(\frac{\partial}{\partial\rho}, \frac{\partial}{\partial\rho}) = -\big(\frac{F''}{F} + 2\frac{G''}{G}\big) \tag{IV.11}$$

$$\mathrm{Ric}(\frac{e_i}{G}, \frac{e_i}{G}) = -\big(\frac{F'G'}{FG} + \frac{G''}{G} - \frac{1}{G^2} + (\frac{G'}{G})^2\big). \tag{IV.12}$$

This finishes the calculation of Ric since this is a basis of eigenvectors.

Now the system Ric = 0 is easily solved: the first equation implies that $F'G^2$ is a constant, say

$$F'G^2 = m.$$

The two first equations together are $\frac{F'G'}{FG} - G''G = 0$, that is $(\frac{F'}{G})' = 0$ so $\frac{F'}{G}$ is a constant, which we can take to be equal to +1, up to multiplying F and the variable ρ by a constant:

$$\frac{G'}{F} = 1$$

and therefore $G'' = F' = \frac{m}{G^2}$, and from the third equation follows $\frac{2m}{G} - 1 + F^2 = 0$, that is $F = \sqrt{1 - \frac{2m}{G}}$. Finally the metric can be written

$$g = (1 - \frac{2m}{G})dt^2 - d\rho^2 - G^2 ds^2_{S^2}.$$

It can seem here that we have a lot of different solutions, corresponding to the different functions G. Actually this is not the case: we can change coordinate by taking $r = G(\rho)$ and we obtain the metric

$$g = (1 - \tfrac{2m}{r})dt^2 - \frac{dr^2}{1 - \frac{2m}{r}} - r^2 ds^2 \tag{IV.13}$$

which depends only on one parameter $m > 0$. This is the Schwarzschild metric.
Remark that when $r \longrightarrow \infty$ then $g \sim dt^2 - dr^2 - r^2 ds^2$, so at large distance the Schwarzschild metric is asymptotic to the flat Minkowski metric on $\mathbb{R}^{1,3}$. The Schwarzschild is a model for the gravitation field of one central star, and $a = 2m$ is called Schwarzschild gravitation radius. For example, $a = 3\,\mathrm{km}$ for the sun and $a = 0,44\,\mathrm{cm}$ for the earth.

18.c Null geodesics

Here we will show that the light is deviated by the gravitation field—one of the first confirmations of general relativity. In general relativity, the light propagates along null geodesics, that is along geodesics c such that $\langle \dot c, \dot c \rangle = 0$. So we have to calculate the geodesics. Certainly, we can restrict to calculate the geodesics lying inside totally geodesic hypersurfaces of the form $\mathbb{R} \times \mathrm{I} \times \mathrm{S}^1$, where $\mathrm{S}^1 \subset \mathrm{S}^2$ is a great circle, on which we take an angular coordinate φ.
We have Killing fields $\frac{\partial}{\partial t}, \frac{\partial}{\partial \varphi}$, leading by lemma 9.8 to first integrals

$$\varepsilon = \langle \frac{\partial}{\partial t}, \dot c \rangle = (1 - \frac{2m}{r})\dot t,$$
$$\mu = -\langle \frac{\partial}{\partial \varphi}, \dot c \rangle = r^2 \dot\varphi.$$

Moreover for a light geodesic we have

$$(1 - \tfrac{2m}{r})\dot t^2 - \frac{1}{1 - \frac{2m}{r}}\dot r^2 - r^2 \dot\varphi^2 = 0.$$

Replacing $\dot t$ and $\dot\varphi$ in this equation by their values in terms of ε and μ, one obtains

$$\dot r^2 = \varepsilon^2 - \frac{1 - \frac{2m}{r}}{r^2}\mu^2.$$

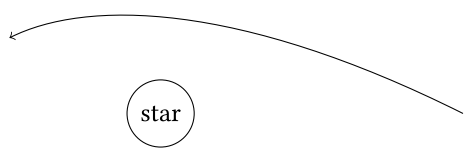
Since $\dot r^2 \geqslant 0$, one must have

$$\frac{1 - \frac{2m}{r}}{r^2} \leqslant \frac{\varepsilon^2}{\mu^2}.$$

The maximum value of the LHS is $\frac{1}{27m^2}$ at $r = 3m$. Then there are three cases:

1. $\frac{\varepsilon^2}{\mu^2} = \frac{1}{27m^2}$, then $r = 3m$, the light ray remains inside the 'photonic sphere' $r = 3m$;
2. $\frac{\varepsilon^2}{\mu^2} < \frac{1}{27m^2}$, then the geodesic remains outside or inside the photonic sphere;
3. $\frac{\varepsilon^2}{\mu^2} > \frac{1}{27m^2}$, then $\dot r^2 > 0$ so r is monotone and the ray goes to the star or comes from the star.

