Differentialgeometry I - Vorlesungs-Script

Prof. Demetrios Christodoulou

19. Oktober 2008

Mitschrift:
Raphael Honegger

Id

Inhaltsverzeichnis

0 Introduction

0.1 Literature

- "Geometry of Manifolds" by Bishop and Crittenden
- "Comprehensive Introduction to Differential Geometry" by Spivak (5 volumes)

0.2 Topics to be covered in this lecture

- Differentiable manifolds
- The tangent space
- The tangent bundle
- Vectorfields
- 1-parameter groups of diffeomorphisms
- Exterior differential forms, exterior derivative
- Pull-back and push-forward
- Lie derivatives

And if there's time left

- Vector bundles
- Parallel transport
- Connection and covariant differentiation
- Curvature

0.3 Colors

Besides the standard color black, you can find sections in this script, which are colored in red or blue. The colors have the following meaning:

- Black: Sections in the standard color black have been revised by me.
- Blue: Blue sections haven't been revised by me, yet.
- Red: Such sections I haven't understood by the time I revised them. If somebody has understood a red section, he's cordially invited to explain or tell me the appropriate matters.

0.4 History

The first time, something was published about Differential Geometry was 1828 by Gauss. This work was called "General Investigations of Curved Surfaces". Gauss considered a general smooth surface in the Euclidean 3-dimensional space in itself an intrinsic geometry. 1854 Riemann published the paper "On the hypotheses which lie at the foundation of geometry". It is in this work, that the first time somebody talks about a geometry of n-dimensional manifolds. Only in the 1930's, the modern notion of a differentiable manifold was finally established.

Note: From Riemann's work onwards, manifolds were considered in and by themselves and not imbedded in the Euclidean space.

1956 John Nash demonstrated that any n-dimensional Rimannian manifold can be isometrically imbedded in an N-dimensional Euclidean space for some suitably large N (that is so that the geomtry in the Euclidean space is the same as given a priori).

Remarks on the development from 1954 to the 1930's: Riemann himself did not discuss at all parallel transport.

Of course he introduced the Rimannian metric an its notion and also introduced the curvature as the abstraction to the existance of a system of local coordinates where the metric takes the standard Euclidean form in rectangular coordinates (an abstraction prevents you from doing something). Covariant differentiation began to be investigated by Christofel, further by Ricci, who introduced the concept of Tensorfields. Finally Levi-Civita understood parallel transport and the notion of connection in the tangent bundle. Simultaneously the general theory of relativity was developed by Einstein. Then, Cartan generalized the theory of connections of general vector bundles. This found application in the gauge theories developed in Physics since the 1950's.

1 Differential Manifolds

1.1 Definitions

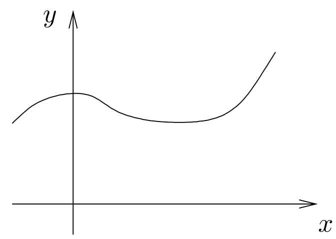
There are two basic notions

- Continuity
- Differentiability

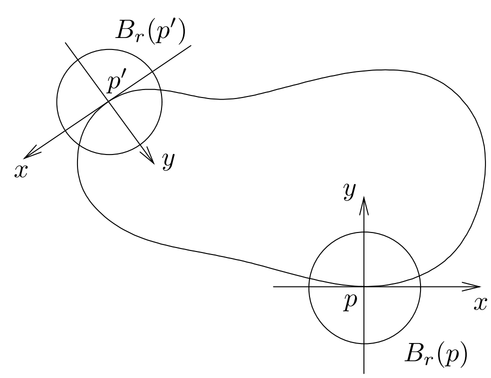
Correspondingly there are two notions of manifolds

- Topological manifolds
- Differentiable manifolds

We begin with the notion of an m-dimensional submanifold of the Euclidean n-dimensional space ($n > m$). The simplest example is the curve on a plane



that is the graph $y = f(x)$ of a function, where f is a continuous curve and differentiable for a differentiable curve. But this definition is two restrictive, because we could also have a curve of the form



If we take any point p on the curve, we can define $B_r(p)$ as the open disk with center at the point $p \in K$ and of radius r, where $r > 0$. Now we consider the part of K which lies in $B_r(p)$, that is $K \cap B_r(p)$. Choosing r suitably small, there should be a system of rectangular axes with origin at p, such that $K \cap B_r(p)$ is in fact the graph of such a continuous or differentiable function, as required.

Definition 1.1: A *continuous curve* on the plane is a subset K of the plane with the following property: For each point $p \in K$ there is an $r > 0$ such that $K \cap B_r(p)$ is represented as the graph $y = f(x)$ of a continuous function f relative to a rectangular system of axes x, y with origin at p.

A *differentiable curve* is defined in a similar way, replacing the requirement of continuity f by the requirement of differentiability. □

These definitions can by immediately generalized to the case of an m-dimensional submanifolds of the n-dimensional Euclidean space, where $n > m$.

Definition 1.2: A *continuous/differentiable m-dimensional submanifold* of the Euclidean n-dimensional space E^n is a subset $K \in E^n$ with the following property: For each point $p \in K$ there is an $r > 0$ such that $K \cap B_r(p)$ can be represented as the graph

$$x^i = f^i(x^1, ..., x^m) \quad : \quad i = m+1, ..., n$$

where the functions f^i, $i = m+1, ..., n$ are all continuous/differentiable. That is, the $n-m$ remaining coordinates are functions of the first m coordinates. □

1.2 Topological Manifolds

We begin with a review of basic notions of topology.

Definition 1.3: A *topological space* is a set X together with a set τ of subsets of X, such that

(1) $\varnothing, X \in \tau$

(2) Arbitrary unions of members of τ are themselves members of τ.

(3) Finite intersections of members of τ are themselves members of τ.

τ is called the *topology* of X. The members of τ are called *open sets*. A *neighborhood* of a point $p \in X$ is an open set containing p. A *closed set* is a set whose complement is open. □

Definition 1.4: A *distance function* d on a set X is a real valued function on $X \times X$ such that

(1) it is symmetric, that is

$$d(y,x) \;=\; d(x,y) \quad \forall x, y \in X$$

(2) it is positive definit with equality if and only if $x = y$, so

$$d(x,y) \;\geq\; 0,\;\; d(x,y) \;=\; 0 \;\Leftrightarrow\; x \;=\; y \quad \forall x, y \in X$$

(3) it satisfies the triangle inequality

$$d(x,z) \;\leq\; d(x,y) + d(y,z) \quad \forall x, y, z \in X$$

□

Definition 1.5: Given a distance function d on X we define the *open ball* $B_r(p)$ with center at $p \in X$ of radius $r > 0$ by

$$B_r(p) \;:=\; \{x \in X : d(x,p) < r\}$$

□

Definition 1.6: Given a distance function d on a set X, we define a topology τ as follows: It is $U \in \tau$ according as to whether

$$p \in U \quad \Rightarrow \quad \exists r > 0 : B_r(p) \subset U$$

This topology is called the *metric topology* induced by d. A topology is called *metric* if it is induced by some distance function. □

Example 1.1: Metric Spaces

1. In the Euclidean n-dimensional space $\mathbb{R}^n$, a point $x, y \in \mathbb{R}^n$ are n-tuplets $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ with the distance function

$$d(x,y) \;=\; \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

2. For any set X and $x, y \in X$ define the distance function as

$$d(x,y) \;=\; \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \end{cases}$$

Then we get that

$$B_r(x) \;=\; \begin{cases} \{x\} & \text{if } r \leq 1 \\ X & \text{if } r > 1 \end{cases}$$

It follows that any subset of X is open. The topology τ induced by this distance function is the so called *power set* of X, namely the set of all subsets of X.

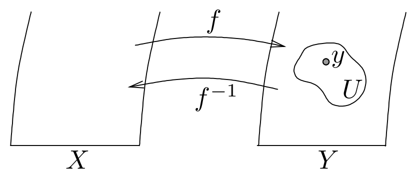
□

1.3 Continuity

Let us recall the notion of continuity.

Definition 1.7: Let X and Y be topological spaces and $f : X \to Y$ a mapping of X into Y. f is called *continuous* if the inverse image in X of each open set in Y is an open set in X, that is

$$U \subset Y, U \text{ open} \quad \Rightarrow \quad f^{-1}(U) \subset X, f^{-1}(U) \text{ open}$$



□

Definition 1.8: Let X and Y be metric spaces. A function f is *continuous* at x, if $\forall \varepsilon > 0 \; \exists \delta > 0$ such that

$$f(B_\delta) \subset B_\varepsilon(f(x))$$

If $y = f(x)$ this condition reads

$$d_X(x', x) < \delta \quad \Rightarrow \quad d_Y(y', y) < \varepsilon$$

where $y' = f(x')$. The function f is called *continuous* if and only if it is continuous at all $x \in X$. □

Remark 1.1: For a continuous function f, let be $U = f^{-1}(V)$. Then, if V is open, U is also open. Take an $x \in U$, then $f(x) = y$ for some $y \in V$ and there is a suitable $\varepsilon > 0$ to satisfy $B_\varepsilon(y) \subset V$, since V is open. By the second definition, there is a $\delta > 0$ such that

$$f(B_\delta(x)) \subset B_\varepsilon(y) \subset V$$

It follows that $B_\delta(x)$ is contained in the inverse image of V, namely in U. Thus, $B_\delta(x) \subset U$ hence U is open. □

1.4 Homeomorphism

Definition 1.9: Let X and Y be topological spaces and $f : X \to Y$ a mapping. f is a *homeomorphism* of X onto Y if

(1) f is one-to-one and onto (that is injective and surjective). Thus f^{-1} is a mapping of Y into X.

(2) f as well as f^{-1} are continuous maps.

Two topological spaces X, Y are called *homeomorphic* if there exists a mapping $f : X \to Y$ which is a homeomorphism. □

Example 1.2: The unit ball with center at the origin in $\mathbb{R}^n$, that is $B_1(0)$, is homeomorphic to $\mathbb{R}^n$ itself. The homeomorphism $f : B_1(0) \to \mathbb{R}^n$ may be defined as follows: For $x \in B_1(0)$ set

$$|x| = d(x, 0) = \sqrt{\sum_{i=1}^{n} x_i^2} < 1$$

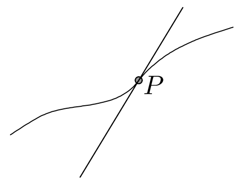
Now we define the function and give its inverse by

$$f(x) := \frac{x}{\sqrt{1 - |x|^2}} \qquad f^{-1}(y) = \frac{y}{\sqrt{1 + |y|^2}}$$

where $y \in \mathbb{R}^n$. The function f is a homeomorphism. □

1.5 Differentiability

This concept requires a linear structure



The tangent line to a plane curve at a point P is the closest approximation by a straight line to the curve in a neighborhood of P.

Definition 1.10: A set V is a real *vector space* and its elements are called *vectors* if it is endowed with the following structure:

(1) It has a vector addition

$$v_1 + v_2 \in V \quad \forall v_1, v_2 \in V$$

which is

(i) associative

$$v_1 + (v_2 + v_3) \;=\; (v_1 + v_2) + v_3 \quad \forall v_1, v_2, v_3 \in V$$

(ii) commutative

$$v_1 + v_2 \;=\; v_2 + v_1 \quad \forall v_1, v_2 \in V$$

(iii) there is a distinguished element $0 \in V$ called the 0 vector such that

$$v + 0 \;=\; v \quad \forall v \in V$$

(2) It has a scalar multiplication

$$\alpha v \in V \quad \forall \alpha \in \mathbb{R}, v \in V$$

such that

(i) $0v = 0 \; \forall v \in V$

(ii) $1v = v \; \forall v \in V$

(iii) $(\alpha\beta)v = \alpha(\beta v) \; \forall \alpha, \beta \in \mathbb{R}, \forall v \in V$

(3) Moreover we have the distributivity properties

(i) $\alpha(v_1 + v_2) = \alpha v_1 + \alpha v_2$

(ii) $(\alpha + \beta)v = \alpha v + \beta v$

□

Remark 1.2: If in (3ii) we take $\alpha = 1$, $\beta = -1$, we get

$$0 \;=\; (1 + (-1))v \;=\; 1v + (-1)v \;=\; v + (-v) \;=\; 0$$

where we applied $\alpha + \beta = 0$ using (2i) and denoted the vector $(-1)v$ simply by $-v$. Thus there is an additive inverse to every vector v, the vector $-v$. □

Example 1.3: The space of continuous real valued functions on E^n is a vector space with vector addition

$$(f + g)(x) \;=\; f(x) + g(x) \quad \forall x \in E^n$$

It satisfies the requirements (1) with the zero function f defined by

$$f(x) \;:=\; 0 \quad \forall x \in E^n$$

And it has the scalar multiplication

$$(\alpha f)(x) \;=\; \alpha f(x) \quad \forall x \in E^n$$

which as well satisfies (2) and (3). □

Definition 1.11: A vector space V is said to be of *dimension* n, and thus *finite dimensional*, if there is a set $(e_1, ..., e_n)$ of n linearly independent vectors, and every set of $n+1$ vectors is linearly dependent. Such a set is called a *basis* for V. Given such a basis, any vector $v \in V$ can be expanded as

$$v = \sum_{i=1}^{n} v_i e_i \quad v_1, ..., v_n \in \mathbb{R}$$

uniquely in the given basis. Thus, given a basis $(e_1, ..., e_n)$, we have a linear isomorphism of V onto $\mathbb{R}^n$ given by

$$V \ni v \quad \mapsto \quad (v_1, ..., v_n) \in \mathbb{R}^n$$

The n-tuplet $(v_1, ..., v_n)$ are the *components* of v in the given basis. □

Definition 1.12: On a vector space V we can define a *norm* $\| \cdot \|$. For $v \in V$, $\|v\|$ is then the *magnitude* of the vector v. The norm is a real valued function on V having the following properties:

1) $\|v\| \geq 0$ with equality if and only if $v = 0$ (the zero vector).
2) $\|\alpha v\| = |\alpha| \|v\|$ (where $|\alpha|$ is the absolute value of the real number α).
3) $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$, which is called the *Minkowski inequality.*

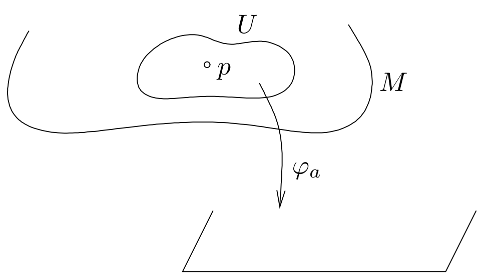
□

Remark 1.3: Given a norm $\| \cdot \|$ in a vector space, there is a distance function d defined by the norm

$$d(u, v) = \|u - v\|$$

where $u - v$ means $u + (-v)$. The Minkowski inequality for the norm then translates to the triangle inequality for d. □

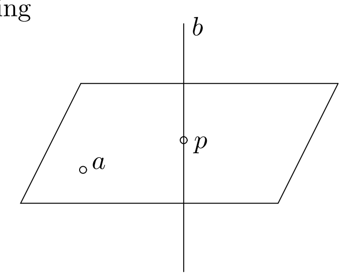
Definition 1.13: A *topological manifold* M is a metric topological space for which there is a positive integer n such that every point $p \in M$ has a neighborhood U homeomorphic to $\mathbb{R}^n$. Thus there is a homeomorphism φ of U onto $\mathbb{R}^n$. n is called the *dimension* of M.



□

Remark 1.4: Requiring U to be homeomorphic to $\mathbb{R}^n$ is equivalent to requiring it to be homeomorphic to an open subset V of $\mathbb{R}^n$, because we can always choose an open ball $B_\varepsilon(\varphi(p)) \subset V$, which is homeomorphic to $\mathbb{R}^n$, and take $\varphi^{-1}(B_\varepsilon(\varphi(p)))$ instead of U. □

Example 1.4: A space like the following



is not a topological manifold. The problem comes from the intersection point p. Points a and b would work fine. □

Definition 1.14: A sequence x_n, $n = 1, 2, 3, ...$ is called a *Cauchy sequence* if $\forall \varepsilon > 0$ $\exists N \in \mathbb{N}$ such that for

$$m, n \in \mathbb{N},\ m, n \geq N \quad \Rightarrow \quad d(x_n, x_m) \ < \ \varepsilon$$

$\square$

Definition 1.15: We say that x_n *converges* to x, or $x_n \to x$ if

$$d(x_n, x) \to 0 \text{ as } n \to \infty$$

$\square$

Definition 1.16: We say that a metric space is *complete* if every Cauchy sequence converges to some point in the space. $\square$

Definition 1.17: A normed linear space is called a *Banach space* if it is complete with respect to the distance function induced by its norm. $\square$

Example 1.5: For the vector space of continuous functions, we could construct the two norms

$$\|f(x)\|_1 \ := \ \sqrt{\int_{E^n} |f(x)|^2 \, \mathrm{d}x} \qquad \|f(x)\|_2 \ := \ \sup_{x \in E^n} |f(x)|$$

The first one would not make the vector space complete, but the second one would. $\square$

Proposition 1.1: A finite dimensional normed linear space is complete and therefore a Banach space. This is because such a space is isomorphic to $\mathbb{R}^n$ and moreover, all norms in finite dimensional linear spaces are equivalent. $\square$

Lemma 1: All norms in a finite dimensional linear space V are equivalent. $\square$

Proof: We claim that there is a positive constant c_M such that

$$\sum_{i=1}^{n} |v_i| \ \leq \ c_M \|v\|$$

To do so, suppose that no such constant c_M exists. That is suppose that there is no upper bound for the ratio

$$\frac{\sum_{i=1}^{n} |v_i|}{\|v\|}$$

for $v \in V$. Therefore, there has to be sequence $(v^{(m)},\ m = 1, 2, ...)$ of non-zero vectors such that

$$\sum_{i=1}^{n} |v_i^{(m)}| \ = \ 1 \quad m = 1, 2, ...$$

while $\|v^{(m)}\| \to 0$ as $m \to \infty$. For each $i = 1, ..., n$ the sequence $(v_i^{(m)},\ m = 1, 2, ...)$ is a numerical sequence contained in the closed interval $[-1, 1]$, because $|v_i^{(m)}| \leq 1$. We can thus apply the Bolzano-Weierstraß theorem. So there is a subsequence $(v^{(m_k)},\ k = 1, 2, ...)$ such that

$$v_i^{(m_k)} \ \overset{k \to \infty}{\longrightarrow} \ v_i^* \quad i = 1, ..., n$$

where $v_i^* \in [-1, 1]$. It follows that

$$1 \ = \ \sum_{i=1}^{n} |v_i^{(m_k)}| \ \overset{k \to \infty}{\longrightarrow} \ \sum_{i=1}^{n} |v_i^*|$$

Hence

$$\sum_{i=1}^{n} |v_i^*| \ = \ 1$$

Then

$$v^* - v^{(m_k)} \;=\; \sum_{i=1}^{n}\left(v_i^* - v_i^{(m_k)}\right)e_i \quad\Rightarrow\quad \|v^* - v^{(m_k)}\| \;\leq\; \sum_{i=1}^{n}|v_i^* - v_i^{(m_k)}| \quad \stackrel{k\to\infty}{\longrightarrow} \quad 0$$

By the Minkowski inequality and $v^* = (v^* - v^{(m_k)}) + v^{(m_k)}$ we get

$$\|v^*\| \;\leq\; \|v^* - v^{(m_k)}\| + \|v^{(m_k)}\| \quad \stackrel{k\to\infty}{\longrightarrow} \quad 0$$

It follows $\|v^*\| = 0$ and therefore $v^* = 0$ which contradicts the fact that

$$\sum_{i=1}^{n}|v_i^*| \;=\; 1$$

This stablishes the claim that

$$\sum_{i=1}^{n}|v_i| \;\leq\; c_M\|v\|$$

To show that we can also find a constant c_m to satisfy

$$\sum_{i=1}^{n}|v_i| \;\geq\; c_m\|v\|$$

recall the Minkowski inequality

$$\|v+u\| \;\leq\; \|v\| + \|u\|$$

Together with finite induction, this yields that if $v^{(1)}, ..., v^{(m)}$ are m vectors, then

$$\left\|\sum_{\alpha=1}^{m} v^{(\alpha)}\right\| \;\leq\; \sum_{\alpha=1}^{m}\|v^{(\alpha)}\|$$

Let us apply this to the n vectors $v_1e_1, ..., v_ne_n$

$$\|v\| \;\leq\; \sum_{i=1}^{n}\|v_ie_i\| \;=\; \sum_{i=1}^{n}|v_i|\,\|e_i\| \;=\; \sum_{i=1}^{n}|v_i|$$

So c_m can just be set to be one. (QED)

Proof of Proposition 1.1: Let V be an n-dimensional linear space and $(e_1, ..., e_n)$ a basis for V. We can assume that the basis is normalized, that is $\|e_i\| = 1$, $i = 1, ..., n$. Otherwise we set

$$\hat{e}_i \;=\; \frac{e_i}{\|e_i\|} \qquad i = 1, ..., n$$

to obtain a normalized basis. Any vector $v \in V$ is expanded as

$$v \;=\; \sum_{i=1}^{n} v_ie_i \qquad (v_1, ..., v_n) \in \mathbb{R}^n$$

in unique way. Thus we have an isomorphism to the components of v

$$V \ni v \;\;\mapsto\;\; (v_1, ..., v_n) \in \mathbb{R}^n$$

In $\mathbb{R}^n$ with $x = (x_1, ..., x_n) \in \mathbb{R}^n$ we have the Euclidean norm

$$|x| \;:=\; \sqrt{\sum_{i=1}^{n} x_i^2}$$

From Lemma 1 we know that in a finite dimensional linear space V, all norms are equivalent. It follows that there are positive constants c_m and c_M such that

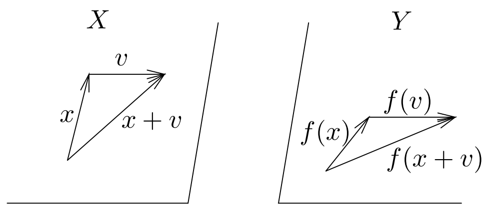
$$c_m|v| \;\leq\; \|v\| \;\leq\; c_M|v|$$

Given a Cauchy sequence in V, the corresponding sequence in components (relative to the basis $(e_1, ..., e_n)$) is a Cauchy sequence in $\mathbb{R}^n$, thus by the completence of $\mathbb{R}^n$ it converges to some n-tuplet in $\mathbb{R}^n$. This n-tuplet defines a limit vector in V and the original sequence must converge to this limit vector. (QED)

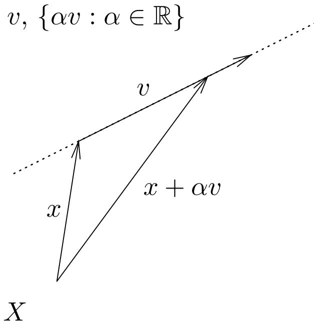
Definition 1.18: Let X, Y be normed linear spaces and f be a continuous mapping $f : X \to Y$. We say that f is *differentiable* at $x \in X$ if there is a linear map $Df(x) : X \to Y$ such that

$$f(x+v) - f(x) \;=\; Df(x) \cdot v + h_x(v) \qquad \frac{\|h_x(v)\|_Y}{\|v\|_X} \overset{v\to 0}{\longrightarrow} 0$$

where $Df(x) \cdot v$ is the notation for linear maps. Then $Df(x)$ is the *derivative* of f at x.



If we fix v and consider the linear span of v, $\{\alpha v : \alpha \in \mathbb{R}\}$



we get

$$f(x+\alpha v) - f(x) \;=\; \alpha\,(Df(x) \cdot v) + h_x(\alpha v) \qquad \frac{\|h_x(\alpha v)\|_Y}{\|\alpha v\|_X} \overset{\|\alpha v\|_X \to 0}{\longrightarrow} 0$$

that is

$$\frac{h_x(\alpha v)}{|\alpha|} \overset{\alpha\to 0}{\longrightarrow} 0$$

□

2 Differentiable Manifolds

2.1 Definition

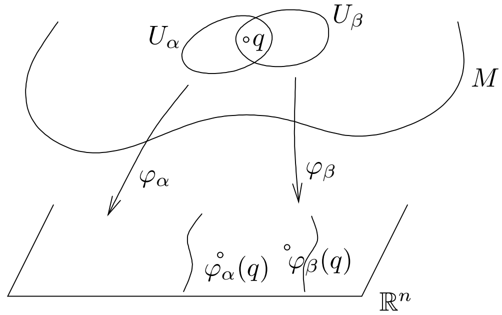
From the definition of topological manifolds, we have a collection of open sets

$$\mathcal{A} \quad := \quad \{U_\alpha : \alpha \in I\}$$

which cover the topological manifold M, where I is the *indexing set* which is completely arbitrary. That is, for each $p \in M$ there is some $U_\alpha \in \mathcal{A}$ such that $p \in U_\alpha$. Each U_α is paired with a homeomorphism φ_α of U_α onto $\mathbb{R}^n$. We therefore write

$$\mathcal{A} \quad = \quad \{(U_\alpha, \varphi_\alpha) : \alpha \in I\}$$

Consider a pair of indices α, β such that $U_\alpha \cap U_\beta \neq \varnothing$.



Since the restriction of a homeomorphism is itself a homeomorphism (onto its image), $\varphi_\alpha|_{U_\alpha \cap U_\beta}$ is a homeomorphism of the open set $U_\alpha \cap U_\beta \in M$ onto its image $U_{\alpha,\beta} := \varphi_\alpha(U_\alpha \cap U_\beta)$, which is an open set in $\mathbb{R}^n$. Similarly, $\varphi_\beta|_{U_\alpha \cap U_\beta}$ is a homeomorphism of $U_\alpha \cap U_\beta$ onto its image $U_{\beta,\alpha} := \varphi_\beta(U_\alpha \cap U_\beta)$ and this is itself an open set in $\mathbb{R}^n$. Therefore, the mapping

$$\varphi_{\alpha,\beta} \quad = \quad \varphi_\beta \circ \varphi_\alpha^{-1} : U_{\alpha,\beta} \to U_{\beta,\alpha}$$

is a homeomorphism of an open set in $\mathbb{R}^n$ onto an other open set in $\mathbb{R}^n$. We can require this mapping to be a continuously differentiable mapping in the standard sense of $\mathbb{R}^n$. Interchanging the roles of α and β and requiring the mapping

$$\varphi_{\beta,\alpha} \quad = \quad \varphi_\alpha \circ \varphi_\beta^{-1} : U_{\beta,\alpha} \to U_{\alpha,\beta}$$

to be also continuously differentiable, we have that

$$\varphi_{\alpha,\beta}^{-1} \quad = \quad \varphi_{\beta,\alpha}$$

is continuously differentiable as well, so in fact both maps are diffeomorphisms.

Definition 2.1: Let M be a metric topological space. The collection

$$\mathcal{A} \quad = \quad \{(U_\alpha, \varphi_\alpha) : \alpha \in I\}$$

is an *atlas* for M, if for each pair of indices $\alpha, \beta \in I$ the mapping

$$\varphi_{\alpha,\beta} : U_{\alpha,\beta} \to U_{\beta,\alpha}$$

is continuously differentiable in the standard sense of $\mathbb{R}^n$, where $U_{\alpha,\beta}, U_{\beta,\alpha} \subset \mathbb{R}^n$ are open. M together with such an atlas $\mathcal{A}$ is a *differentiable manifold* $M_{\mathcal{A}}$. Each pair $(U_\alpha, \varphi_\alpha)$ is called a *chart*. The U_α is called the *domain* of the chart. □

Alternatively, we could define a differentiable manifold like this:

Definition 2.2: A *differentiable manifold* is a topological manifold, whose *transition maps*

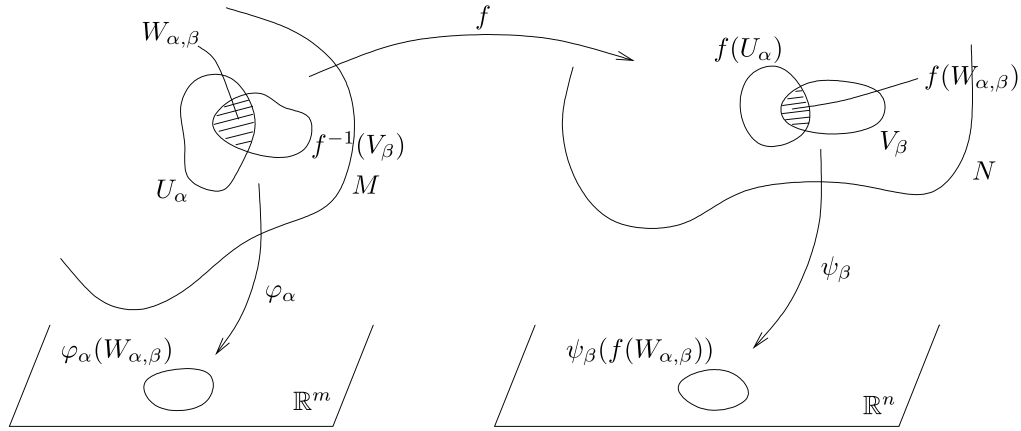
$$\varphi_{\alpha,\beta} \quad = \quad \varphi_\beta \circ \varphi_\alpha^{-1} : U_{\alpha,\beta} \to U_{\beta,\alpha} \quad \alpha, \beta \in I$$

are all differentiable. □

2.2 Differentiable Functions

Definition 2.3: Let M be a differentiable manifold. A continuous function f on M is *continuously differentiable*, if $\forall \alpha \in I$ the function $f \circ \varphi_\alpha^{-1}$ is a continuously differentiable function on $\mathbb{R}^n$.

Let M and N be differentiable manifolds with $\dim M = m$ and $\dim N = n$. Let $\mathcal{A} = \{(U_\alpha, \varphi_\alpha) : \alpha \in I\}$ be the atlas of M and $\mathcal{B} = \{(V_\beta, \psi_\beta) : \beta \in J\}$ be the atlas of N. Consider a continuous mapping $f : M \to N$.



Let be $f(U_\alpha) \cap V_\beta \neq \varnothing$ for some pair α, β, $\alpha \in I$, $\beta \in J$. f being a continuous mapping and V_β being an open set in N, $f^{-1}(V_\beta)$ is an open set in M. Thus $W_{\alpha,\beta} := U_\alpha \cap f^{-1}(V_\beta)$ is a non-empty open set in M. $\varphi_\alpha|_{W_{\alpha,\beta}}$ is a homeomorphism of $W_{\alpha,\beta}$ onto its image, an open set in $\mathbb{R}^m$. We say that the mapping f is *continuously differentiable*, if for every such pair of indices α, β

$$\psi_\beta \circ f \circ \varphi_\alpha^{-1} : \varphi_\alpha(W_{\alpha,\beta}) \to \psi(f(W_{\alpha,\beta}))$$

is a continuously differentiable mapping of a domain in $\mathbb{R}^m$ (namely the domain $\varphi_\alpha(W_{\alpha,\beta})$) into $\mathbb{R}^n$. □

Definition 2.4: A *diffeomorphism* of a differentiable manifold M onto a differentiable manifold N is a homeomorphism f of M onto N such that both f and f^{-1} are continuously differentiable mappings.

Two differentiable manifolds M, N are called *diffeomorphic* if there exists a diffeomorphism f of M onto N. □

Definition 2.5: Let M be a topological manifold and let

$$\mathcal{A} = \{(U_\alpha, \varphi_\alpha : \alpha \in I\} \qquad \mathcal{B} = \{(V_\beta, \varphi_\beta) : \beta \in J\}$$

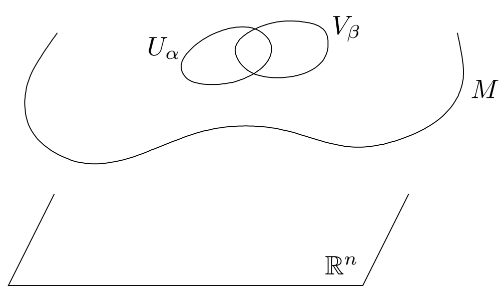
be two atlases for M. We write $M_\mathcal{A}$ and $M_\mathcal{B}$ for the corresponding differentiable manifold. We say that the two atlases are *compatible* if the identity mapping

$$id : M_\mathcal{A} \to M_\mathcal{B} \quad : \quad id(p) \;=\; p \quad \forall p \in M$$

is a diffeomorphism. □

Proposition 2.1: An atlas $\mathcal{A}$ is compatible with $\mathcal{B}$, if and only if $\mathcal{A} \cup \mathcal{B}$ is also an atlas for M, where $\mathcal{A} \cup \mathcal{B}$ consists of all the charts of $\mathcal{A}$ together with al the charts of $\mathcal{B}$. □

Proof: Obviously the identity is a homeomorphism. Since also $id^{-1} = id$, it follows that id is a diffeomorphism, if and only if id is continuously differentiable. Apply the previous definition of a continuously differentiable mapping between two manifolds.



We have $id(U_\alpha) = U_\alpha$ and consider all pairs of indices α, β, where $\alpha \in I$ and $\beta \in J$, such that $U_\alpha \cap V_\beta \neq \varnothing$. With $W_{\alpha,\beta} := U_\alpha \cap V_\beta$, the mapping

$$\psi_\beta \circ id \circ \varphi_\alpha^{-1} = \psi_\beta \circ \varphi_\alpha^{-1} \qquad \psi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(W_{\alpha,\beta}) \to \mathbb{R}^n$$

must be continuously differentiable, where $\varphi_\alpha(W_{\alpha,\beta}) \subset \mathbb{R}^n$ is an open set. It follows by definition, that $\mathcal{A} \cup \mathcal{B}$ is an atlas for M. On the other hand, if $\mathcal{A} \cup \mathcal{B}$ is an atlas for M, $\psi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(W_{\alpha,\beta})$ is continuously differentiable, so id is a diffeomorphism. (QED)

Remark 2.1: Compatibility is an equivalence relation, where reflexivity and symmetry are obvious and transitivity can be checked, applying the initial definition

$$M_\mathcal{A} \xrightarrow{\text{id}} M_\mathcal{B} \xrightarrow{\text{id}} M_C \qquad \text{id} \circ \text{id} \;=\; \text{id}$$

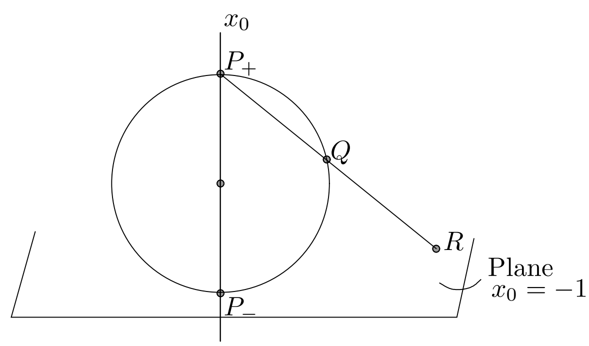
and using the fact, that the composition of two diffeomorphisms is a diffeomorphism. □

Definition 2.6: Given an atlas $\mathcal{A}$ for M, we consider the union of all atlases compatible with $\mathcal{A}$. This is itself an atlas for M, which is called the *maximal atlas* induced by $\mathcal{A}$. It is also called the *differential structure* induced by $\mathcal{A}$. □

Example 2.1: The n-dimensional sphere S^n may be considered as a subset of $\mathbb{R}^{n+1}$. Writing $\mathbb{R}^{n+1} = \mathbb{R} \times \mathbb{R}^n$ and pairs (x_0, x), where $x_0 \in \mathbb{R}$ and $x \in \mathbb{R}^n$, we get

$$S^n \;=\; \{(x_0, x) : x_0^2 + |x|^2 = 1\} \qquad |x| \;=\; \sqrt{\sum_{i=1}^n x_i^2}$$

We distinguish the two points $P_+ = (1, 0)$ $(0 \in \mathbb{R}^n)$, which we call *North Pole* of S^n and the *South Pole* of S^n $P_- = (-1, 0)$.



The *stereographic projection* maps $S^n \setminus P_+$ onto the plane $x_0 = -1$, which may be identified with $\mathbb{R}^n$. The condition that P_+, Q, R are collinear reads

$$P_+R \;=\; \lambda P_+Q$$

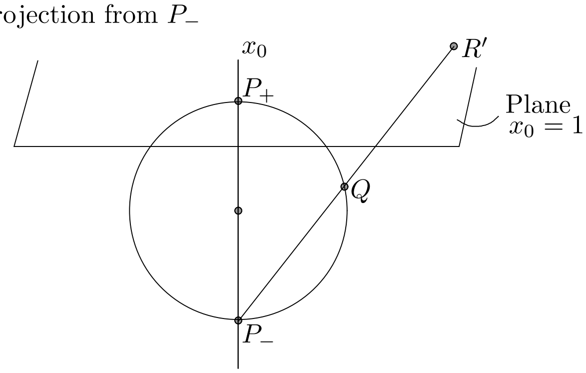
where we have

$$P_+Q \;=\; (x_0, x) - (1, 0) \;=\; (x_0 - 1, x) \qquad P_+R \;=\; (-1, y) - (1, 0) \;=\; (-2, y)$$

The condition then is

$$\left.\begin{array}{rcl} -2 &=& \lambda(x_0 - 1) \\ y &=& \lambda x \end{array}\right\} \;\Rightarrow\; y \;=\; \left(\frac{2}{1 - x_0}\right) x$$

This defines a homeomorphism $\varphi_+ : U_+ := S^n \setminus P_+ \to \mathbb{R}^n$. Note that the image of the South Pole P_- is $0 \in \mathbb{R}^n$. Then take the stereographic projection from P_-



The condition that P_-, Q, R' are collinear is

$$P_-R' \;=\; \lambda P'_- Q$$

where we have

$$P_-Q \;=\; (x_0, x) - (-1, 0) \;=\; (x_0 + 1, x) \qquad P_-R' \;=\; (1, y') - (-1, 0) \;=\; (2, y')$$

The condition then is

$$\left.\begin{array}{rcl} 2 & = & \lambda'(x_0+1) \\ y' & = & \lambda' x \end{array}\right\} \quad \Rightarrow \quad y' \;=\; \left(\frac{2}{1+x_0}\right) x$$

This defines a homeomorphism $\varphi_- : U_- := S^n \setminus P_- \to \mathbb{R}^n$. The image of P_+ is $0 \in \mathbb{R}^n$. We have an atlas with two charts

$$\mathcal{A} \;=\; \{(U_+, \varphi_+), (U_-, \varphi_-)\}$$

In fact, $U_+ \cap U_- = S^n \setminus \{P_+, P_-\}$ and $\varphi_+(U_+ \cap U_-) = \mathbb{R}^n \setminus \{0\} = \varphi_-(U_+ \cap U_-)$. The mappings

$$\varphi_{+,-} \;=\; \varphi_- \circ \varphi_+^{-1} \qquad \varphi_{-,+} \;=\; \varphi_+ \circ \varphi_-^{-1}$$

where $\varphi_{+,-}^{-1} = \varphi_{-,+}$, are given by

$$y' \;=\; \varphi_{+,-}(y) \qquad y \;=\; \varphi_{-,+}(y')$$

where we used that $Q = \varphi_+^{-1}(y)$, $\varphi_-(Q) = y'$. These mappings are homeomorphisms of $\mathbb{R}^n \setminus 0$ onto itself. We see from the above formulas that $y' = \alpha y$ with $\alpha > 0$. To find α, consider

$$|y|\,|y'| \;=\; \frac{4|x|^2}{1 - x_0^2} \;=\; \frac{4|x|^2}{|x|^2} \;=\; 4 \qquad \alpha \;=\; \frac{|y'|}{|y|}$$

Thus we obtain simply

$$y' \;=\; \frac{4y}{|y|^2} \qquad y \;=\; \frac{4y'}{|y'|^2}$$

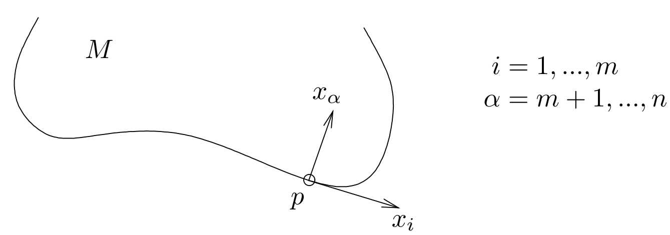
These mappings being continuously differentiable to all order, $\varphi_{+,-}$ and $\varphi_{-,+}$ are in fact diffeomorphisms. □

Proposition 2.2: Consider an m-dimensional submanifold M of E^n with $m < n$. Then, M is in fact a differentiable manifold. □

Proof: At each point $p \in M$, M is locally (in some neighborhood U_p) represented as a graph

$$x_\alpha \;=\; f_\alpha(x_1, ..., x_m)$$

with respect to a system of rectangular axes with origin p.



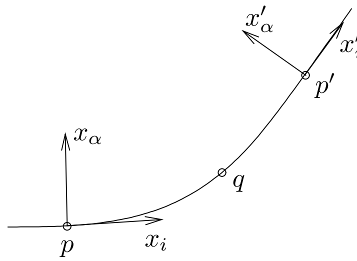
We consider the collection

$$\mathcal{A} \;=\; \{(U_p, \varphi_p) : p \in M\}$$

where

$$\varphi_p(q) \;=\; (x_1, ..., x_m) \quad q \in U_p$$

is a homeomorphism of U_p onto an open set in $\mathbb{R}^n$ which is homeomorphic to $\mathbb{R}^n$ itself. $(x_1, ..., x_m)$ are the first m coordinates of q with respect to the axes based on p. We must show that the above collection constitues an atlas for M.



Consider then a pair $p, p' \in M$ such that $U_p \cap U_{p'} \neq \varnothing$. Let be $g = \varphi_{p'} \circ \varphi_p^{-1}$, defined on $\varphi_p(U_p \cap U_{p'})$. Let be $q \in U_p \cap U_{p'}$. Then q has a representation in both the primed and unprimed axes. But the coordinates of any point, in particular of q, with respect to the two systems of (right handed) axes, differ only by a translation and a rotation. Thus for q, we have

$$x'_i \;=\; \sum_{j=1}^{n} O_{ij} x_j + c_i \quad i = 1, ..., n$$

where $O \in \mathrm{SO}(n)$. Hence g is given by

$$x'_i \;=\; \sum_{j=1}^{n} O_{ij} x_j + c_i \;=\; \sum_{j=1}^{m} O_{ij} x_j + \sum_{\alpha=m+1}^{n} O_{i\alpha} x_\alpha + c_i$$

That is

$$x'_i \;=\; \sum_{j=1}^{m} O_{ij} x_j + \sum_{\alpha=m+1}^{n} O_{i\alpha} f_\alpha(x_1, ..., x_m) + c_i \quad i = 1, ..., m$$

Since the functions f_α, $\alpha = m+1, ..., n$ are continuously differentiable, the mapping $g = g_{p',p}$ is itself continuously differentiable. Thus the collection $\mathcal{A}$ is in fact an atlas for M and M is indeed a differentiable manifold. (QED)

2.3 Quotient Spaces

Remark 2.2: Let $\sim$ be an equivalence relation on a set X. We write $x \sim y$ to denote $(x, y) \in \sim$. The *equivalence class* C_x of x is defined as

$$C_x \;:=\; \{y \in X : y \sim x\}$$

It follows that for any pair $x, y \in X$, it is

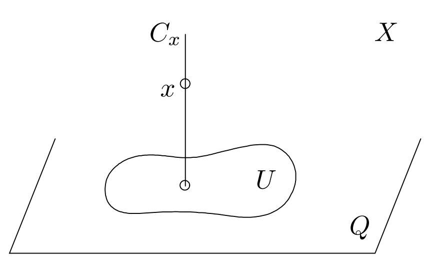
$$x \sim y \quad \Leftrightarrow \quad C_x \;=\; C_y \qquad x \not\sim y \quad \Leftrightarrow \quad C_x \cap C_y \;=\; \varnothing$$

The *quotient* $Q = X/_\sim$ is the set of equivalence classes

$$Q \;=\; \{C_x : x \in X\}$$

We have the *projection* $\pi : X \to Q$, defined by $\pi(x) = C_x$. □

Definition 2.7: If X is a topological space with an equivalence relation $\sim$, there is a natural topology induced on $Q := X/_\sim$, the *quotient topology.* It's the family of open subsets of Q, where a subset $U \subset Q$, which is a set of equivalence classes, is open in Q, if and only if $\pi^{-1}(U)$ is an open set in X.



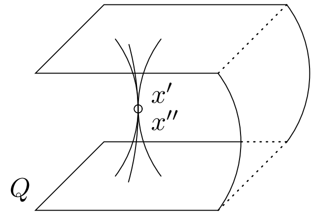
□

Remark 2.3: In the case that X is a topological manifold, $X/_\sim$ is not, in general, a topological manifold. □

Example 2.2: Take $X = \mathbb{R}^2$ and x', x'' two distinct points in X. Let be

$$\sim \ := \ \Delta \cup \{(x', x''), (x'', x')\} \qquad \Delta \ = \ \{(x, x) : x \in X\}$$

where Δ is the diagonal for any set X. Then $Q := X/_\sim$ is not a topological manifold



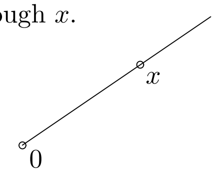
Q is a folded plane which intersects itself at one point. There is no neighborhood of the equivalence class $\{x', x''\}$, which is homeomorphic to $\mathbb{R}^2$. □

We would like to find out, which equivalence relations lead to quotients of topological manifolds, which are themselves topological manifolds.