The deviation of light rays we are interested in is the second case.



We define $\varphi_c = \varphi(+\infty) - \varphi(-\infty)$. For a straight line, $\varphi_c = \pi$. We want to prove that there is a deviation, that is $\varphi_c > \pi$. One has $\frac{dr}{d\varphi} = \frac{\dot r}{\dot\varphi} = \frac{\dot r r^2}{\mu}$, and

$$\left(\frac{dr}{d\varphi}\right)^2 = \frac{r^4}{\mu^2}\left(\varepsilon^2 - \frac{1-\frac{2m}{r}}{r^2}\mu^2\right)$$
$$= \frac{m^2}{v^2\mu^2}\left(\varepsilon^2\frac{m^2}{v^2} - (1-2v)\mu^2\right) \qquad \text{with } v = \frac{m}{r}.$$

It follows that

$$\left(\frac{dv}{d\varphi}\right)^2 = \frac{m^2\varepsilon^2}{\mu^2} - v^2 + 2v^3.$$

The variable v goes from 0 to a maximum value at, say, v_0, such that $v_0 - 2v_0^3 = \frac{m^2\varepsilon^2}{\mu^2}$, and then decreases back to 0; we can rewrite

$$\left(\frac{dv}{d\varphi}\right)^2 = v_0^2 - 2v_0^3 - v^2 + 2v^3,$$

and

$$\varphi_c = 2\int_0^{v_0}\frac{d\varphi}{dv}dv = 2\int_0^{v_0}\frac{dv}{\sqrt{v_0^2-v^2}\sqrt{1-2\frac{v_0^3-v^3}{v_0^2-v^2}}}$$
$$> 2\int_0^{v_0}\frac{dv}{\sqrt{v_0^2-v^2}}\left(1+\frac{v_0^3-v^3}{v_0^2-v^2}\right) = \pi + 4v_0.$$

This indeed proves that $\varphi_c > \pi$, so there is a deviation. □

For other applications of the Schwarzschild metric to general relativity, see for example the book [Bes87].

Chapter V

Some exercises

Complex projective space

Let us see the unit sphere $S^{2n+1} \subset \mathbb{C}^{n+1}$ equipped with the standard Hermitian product $\langle z, t\rangle = \sum \overline{z^i} t^i$. We define a vector field T and a 1-form η on S^{2n+1} by: for any point $z \in S^{2n+1}$, one takes $T_z = iz \in T_z S^{2n+1}$ and $\eta_z(V) = \langle T, V\rangle$ for any $V \in T_z S^{2n+1}$ (why is this real ?).

1° What is the flow of the vector field T ? Prove that $\mathscr{L}_T \eta = \iota_T d\eta = 0$.

2° Let $p : S^{2n+1} \longrightarrow \mathbb{C}P^n$ be the natural projection. Prove that there is a unique 2-form Ω on $\mathbb{C}P^n$ such that $d\eta = p^*\Omega$. Prove that Ω is symplectic, that is Ω is closed and nondegenerate, cf. example 6.8, 3°. The form Ω is known as the ***Fubiny-Study form***.

3° Deduce that Ω^k defines a nonzero cohomology class in $H^{2k}(\mathbb{C}P^n)$ for $k \leqslant n$. Actually one can show that this gives the whole cohomology, $b_{2k}(\mathbb{C}P^n) = 1$ for $k \leqslant n$ and $b_{2k+1}(\mathbb{C}P^n) = 0$.

Isotopies

1° Let M be a manifold and $(\psi_t)_{0\leqslant t\leqslant 1}$ denote a smooth one-parameter family of diffeomorphisms ($(t, x) \longmapsto \psi_t(x)$ is smooth). Prove that for any form α, one has $\frac{d}{dt}\psi_t^*\alpha = \psi_t^*\mathscr{L}_{X_t}\alpha$, where X_t is the time dependent vector field defined by $X_t = \frac{d\psi_t}{dt}\circ\psi_t^{-1}$. Prove that if α is closed, then

$$\psi_1^*\alpha - \psi_0^*\alpha = d\left(\int_0^1 \psi_t^*(\iota_{X_t}\alpha)dt\right).$$

2° Prove the Poincaré lemma: any closed form of positive degree is locally exact. *Hint:* use 1° on a small ball of $\mathbb{R}^n$, with $\psi_t(x) = tx$ and check that the apparent problem at $t = 0$ is irrelevant.