Example 2.3: With $X = \mathbb{R}^{n+1} \setminus \{0\}$ we introduce the equivalence relation

$$y \ \sim \ x \quad \Leftrightarrow \quad y \ = \ \alpha x, \, \alpha > 0$$

The equivalence class C_x is then the way through x.



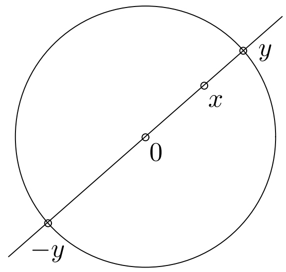
The quotient $\mathbb{R}^{n+1} \setminus \{0\}/_\sim$ is in fact S^n. In each ray there is exactly one representative y, with $|y| = 1$.

This in fact is san example of the ancient astronomers. When they saw an object, they couldn't tell, how far it was, but they could tell, that it wasn't where they stayed. That's why the origin is excluded. This is how the notion of the *celestial sphere* came about. □

Example 2.4: For $X = \mathbb{R}^{n+1} \setminus \{0\}$ we introduce the equivalence relation

$$y \ \sim \ x \quad \Leftrightarrow \quad y \ = \ \alpha x, \, \alpha \in \mathbb{R} \setminus \{0\}$$

Then C_x is the line through x.



The quotient $\mathbb{R}^{n+1} \setminus \{0\}/_\sim$ is called the *real projective n-dimensional space* $\mathbb{R}P^n$. In each equivalence class there are exactly two representatives $y, -y$ lying on S^n such that $|y| = 1$. Thus we can also consider $\mathbb{R}P^n$ to be the quotient

$$\mathbb{R}P^n \ = \ S^n/_\sim$$

where $\sim$ is the equivalence relation

$$y \ \sim \ x \quad \Leftrightarrow \quad \text{either } y = x \text{ or } y = -x$$

An eqivalenc class is then a pair $\{x, -x\}$ of antipodal points on S^n. So $\mathbb{R}P^n$ is the space of diameters. We have the projection

$$\pi : S^n \to \mathbb{R}P^n \ : \ \pi(x) \ := \ \{x, -x\} \ := \ y$$

Let U_y be a neighborhood of $y \in \mathbb{R}P^n$. Then $\pi^{-1}(U_y)$ is an open set in S^n containing x and such that

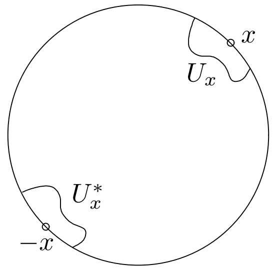
$$x' \in \pi^{-1}(U_y) \quad \Rightarrow \quad -x' \in \pi^{-1}(U_y)$$

In fact, any open set U in S^n with the properties ($x \in U$ and $x' \in U \Rightarrow -x' \in U$, projects to a neighborhood of $y \in \mathbb{R}P^n$. □

Example 2.5: For any subset $K \subset S^n$ let us denote

$$K^* \; := \; \{-x : x \in K\}$$

For any $x \in S^n$ we can find a neighborhood U_x of x, homeomorphic to $\mathbb{R}^n$ and such that $U_x^* \cap U_x = \varnothing$.



Let then $\varphi_x : U_x \to \mathbb{R}^n$ be a homeomorphism. The corresponding neighborhood of $y = \{x, -x\}$ in $\mathbb{R}P^n$ is $W_y = \{\{x', -x'\} : x' \in U_x\}$. Given a pair $\{x', -x'\} \in W_y$, exactly one of $x', -x'$ belongs to U_x and the other to U_x^*. We can thus define the mapping

$$\psi_y : W_y \to \mathbb{R}^n \quad : \quad \psi_y(\{x', -x'\}) \; := \; \begin{cases} \varphi_x(x') & x' \in U_x \\ \varphi_x(-x') & -x' \in U_x \end{cases}$$

ψ_y then is a homeomorphism onto $\mathbb{R}^n$ and $\mathbb{R}P^n$ is a topological manifold. Moreover, if $\{(U_x, \varphi_x) : x \in S^n\}$ is an atlas for S^n, then $\{(W_y, \psi_y) : y \in \mathbb{R}P^n\}$ is an atlas for $\mathbb{R}P^n$ and $\mathbb{R}P^n$ is a differentiable manifold. □

Remark 2.4: The example of $\mathbb{R}P^n$ is an example of a quotient of a differentiable manifold M by a discrete group G of diffeomorphisms of M. In the case of $\mathbb{R}P^n$, $M = S^n$ and $G = \{id, a\}$, where a is the antipodal map, defined by $a(x) := -x$, $x \in S^n$. So $a \circ a = \text{id}$. More generally, such a *discrete group* is a countable set of diffeomorphisms of M onto itself, which forms a group under composition

$$f_1, f_2 \in G \;\Rightarrow\; f_1 \circ f_2 \in G \qquad f \in G \;\Rightarrow\; f^{-1} \in G \qquad id \in G$$

A discrete group of homeomorphisms of a topological manifold M is defined in a similar way. □

Proposition 2.3: If for each $p \in M$ there is a neighborhood U of p such that

$$f(U) \cap U \neq \varnothing, \, f \in G \;\Rightarrow\; f \;=\; id$$

then M/G is a topological (differentiable) manifold. M/G means the quotient of M by the equivalence relation

$$q \;\sim\; p \quad\Leftrightarrow\quad \exists f \in G : q \;=\; f(p)$$

□

Proof: See example 2.5. (QED)

Example 2.6: The n-dimensional torus Π^n arises as the quotient

$$\Pi^n \;=\; \mathbb{R}^n/\mathbb{Z}^n \qquad \mathbb{Z}^n \;:=\; \{f_k : k = (k_1, ..., k_n) \in \mathbb{Z}^n\}$$

with

$$f_k(x) \;:=\; x + k \;=\; (x_1 + k_1, ..., x_n + k_n)$$

We check the sufficient condition, showing that for every $x \in \mathbb{R}^n$, $B_{\frac{1}{2}}(x)$ has the required property. So, let be $x', x'' \in B_{\frac{1}{2}}(x)$ with $x'' = x' + k$, $0 \neq k \in \mathbb{Z}^n$. So $|k| \geq 1$ and therefore $|x'' - x'| \geq 1$, but this is impossible because $\operatorname{diam} B_{\frac{1}{2}}(x) = 1$. That is

$$B_{\frac{1}{2}}(x) \cap f_k(B_{\frac{1}{2}}(x)) \;=\; \varnothing$$

A continuous (or continuously differentiable) function φ on Π^n is induced by a continuous (or continuously differentiable) periodic function $\widehat{\varphi}$ on $\mathbb{R}^n$, satisfying

$$\widehat{\varphi}(f_k(x)) \;=\; \widehat{\varphi}(x) \quad \forall k \in \mathbb{Z}^n$$

□

Remark 2.5: Π^n is diffeomorphic to the product $S^1 \times ... \times S^1$ with n factors. $\square$

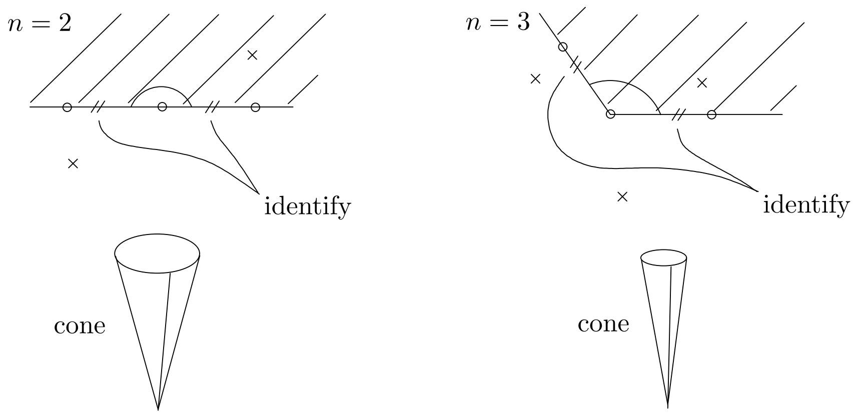
Example 2.7: The above sufficient condition is not always necessary. Let us consider the quotient of $\mathbb{R}^2$ by a cyclic group of rotations about the origin, identifying $\mathbb{R}^2$ with the complex plane $\mathbb{C}$. The group is

$$G_n = \{f_k : k = 0, ..., n-1\} \qquad f_k = k \text{ fold composition}$$

where

$$f_0 = \text{id} \qquad f_1(z) = e^{i2\pi/n} z \qquad f_k(z) = e^{i2\pi k/n} z$$

The orbit of z by G_n consists of the points $z, e^{i\pi/n}z, ..., e^{i\pi(n-1)/n}z$. The cone is homeomorphic to $\mathbb{R}^2$ itself.



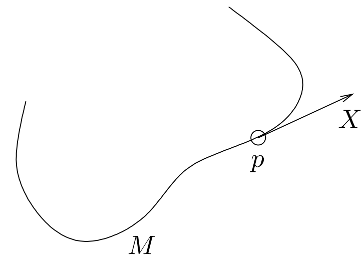
However the origin does not move by any rotation. Therefore if U is any neighborhood of the origin, $f_k(U)$ must intersect U for any $k = 0, ..., n-1$. But despite the failure of the sufficient condition, the quotient $\mathbb{R}^2/G_n$ is a topological manifold, although not a differentiable manifold. $\square$

3 The Tangent Bundle

3.1 Tangent Vector

Remark 3.1: From now on we consider C^∞ differentiable manifolds and C^∞-functions on them. □

Definition 3.1: For a submanifold M with $\dim M = m$ of the Euclidean space E^n $(n > m)$, the *tangent space* to M at a point $p \in M$, denoted by T_pM, can be defined to be the subspace of all vectors $X \in E^n$ attached to p, such that an infinitesimal displacement along X keeps us on M.



That is, X is required to satisfy the condition (see remark 3.2)

$$X \cdot g_\alpha \; = \; 0 \qquad \alpha \; = \; m+1, ..., n$$

where the g_α are the function

$$g_\alpha \; = \; x_\alpha - f_\alpha(x_1, ..., x_m) \qquad \alpha \; = \; m+1, ..., n$$

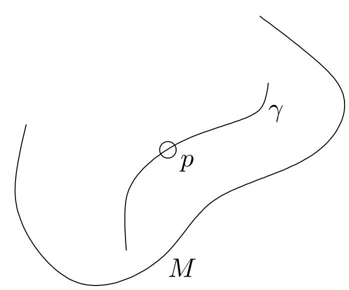
M being represented in a neighborhood of p as the subset of E^n where these functions vanish. □

Remark 3.2: For any continuously differentiable function f on E^n and a vector $X \in E^n$, we denote

$$X \cdot f \; = \; \lim_{t \to 0} \frac{f(p + tX) - f(p)}{t}$$

□

We want to discuss now, how one can define the tangent space at a point to a manifold intrinsically. We must define the concept of a tangent vector v at a point to a manifold M. A tangent vector is associated to a continuously differentiable parametrized curve through p.



Definition 3.2: A *continuously differentiable parametrized curve* γ through $p \in M$ is a continuously differentiable mapping $\gamma : I \to M$, where I is an interval of $\mathbb{R}$ containing 0 and satisfying $\gamma(0) = p$. □

Remark 3.3: Intuitively, I can be thought of as a time interval and γ then represents the motion of a particle which at $t = 0$ is at p. What we want is to define the velocity v of the particle at time $t = 0$. The idea is to consider v as a directional derivative. Let f be a continuously differentiable (real valued) function on M. We consider f along the curve, that is we consider $f \circ \gamma : I \to \mathbb{R}$. Then the directional derivative to f along γ at p is

$$\left. \frac{\mathrm{d} f \circ \gamma}{\mathrm{d}t} \right|_{t=0}$$

a real number. Consider now the space $C^\infty(M)$ of all C^∞-functions on M. This is a linear space since for $f, g \in C^\infty(M)$ we have the addition rule

$$(f+g)(q) \;=\; f(q)+g(q) \quad \forall q \in M$$

and the scalar multiplication rule

$$(\alpha f)(q) \;=\; \alpha f(q) \quad \forall q \in M, \forall \alpha \in \mathbb{R}$$

This is an infinite dimensional linear space. Moreover, $C^\infty(M)$ is a commutative ring with respect to the multiplication of functions, that is for any pair $f, g \in C^\infty(M)$, fg is defined by

$$(fg)(q) \;=\; f(q)g(q) \quad \forall q \in M$$

□

Definition 3.3: We first define the *tangent vector* v to γ at p to be the linear function on $C^\infty(M)$ given by

$$v \cdot f \;=\; \left.\frac{\mathrm{d}(f\circ\gamma)}{\mathrm{d}t}\right|_{t=0}$$

We can consider v as a linear map $v : C^\infty(M) \to \mathbb{R}$. □

Remark 3.4: v is linear, because

$$\begin{aligned} v\cdot(f_1+f_2) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}(f_1\circ\gamma + f_2\circ\gamma)\right|_{t=0} = v\cdot f_1 + v\cdot f_2 \\ v\cdot(\alpha f) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\alpha f\circ\gamma\right|_{t=0} = \alpha v\cdot f \end{aligned}$$

□

Remark 3.5: Consider now the commutative ring structure of $C^\infty(M)$

$$\begin{aligned} v\cdot(fg) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}(fg)\circ\gamma\right|_{t=0} = \left.\frac{\mathrm{d}}{\mathrm{d}t}((f\circ\gamma)(g\circ\gamma))\right|_{t=0} \\ &= (f\circ\gamma)(0)\left.\frac{\mathrm{d}}{\mathrm{d}t}(g\circ\gamma)\right|_{t=0} + (g\circ\gamma)(0)\left.\frac{\mathrm{d}}{\mathrm{d}t}(f\circ\gamma)\right|_{t=0} \\ &= f(p)v\cdot g + g(p)v\cdot f \end{aligned}$$

what we call the *Leibnitz rule.* □

Two different curves through p may have the same tangent vector at p. The above motivates the following definition.

Definition 3.4: A *tangent vector* v at $p \in M$ is a linear (real valued) function on $C^\infty(M)$ satisfying the Leibnitz rule

$$v\cdot(fg) \;=\; f(p)v\cdot g + g(p)v\cdot f \quad \forall f, g \in C^\infty(M)$$

□

Definition 3.5: A *bump function* on $\mathbb{R}^n$ is a C^∞ function ρ on $\mathbb{R}^n$ which is spherically symmetric, that is

$$|x'| \;=\; |x| \quad\Rightarrow\quad \rho(x') \;=\; \rho(x)$$

and ρ is non-increasing with $|x|$, that is

$$|x'| \;>\; |x| \quad\Rightarrow\quad \rho(x') \;\le\; \rho(x)$$

□

Proposition 3.1: The basic properties of tangent vectors as given in definition 3.4 are:

(1) $v \cdot f = 0$, if f is a constant function, that is $f(q) = \alpha \ \forall q \in M$, with some fixed $\alpha \in \mathbb{R}$.

(2) $v \cdot f = 0$, if f vanishes in a neighborhood of p. Thus, $v \cdot f$ depends only on f in a neighborhood of p.

□

Proof:

(1) We can assume that $\alpha \neq 0$. Again by linearity we can assume that $\alpha = 1$. Then $f^2 = f$ and it follows by the Leibnitz rule

$$v \cdot f \;=\; v \cdot f^2 \;=\; 2f(p)v \cdot f \;=\; 2v \cdot f$$

hence $v \cdot f = 0$.

(2) We begin by constructing a bump function on $\mathbb{R}^n$ with the property

$$\begin{cases} \rho(0) = 1 \\ \rho(x) = 0 \quad x \notin B_1(0) \end{cases}$$

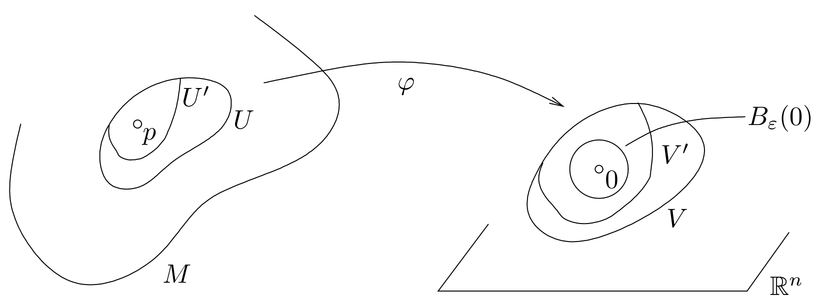
It follows that the range of possible values for ρ is $[0, 1]$. For example, we could define

$$\rho(x) \;=\; \begin{cases} \exp\left(1 - \frac{1}{1-|x|^2}\right) & |x| < 1 \\ 0 & |x| \geq 1 \end{cases}$$

We now set, for $\varepsilon > 0$

$$\rho_\varepsilon(x) \;=\; \rho(x/\varepsilon)$$

Then $\rho_\varepsilon(x)$ is a bump function but with $B_1(0)$ replaced by $B_\varepsilon(0)$. Given now a C^∞-function f on M which vanishes in a neighborhood W of p, take a chart (U, φ), with $p \in U$. Then $U' := U \cap W$ is a neighborhood of p contained in U where f vanishes. Set now $V := \varphi(U)$ and $V' := \varphi(U')$.



We can assume (by translation) that $\varphi(p) = 0$. Since V' is open, there is an $\varepsilon > 0$ such that $B_\varepsilon(0) \subset U'$. We then define a C^∞-function g on M by

$$g(q) \;:=\; \begin{cases} (\rho_\varepsilon \circ \varphi)(q), & q \in U' \\ 0, & q \notin U' \end{cases}$$

We then have $g(p) = 1$. Consider now the product fg. Since f vanishes in U' while g vanishes in the complement of U', the product fg vanishes identically. Therefore

$$0 \;=\; v \cdot (fg) \;=\; \underbrace{f(p)}_{=0} v \cdot g + \underbrace{g(p)}_{=1} v \cdot f \;=\; v \cdot f$$

that is $v \cdot f = 0$.

(QED)

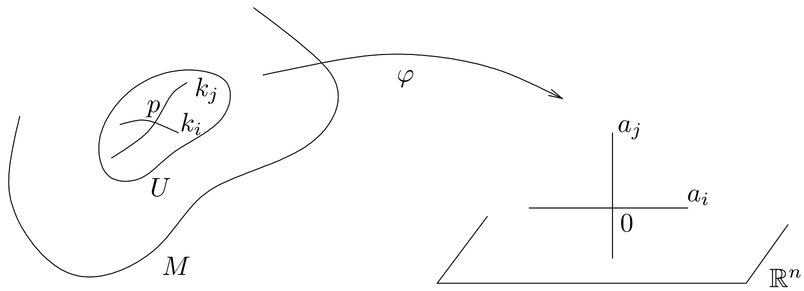
Corollary 3.1: Combining the properties (1) and (2) of Proposition 3.1 we obtain that $v \cdot f = 0$ if f is constant in a neighborhood of p. □

3.2 Tangent Space

Let M be a differentiable manifold of dimension n.

Definition 3.6: The space of all tangent vectors at $p \in M$ is called the *tangent space* at p to M, denoted T_pM. □

Remark 3.6:



Let $p \in U$ be the domain of a chart (U, φ). By suitable translation we may assume that $\varphi(p) = 0$, the origin in $\mathbb{R}^n$. Consider the coordinate lines a_i, $i = 1, .., n$ in $\mathbb{R}^n$

$$a_i : \mathbb{R} \to \mathbb{R}^n \quad : \quad t \mapsto (0, ..., t, ..., 0)$$

where the t stays at the i-th component. Consider then the curves $k_i = \varphi^{-1} \circ a_i$ in M through p. Let ℓ_i be the tangent vector to k_i at p

$$\ell_i \cdot f \quad = \quad \left. \frac{\mathrm{d} f \circ k_i}{\mathrm{d}t} \right|_{t=0} \qquad \forall f \in C^\infty(M)$$

Now the real function $\tilde{f} = f \circ \varphi^{-1}$ on $\mathbb{R}^n$ represents f in the chart (U, φ). We have

$$f \circ k_i \quad = \quad \tilde{f} \circ a_i \quad = \quad \tilde{f}(0, ..., t, ..., 0)$$

where the t again stays at the i-th component. So

$$\ell_i \cdot f \quad = \quad \left. \frac{\mathrm{d} \tilde{f} \circ a_i}{\mathrm{d}t} \right|_{t=0} \quad = \quad \left. \frac{\partial \tilde{f}}{\partial x^i} \right|_0$$

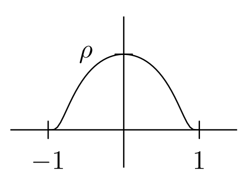
We denote

$$\ell_i \quad = \quad \left. \frac{\partial}{\partial x^i} \right|_p \qquad i = 1, ..., n$$

□

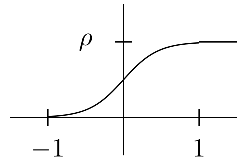
Remark 3.7: Digression on Cutoff functions
Let us define the *bump function* on $\mathbb{R}$

$$\rho(t) \quad := \quad \begin{cases} \exp\left(1 - \frac{1}{1-t^2}\right) & t \in (-1, 1) \\ 0 & t \notin (-1, 1) \end{cases}$$



$\rho(t)$ is a C^∞-function. We then define the C^∞-function σ by

$$\sigma(t) \quad := \quad \begin{cases} 0, & t \leq -1 \\ \frac{\int_{-1}^{t} \rho(t')\,\mathrm{d}t'}{\int_{-1}^{1} \rho(t')\,\mathrm{d}t'}, & -1 < t < 1 \\ 1, & t \geq 1 \end{cases}$$



Given $a, b \in \mathbb{R}$, $a > 0$, we then define $\sigma_{a,b}$ by

$$\sigma_{a,b}(t) \quad := \quad \sigma\left(\frac{t-b}{a}\right)$$

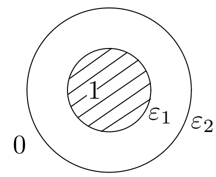
Then $\sigma_{a,b}$ is C^∞, non-decreasing, with range $[0,1]$, $\sigma_{a,b}(t) = 0$ for $t \le b-a$ and $\sigma_{a,b}(t) = 1$ for $t \ge b+a$. Then, given $\varepsilon_1, \varepsilon_2 > 0$ with $\varepsilon_2 > \varepsilon_1$ we define on $\mathbb{R}^n$ the real C^∞-function $\eta_{\varepsilon_1,\varepsilon_2}$ by

$$\eta_{\varepsilon_1,\varepsilon_2}(x) \quad := \quad 1 - \sigma_{a,b}(|x|) \quad x \in \mathbb{R}^n$$

where a, b are defined by

$$b - a \quad = \quad \varepsilon_1 \qquad b + a \quad = \quad \varepsilon_2$$

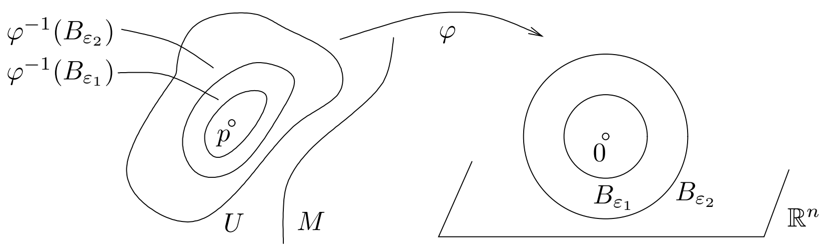
η is spherically symmetric, non-increasing with $|x|$, of range $[0,1]$, such that $\eta = 1$ on $\overline{B_{\varepsilon_1}}(0)$ and $\eta = 0$ on $\mathbb{R}^n \setminus B_{\varepsilon_2}(0)$. Such a function we call a *Cutoff function.*



We now introduce Cutoffs of the coordinate functions on $\mathbb{R}^n$

$$g^i \quad := \quad \eta_{\varepsilon_1,\varepsilon_2}(x)\, x^i \quad i = 1, ..., n$$

These are C^∞-functions with support in $B_{\varepsilon_2}(0)$ and agreeing with the coordinate functions in $B_{\varepsilon_1}(0)$. Consider now $g^i \circ \varphi$, functions defined on U, the domain of the chart. $g^i \circ \varphi$ is a real function on U, vanishes on $U \setminus \varphi^{-1}(B_{\varepsilon_2}(0))$ and represents the i-th componenent of φ on $\varphi^{-1}(B_{\varepsilon_1}(0))$. So we write $\varphi^i := g^i \circ \varphi$.