3° Prove the Hairy Ball theorem: there is no nonvanishing vector field on the sphere S^{2n}. *Hint:* if such a vector field X existed, you could associate to any point $x \in S^{2n}$ the point $\psi_t(x)$ obtained by rotating x of an angle πt in the plane containing x and $X(x)$; deduce that the standard volume form of S^{2n} would be preserved by the antipodal map. (Observe the previous exercise yields a nonvanishing vector field on the sphere S^{2n+1}).

4° Prove the symplectic Darboux theorem: any symplectic form is locally isomorphic to the standard symplectic form

$$\Omega_0 = \sum_1^n dx^i \wedge dx^{i+n}$$

on $\mathbb{R}^{2n}$, namely $\Omega = \psi^*\Omega_0$ for some local diffeomorphism ψ. To prove this, you may work in coordinates, on a small ball around 0 in $\mathbb{R}^{2n}$, assume Ω and Ω_0 coincide at the origin and consider $\Omega_t = t\Omega + (1-t)\Omega_0$. The trick (known as ***Moser's trick***) consists in finding a family of diffeomorphisms ψ_t such that ψ_0 is the identity and $\psi_t^*\Omega_t$ is constant in time, so that ψ_1 will do the job. In order to build this family of diffeomorphisms, you will first find the vector fields X_t by computing the derivative of $\psi_t^*\Omega_t$ and then integrate the time-dependent ordinary differential equation $\frac{d\psi_t}{dt} = X_t \circ \psi_t$.

5° A distribution D of rank $2n$ on a manifold M^{2n+1} is called a ***contact distribution*** if it can be written locally as $D = \ker\alpha$ some one-form α such that $\alpha \wedge (d\alpha)^n$ does not vanish; it is equivalent to require that the restriction of $d\alpha$ on D is non-degenerate. Prove the contact Darboux theorem: any contact distribution D on M^{2n+1} is locally isomorphic to the standard contact structure on $\mathbb{R}^{2n+1}$, given by the kernel of

$$\alpha_0 = dx^{2n+1} - \sum_1^n x^i dx^{i+n}.$$

Hint: Moser's trick.

Gauss map

Let Σ be a connected compact surface embedded in $\mathbb{R}^3$, with sectional curvature $K > 0$. We do not suppose Σ orientable.

1° Prove that the second fundamental form of Σ is definite positive or definite negative at every point.

2° For every $x \in \Sigma$ we denote π the orthogonal projection $\mathbb{R}^3 \longrightarrow (T_x\Sigma)^\perp$. Prove that the vector $\nu_x = -\frac{\pi(\nabla_X X)}{|\pi(\nabla_X X)|}$ does not depend on the non zero vector $X \in T_x\Sigma$. Prove that Σ is orientable.

3° The application $\Gamma : \Sigma \longrightarrow S^2$, given by $\Gamma(x) = \nu_x \in \mathbb{R}^3$, is called the ***Gauss map***. Prove that Γ is a diffeomorphism. Calculate Γ in case $\Sigma = S^2$.

4° Prove that $\Gamma^* \mathrm{vol}^{S^2} = K\,\mathrm{vol}^\Sigma$ (how general is this formula ?) Deduce in that case a proof of the Gauss-Bonnet formula,

$$\frac{1}{2\pi}\int_\Sigma K\,\mathrm{vol}^\Sigma = 2.$$

Umbilic submanifolds

1° Let N^{n-1} be a submanifold of (M^n, g), oriented by the normal $\vec{n}$, with second fundamental form $\mathbb{I}$. Prove that, for all vectors X, Y and Z of N one has

$$\langle R^M_{X,Y}Z, \vec{n}\rangle = \nabla_X \mathbb{I}(Y,Z) - \nabla_Y \mathbb{I}(X,Z). \tag{V.1}$$

(The RHS is $(d^\nabla \mathbb{I})(X,Y,Z)$ if one considers $\mathbb{I}$ as a TN-valued 1-form on N, and this formula is the Gauss-Codazzi equation, see also remark 18.3.)