We extend the $g^i \circ \varphi$ by 0 outside U, defining in this way the functions $\overline{g}^i$ on M. That is

$$\overline{g}^i \quad := \quad \begin{cases} g^i \circ \varphi & \text{on } U \\ 0 & \text{on } M \setminus U \end{cases}$$

These are C^∞-functions on M. Therefore, we can apply any given tangent vector, in particular a given $v \in T_pM$, to them to obtain the real numbers

$$v^i \quad = \quad v \cdot \overline{g}^i \quad i = 1, ..., n$$

the components of v in the chart (U, φ). □

Proposition 3.2: The tangent space T_pM is a vector space and we can write

$$v \quad = \quad \sum_{i=1}^{n} v^i \left.\frac{\partial}{\partial x^i}\right|_p \qquad \forall v \in T_pM$$

Thus the vectors $\left.\frac{\partial}{\partial x^i}\right|_p$, $i = 1, ..., n$ form a basis for T_pM and T_pM is n-dimensional. □

Proof: We have the vector addition

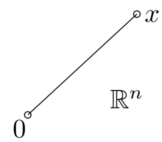
$$(v_1 + v_2) \cdot f \;=\; v_1 \cdot f + v_2 \cdot f \quad \forall f \in C^\infty(M)$$

and the scalar multiplication

$$(\alpha v) \cdot f \;=\; \alpha v \cdot f \quad \forall \alpha \in \mathbb{R}, \forall f \in C^\infty(M)$$

which satisfy the Leibnitz rule. So T_pM is a vector space.

Consider an arbitrary $f \in C^\infty(M)$. First, since v vanishes on constant functions, we may subtract a suitable constant from f to acheive $f(p) = 0$. Then we set $\tilde{f} := f \circ \varphi^{-1}$, the representation of f in the chart (U, φ), which is a C^∞-function on $\mathbb{R}^n$ vanishing at 0 (because we chose φ such that $\varphi(p) = 0$).



For a given $x \in \mathbb{R}^n$ consider $\tilde{f}$ along the line segment $t \mapsto tx$, $t \in [0,1]$, joining 0 to x. At $t = 0$, $\tilde{f}(tx) = \tilde{f}(0) = 0$ and at $t = 1$, $\tilde{f}(tx) = \tilde{f}(x)$. So

$$\tilde{f}(x) \;=\; \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\,\tilde{f}(tx)\,\mathrm{d}t \;=\; \int_0^1 \sum_{i=1}^{n} \frac{\partial \tilde{f}}{\partial x^i}(tx)\,x^i\,\mathrm{d}t \;=\; \sum_{i=1}^{n} x^i \tilde{h}_i(x)$$

where we set

$$\tilde{h}_i(x) \;:=\; \int_0^1 \frac{\partial \tilde{f}}{\partial x^i}(tx)\,\mathrm{d}t$$

This can be thought of as the mean value of $\frac{\partial \tilde{f}}{\partial x^i}$ on the line segment, joining 0 to x. Since the $\frac{\partial \tilde{f}}{\partial x^i}$ are continuous functions, we have

$$\tilde{h}_i(x) \;\overset{x \to 0}{\longrightarrow}\; \left.\frac{\partial \tilde{f}}{\partial x^i}\right|_0$$

In $U \subset M$ we have $f = \tilde{f} \circ \varphi$. That is, if $q \in U$ and $x = \varphi(q)$, $x^i = \varphi^i(q)$, $i = 1, ..., n$, we have

$$f(q) \;=\; \tilde{f}(x) \;=\; \sum_{i=1}^{n} x^i \tilde{h}_i(x)$$

We now define $\tilde{h}'_i := \eta_{\varepsilon_1,\varepsilon_2}\tilde{h}_i$. Then the $\tilde{h}'_i$ agrees with the $\tilde{h}_i$ on $B_{\varepsilon_1}(0)$ and vanishes in $\mathbb{R}^n \setminus B_{\varepsilon_2}(0)$. We then define the C^∞-functions h_i on M by

$$h_i(q) \;:=\; \begin{cases} \tilde{h}'_i(\varphi(q)), & q \in U \\ 0, & q \in M \setminus U \end{cases}$$

We then define a new C^∞-function f' on M by

$$f' \;=\; \sum_{i=1}^{n} \overline{g}^{\,i} h_i$$

Then f' vanishes on $M \setminus U$ and agrees with f on $\varphi^{-1}(B_{\varepsilon_1}(0)) \subset U$. This is because, for $q \in \varphi^{-1}(B_{\varepsilon_1}(0))$ we have $\varphi(q) = x \in B_{\varepsilon_1}(0)$, hence $\eta_{\varepsilon_1,\varepsilon_2}(x) = 1$, so $\overline{g}^{\,i}(q) = g^i(x) = x^i$ and $h_i(q) = \tilde{h}'_i(x) = \tilde{h}_i(x)$. So, the function $f - f'$ vanishes on $\varphi^{-1}(B_{\varepsilon_1}(0))$, a neighborhood of p in M. By property (1) of tangent vectors at p from above, we get

$$v \cdot (f - f') \;=\; 0 \quad \Rightarrow \quad v \cdot f \;=\; v \cdot f'$$

Now

$$v \cdot f' \;=\; v \cdot \left(\sum_{i=1}^{n} \overline{g}^{\,i} h_i \right) \;=\; \sum_{i=1}^{n} v \cdot (\overline{g}^{\,i} h_i)$$

and so by the Leibnitz rule

$$v \cdot f' \;=\; \sum_{i=1}^{n} \left(\overline{g}^{\,i}(p) v \cdot h_i + h_i(p) v \cdot \overline{g}^{\,i} \right)$$

But now, $\overline{g}^{\,i}(p) = x_i\big|_0 = 0$, so the first term within the sum above vanishes, while

$$h_i(p) \;=\; \tilde{h}_i(0) \;=\; \left.\frac{\partial \tilde{f}}{\partial x^i}\right|_0 \;=\; \left.\frac{\partial}{\partial x^i}\right|_p \cdot f$$

and

$$v \cdot \overline{g}^{\,i} \;=\; v^i$$

the components of v in the chart φ. We conclude that

$$v \cdot f \;=\; v \cdot f' \;=\; \sum_{i=1}^{n} v_i \left.\frac{\partial}{\partial x_i}\right|_p \cdot f$$

and this proves the proposition. (QED)

3.3 Tangent Bundle

Definition 3.7: The *tangent bundle* TM of a differentiable manifold M is given by

$$TM \;=\; \bigcup_{p \in M} T_p M$$

$\square$

Proposition 3.3: The tangent bundle TM is a differentiable manifold of dimension $2n$, where $\dim M = n$. $\square$

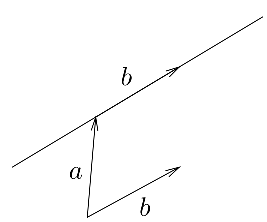
Remark 3.8: In fact, TM is a differentiable *vector bundle* over M, as we shall describe later on. $\square$

Proof of Proposition 3.3: Let $\mathcal{A} = \{(U_\alpha, \varphi_\alpha \mid \alpha \in I\}$ be an atlas for M. At each $p \in U_\alpha$, the $\frac{\partial}{\partial x^i}|_p$, $i = 1, ..., n$, defined by φ_α, constitutes a basis for T_pM. For any $v \in T_pM$ there is a unique n-tuplet of real numbers $(v^1, ..., v^n)$, the components of v in the chart φ_α, such that

$$v \;=\; \sum_{i=1}^{n} v^i \left.\frac{\partial}{\partial x^i}\right|_p$$

The v^i, $i = 1, ..., n$ are linear coordinates in the vector space T_pM. This means that a straight line

$$t \mapsto a + tb \quad a, b \in T_pM$$



in T_pM is represented in these coordinates by the linear equations

$$v^i \;=\; a^i + tb^i \quad i = 1, ..., n$$

Let us write

$$v^i \;=\; \chi^i_{\alpha,p} \cdot v$$

where $\chi_{\alpha,p} = (\chi^1_{\alpha,p}, ..., \chi^n_{\alpha,p})$ is a linear isomorphism between T_pM and $\mathbb{R}^n$, so

$$\chi_{\alpha,p} : T_pM \to \mathbb{R}^n$$

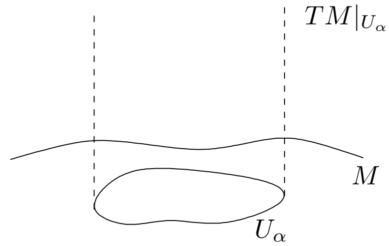
It is associated to φ_α and each $p \in U_\alpha$. Consider

$$TM|_{U_\alpha} := \bigcup_{p \in U_\alpha} T_pM$$

On $TM|_{U_\alpha}$ we define a chart ψ_α by

$$\psi_\alpha(v) = (\varphi_\alpha(p), \chi_{\alpha,p} \cdot v) = ((x^1, ..., x^n), (v^1, ..., v^n))$$

where $v \in T_pM$, $p \in U_\alpha$. Here $(x^1, ..., x^n)$ are the coordinates of p and $(v^1, ..., v^n)$ the components of v. ψ_α maps $TM|_{U_\alpha}$ onto $\mathbb{R}^{2n}$.



The collection $\{(TM|_{U_\alpha}, \psi_\alpha) | \alpha \in I\}$ forms an atlas for TM. To show this, consider $p \in U_\alpha \cap U_\beta$. Then a tangent vector $v \in T_pM$ has two sets of components, $v_\alpha^i = \chi_{\alpha,p}^i \cdot v$ in the chart φ_α and $v_\beta^i = \chi_{\beta,p}^i \cdot v$ in the chart φ_β. It holds

$$v_\beta^i = \sum_{j=1}^{n} M_j^i v_\alpha^j \quad i = 1, ..., n$$

We denote

$$f = \varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$$

where f is continuously differentiable and $\varphi_\alpha(U_\alpha \cap U_\beta)$, and $\varphi_\beta(U_\alpha \cap U_\beta)$ are both open subsets of $\mathbb{R}^n$. Let $x_\alpha^i = \varphi_\alpha^i(q)$ be the coordinates of a point $q \in U_\alpha \cap U_\beta$ with respect to the chart φ_α and let $x_\beta^i = \varphi_\beta^i(q)$ be the coordinates of q with respect to φ_β. Then

$$x_\beta^i = f^i(x_\alpha^1, ..., x_\alpha^n) \quad i = 1, ..., n$$

and we have (see Explanation A.1 on page 85)

$$M_j^i = \left. \frac{\partial f^i}{\partial x_\alpha^j} \right|_{\varphi_\alpha(p)} \tag{3.1}$$

Both transformations, that is the one for the coordinates x_β^i and the one for the components v_β^i, are continuously differentiable. Therefore, $\{(TM|_{U_\alpha}, \psi_\alpha) | \alpha \in I\}$ does satisfy the requirement for an atlas, where the ψ_α map to $\mathbb{R}^{2n}$, which means that $\dim TM = 2n$. (QED)

3.4 The Structure of TM

Definition 3.8: We define the *projection map*

$$\pi : TM \to M \quad : \quad \pi(v) = p$$

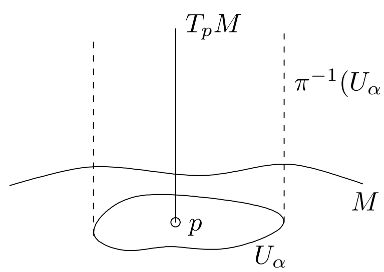
if $v \in T_pM$. π tells us at which point a tangent vector's attached. □

Remark 3.9: We get that

$$T_pM = \pi^{-1}(p)$$

which is the set of all vectors attached at p and

$$TM|_{U_\alpha} = \pi^{-1}(U_\alpha)$$



□

Remark 3.10: Consider now a covering of M by open sets U_α

$$M = \bigcup_{\alpha \in I} U_\alpha$$

which are not necessarily the domains of charts, but which have the property, that every U_α is associated to a diffeomorphism

$$\omega_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times \mathbb{R}^n$$

such that $\omega_{\alpha,p} := \omega_\alpha|_{T_pM}$ is a linear isomorphism of T_pM onto $\mathbb{R}^n$, for each $p \in U_\alpha$ and we have

$$\omega_\alpha(v) = (\pi(v), \omega_{\alpha,p} \cdot v) = (p, \omega_{\alpha,p} \cdot v)$$

In this case, if we denote by e_i, $i = 1, ..., n$ the standard basis of $\mathbb{R}^n$, then for each $p \in U_\alpha$ the vectors $\chi_{\alpha,i}(p) := \omega_{\alpha,p}^{-1} \cdot e_i$, $i = 1, ..., n$ form a basis for T_pM. If the U_α are domains of charts of M and ω_α is what we denoted ψ_α, then $\chi_{\alpha,i}(p) = \frac{\partial}{\partial x^i}|_p$. □

Example 3.1: An atlas for S^1 consists of at least two charts. However, in the above definition we can take in the case $M = S^1$, one domain U_α which is the whole of S^1. In fact, the same is true when M is any orientable 3-dimensional manifold. So for any such manifold there is a diffeomorphism

$$\omega : TM \to M \times \mathbb{R}^n \quad n = 3$$

such that $\omega(v) = (p, \omega_p \cdot v)$, $v \in T_pM$, where ω_p is a linear isomorphism of T_pM onto $\mathbb{R}^n$, (for $n = 3$). This is not true for 2-dimensional manifolds. The simplest counter-example is the two-dimensional sphere S^2. □

4 Vector Bundles

Definition 4.1: A differentiable *vector bundle* $\mathcal{B}$ over a differentiable manifold M is a differentiable manifold with the following structure:

1. There is a differentiable surjective map
$$\pi : \mathcal{B} \to M$$
called the *projection map* of the bundle.

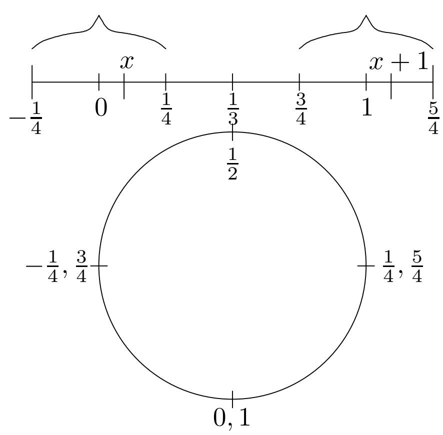
2. Each $\mathcal{B}_p = \pi^{-1}(p)$ with $p \in M$ is a vector space, which is isomorphic to a fixed finite dimensional vector space V. Note that
$$\mathcal{B} \;=\; \bigcup_{p\in M} B_p$$

3. There is an open covering $\{U_\alpha | \alpha \in I\}$, such that to each U_α is associated a diffeomorphism
$$\omega_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times V$$
where $\omega_\alpha(v) = (p, \omega_{\alpha,p} \cdot v)$, $p = \pi(v)$ and for each $p \in U_\alpha$, $\omega_{\alpha,p}$ is a linear isomorphism of the vector space $\mathcal{B}_p$ onto V.

□

Remark 4.1: The model vector space V in the above definition may be real or complex. The vector bundle $\mathcal{B}$ is then accordingly called *real* or *complex vector bundle* over M. In the real case, if $\dim V = m$, then $\dim \mathcal{B}_p = m$ for each $p \in M$, and $\dim \mathcal{B} = n + m$. In the simplest realization in quantum theory $V = \mathbb{C}$ and $\mathcal{B}$ is called a *complex line bundle* over M. □

Example 4.1: The Möbius line bundle
The Möbius line bundle is the simplest non-trivial example of a vector bundle. In this case $M = S^1$, $V = \mathbb{R}$. We begin with S^1, which we consider as the quotient
$$S^1 \;=\; \left(-\tfrac{1}{4}, \tfrac{5}{4}\right) / \sim$$
where $\sim$ is the equivalence relation
$$\sim \;=\; \Delta \cup \left\{(x, x+1) | x \in \left(-\tfrac{1}{4}, \tfrac{1}{4}\right)\right\} \cup \left\{(x, x-1) | x \in \left(\tfrac{3}{4}, \tfrac{5}{4}\right)\right\}$$



The equivalence classes are
$$C_x \;=\; \begin{cases} \{x\} & x \in \left[\tfrac{1}{4}, \tfrac{3}{4}\right] \\ \{x, x+1\} & x \in \left(-\tfrac{1}{4}, \tfrac{1}{4}\right) \\ \{x, x-1\} & x \in \left(\tfrac{3}{4}, \tfrac{5}{4}\right) \end{cases}$$

We have a cover $\{U_-, U_+\}$ of S^1, where
$$U_- \;:=\; \left\{C_x | x \in \left(-\tfrac{1}{4}, \tfrac{3}{4}\right)\right\} \qquad U_+ \;=\; \left\{C_x | x \in \left(\tfrac{1}{4}, \tfrac{5}{4}\right)\right\}$$

We define the charts $\varphi_\pm$ as homeomorphism of $U_\pm$ onto an open interval in $\mathbb{R}$ by

$$\varphi_-(C_x) \;=\; x$$

if x is the unique representative of C_x lying in $(-\frac{1}{4}, \frac{3}{4})$ (there is a unique such representative for each equivalence class in U_-) and

$$\varphi_+(C_x) \;=\; x$$

if x is the unique representative of C_x lying in $\left(\frac{1}{4}, \frac{5}{4}\right)$ (there is a unique such representative for each equivalence class in U_+). Then

$$\varphi_-(U_-) \;=\; \left(-\tfrac{1}{4}, \tfrac{3}{4}\right) \qquad \varphi_+(U_+) \;=\; \left(\tfrac{1}{4}, \tfrac{5}{4}\right)$$

$U_- \cap U_+$ consists of the components I_-, I_+, where

$$\begin{aligned} I_+ &= \left\{C_x | x \in \left(\tfrac{1}{4}, \tfrac{3}{4}\right)\right\} &=& \left\{\{x\} | x \in \left(\tfrac{1}{4}, \tfrac{3}{4}\right)\right\} \\ I_- &= \left\{C_x | x \in \left(-\tfrac{1}{4}, \tfrac{1}{4}\right)\right\} &=& \left\{\{x, x+1\} | x \in \left(-\tfrac{1}{4}, \tfrac{1}{4}\right)\right\} \end{aligned}$$

We have

$$\varphi_-(I_+) \;=\; I_+ \qquad \varphi_+(I_+) \;=\; I_+$$

and $\varphi_+ \circ \varphi_-^{-1}$ is the identity on I_+ while

$$\varphi_-(I_-) \;=\; I_- \;=\; \left(-\tfrac{1}{4}, \tfrac{1}{4}\right) \qquad \varphi_+(I_-) \;=\; \left(\tfrac{3}{4}, \tfrac{5}{4}\right)$$

and $\varphi_+ \circ \varphi_-^{-1}$ for $x \in (-\frac{1}{4}, \frac{1}{4})$ is a function that maps $x \mapsto x+1$. We now consider the quotient

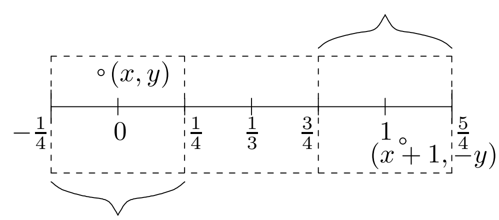
$$\mathcal{B} \;=\; \left(-\tfrac{1}{4}, \tfrac{5}{4}\right) \times \mathbb{R} \;/\approx$$

where $\approx$ is the equivalence relation

$$\approx \;=\; \Delta \cup \left\{((x,y),(x+1,-y)) \mid x \in \left(-\tfrac{1}{4}, \tfrac{1}{4}\right), y \in \mathbb{R}\right\} \cup \left\{((x,y),(x-1,-y)) \mid x \in \left(\tfrac{3}{4}, \tfrac{5}{4}\right), y \in \mathbb{R}\right\}$$

$\mathcal{B}$ is the Möbius bundle.

We can picture the Möbius band, which is obtained by replacing $\mathbb{R}$ by $(-1, 1)$.



We first show that $\mathcal{B}$ is a differentiable manifold. In fact

$$\mathcal{A} \;=\; \{(U_-, \psi_-), (U_+, \psi_+)\}$$

is an atlas for $\mathcal{B}$. Let us denote by $\varepsilon_{(x,y)}$ the equivalence class of (x, y). We have

$$\varepsilon_{(x,y)} \;=\; \begin{cases} \{(x,y)\} & x \in \left(\frac{1}{4}, \frac{3}{4}\right), y \in \mathbb{R} \\ \{(x,y),(x+1,-y)\} & x \in \left(-\frac{1}{4}, \frac{1}{4}\right), y \in \mathbb{R} \\ \{(x,y),(x-1,-y)\} & x \in \left(\frac{3}{4}, \frac{5}{4}\right), y \in \mathbb{R} \end{cases}$$

Then

$$V_- \;=\; \left\{\varepsilon_{(x,y)} \mid x \in \left(-\tfrac{1}{4}, \tfrac{3}{4}\right), y \in \mathbb{R}\right\} \qquad V_+ \;=\; \left\{\varepsilon_{(x,y)} \mid x \in \left(\tfrac{1}{4}, \tfrac{5}{4}\right), y \in \mathbb{R}\right\}$$

Note that for $\varepsilon_{(x,y)} \in V_-$, there is a unique representative in $\varepsilon_{(x,y)}$ such that its first coordinate belongs to $\left(-\frac{1}{4}, \frac{3}{4}\right)$. And for $\varepsilon_{(x,y)} \in V_+$, there is a unique representative in $\varepsilon_{(x,y)}$ such that its first coordinate belongs to $\left(\frac{1}{4}, \frac{5}{4}\right)$. Taking in each case this unique representative we define

$$\psi_-(\varepsilon_{(x,y)}) \;=\; (x,y) \quad x \in \left(-\tfrac{1}{4}, \tfrac{3}{4}\right) \qquad \psi_+(\varepsilon_{(x,y)}) \;=\; (x,y) \quad x \in \left(\tfrac{1}{4}, \tfrac{5}{4}\right)$$

Note that

$$\psi_-(V_-) \;=\; \left(-\tfrac{1}{4}, \tfrac{3}{4}\right) \times \mathbb{R} \qquad \psi_+(V_+) \;=\; \left(\tfrac{1}{4}, \tfrac{5}{4}\right) \times \mathbb{R}$$

Consider $V_- \cap V_+$. This consists of two components J_-, J_+, where

$$\begin{array}{rcccl} J_+ & = & \left\{\varepsilon_{(x,y)} \mid x \in \left(\frac{1}{4}, \frac{3}{4}\right), y \in \mathbb{R}\right\} & = & \left\{\{(x,y)\} \mid x \in \left(\frac{1}{4}, \frac{3}{4}\right), y \in \mathbb{R}\right\} \\ J_- & = & \left\{\varepsilon_{(x,y)} \mid x \in \left(-\frac{1}{4}, \frac{1}{4}\right), y \in \mathbb{R}\right\} & = & \left\{\{(x,y), (x+1,-y)\} \mid x \in \left(-\frac{1}{4}, \frac{1}{4}\right), y \in \mathbb{R}\right\} \end{array}$$

We have

$$\psi_-(J_+) = \left(\tfrac{1}{4}, \tfrac{3}{4}\right) \times \mathbb{R} \qquad \psi_+(J_+) = \left(\tfrac{1}{4}, \tfrac{3}{4}\right) \times \mathbb{R}$$

and $\psi_+ \circ \psi_-^{-1}$ is the identity on $\left(\frac{1}{4}, \frac{3}{4}\right) \times \mathbb{R}$, while

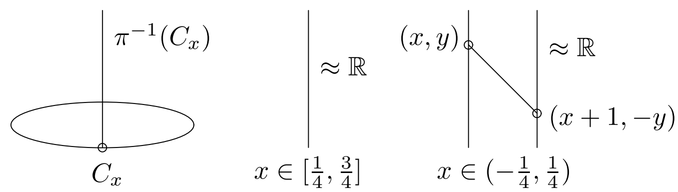
$$\psi_-(J_-) = \left(-\tfrac{1}{4}, \tfrac{1}{4}\right) \times \mathbb{R} \qquad \psi_+(J_-) = \left(\tfrac{3}{4}, \tfrac{5}{4}\right) \times \mathbb{R}$$

and $\psi_+ \circ \psi_-^{-1}$ on $(-\frac{1}{4}, \frac{1}{4}) \times \mathbb{R}$ is given by $(x,y) \mapsto (x+1,-y)$. Thus, both (ψ_+ and ψ_-) are C^∞ with C^∞-inverses and $\mathcal{A}$ is an atlas for $\mathcal{B}$. $\mathcal{B}$ is in fact a non-trivial line bundle over S^1 (non-trivial means not diffeomorphic to $S^1 \times \mathbb{R}$). We define the projection $\pi : \mathcal{B} \to S^1$ by

$$\pi(\varepsilon_{(x,y)}) = C_x$$

We have

$$\pi^{-1}(C_x) = \{\varepsilon_{(x,y)} \mid y \in \mathbb{R}\} = \begin{cases} \{\{(x,y) \mid y \in \mathbb{R}\} & x \in \left[\frac{1}{4}, \frac{3}{4}\right] \\ \{\{(x,y), (x+1,-y)\} \mid y \in \mathbb{R}\} & x \in \left(-\frac{1}{4}, \frac{1}{4}\right) \\ \{\{(x,y), (x-1,-y)\} \mid y \in \mathbb{R}\} & x \in \left(\frac{3}{4}, \frac{5}{4}\right) \end{cases}$$



So we have a vector space $V = \mathbb{R}$ which is a line bundle. In fact, we have

$$\pi^{-1}(U_-) = V_- \qquad \pi^{-1}(U_+) = V_+ \qquad \pi^{-1}(I_-) = J_- \qquad \pi^{-1}(I_+) = J_+$$

Moreover, we have the diffeomorphisms

$$\omega_- : V_- \to U_- \times \mathbb{R} \qquad \omega_+ : V_+ \to U_+ \times \mathbb{R}$$

given by the rule that

$$\omega_-(\varepsilon_{(x,y)}) = (C_x, y)$$

where (x,y) is the unique representative of $\varepsilon_{(x,y)}$ such that its first coordinate lies in $\left(-\frac{1}{4}, \frac{3}{4}\right)$ and

$$\omega_+(\varepsilon_{(x,y)}) = (C_x, y)$$

where (x,y) is the unique representative of $\varepsilon_{(x,y)}$ such that its first coordinate lies in $\left(\frac{1}{4}, \frac{5}{4}\right)$. Then, $\omega_-|_{\pi^{-1}(C_x)}$ is the linear isomorphism of $\pi^{-1}(C_x)$ onto $\mathbb{R}$ given by

$$\omega_-|_{\pi^{-1}(C_x)} \cdot \varepsilon_{(x,y)} = y \quad x \in \left(-\tfrac{1}{4}, \tfrac{3}{4}\right)$$

and similarly

$$\omega_+|_{\pi^{-1}(C_x)} \cdot \varepsilon_{(x,y)} = y \quad x \in \left(\tfrac{1}{4}, \tfrac{5}{4}\right)$$

We see that for $C_x \in I_+$

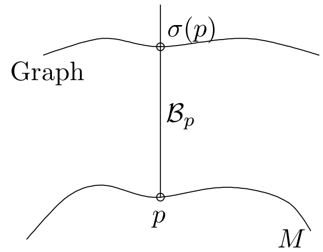
$$\omega_+|_{\pi^{-1}(C_x)} \cdot \left(\omega_-|_{\pi^{-1}(C_x)}\right)^{-1}$$

is the linear isomorphism of $\mathbb{R}$ onto itself ($y \mapsto y$, simply the identity), while for $C_x \in I_-$

$$\omega_+|_{\pi^{-1}(C_x)} \cdot \left(\omega_-|_{\pi^{-1}(C_x)}\right)^{-1}$$

is the linear isomorphism of $\mathbb{R}$ onto itself ($y \mapsto -y$, orientation reversing). So $\mathcal{B}$ is non-orientable (see later). $\square$

Definition 4.2: A continuously differentiable *section* of a vector bundle $\mathcal{B}$ over a differentiable manifold M is a continuously differentiable mapping $\sigma : M \to \mathcal{B}$ such that $\pi \circ \sigma = \mathrm{id}_M$



That is, $\sigma(p) \in \mathcal{B}_p$ for each $p \in M$. The Graph of σ is $\{\sigma(p) \in \mathcal{B} \mid p \in M\}$. □

Remark 4.2: If V, the model vector space, is $\mathbb{R}^m$, we get the local *basis sections* $\sigma_{\alpha,a}$, $a = 1, ..., m$ over U_α, using the diffeomorphisms $\omega_\alpha : \pi^{-1}(U_\alpha) \to U_\alpha \times V$, by

$$\sigma_{\alpha,a}(p) \;=\; \omega_{\alpha,p}^{-1} \cdot e_a$$

where $\omega_{\alpha,p} = \omega_\alpha|_{\mathcal{B}_p}$ is the linear isomorphism of $\mathcal{B}_p$ onto V and where $e_a = (0, ..., 1, ..., 0)$ with the 1 at the a-th position, so the e_a, $a = 1, ..., m$ are the standard basis of $\mathbb{R}^m$. These are *basis sections* over U_α, because the set of $(\sigma_{\alpha,a}(p),\ a = 1, ..., m)$ forms a basis for $\mathcal{B}_p$ at each $p \in U_\alpha$. Therefore, any section ψ of $\mathcal{B}$ over M can be expanded over U_α as

$$\psi \;=\; \sum_{\alpha=1}^{m} \psi_\alpha^a \sigma_{\alpha,a}$$

where the ψ_α^a, $a = 1, ..., m$ are real valued functions defined on U_α, which are continuously differentiable if ψ is continuously differentiable. This works similarly, if $V = \mathbb{C}^m$. In particular, when $V = \mathbb{C}$, there is one local basis section over U_α, given by

$$\sigma_\alpha(p) \;=\; \omega_{\alpha,p}^{-1} \cdot 1$$

Then, any section ψ of this complex line bundle can be expanded over U_α in the form

$$\psi \;=\; \psi_\alpha \sigma_\alpha$$

where ψ_α is a complex valued function defined on U_α, which is continuously differentiable if ψ is continuously differentiable. □

We now return to TM.

Definition 4.3: A continuously differentiable section X of TM is called a continuously differentiable *vectorfield* on M. It is $X(p) \in T_pM$ for each $p \in M$. □

Definition 4.4: We may also define a C^∞-vectorfield X on M as a *linear operator* on $C^\infty(M)$, the space of C^∞-functions on M. If $f \in C^\infty(M)$ then the function Xf is defined by

$$(Xf)(p) \;=\; X(p) \cdot f \quad \forall p \in M$$

So $Xf \in C^\infty(M)$. □

Remark 4.3: However, not all linear operators on $C^\infty(M)$ correspond to vectorfields. Vectorfields are distinguished or characterized by the fact, that they satisfy the Leibnitz rule. That is

$$X(fg) \;=\; f(Xg) + g(Xf) \quad \forall f, g \in C^\infty(M)$$

Let X, Y be C^∞-vectorfields on M. Then XY and YX are both linear operators on $C^\infty(M)$

$$(XY)f \;=\; X(Yf) \qquad (YX)f \;=\; Y(Xf) \qquad \forall f \in C^\infty(M)$$

but not vectorfields. However, the commutator is indeed a vectorfield. □

Proposition 4.1: The commutator

$$[X,Y] \;:=\; XY - YX$$

is a vectorfield. □

Proof: To show this, we have to verify the Leibnitz rule

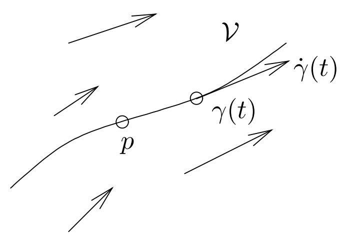
$$\begin{aligned}
&(XY-YX)(fg) \\
&\quad= X(f(Yg)+g(Yf)) - Y(f(Xg)+g(Xf)) \\
&\quad= fXYg + (Yg)(Xf) + gXYf + (Yf)(Xg) - fYXg - (Xg)(Yf) - gYXf - (Xf)(Yg) \\
&\quad= f(XY-YX)g + g(XY-YX)f
\end{aligned}$$

(QED)

4.1 Integral Curves

Definition 4.5: Suppose that M is a differentiable manifold and $\mathcal{V}$ a continuously differentiable vectorfield on M. An *integral curve* of $\mathcal{V}$ through the point $p \in M$ is a continuously differentiable curve $\gamma : I \to M$, where I is an open interval containing 0, such that $\gamma(0) = p$ and for each $t \in I$, the tangent vector $\dot\gamma(t)$ to γ at $\gamma(t)$ is the vector $\mathcal{V}(\gamma(t))$

$$\dot\gamma(t) \;=\; \mathcal{V}(\gamma(t)) \qquad \dot\gamma(t)\cdot f \;=\; \left.\frac{\mathrm{d}(f\circ\gamma)}{\mathrm{d}t'}\right|_{t'=t}$$



□

Remark 4.4: Let (U,φ) be a local chart at p. Then γ is represented by the curve

$$\tilde\gamma = \varphi\circ\gamma : J \to \mathbb{R}^n$$

a curve in $\mathbb{R}^n$, where $J \subset \gamma^{-1}(U)$ is an open subinterval of I containing 0. We have $\tilde\gamma = (\tilde\gamma^1, ..., \tilde\gamma^n)$, where the $\tilde\gamma^i$ are real functions defined on J. Moreover, a vectorfield $\mathcal{V}(q)$ for $q \in U$ is expressed in terms of the chart φ by

$$\mathcal{V}(q) \;=\; \sum_{i=1}^{n} \mathcal{V}^i(q) \left.\frac{\partial}{\partial x^i}\right|_q$$

where the $\mathcal{V}^i$, the components of $\mathcal{V}$ in the chart, are functions defined on U. These functions are represented by the functions

$$\tilde{\mathcal{V}}^i \;=\; \mathcal{V}^i \circ \varphi^{-1}$$

real functions defined on the open set $\varphi(U) = V \subset \mathbb{R}^n$. Thus, $\mathcal{V}$ is represented in U by the vectorfield $\tilde{\mathcal{V}}$ on V, given by

$$\tilde{\mathcal{V}}(x) \;=\; \sum_{i=1}^{n} \tilde{\mathcal{V}}^i(x) \left.\frac{\partial}{\partial x^i}\right|_x \qquad x \in V$$

where $V \subset \mathbb{R}^n$ is open. Here $\frac{\partial}{\partial x^i}\big|_x$ is the standard partial derivative with respect to x^i at $x \in \mathbb{R}^n$. Then, $\tilde\gamma$ is an integral curve of $\tilde{\mathcal{V}}$, if

$$\dot{\tilde\gamma}(t) \;=\; \tilde{\mathcal{V}}(\tilde\gamma(t)) \quad \forall t \in J$$

Since

$$\dot{\tilde\gamma}(t) \;=\; \sum_{i=1}^{n} \frac{\mathrm{d}\tilde\gamma^i}{\mathrm{d}t}(t) \left.\frac{\partial}{\partial x^i}\right|_{\tilde\gamma(t)}$$

we can express the above in components as

$$\frac{\mathrm{d}\tilde{\gamma}^i}{\mathrm{d}t}(t) \;=\; \tilde{\mathcal{V}}^i(\tilde{\gamma}(t)) \quad i = 1, ..., n$$

This is a system of ordinary differential equations (o.d.e.). The initial condition is

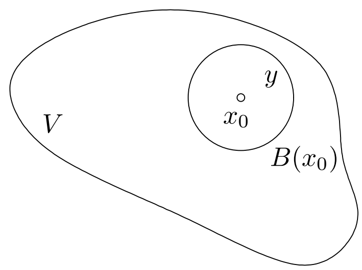
$$\tilde{\gamma}(0) \;=\; \varphi(\gamma(0)) \;=\; \varphi(p)$$

So the problem of finding the integral curves of a vector field is the problem of finding the solutions of a system of ordinary differential equations. □

4.2 Digression into Ordinary Differential Equations

Theorem 4.1: Local Existence Theorem
Let $V \subset \mathbb{R}^n$ be an open subset and let $\mathcal{V}$ be a C^1 vectorfield on V. Consider the problem of finding the integral curve of $\mathcal{V}$ through y

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) \;=\; V(x(t)) \qquad x(0) \;=\; y \in V$$



Then for any $x_0 \in V$ there is a ball $B(x_0) \subset V$ and an open interval $I(x_0)$, $0 \in I(x_0)$, such that for every $y \in B(x_0)$ the problem has a solution $x(t; y)$ defined on $I(x_0)$, which is C^2 in t at each y and C^1 in y at each t. □

Theorem 4.2: Global Uniqueness Theorem
For every $y \in V$ the problem has a unique maximal solution, defined on a maximal open interval $J(y)$, $0 \in J(y)$. □

Remark 4.5: $J(y)$ may be of the form $(-a, \infty)$ or $(-\infty, a)$ with $a > 0$ or also $\mathbb{R}$, where we have a global solution for the initial condition y. If $y \in B(x_0)$ as in the local existence theorem, then $J(x_0) \subset I(y)$. □

Theorem 4.3: Key Theorem
Set $T_+(y) = \sup J(y)$, which is the right end point of $J(y)$ or ∞. Then the following alternative holds. Either $T_+(y) = \infty$ or there is no compact subset $K \subset V$ such that the motion in $[0, T_+(y))$ is contained in K. A similar statement holds with respect to $T_-(y) = \inf J(y)$, the left end point of $J(y)$ or $-\infty$. □

Example 4.2: Take $n = 1$ and $V = \mathbb{R}$. Define $\mathcal{V}(x) = x^2$. The problem to solve may be

$$\frac{\mathrm{d}x}{\mathrm{d}t} \;=\; x^2 \qquad x(0) \;=\; y$$

We have two cases:

1) $y = 0$: Then $x(t) = 0 \;\forall t$, which is a global solution, that is $J(y) = \mathbb{R}$.

2) $y \neq 0$: Then

$$x(t) \;=\; \frac{1}{\frac{1}{y} - t}$$

You see that this solution is not defined for $t = \frac{1}{y}$. So for $y > 0$, we get $J(y) = (-\infty, \frac{1}{y})$, $T_+(y) = \frac{1}{y}$. On the other hand, for $y < 0$ we get $J(y) = (\frac{1}{y}, \infty)$, $T_-(y) = \frac{1}{y}$. The image of $[0, \frac{1}{y}]$ here is $[y, \infty)$, which is not compact.

□

Theorem 4.4: Let $\mathcal{V}$ be a C^1 vectorfield on $\mathbb{R}^n$ such that there are positive constants A and B with

$$|\mathcal{V}(x)| \leq A|x| + B$$

Then we have a global solution for all initial conditions $y \in \mathbb{R}^n$. □

Proof: According to the Key Theorem, either $T_+(y) = \infty$ or it is finite in which case the motion in $[0, T_+(y)]$ is not contained in any compact subset $K \subset \mathbb{R}^n$. Consider

$$|x|^2 = \sum_{i=1}^{n} (x^i)^2$$

We have

$$2|x| \frac{\mathrm{d}}{\mathrm{d}t}|x| = \sum_{i=1}^{n} 2x^i \frac{\mathrm{d}x^i}{\mathrm{d}t} = \sum_{i=1}^{n} 2x^i \mathcal{V}^i(x) = 2\langle x, \mathcal{V}(x)\rangle \leq 2|x|\,|\mathcal{V}(x)| \leq 2|x|\,(A|x| + B)$$

Hence

$$\frac{\mathrm{d}|x|}{\mathrm{d}t} \leq A|x| + B \quad \rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-At}|x|\right) \leq Be^{-At}$$

Now integrate on $[0, t)$ with $t > 0$ to get

$$e^{-At}|x(t)| - |y| \leq \frac{B}{A}\left(1 - e^{-At}\right)$$

That is

$$|x(t)| \leq |y|e^{At} + \frac{B}{A}\left(e^{At} - 1\right) \leq |y|e^{AT_+(y)} + \frac{B}{A}\left(e^{AT_+(y)} - 1\right) =: R$$

So the motion in $[0, T_+(y)]$ is contained in the closed ball of radius R in $\mathbb{R}^n$. This is a compact set. So we have a contradiction. This shows that only the first alternative is possible, hence $T_+(y) = \infty$. Similarly, we establish that $T_y(y) = -\infty$ (see Explanation A.2 on page 85). (QED)

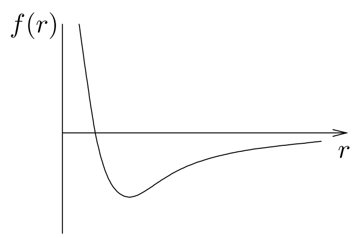
Example 4.3: Take N particles in $E^3 = \mathbb{R}^3$. We have the configuration $(\mathbf{x}_1, ..., \mathbf{x}_N) \in (E^3)^N = \mathbb{R}^{3N}$, where $\mathbf{x}_\alpha \in \mathbb{R}^3$, $\alpha = 1, ..., N$. Let $m_\alpha > 0$ be the mass of the paricle α, $\mathbf{y}_\alpha \in \mathbb{R}^3$ its momentum. The Hamiltonian is the following function on the phase space $(E^3)^N \times \mathbb{R}^{3N} = \mathbb{R}^{6N}$

$$H = \sum_{\alpha=1}^{N} \frac{|\mathbf{y}_\alpha|^2}{2m_\alpha} + V(\mathbf{x}_1, ..., \mathbf{x}_N)$$

where V is the potential, a function on the configuration space which is invariant under the Euclidean group. We can get a class of examples, by taking

$$V(\mathbf{x}_1, ..., \mathbf{x}_N) = \sum_{\alpha<\beta} q_\alpha q_\beta f(|\mathbf{x}_\alpha - \mathbf{x}_\beta|)$$

where the q_α are generalized charges. Take $q_\alpha = m_\alpha$ and $f(r) = -\frac{1}{r}$, which is the Newtonian problem. Or take the case where the q_α are electric charges, in which we have $f(r) = \frac{1}{r}$. We could also think of the example $q_\alpha = 1$, $m_\alpha = m > 0$, with



where the configuration space is $C := \mathbb{R}^3 \setminus \Delta$ with

$$\Delta := \{(\mathbf{x}_1, ..., \mathbf{x}_N) \in \mathbb{R}^{3N} | \mathbf{x}_\alpha = \mathbf{x}_\beta \text{ for some } \alpha \neq \beta\}$$