2° The submanifold N is said ***totally umbilic*** if for every $x \in N$ the second fundamental form is a multiple of the metric: $\mathbb{I} = \lambda(x) g$. If N is a totally umbilic submanifold of the flat $\mathbb{R}^n$, show that λ is a constant function. Deduce that the application $\phi : N \longrightarrow S^{n-1}$ given by $\phi(x) = \frac{\vec{n}_x}{\lambda}$ is an isometry, and that N is a sphere (or an open set of a sphere).

3° Find all submanifolds $N \subset \mathbb{R}^n$ with constant positive sectional curvature.

Submanifolds of the hyperbolic space

We choose coordinates in $\mathbb{R}^{1,n}$ such that the quadratic form is

$$h(x,x) = 2x^0x^1 - (x^2)^2 - \cdots - (x^n)^2.$$

The hyperbolic space is the submanifold $H^n = \{h(x,x) = 1, x^0 > 0\}$ with the induced metric $g = -h$.

1° Choose on H^n coordinates $(y^1, \dots, y^n) \in \mathbb{R}_+^* \times \mathbb{R}^{n-1}$ by taking $(x^1 = 1/y^1, x^2 = y^2/y^1, \dots, x^n = y^n/y^1)$. Prove that

$$g = \frac{(dy^1)^2 + (dy^2)^2 + \cdots + (dy^n)^2}{(y^1)^2}$$

(half-space model).

2° Prove that for $\lambda > 0$ the homothety

$$(y^1, \dots, y^n) \longrightarrow (\lambda y^1, \dots, \lambda y^n)$$

and the inversion

$$(y^1, \dots, y^n) \longrightarrow \frac{(y^1, \dots, y^n)}{(y^1)^2 + \cdots + (y^n)^2}$$

are isometries of H^n.

Prove that the intersection with $H^n = \{y^1 > 0\}$ of a sphere of $\mathbb{R}^n$ centered on the hyperplace $\{y^1 = 0\}$ is totally geodesic.

3° If g_2 and g_1 are Riemannian metrics on a manifold M^n such that $g_2 = e^{2f} g_1$, where f is a function, prove that their Levi-Civita connections satisfy

$$g_1(\nabla^{g_2}_X Y, Z) = g_1(\nabla^{g_1}_X Y, Z) + df(X) g_1(Y,Z) + df(Y) g_1(X,Z) - g_1(X,Y) df(Z).$$

Using the previous exercise, deduce that a submanifold $N^{n-1} \subset M$ is totally umbilic for g_1 if and only if it is totally umbilic for g_2. Deduce all totally umbilic submanifolds of H^n in the half-space model.

Toral black hole Einstein metrics

Let $M = I \times S^1 \times T^{n-2}$ the product of an interval I of $\mathbb{R}$, a circle S^1 and a $(n-2)$-dimensional torus T^{n-2}.

1° Find all metrics g on M, satisfying the Einstein equation

$$\mathrm{Ric}(g) = -(n-1)g,$$

of the form $g = \phi(r)dr^2 + \psi(r)d\theta^2 + r^2 g_T$, where r and θ are coordinates on I and S^1, g_T is a flat metric on the torus T, and ϕ and ψ are functions of r only.

Hint: first, look for a metric of the form $d\rho^2 + F^2(\rho)d\theta^2 + G^2(\rho)g_T$.

2° These metrics are defined for an interval of the form $I = (r_+, +\infty)$. Which metric do we get when $r_+ = 0$? prove that it is complete, with constant sectional curvature equal to -1 (real hyperbolic cusp).

3° Prove that in some cases, one can add in $r = r_+$ a torus T^{n-2}, so that the metric extends to a complete smooth Einstein metric on $\bar{M} = M \cup T^{n-2}$.

Bibliography

[Bes87] Arthur L. Besse. *Einstein manifolds.* Springer-Verlag, Berlin, 1987.

[BGV04] Nicole Berline, Ezra Getzler, and Michèle Vergne. *Heat kernels and Dirac operators.* Grundlehren Text Editions. Springer-Verlag, Berlin, 2004. Corrected reprint of the 1992 original.

[GHL04] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine. *Riemannian geometry.* Universitext. Springer-Verlag, Berlin, third edition, 2004.

[Kob95] Shoshichi Kobayashi. *Transformation groups in differential geometry.* Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1972 edition.

[Lee13] John M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics.* New York, NY: Springer, 2nd revised edition, 2013.

Index