The canonical equations are

$$\frac{\mathrm{d}\mathbf{x}_\alpha}{\mathrm{d}t} = \frac{\partial H}{\partial \mathbf{y}_\alpha} \qquad \frac{\mathrm{d}\mathbf{y}_\alpha}{\mathrm{d}t} = -\frac{\partial H}{\partial \mathbf{x}_\alpha}$$

for any Hamiltonian. We are considering the integral curves of the Hamiltonian vectorfield

$$\mathcal{V} = \sum_\alpha \left(\frac{\partial H}{\partial \mathbf{y}_\alpha} \frac{\partial}{\partial \mathbf{x}_\alpha} - \frac{\partial H}{\partial \mathbf{x}_\alpha} \frac{\partial}{\partial \mathbf{y}_\alpha} \right)$$

where we get with the Hamiltonian from above

$$\frac{\partial H}{\partial \mathbf{y}_\alpha} = \frac{\mathbf{y}_\alpha}{m_\alpha} = \frac{\mathrm{d}\mathbf{x}_\alpha}{\mathrm{d}t} \qquad \frac{\partial H}{\partial \mathbf{x}_\alpha} = \frac{\partial V}{\partial \mathbf{x}_\alpha} = \frac{\mathrm{d}\mathbf{y}_\alpha}{\mathrm{d}t}$$

□

Theorem 4.5: Let the potential V in the Hamiltonian from above be a C^2 function on the configuration space $(E^3)^N = \mathbb{R}^{3N}$ which is bounded from below, that means that there is a positive constant B such that

$$V(\mathbf{x}_1, ..., \mathbf{x}_N) \geq -B$$

Then there is a global solution of the canonical equations for all initial data. □

Proof: We apply the key theorem. Suppose that for some initial condition, T_+ is finite. Consider the Hamiltonian function along the integral curve on $[0, T_+)$. For any Hamiltonian system, H is constant along the integral curves, that is

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \mathcal{V}H = \sum_\alpha \left(\frac{\partial H}{\partial \mathbf{y}_\alpha} \frac{\partial H}{\partial \mathbf{x}_\alpha} - \frac{\partial H}{\partial \mathbf{x}_\alpha} \frac{\partial H}{\partial \mathbf{y}_\alpha} \right) = 0$$

where $\mathcal{V}H$ is the vectorfield from above applied to H. Therefore, we have $H(t) = H(0)\ \forall t \in [0, T_+)$, which means that it is a real number depending only on the initial conditions. We can write

$$\sum_\alpha \frac{|\mathbf{y}_\alpha|^2}{2m_\alpha} = H(0) - V \leq H(0) + B$$

because $-V \leq B$. So it is smaller or equal to a fixed real number. Let $m_M = \max\{m_\alpha | \alpha = 1, ..., N\}$ be the largest mass. Then we get that

$$|\mathbf{y}| \leq \sqrt{2m_M\,(H(0) + B)}$$

where we defined

$$|\mathbf{y}| := |(\mathbf{y}_1, ..., \mathbf{y}_N)| = \sqrt{\sum_\alpha |\mathbf{y}_\alpha|^2}$$

To bound the velocities, let $m_m = \min\{m_\alpha | \alpha = 1, ..., N\} > 0$ be the smallest mass. Then

$$\frac{|\mathbf{y}_\alpha|}{m_\alpha} \leq \frac{|\mathbf{y}_\alpha|}{m_m}$$

Considering

$$|\mathbf{x}|^2 = \sum_\alpha |\mathbf{x}_\alpha|^2$$

we have

$$2|\mathbf{x}| \frac{\mathrm{d}}{\mathrm{d}t}|\mathbf{x}| = \frac{\mathrm{d}}{\mathrm{d}t}|\mathbf{x}|^2 = \sum_\alpha 2 \left\langle \mathbf{x}_\alpha, \frac{\mathrm{d}\mathbf{x}_\alpha}{\mathrm{d}t} \right\rangle \leq 2\sqrt{\sum_\alpha |\mathbf{x}_\alpha|^2} \cdot \sqrt{\sum_\alpha \left| \frac{\mathrm{d}\mathbf{x}_\alpha}{\mathrm{d}t} \right|^2}$$

It follows that

$$\frac{\mathrm{d}|\mathbf{x}|}{\mathrm{d}t} \leq \sqrt{\sum_\alpha \left| \frac{\mathrm{d}\mathbf{x}_\alpha}{\mathrm{d}t} \right|^2} \leq \frac{1}{m_m} \sqrt{\sum_\alpha |\mathbf{y}_\alpha|^2} = \frac{|\mathbf{y}|}{m_m} \leq \frac{\sqrt{2m_M(H(0) + B)}}{m_m}$$

We integrate this over $[0, t]$ with $t \in (0, T_+)$ to obtain

$$|x(t)| \leq |x(0)| + \frac{\sqrt{2m_M(H(0)+B)}}{m_m} t \leq |x(0)| + \mathcal{C}T_+ \qquad \mathcal{C} := \frac{\sqrt{2m_M(H(0)+B)}}{m_m}$$

This shows that on $[0, T_+)$ the motion is contained in a compact set

$$K = \overline{B}_R \times \overline{B}_Q$$

where $R := |x(0)| + \mathcal{C}T_+$ and $Q := \mathcal{C}m_m$. This contradicts the Key theorem, so $T_+ = \infty$. With a similar argument, it follows that $T_- = -\infty$, what makes the existance interval to be whole $\mathbb{R}$. (QED)

Remark 4.6: We can show that with $r_m(t) = \min_{\alpha<\beta} |x_\alpha(t) - x_\beta(t)|$, $r_m(t) \to 0$ as $t \to T_+$. So, it somehow collides. There is still a discussion, whether the probability for a collision is bigger than zero or actually zero. □

Definition 4.6: A vectorfield $\mathcal{V}$ on a manifold M is *complete* if each integral curve of $\mathcal{V}$ is defined for all time. □

Remark 4.7: Given a point $p \in M$, let us denote by $I(p)$ the maximum interval of existence with initial condition p. Let then $\mathcal{V}$ be a C^1 vectorfield on M and let

$$\mathcal{A} := \{(U_\alpha, \varphi_\alpha) | \alpha \in I\}$$

be an atlas for M. For $p \in U_\alpha$ we then obtain the integral curve of $\mathcal{V}$ through p by considering the integral curve of the vectorfield $\tilde{\mathcal{V}}$, representing $\mathcal{V}$ in the chart φ_α (see Remark 4.4), which is a vectorfield defined on the open set $V_\alpha = \varphi_\alpha(U_\alpha) \subset \mathbb{R}^n$, through $y = \varphi_\alpha(p) \in V_\alpha$. The corresponding integral curve $\tilde{\gamma}$ is defined on and contained in V_α for $t \in I(y)$, the maximal interval of existance with initial condition y. Then

$$\gamma = \varphi_\alpha^{-1} \circ \tilde{\gamma}$$

is an integral curve of $\mathcal{V}$ through p defined in $I(y)$ and contained in U_α. Therefore $I(y) \subset I(p)$. □

Remark 4.8: By the Local Existence Theorem, there is a ball $B(y) \subset V_\alpha$ for each $y \in V_\alpha$ such that $\forall y' \in B(y)$, we have that $[-\varepsilon(y), \varepsilon(y)] \subset I(y')$ where $\varepsilon(y) > 0$ and depends only on y, not on y'. Now, any point $p \in M$ belongs to some U_α. Consider $W_p = \varphi_\alpha^{-1}(B(y))$ and recall that $y = \varphi_\alpha(p)$. Then W_p is a nieghborhood of p in M and for every $q \in W_p$, the motion with initial condition q is defined $\forall t \in [-\varepsilon_p, \varepsilon_p]$, where $\varepsilon_p = \varepsilon(y)$. Here, $\varphi_\alpha(q)$ is what we denoted above by y'. □

Definition 4.7: A 1-parameter family $\{\phi_t : M \to M | t \in \mathbb{R}\}$ of diffeomorphisms of M having the property

$$\phi_0 = \text{id} \qquad \phi_s \circ \phi_t = \phi_{s+t} \quad \forall s, t \in \mathbb{R}$$

is called a 1-*parameter group* of diffeomorphisms of M. The path $\{\phi_t(p) | t \in \mathbb{R}\} \subset M$ is called the *orbit* of the group through p. □

Theorem 4.6: If M is a compact manifold, then any vectorfield $\mathcal{V}$ on M is complete. □

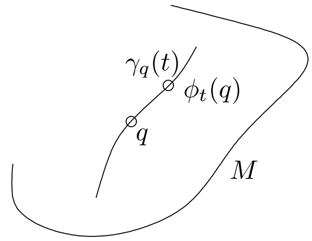
Proof: Our manifold M is compact, so for every open cover of M there is a finite subcover. The collection of open sets $\{W_p | p \in M\}$ covers M. By the compactness of M, a finite subcollection suffices to cover M, so take $\{W_{p_i} | i = 1, ..., N\}$. Set

$$\varepsilon := \min_{i=1,\dots,N} \{\varepsilon_{p_i}\} > 0$$

Then for each $q \in M$ the motion with initial condition q is defined $\forall t \in [-\varepsilon, \varepsilon]$. Let us then set

$$\phi_t(q) = \gamma_q(t) \quad t \in [-\varepsilon, \varepsilon],\ q \in M$$

the integral curve through q, that is $\gamma_q(0) = q$.



Then, following from the local existence theorem, ϕ_t is a C^1 mapping of M onto itself. It is also $\phi_{-t} = \phi_t^{-1}$, which means that ϕ_t is a diffeomorphism of M onto itself. More generally, if $s, t, s+t \in [-\varepsilon, \varepsilon]$ then

$$\phi_s \circ \phi_t(q) \quad = \quad \phi_s(\phi_t(q)) \quad = \quad \gamma_{\phi_t(q)}(s) \quad = \quad \gamma_{\gamma_q(t)}(s) \quad = \quad \gamma_q(s+t) \quad = \quad \phi_{s+t}(q)$$

We now extend ϕ_t so that t takes all real values. Given any $t \in \mathbb{R}$, we write

$$t \quad = \quad k\varepsilon + s \quad k \quad = \quad \left[\frac{t}{\varepsilon}\right]$$

where by $[x]$ we denote the integral part of the real number x. Then $s \in [0, \varepsilon)$ and ϕ_s is already defined. If $k > 0$, we set

$$\phi_{k\varepsilon} \quad = \quad \underbrace{\phi_\varepsilon \circ \ldots \circ \phi_\varepsilon}_{k \text{ folds}} \qquad \phi_{k\varepsilon} \quad = \quad \underbrace{\phi_{-\varepsilon} \circ \ldots \circ \phi_{-\varepsilon}}_{-k \text{ folds}}$$

We then set

$$\phi_t \quad = \quad \phi_{k\varepsilon} \circ \phi_s$$

in order to define ϕ_t for all $t \in \mathbb{R}$. Moreover ϕ_t, defined like this, satisfies

$$\phi_0 \quad = \quad \text{id} \qquad \phi_s \circ \phi_t \quad = \quad \phi_{s+t} \quad \forall s, t \in \mathbb{R}$$

So we have a 1-*parameter group* of diffeomorphisms of M. For any point $p \in M$, the integral curve starting at p is given by

$$\gamma_p(t) \quad = \quad \phi_t(p) \quad \forall t \in \mathbb{R}$$

that is by the *orbit* of the group through p. Thus the motion is defined for all time and for any initial condition. Hence the vectorfield $\mathcal{V}$ is complete. (QED)

Remark 4.9: This is not true for non-compact manifolds. An example would be $M = \mathbb{R}$ with the vectorfield $V(x) = x^2$. However, a complete vectorfield $\mathcal{V}$ on a manifold M (compact or not) generates a 1-parameter group ϕ_t of diffeomorphisms of M called the *flow* of $\mathcal{V}$. □

4.3 The cotangent bundle

Let M and N be differentiable manifolds (not necessarily of the same dimension) and $\psi : M \to N$ a differentiable mapping. We assume for simplicity that ψ is C^∞.

Definition 4.8: The *differential* of ψ at $p \in M$ denoted $\mathrm{d}\psi(p)$ is the linear mapping of T_pM into T_qN (where $q = \psi(p)$) given by the property that if $u \in T_pM$, then $v = \mathrm{d}\psi(p) \cdot u$ is defined by

$$v \cdot f \quad = \quad u \cdot (f \circ \psi) \quad \forall f \in C^\infty(N)$$

where $f \circ \psi \in C^\infty(M)$. □

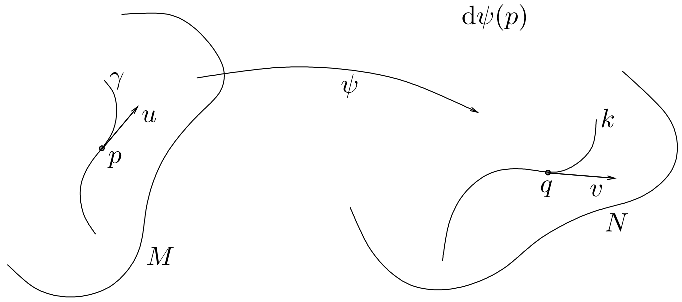
Remark 4.10: Let u be the tangent vector at p to a curve γ in M through p. Thus $\forall g \in C^\infty(M)$ we have

$$u \cdot g \quad = \quad \frac{\mathrm{d}}{\mathrm{d}t} g \circ \gamma \bigg|_{t=0}$$

Setting $g = f \circ \psi$ in the definition above, we obtain $\forall t \in C^\infty(N)$

$$v \cdot f \quad = \quad \frac{\mathrm{d}}{\mathrm{d}t} f \circ \psi \circ \gamma \bigg|_{t=0} \quad = \quad \frac{\mathrm{d}}{\mathrm{d}t} f \circ k \bigg|_{t=0}$$

where $k = \psi \circ \gamma$ is a curve in N through q and therefore the image by ψ of the curve γ. Thus v is the tangent vector at q to k.



In the special case that $N = \mathbb{R}$, the differential of a C^∞-function ψ on M at $p \in M$, denoted $\mathrm{d}\psi(p)$, is a linear function on T_pM given by

$$\mathrm{d}\psi(p) \cdot u \;=\; u \cdot \psi \quad \forall u \in T_pM \tag{4.1}$$

and $\mathrm{d}\psi$ is in fact a differentiable section of T^*M, the cotangent bundle of M. See also Explanation A.3 on page 85. □

Definition 4.9: The *cotangent bundle* of M is defined by

$$T^*M \;=\; \bigcup_{p \in M} T_p^*M$$

where T_p^*M, the *cotangent space* at p, is the dual space of T_pM, that is

$$T_p^*M \;=\; (T_pM)^*$$

□

Remark 4.11: In general, if V is a real finite dimensional vector space, the dual space V^* is the space of real linear functions on V. Similarly in the complex case. □

Definition 4.10: In general, if $\mathcal{B}$ is a vector bundle over M, the bundle

$$\mathcal{B}^* \;=\; \bigcup_{p \in M} B_p^*$$

where $\mathcal{B}_p^*$ is the dual vector space to $\mathcal{B}_p$, is called the *dual bundle* to $\mathcal{B}$. □

Definition 4.11: Let ι_a, $a = 1, ..., n = \dim \mathcal{B}_p$ be local basis sections of $\mathcal{B}$ over $U \subset M$. Then the *dual local basis sections* of $\mathcal{B}^*$ over U are defined by the conditions $\iota^{*a} \cdot \iota_b = \delta^a_b$ in U. □

Definition 4.12: A *continuously differentiable* 1*-form* on M is a continously differentiable section of T^*M. □

Remark 4.12: If θ is a 1-form on M and X a vectorfield on M, then $\theta \cdot X$ is a function on M. If f is a continuously differentiable function on M, then $\mathrm{d}f$, the differential of f, is a 1-form on M. □

Definition 4.13: We use the following notations, to simplify some expressions

- $C^\infty(M)$ denotes the *space of smooth functions* on M.
- $\mathscr{X}^\infty(M)$ denotes the *space of smooth vectorfields* on M.
- $\Omega_1^\infty(M)$ denotes the *space of smooth* 1*-forms* on M.

□

Remark 4.13: Given an $f \in C^\infty(M)$ and an $X \in \mathscr{X}^\infty(M)$, we get that $fX \in \mathscr{X}^\infty(M)$ (where fX is defined pointwise) and given a $\theta \in \Omega_1^\infty(M)$, we get that $f\theta \in \Omega_1^\infty(M)$ (where $f\theta$ is also defined pointwise). □

Remark 4.14: If U is the domain of a chart, then $\iota_\mu = \frac{\partial}{\partial x^\mu}$, $\mu = 1, ..., m$, $m = \dim M$ are local vectorfields which are the local basis sections of TM over U. The dual basis sections of T^*M is given by the $i^{*\nu}$, $\nu = 1, ..., m$, where $\iota^{*\nu} = \mathrm{d}x^\nu$ (here we defined $x^\nu := \varphi^\nu$ as the components of the chart φ). We have

$$\mathrm{d}x^\nu \cdot \frac{\partial}{\partial x^\mu} \quad = \quad \frac{\partial}{\partial x^\mu} \cdot x^\nu \quad = \quad \frac{\partial x^\nu}{\partial x^\mu} \quad = \quad \delta^\nu_\mu \tag{4.2}$$

See also Explanation A.4 on page 86. Every 1-form θ can then be locally expanded as

$$\theta \quad = \quad \sum_{\nu=1}^{m} \theta_\nu \, \mathrm{d}x^\nu$$

where the coefficients θ_ν are functions defined in U. Recalling that given a vectorfield X we have the expansion

$$X \quad = \quad \sum_{\mu=1}^{m} X^\mu \frac{\partial}{\partial x^\mu}$$

in U, then the function $\theta \cdot X$ is expressed in U as

$$\theta \cdot X \quad = \quad \sum_{\mu=1}^{m} \theta_\mu X^\mu \quad = \quad \theta_\mu X^\mu$$

□

5 Lie Derivatives

5.1 Pull-Back and Push-Forward

Pull-back applies to covariant objects. Let M and N be differentiable manifolds with $\dim M = m$, $\dim N = n$ and let $\phi : M \to N$ be a continuously differentiable mapping. The simplest covariant object is a function. Consider a continuously differentiable function f on N (the target). Then $f \circ \phi$ is a continuously differentiable function on M (the domain).

Definition 5.1: We call $f \circ \phi$ the *pull-back* by ϕ of f and write

$$\phi^* f \;=\; f \circ \phi$$

□

A 1-form ist the next simplest covariant structure.

Definition 5.2: Let θ be a 1-form on N (the target). Then, the *pull-back* of θ by ϕ, denoted $\phi^*\theta$, is the 1-form on M (the domain) defined by

$$(\phi^*\theta) \cdot v \;=\; \theta \cdot (\mathrm{d}\phi \cdot v) \quad \forall v \in T_pM,\, \forall p \in M$$

□

Remark 5.1: Recall that $\mathrm{d}\phi(p)$ is a linear map of T_pM into $T_{\phi(p)}N$. In terms of local coordinates x^μ, $\mu = 1, ..., m$ for M and y^a, $a = 1, ..., n$ for N

$$\theta \;=\; \sum_{a=1}^{n} \theta_a \, \mathrm{d}y^a \qquad \phi^*\theta \;=\; \sum_{\mu=1}^{m} (\phi^*\theta)_\mu \, \mathrm{d}x^\mu$$

where with ϕ given by $y^a = \phi^a(x^1, ..., x^m)$, we have

$$(\phi^*\theta)_\mu(p) \;=\; \sum_{a=1}^{n} \theta_a(\phi(p)) \left.\frac{\partial y^a}{\partial x^\mu}\right|_p \tag{5.1}$$

See Explanation A.5 on page 86. □

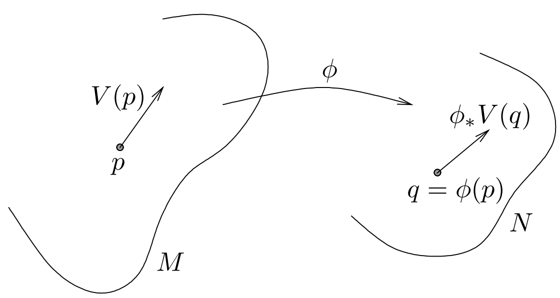
Push-Forward applies to contravariant objects. The simplest contravariant object is a vectorfield. We have again a continuously differentiable $\phi : M \to N$. We now require ϕ to be one to one. Then for each $q \in \phi(M) \subset N$, there is a unique $p \in M$ such that $\phi(p) = q$.

Definition 5.3: Consider a vectorfield V on M (the domain). The *push-forward* $\phi_* V$ of V by ϕ is the vectorfield defined along the image $\phi(M) \subset N$ (the target) by

$$(\phi_* V)(\phi(p)) \;=\; \mathrm{d}\phi \cdot V(p) \quad \forall p \in M$$

That is

$$(\phi_* V)(q) \;=\; \mathrm{d}\phi \cdot V(\phi^{-1}(q)) \quad \forall q \in \phi(M) \subset N$$



□

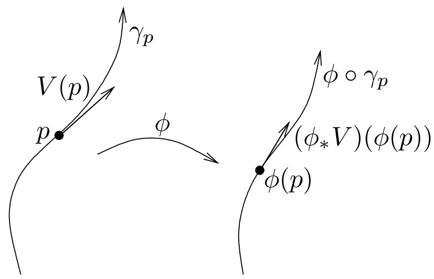
Remark 5.2: To obtain the analogue of a pull-back for a vectorfield (more generally, for a contravariant object) we consider the push-forward by ϕ^{-1}

$$(\phi_*^{-1}V)(p) \;=\; \mathrm{d}\phi^{-1}\cdot V(\phi(p)) \quad \forall p\in M$$

□

Remark 5.3: Note that the definition requires ϕ to be one to one. From now on we restrict ourselves in discussing the push-forward, to the case $M=N$ and ϕ being a diffeomorphism of M onto itself. Then V being a smooth vectorfield on M implies that ϕ_*V is also a smooth vectorfield on M □

Remark 5.4: If γ_p is the integral curve of V through p, then $\phi\circ\gamma_p$ is the integral curve of ϕ_*V through $\phi(p)$ (by definition).



Suppose now that V is a complete voctorfield. Then V generates a 1-parameter group ψ_t of diffeomorphisms of M and

$$\gamma_p(t) \;=\; \psi_t(p)$$

Let $\tilde{\psi}_t$ be the 1-parameter group of diffeomorphisms generated by ϕ_*V. Then we have

$$\tilde{\psi}_t(\phi(p)) \;=\; (\phi\circ\gamma_p)(t) \;=\; \phi(\psi_t(p)) \quad \forall p\in M,\,\forall t\in\mathbb{R}$$

That is

$$\tilde{\psi}_t\circ\phi \;=\; \phi\circ\psi_t \quad \forall t\in\mathbb{R}$$

or

$$\tilde{\psi}_t \;=\; \phi\circ\psi_t\circ\phi^{-1} \quad \forall t\in\mathbb{R}$$

That is, the 1-parameter group generated by ϕ_*V is obtained from the 1-parameter group generated by V through conjugation by ϕ. In particular

$$\phi_*V \;=\; V \quad\Leftrightarrow\quad \psi_t\circ\phi \;=\; \phi\circ\psi_t \quad \forall t\in\mathbb{R}$$

□

5.2 Lie Deriatives

Definition 5.4: Let X be a complete vectorfield on M generating the 1-parameter group ϕ_t of diffeomorphisms on M. Given a function f or a 1-form θ on M we consider the pull-backs ϕ_t^*f, $\phi_t^*\theta$ and define the *Lie derivative* of f and θ with respect to X by

$$\begin{aligned}
\pounds_X f &= \lim_{t\to 0}\left(\frac{\phi_t^* f - f}{t}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\,\phi_t^* f\bigg|_{t=0}\\
\pounds_X \theta &= \lim_{t\to 0}\left(\frac{\phi_t^* \theta - \theta}{t}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\,\phi_t^* \theta\bigg|_{t=0}
\end{aligned}$$

□

Remark 5.5: Now, since $\phi_t^*f = f\circ\phi_t$, we have that

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\phi_t^* f\bigg|_{t=0}(p) \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\,f(\phi_t(p))\bigg|_{t=0} \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\,f(\gamma_p(t))\bigg|_{t=0} \;=\; X(p)\cdot f \quad \forall p\in M$$

where the tangent vector to γ_p at p is $X(p)$. We conclude that

$$\pounds_X f \;=\; Xf$$

□

Definition 5.5: Let now Y be another vectorfield on M. We define the *Lie derivative* of Y with respect to X by

$$\pounds_X Y \;=\; \lim_{t\to\infty}\left(\frac{\phi_{-t*}Y - Y}{t}\right) \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\,\phi_{-t*}Y\Big|_{t=0}$$

where ϕ_t is generated by X. □

Proposition 5.1: The Lie derivative of a vectorfield Y with respect to X satisfies

$$\pounds_X Y \;=\; [X,Y]$$

So it's nothing else than the usual commutator. □

Proof: Given any function $f \in C^\infty(M)$ we have

$$(\pounds_X Y)\cdot f \;=\; \lim_{t\to 0}\,(t^{-1}E_t)\cdot f \qquad E_t \;:=\; \phi_{-t*}Y - Y$$

so we get that $\forall p \in M$

$$E_t(p)\cdot f \;=\; \mathrm{d}\phi_{-t}\cdot Y(\phi_t(p))\cdot f - Y(p)\cdot f \;=\; Y(\phi_t(p))\cdot(f\circ\phi_{-t}) - Y(p)\cdot f$$

We will use this equation later. Now we write

$$f\circ\phi_{-t} - f \;=\; t\,\frac{1}{t}\int_0^t \frac{\mathrm{d}}{\mathrm{d}t'}\,f\circ\phi_{-t'}\,\mathrm{d}t' \;=:\; t g_t$$

where we defined g_t to be the mean value of $\frac{\mathrm{d}}{\mathrm{d}t'}\,f\circ\phi_{-t'}$ on $[0,t]$. Then $g_t \in C^\infty(M)$ as well as continuous in t and

$$g_0 \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\,f\circ\phi_{-t}\Big|_{t=0} \;=\; -Xf$$

We can then write

$$E_t(p)\cdot f \;=\; Y(\phi_t(p))\cdot f - Y(p)\cdot f + tY(\phi_t(p))\cdot g_t$$

where we get that

$$\lim_{t\to 0} t^{-1}\big(Y(\phi_t(p))\cdot f - Y(p)\cdot f\big) \;=\; \lim_{t\to 0} t^{-1}[(Yf)(\phi_t(p)) - (Yf)(p)] \;=\; (XYf)(p)$$

while

$$\lim_{t\to 0} Y(\phi_t(p))\cdot g_t \;=\; Y(p)\cdot g_0 \;=\; (Yg_0)(p) \;=\; -(YXf)(p)$$

We conclude that

$$\lim_{t\to 0}\,(t^{-1}E_t(p))\cdot f \;=\; (XYf)(p) - (YXf)(p) \;=\; ([X,Y]f)(p)$$

(QED)

Proposition 5.2: Let θ be a 1-form on M, Y an arbitrary vectorfield on M and ϕ a diffeomorphism of M onto itself. Furthermore let X be a complete vectorfield on M generating the 1-parameter group ϕ_t. Then we have the following formula for the Lie derivative of a 1-form

$$(\pounds_X\theta)\cdot Y \;=\; X(\theta\cdot Y) - \theta\cdot[X,Y]$$

□

Proof: By the definition of pull-back and push-forward, we have

$$\phi^*\theta \cdot Y \;=\; \theta \cdot (\mathrm{d}\phi \cdot Y) \;=\; (\theta \cdot \phi_* Y) \circ \phi$$

This also holds for $\phi_*^{-1} Y$, which leads to

$$\phi^*\theta \cdot \phi_*^{-1} Y \;=\; (\theta \cdot Y) \circ \phi$$

Let now X be a complete vectorfield on M generating the 1-parameter group ϕ_t. Then we have $\forall t \in \mathbb{R}$

$$\phi_t^*\theta \cdot \phi_{-t*} Y \;=\; \phi_t^*(\theta \cdot Y)$$

Taking $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0}$ then yields

$$(\pounds_X \theta) \cdot Y + \theta \cdot (\pounds_X Y) \;=\; \pounds_X (\theta \cdot Y)$$

which is nothing else than the *Leibnitz rule* for Lie derivatives. Substituting

$$\pounds_X(\theta \cdot Y) \;=\; X(\theta \cdot Y) \qquad \pounds_X Y \;=\; [X, Y]$$

we obtain the formula

$$(\pounds_X \theta) \cdot Y \;=\; X(\theta \cdot Y) - \theta \cdot [X, Y]$$

(QED)

Remark 5.6: General Principle
A 1-form $\theta \in \Omega_1^\infty(M)$ can be thought of as an assignment of a function in $C^\infty(M)$ to each vectorfield $Y \in \mathscr{X}^\infty(M)$, which is linear with respect to multiplication by an element f of the ring $C^\infty(M)$ of C^∞-functions on M. That is, we have

$$\theta \cdot (fY) \;=\; f\theta \cdot Y \quad \forall f \in C^\infty(M), \forall Y \in \mathscr{X}^\infty(M)$$

Consider then

$$\begin{aligned}
(\pounds_X \theta) \cdot (fY) &= X(\theta \cdot (fY)) - \theta \cdot [X, fY] \\
&= X(f\theta \cdot Y) - \theta \cdot (f[X,Y] + (Xf)Y) \\
&= (Xf)(\theta \cdot Y) + fX(\theta \cdot Y) - f\theta \cdot [X,Y] - (Xf)(\theta \cdot Y) \\
&= f(\pounds_X \theta) \cdot Y
\end{aligned}$$

Where we used that

$$\begin{aligned}
[X, fY]g &= X(fYg) - fYXg \;=\; (Xf)(Yg) + fXYg - fYXg \\
&= f[Xr, Y]g + (Xf)(Yg)
\end{aligned}$$

□

Proposition 5.3: Let X and Y be complete vectorfields on M, generating the 1-parameter groups ϕ_t and ψ_s respectively. Suppose that X and Y commute, that is that $[X, Y] = 0$. Then we have

$$\phi_t \circ \psi_s \;=\; \psi_s \circ \phi_t \quad \forall t, s \in \mathbb{R}$$

□

Proof: Consider

$$\frac{\mathrm{d}}{\mathrm{d}t'} \phi_{-t'*} Y \bigg|_{t'=t} \;=\; \lim_{r \to 0} \left(\frac{\phi_{-t-r*} Y - \phi_{-t*} Y}{r} \right)$$

Now

$$\phi_{-t-r*} Y \;=\; \phi_{-r*} \phi_{-t*} Y$$

In general, if X is a vectorfield on M and ϕ, ψ any two diffeomorphisms of M onto itself, then

$$(\phi \circ \psi)_* X \;=\; \phi_* \psi_* X \tag{5.2}$$

(see Explanation A.6 on page 86). Thus

$$\frac{\mathrm{d}}{\mathrm{d}t'} \phi_{-t'*} Y \bigg|_{t'=t} \;=\; \pounds_X(\phi_{-t*} Y) \;=\; [X, \phi_{-t*} Y]$$

We now use the following two facts:

1) If X generates ϕ_t, then $\phi_{s*}X = X \ \forall s \in \mathbb{R}$.

2) If X, Y are any two vector fields and ϕ any diffeomorphism, then

$$\phi_*[X,Y] \;=\; [\phi_* X, \phi_* Y]$$

Proof of fact 1: If γ_p is the integral curve of X through p, then the integral curve of $\phi_{s*}X$ with fixed s through $\phi_s(p)$ is $\phi_s \circ \gamma_p$. But

$$(\phi_s \circ \gamma_p)(t) \;=\; \phi_s(\gamma_p(t)) \;=\; \gamma_p(t+s) \;=\; \gamma_{\phi_s(p)}(t)$$

This is the integral curve of X itself though $\phi_s(p)$. Therefore the integral curves of the vectorfields $\phi_{s*}X$ and X coincide. Hence the vectorfields themselves coincide.

Proof of fact 2: We first note that the definition of push-forward of a vectorfield V by a diffeomorphism ϕ can be expressed in the form

$$(\phi_* V)f \;=\; (V(f \circ \phi)) \circ \phi^{-1} \quad \forall f \in C^\infty(M)$$

Then

$$\begin{aligned} [\phi_* X, \phi_* Y]f &= (\phi_* X)((\phi_* Y)f) - (\phi_* Y)((\phi_* X)f) \\ &= (X((\phi_* Y)f) \circ \phi) \circ \phi^{-1} - (Y((\phi_* X)f) \circ \phi) \circ \phi^{-1} \\ &= (XY(f \circ \phi)) \circ \phi^{-1} - (YX(f \circ \phi)) \circ \phi^{-1} \\ &= ([X,Y](f \circ \phi)) \circ \phi^{-1} \;=\; \phi_*[X,Y] \end{aligned}$$

Here take X in the role of V, $(\phi_* Y)f$ in the role of f then interchange X, Y and subtract. Also take Y in the role of V and $(\phi_* X)f$ in the role of f and do the same with X and Y interchanged. Then apply the definition taking $V = [X,Y]$.

By 1) and 2) follows

$$[X, \phi_{-t*}Y] \;=\; [\phi_{-t*}X, \phi_{-t*}Y] \;=\; \phi_{-t*}[X,Y] \;=\; 0$$

We thus obtain

$$\frac{\mathrm{d}}{\mathrm{d}t'}\,\phi_{-t'*}Y\bigg|_{t'=t} \;=\; \frac{\mathrm{d}}{\mathrm{d}t}\,\phi_{-t*}(\phi_{-s*}Y)\bigg|_{t=0} \;=\; \pounds_X(\phi_{-s*}Y) \;=\; [X, \phi_{-t*}Y] \;=\; 0$$

Hence

$$\phi_{-t*}Y \;=\; \phi_{-0*}Y \;=\; Y \quad \forall t \in \mathbb{R}$$

because $\phi_0 = \mathrm{id}$. We conclude that

$$\phi_{t*}Y \;=\; Y \quad \forall t \in \mathbb{R}$$

Taking any fixed t we recall that if ψ_s is the one-parameter group generated by Y, the group generated by $\phi_{t*}Y$ is obtained by conjugation with ϕ_t. It follows

$$\phi_t \circ \psi_s \circ \phi_t^{-1} \;=\; \phi_s$$

Hence

$$\phi_t \circ \psi_s \;=\; \psi_s \circ \phi_t \quad \forall t, s \in \mathbb{R}$$

(QED)

Remark 5.7: We may consider ϕ_* as a linear operator on the space $\mathscr{X}^\infty(M)$ of C^∞ vectorfields on M. The following holds

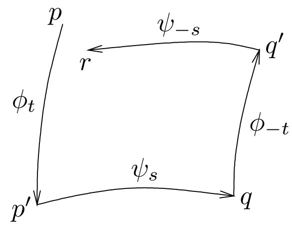
$$(\phi \circ \psi)_* \;=\; \phi_* \psi_*$$

(see Explanation A.6 on page 86). We may also consider ϕ^* as a linear operator on the space $\Omega_1^\infty(M)$ of C^∞ 1-forms on M. The following holds

$$(\phi \circ \psi)^* \;=\; \psi^* \phi^* \tag{5.3}$$

(see Explanation A.7 on page 86). □

Proposition 5.4: Let X, Y be complete vectorfields on M. Let X generate the 1-parameter group ϕ_t and Y generate the 1-parameter group ψ_t. Then $\psi_{-t} \circ \phi_{-t} \circ \psi_t \circ \phi_t$ coincides to $\mathcal{O}(t^2)$ with the 1-parameter group χ_r, generated by $[X, Y]$ for $r = t^2$. More generally, consider



with

$$p' = \phi_t(p) \qquad q = \psi_s(p') \qquad q' = \phi_{-t}(q) \qquad r = \psi_{-s}(q')$$

Then, we can write

$$\psi_{-s} \circ \phi_{-t} \circ \psi_s \circ \phi_t = \chi_r + \mathcal{O}\left((s+t)^3\right) \qquad r = st$$

□

Proof: Since this is a local statement, we may confine attention to a coordinate neighborhood of p. We choose the origin of the coordinates to correspond to p, that is $x^\mu|_p = 0$. The integral curves of X are represented by the solutions of

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = X^\mu(x(t))$$

We have

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}t^2} = \frac{\partial X^\mu}{\partial x^\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}t} = \left(X^\nu \frac{\partial X^\mu}{\partial x^\nu}\right)(x(t))$$

Thus, Taylor expanding

$$\begin{aligned} x^\mu(t) &= x^\mu(0) + t\,\frac{\mathrm{d}x^\mu}{\mathrm{d}t}(0) + \frac{1}{2}\,t^2\,\frac{\mathrm{d}^2x^\mu}{\mathrm{d}t^2}(0) + \mathcal{O}(t^3) \\ &= x^\mu(0) + tX^\mu(x(0)) + \frac{1}{2}\,t^2\left(X^\nu\,\frac{\partial X^\mu}{\partial x^\nu}\right)(x(0)) + \mathcal{O}(t^3) \end{aligned} \tag{5.4}$$

The integral curves of Y are represented by the solutions of

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}s} = Y^\mu(x(s))$$

We have

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}s^2} = \frac{\partial Y^\mu}{\partial x^\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}s} = \left(Y^\nu \frac{\partial Y^\mu}{\partial x^\nu}\right)(x(s))$$

Again, Taylor expanding

$$\begin{aligned} x^\mu(s) &= x^\mu(0) + s\,\frac{\mathrm{d}x^\mu}{\mathrm{d}s}(0) + \frac{1}{2}\,s^2\,\frac{\mathrm{d}^2x^\mu}{\mathrm{d}s^2}(0) + \mathcal{O}(s^3) \\ &= x^\mu(0) + sY^\mu(x(0)) + \frac{1}{2}\,s^2\left(Y^\nu\,\frac{\partial Y^\mu}{\partial x^\nu}\right)(x(0)) + \mathcal{O}(s^3) \end{aligned} \tag{5.5}$$

Consider the arc from p to p'. This is an integral curve of X. Applying 5.4 with $x(0) = x|_p = 0$ we obtain

$$x^\mu|_{p'} = x^\mu(t) = t\,X^\mu|_p + \frac{1}{2}\,t^2\,X^\nu|_p\,\left.\frac{\partial X^\mu}{\partial x^\nu}\right|_p + \mathcal{O}(t^3)$$

Next we have the arc from p' to q. This is an integral curve of Y. Applying 5.5 with $x(0) = x|_{p'}$ we obtain

$$x^\nu|_q = x^\mu(s) = x^\mu|_{p'} + s\,Y^\mu|_{p'} + \frac{1}{2}\,s^2\,Y^\nu|_{p'}\,\left.\frac{\partial Y^\mu}{\partial x^\nu}\right|_{p'} + \mathcal{O}(s^3)$$

and we have

$$\begin{aligned} Y^\mu\big|_{p'} &= Y^\mu\big|_p + \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p \left(x^\nu\big|_{p'} - x^\nu\big|_p \right) + \mathcal{O}\left(\left| x\big|_{p'} - x\big|_p \right|^2 \right) \\ &= Y^\mu\big|_p + t\, X^\nu\big|_p \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p + \mathcal{O}(t^2) \end{aligned}$$

Substituting for $Y^\mu\big|_{p'}$ as well as for $x^\mu\big|_{p'}$ in the expression for $x^\mu\big|_q$ we obtain

$$x^\mu\big|_q = t\, X^\mu\big|_p + s\, Y^\mu\big|_p + \frac{1}{2} t^2\, X^\nu\big|_p \frac{\partial X^\mu}{\partial x^\nu}\bigg|_p + st\, X^\nu\big|_p \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p + \frac{1}{2} s^2\, Y^\nu\big|_p \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p + \mathcal{O}\left((s+t)^3\right)$$

Next we have the arc from q to q'. This is an integral curve of X. Applying 5.4 with $x(0) = x\big|_q$ and t replaced by $-t$, we obtain

$$x^\mu\big|_{q'} = x^\mu\big|_q - t\, X^\mu\big|_q + \frac{1}{2} t^2\, X^\nu\big|_q \frac{\partial X^\mu}{\partial X^\nu}\bigg|_q + \mathcal{O}(t^3)$$

and we have

$$\begin{aligned} X^\mu\big|_q &= X^\mu\big|_p + \frac{\partial X^\mu}{\partial x^\nu}\bigg|_p \left(x^\nu\big|_q - x^\nu\big|_p \right) + \mathcal{O}\left(\left| x\big|_q - x\big|_p \right|^2 \right) \\ &= X^\mu\big|_p + \frac{\partial X^\mu}{\partial x^\nu}\bigg|_p \left(t\, X^\nu\big|_p + s\, Y^\nu\big|_p \right) + \mathcal{O}\left((s+t)^2\right) \end{aligned}$$

Substituting for $X^\mu\big|_q$ as well as for $x^\mu\big|_q$ into the expression for $x^\mu\big|_{q'}$ we obtain

$$\begin{aligned} x^\mu\big|_{q'} &= t\, X^\mu\big|_p + s\, Y^\mu\big|_p + \frac{1}{2} t^2\, X^\nu\big|_p \frac{\partial X^\mu}{\partial x^\nu}\bigg|_p + st\, X^\nu\big|_p \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p + \frac{1}{2} s^2\, Y^\nu\big|_p \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p \\ &\quad -t\, X^\mu\big|_p - t^2\, X^\nu\big|_p \frac{\partial X^\mu}{\partial x^\nu}\bigg|_p - st\, Y^\nu\big|_p \frac{\partial X^\mu}{\partial x^\nu}\bigg|_p + \frac{1}{2} t^2\, X^\nu\big|_p \frac{\partial X^\mu}{\partial x^\nu}\bigg|_p + \mathcal{O}\left((s+t)^3\right) \\ &= s\, Y^\mu\big|_p + st \left(X^\nu\big|_p \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p - Y^\nu\big|_p \frac{\partial X^\mu}{\partial x^\nu}\bigg|_p \right) + \frac{1}{2} s^2\, Y^\nu\big|_p \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p + \mathcal{O}\left((s+t)^3\right) \end{aligned}$$

Finally, we have the arc from q' to r. This is an integral curve of Y. Applying (5.5) with $x(0) = x^\mu\big|_{q'}$ and s replaced by $-s$, we obtain

$$x^\mu\big|_r = x^\mu\big|_{q'} - s\, Y^\mu\big|_{q'} + \frac{1}{2} s^2\, Y^\nu\big|_{q'} \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_{q'} + \mathcal{O}(s^3)$$

and we have

$$\begin{aligned} Y^\mu\big|_{q'} &= Y^\mu\big|_p + \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p \left(x^\nu\big|_{q'} - x^\nu\big|_p \right) + \mathcal{O}\left(\left| x\big|_{q'} - x\big|_p \right|^2 \right) \\ &= Y^\mu\big|_p + s\, Y^\nu\big|_p \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p + \mathcal{O}\left((s+t)^2\right) \end{aligned}$$

Finally, we substitue for $x^\mu\big|_{q'}$ as well as for $Y^\mu\big|_{q'}$ into the expression for $x^\mu\big|_r$ to obtain

$$\begin{aligned} x^\mu\big|_r &= s\, Y^\mu\big|_p + st \left(X^\nu\big|_p \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p - Y^\nu\big|_p \frac{\partial X^\mu}{\partial x^\nu}\bigg|_p \right) + \frac{1}{2} s^2\, Y^\nu\big|_p \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p \\ &\quad -s\, Y^\mu\big|_p - s^2\, Y^\nu\big|_p \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p + \frac{1}{2} s^2\, Y^\nu\big|_p \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p + \mathcal{O}\left((s+t)^3\right) \\ &= st \left(X^\nu\big|_p \frac{\partial Y^\mu}{\partial x^\nu}\bigg|_p - Y^\nu\big|_p \frac{\partial X^\mu}{\partial x^\nu}\bigg|_p \right) + \mathcal{O}\left((s+t)^3\right) \end{aligned}$$

We write

$$X = X^\nu \frac{\partial}{\partial x^\nu} \qquad Y = Y^\nu \frac{\partial}{\partial x^\nu}$$

in a coordinate nieghborhood of p, as therefore

$$Xf = X^\nu \frac{\partial f}{\partial x^\nu} \qquad Yf = Y^\nu \frac{\partial f}{\partial x^\nu}$$

So

$$\begin{aligned} XYf &= X^\nu \frac{\partial (Yf)}{\partial x^\nu} = X^\nu \frac{\partial}{\partial x^\nu}\left(Y^\mu \frac{\partial f}{\partial x^\mu}\right) \\ &= X^\nu Y^\mu \frac{\partial^2 f}{\partial x^\nu \partial x^\mu} + X^\nu \frac{\partial Y^\mu}{\partial x^\nu}\frac{\partial f}{\partial x^\mu} \\ YXf &= Y^\nu X^\mu \frac{\partial^2 f}{\partial x^\nu \partial x^\mu} + Y^\nu \frac{\partial X^\mu}{\partial x^\nu}\frac{\partial f}{\partial x^\mu} \end{aligned}$$

Hence

$$[X,Y]f = XYf - YXf = \sum_{\nu,\mu}\left(X^\nu \frac{\partial Y^\mu}{\partial x^\nu} - Y^\nu \frac{\partial X^\mu}{\partial x^\nu}\right)\frac{\partial f}{\partial x^\mu}$$

Therefore

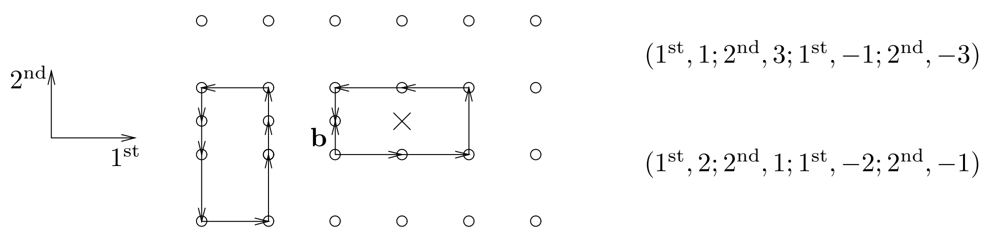
$$[X,Y]^\mu\big|_p = X^\nu\big|_p \left.\frac{\partial Y^\mu}{\partial x^\nu}\right|_p - Y^\nu\big|_p \left.\frac{\partial X^\mu}{\partial x^\nu}\right|_p$$

We conclude that

$$x^\mu\big|_r = st\,[X,Y]^\mu\big|_p + \mathcal{O}\left((s+t)^3\right)$$

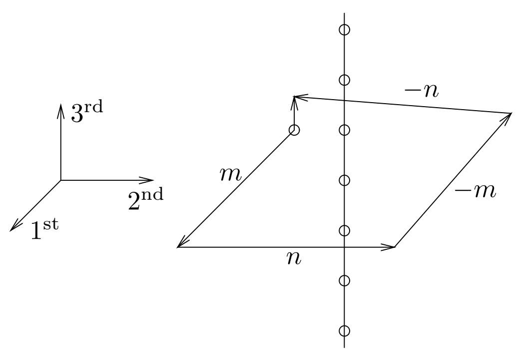
(QED)

Example 5.1: Crystal Dislocations



Suppose that in aperfect lattice an extra half-line of atoms has been inserted along the negative 1st -axis. In the real 3 dimensional world an extra half-plane has been iserted. A circuit of translations alternately along the first and the second axes (1^{st}, m; 2^{nd}, n; 1^{st}, $-m$; 2^{nd}, $-n$) which does not include the origin x closes. However, a circuit of translations alternately along the first and the second axes does not close if the circuit contains the origin x. We in fact arive at an atom which can be reached in a single step along the second axis by 1. The origin is a dislocation point. It is actualla dislocation line in 3 dimensions, the edge of the extra half-plane. The single translation **b** by which the final point is reached from the initial point is called *Burgers vector*. □

Example 5.2: Screw Dislocation
Screw dislocation appears in 3-dimensional cristals



Here, a circuit of translations alternately along the first and the second axes which includes the third axis, does not close. In fact we arrive at an atom which can be reached in a single step along the third axis by 1. □

Remark 5.8: The burgers vector is, what we obtain by $\psi_{-s} \circ \phi_{-t} \circ \psi_s \circ \phi_t$. In the case of crystal dislocation, using Proposition 5.4, up to third order, we get

$$[X, Y] = Y$$

and for screw dislocations

$$[X, Y] = Z$$

□

5.3 2-forms

Definition 5.6: Let us define the bundle

$$\Lambda_2 M = \bigcup_{p \in M} \Lambda_2(T_p M)$$

where for any vector space V we denote by $\Lambda_2(V)$ the space of anti-symmetric bilinear forms on V. □

Definition 5.7: A *2-form* ω on M is a continuously differentiable section of $\Lambda_2 M$. □

Definition 5.8: If α and β are two elements of V^*, we define the *outer product* $\alpha \wedge \beta \in \Lambda_2(V)$ by

$$(\alpha \wedge \beta) \cdot (u, v) = (\alpha \cdot u)(\beta \cdot v) - (\alpha \cdot v)(\beta \cdot u) \quad \forall u, v \in V$$

□

Remark 5.9: Now if $(e_1, ..., e_n)$ is a basis for V and $(e^{*1}, ..., e^{*n})$ is the dualbasis for V^*, then the $n\left(\frac{n-1}{2}\right)$ elements of $\Lambda_2(V)$, namely $e^{*i} \wedge e^{*j}$ with $i < j = 1, ..., n$, form a basis for $\Lambda_2(V)$. This applys in particular to the case $V = T_p M$. Suppose that (U, φ) is a chart of M. Then the vectors $\frac{\partial}{\partial x^1}|_p, ..., \frac{\partial}{\partial x^n}|_p$ form a basis for $T_p M$, where $p \in U$ and the corresponding dual basis for $T_p^* M$ is $\mathrm{d}x^1|_p, ..., \mathrm{d}x^n|_p$. Therefore any 2-form ω can be locally (in U) expanded as

$$\omega = \sum_{\mu < \nu} \omega_{\mu\nu} \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$$

where the coefficients

$$\omega_{\mu\nu} = \omega\left(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right) = -\omega_{\nu\mu}$$

form an anti-symmetric matrix. We can write

$$\omega = \frac{1}{2} \sum_{\mu, \nu} \omega_{\mu\nu} \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$$

□

Definition 5.9: Given a 1-form θ on M, we define a 2-form $\mathrm{d}\theta$ on M, the *exterior derivative* of θ, as follows. For any pair X, Y of vectorfields on M we set

$$\mathrm{d}\theta \cdot (X, Y) = X(\theta \cdot Y) - Y(\theta \cdot X) - \theta \cdot [X, Y]$$

□

Remark 5.10: This defines a 2-form, because it is bilinear with respect to the multiplication by the ring of C^∞-functions on M. Verify that

$$\mathrm{d}\theta \cdot (fX, Y) = f \mathrm{d}\theta \cdot (X, Y) \quad \forall f \in C^\infty(M) \tag{5.6}$$

(see Explanation A.8 on page 86). □

Remark 5.11: Consider a chart (U, φ). Take X, Y to be the local (defined in U) vectorfiels

$$X = \frac{\partial}{\partial x^\mu} \qquad Y = \frac{\partial}{\partial x^\nu}$$

Then

$$\theta \cdot X = \theta_\mu \qquad \theta \cdot Y = \theta_\nu$$

and the definition reduces to

$$(\mathrm{d}\theta)_{\mu\nu} = \frac{\partial \theta_\nu}{\partial x^\mu} - \frac{\partial \theta_\mu}{\partial x^\nu}$$

□

Remark 5.12: Comparing the formulas for $\pounds_X \theta$ and $\mathrm{d}\theta$, we see

$$\pounds_X \theta = i_X \mathrm{d}\theta + \mathrm{d} i_X \theta \qquad \text{or} \qquad \pounds_X \theta \cdot Y = \mathrm{d}\theta \cdot (X, Y) + Y(\theta \cdot X)$$

Here i_X denotes contraction on the left by X. So if θ is a 1-form, then $i_X \theta$ is simply the function $\theta \cdot X$. If ω is a 2-form then $i_X \omega$ is the 1-form defined by

$$i_X \omega \cdot Y = \omega \cdot (X, Y) \quad \forall Y \in \mathscr{X}^\infty(M)$$

□

5.4 Lie Groups

Definition 5.10: A *Lie group* G is a group which is also a differentiable manifold and the group operations

$$G \times G \to G \ : \ (a, b) \mapsto ab \qquad G \to G \ : \ a \mapsto a^{-1}$$

thet is the multiplication and the inversion-map, are C^∞-mappings. □

Remark 5.13: We have two distinguished groups of diffeomorphisms of G onto itself.

1) Right multiplications: $\{r_a : a \in G\}$

$$r_a(b) = ba \quad \forall b \in G$$

2) Left multiplications: $\{\ell_a : a \in G\}$

$$\ell_a(b) = ab \quad \forall b \in G$$

Note that

$$r_a \circ \ell_b = \ell_b \circ r_a \quad \forall a, b \in G$$

□

Definition 5.11: The *Lie algebra* is the following distinguished subspace of $\mathscr{X}^\infty(G)$

$$\mathcal{Y} := \{X \in \mathscr{X}^\infty(G) : \ell_{a*} X = X, \, \forall a \in G\}$$

In other words, $\mathcal{Y}$ is the space of left invariant vectorfields on G. □

Proposition 5.5:

$$X, Y \in \mathcal{Y} \quad \Rightarrow \quad [X, Y] \in \mathcal{Y}$$

□

Proof: Let be $Z := [X, Y] \in \mathscr{X}^\infty(G)$. Recall the proposition

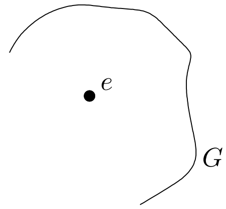
$$\phi_*[X, Y] = [\phi_* X, \phi_* Y]$$

Apply this taking $\phi = \ell_a$, $a \in G$

$$\ell_{a*} Z = [\ell_{a*} X, \ell_{a*} Y] = [X, Y] = Z \quad \forall a \in G$$

Therefore, $Z \in \mathcal{Y}$. (QED)

Remark 5.14: Construction of the left invariant vectorfields



Let us denote by e the identity element in G. Let be $X \in \mathcal{Y}$. Then $\ell_{a*}X = X \ \forall a \in G$. We have

$$(\ell_{a*}X)(\ell_a(b)) \quad = \quad \mathrm{d}\ell_a \cdot X(b) \quad \forall b \in G$$

Set $b = e$ to obtain

$$(\ell_{a*}X)(a) \quad = \quad \mathrm{d}\ell_a \cdot X(e) \quad \forall a \in G$$

But $\ell_{a*}X = X$. We conclude that

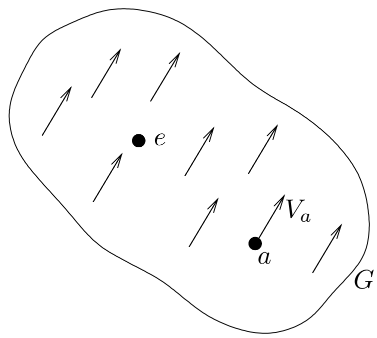
$$X(a) \quad = \quad \mathrm{d}\ell_a \cdot X(e) \quad \forall a \in G$$

This defines X on G given its value at e. □

Remark 5.15: Consider, in general, the evaluation map at $p \in M$

$$\varepsilon_p : \mathscr{X}^\infty(M) \to T_pM \qquad \varepsilon_p(X) \quad = \quad X(p)$$

Now consider ε_a with $a \in G$, restricted to $\mathcal{Y}$, ℓ_a being a diffeomorphism. $\mathrm{d}\ell_a(e)$ is an isomorphism of T_eG onto T_aG (see Explanation A.9 on page 86). It follows that ε_a restricted to $\mathcal{Y}$ is an isomorphism of $\mathcal{Y}$ onto T_aG. Therefore at each $a \in G$, $\varepsilon_a^{-1}(V_a)$ is some element of $\mathcal{Y}$ for each tangent vector $V_a \in T_aG$.

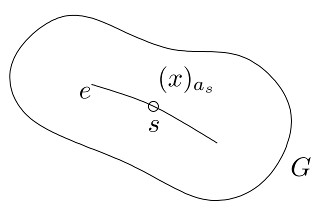


In other words, given any tangent vector $V_a \in T_aG$ (where a is an arbitrary element of G), there is a unique $X \in \mathcal{Y}$ such that $X(a) = V_a$. This implies that any Lie group is *parallelizable*. That means, given a basis for T_eG, the corresponding left invariant vectorfields define a basis for any $\mathcal{Y}$, hence (since ε_a is an isomorphism at each $a \in G$) a *frame* at each $a \in G$. □

Definition 5.12: Given $X \in \mathcal{Y}$, let us denote by $t \mapsto (x)_{a_t}$ the integral curve of X through e. □

Proposition 5.6: $(x)_{a_t}$ is a 1-parameter subgroup of G. □

Proof: Recall that if γ_p is the integral curve of X through p, then $\phi \circ \gamma_p$ is the integral curve of ϕ_*X through $\phi(p)$. Let us fix s.



Then, taking $\phi = \ell_{(x)_{a_s}}$, the integral curve of $\ell_{(x)_{a_s}*}X = X$ through $\ell_{(x)_{a_s}}(e) = (x)_{a_s}$ is

$$\ell_{(x)_{a_s}} \circ (x)_{a_t} \quad : \quad t \mapsto \ell_{(x)_{a_s}}((x)_{a_t}) \quad = \quad (x)_{a_s}(x)_{a_t}$$

It follows that

$$(x)_{a_{s+t}} \quad = \quad (x)_{a_s}(x)_{a_t}$$

So $\{(x)_{a_t} : t \in \mathbb{R}\}$ is indeed a 1-parameter subgroup. (QED)

Remark 5.16: Furthermore, from the proof of the above, taking $\phi = \ell_b$, $b \in G$, we have the integral curve of $\ell_{b*}X = X$ through $\ell_b(e) = b$, which is

$$\ell_b \circ (x)_{a_t} \quad : \quad t \mapsto \ell_b\left((x)_{a_t}\right) \quad = \quad b(x)_{a_t}$$

We have thus proved the following. □

Proposition 5.7: The integral curve of X through $b \in G$ is $t \mapsto b(x)_{a_t}$, for any $b \in G$. Here $(x)_{a_t}$ is the 1-parameter subgroup of G which is the integral curve of X through e. □

Remark 5.17: The 1-parameter group of diffeomorphisms generated by X is

$$\{r_{(x)_{a_t}} : t \in \mathbb{R}\}$$

Thus the left invariant vectorfields generate right multiplications. □

Example 5.3: The Affine Group of the Real Line
The simplest non-trivial example is the affine group of the real line

$$s \mapsto e^x s + y$$

which corresponds to scaling by e^x and translation by y. The group manifold is $G = \mathbb{R}^2$ and the pair $(x, y) \in G$ acts as

$$(x_2, y_2)s \quad = \quad s' \quad = \quad e^{x_2} s + y_2$$

The product of two paris $(x_1, y_1), (x_2, y_2) \in G$ acts as

$$(x_1, y_1)(x_2, y_2)s \quad = \quad (x_1, y_1)s' \quad = \quad e^{x_1} s' + y_1 \quad = \quad e^{x_1+x_2} s + y_1 + e^{x_1} y_2$$

Hence, we can write

$$(x_1, y_1)(x_2, y_2) \quad = \quad (x_1 + x_2, y_1 + e^{x_1} y_2)$$

Two 1-parameter subgroups are

1) $s \mapsto e^t s$, the scaling subgroup $\{(t, 0) : t \in \mathbb{R}\} \subset G$

2) $s \mapsto s + t$, the translation subgroup $\{(0, t) : t \in \mathbb{R}\} \subset G$

Right multiplications by the scaling subgroup gives

$$(x, y) \mapsto (x, y)(t, 0) \quad = \quad (x + t, y)$$

The corresponding left invariant vectorfield is

$$X \quad = \quad \frac{\partial}{\partial x}$$

Right multiplications by the translational subgroup gives

$$(x, y) \mapsto (x, y)(0, t) \quad = \quad (x, y + e^x t)$$

The corresponding left invariant vectorfield is

$$Y \quad = \quad e^x \frac{\partial}{\partial y}$$

Then

$$[X, Y] = Y$$

This corresponds to a uniform distribution of edge dislocations on the plane. □

Example 5.4: The Heisenberg Group
We have $G = \{(x, y, z)\} = \mathbb{R}^3$. We let G act on 1-dimensional square integrable complex valued functions on $\mathbb{R}$ (1-dimensional quantum mechanics)

$$((x, y, z)\psi)(x) \quad = \quad e^{iys + iz} \psi(s + x) \quad s \in \mathbb{R}$$

There are three left invariant vector fields X, Y, Z with $[X, Y] = Z$ (see exercice sheet 8). So we have a uniform distribution of screw dislocations. □

6 Length and Volume

6.1 Metric

Definition 6.1: Given a real vector space V, we denote by $S_2(V)$ the *space of symmetric bilinear forms on* V (also called *quadratic forms*). If $(e_1, ..., e_n)$ is a basis for V, then given $h \in S_2(V)$,

$$h_{ij} \quad = \quad h(e_i, e_j) \quad = \quad h_{ji}$$

are the components of h in this basis. □

Definition 6.2: For $\alpha, \beta \in V^*$, we define the *tensor product* $\alpha \otimes \beta$, a bilinear form on V, by

$$(\alpha \oplus \beta) \cdot (u, v) \quad = \quad \alpha(u)\beta(v) \quad \forall u, v \in V$$

□

Remark 6.1: If $(e^{*1}, ..., e^{*n})$ is the basis for V^*, which is dual to the basis $(e_1, ..., e_n)$ for V, we have the expansion

$$h \quad = \quad \sum_{i<j} h_{ij} e^{*i} \otimes e^{*j} \quad = \quad \frac{1}{2} \sum_{i,j} h_{ij} (e^{*i} \otimes e^{*j} + e^{*j} \otimes e^{*i})$$

The symmetric tensor product fo α and β is simply given by

$$\frac{1}{2}(\alpha \otimes \beta + \beta \otimes \alpha) \in S_2(V)$$

If now

$$\mathcal{B} \quad = \quad \bigcup_{p \in M} \mathcal{B}_p$$

is a vector bundle over a manifold M, then

$$S_2(\mathcal{B}, M) \quad = \quad \bigcup_{p \in M} S_2(\mathcal{B}_p)$$

is also a vector bundle over M. □

Definition 6.3: Going back to $S_2(V)$, we denoty by $S_2^+(V)$ the subset of $S_2(V)$ consisting of those quadratic forms which are *positive definite*, that is

$$h \in S_2^+(V) \quad \Leftrightarrow \quad \big(h \in S_2(V),\, h(v,v) \geq 0,\, h(v,v) = 0 \quad \Leftrightarrow \quad v = 0\big)$$

Thus $S_2^+(V)$ is the *space of inner products* on V. □

Proposition 6.1: $S_2^+(V)$ is an open positive convex cone in $S_2(V)$. □

Proof: Openess is obvious. Positive cone means that

$$h \in S_2^+(V),\, \lambda > 0 \quad \Rightarrow \quad \lambda h \in S_2^+(V)$$

So this is also obvious. Convex means, that if $h_1, h_2 \in S_2^+(V)$, then each point in the straight line segment in $S_2(V)$ joining h_1 to h_2 also belongs to $S_2^+(V)$. In other words

$$(1 - \lambda)h_1 + \lambda h_2 \in S_2^+(V) \quad \forall \lambda \in [0, 1]$$

This is also obvious. (QED)

Definition 6.4: A *metric* m on a real vector bundle $\mathcal{B}$ is a continuously differentiable section of

$$S_2^+(\mathcal{B}, M) \quad = \quad \bigcup_{p \in M} S_2^+(\mathcal{B}_p)$$

which is an open subbundle of $S_2(\mathcal{B}, M)$. Thus m is a continuously differentiable assignment of an inner product m_p in $\mathcal{B}_p$ at each $p \in M$. □

Remark 6.2: Given two continuously differentiable sections σ, τ of $\mathcal{B}$, $m(\sigma, \tau)$ defined pointwise

$$m(\sigma,\tau)(p) \;\; = \;\; m_p(\sigma(p), \tau(p))$$

is a continuously differentiable function on M. □

Definition 6.5: In the case $\mathcal{B} = TM$, $m =: g$ is called a *Riemannian metric* on M. So g is a continuously differentiable section of

$$S_2^+ M \;\; = \;\; \bigcup_{p \in M} S_2^+(T_p M)$$

That is a continuously differentiable assignment of an inner product g_p in T_pM, at each $p \in M$. □

Remark 6.3: Note that the bundle

$$S_2 M \;\; = \;\; \bigcup_{p \in M} S_2(T_p M)$$

is a *tensor bundle* over M (the bundle of symmetric 2-covariant tensors on M), like the bundles $\Lambda_1 M = T^*M$ (the bundle of 1-forms or cotangent bundle on M) and $\Lambda_2 M$ (the bundle of 2-forms on M), that we have already introduced. □

Remark 6.4: Recall that for a real vector bundle $\mathcal{B}$ on M, we have the bundle charts (U, ω), where U is an open domain in M and ω is a diffeomorphism $\omega : \pi^{-1}(U) \to U \times \mathbb{R}^n$ (with π, the projection of the bundle) and where for $b \in \mathcal{B}_p = \pi^{-1}(p)$, $\omega(b) = (p, \chi_p(b))$ and for each $p \in U$, $\chi_p : \mathcal{B}_p \to \mathbb{R}^n$ is a linear isomorphism. Suppose now that $\mathcal{B}$ is equipped with a metric m. $\mathbb{R}^n$ is endowed with the standard inner product $\langle \cdot, \cdot \rangle$. We denote by $\| \cdot \|$ the corresponding norm, that is

$$\|a\| \;\; = \;\; \sqrt{\langle a, a \rangle} \;\; = \;\; \sqrt{\sum_{\mu=1}^{n} (a^\mu)^2} \quad \forall a \in \mathbb{R}^n$$

We require χ_p to preserve the norms, so

$$\sqrt{m_p(b,b)} \;\; = \;\; \|\chi_p(b)\|$$

We say that χ_p is a *linear isometry* for each $p \in U$. It follows that

$$m_p(u,v) \;\; = \;\; \langle \chi_p(u), \chi_p(v) \rangle \quad \forall u, v \in \pi^{-1}(p) = \mathcal{B}_p$$

Let $e_\mu = (0_1, ..., 1_\mu, ..., 0_n)$, $\mu = 1, ..., n$ be the standard basis for $\mathbb{R}^n$. This is an orthonormal basis relative to the standard inner product $\langle \cdot, \cdot \rangle$, that is $\langle e_\mu, e_\nu \rangle = \delta_{\mu\nu}$. We then define the local (over U) basis sections σ_μ of $\mathcal{B}$ by

$$\sigma_\mu(p) \;\; = \;\; \chi_p^{-1} \cdot e_\mu \quad \forall p \in U$$

We then have

$$m_p(\sigma_\mu(p), \sigma_\nu(p)) \;\; = \;\; \langle e_\mu, e_\nu \rangle \;\; = \;\; \delta_{\mu\nu} \quad \forall p \in U$$

Thus the sections σ_μ, $\mu = 1, ..., n$ form an orthonormal basis for $\mathcal{B}_p$, at each $p \in U$. In the case $\mathcal{B} = TM$, a section σ of $\mathcal{B}$ is a vectorfield X and the set $(\sigma_\mu,\ \mu = 1, ...n)$ is an *orthonormal frame field* $(X_1, ..., X_M)$. Here $m = \dim M$ and $n = \dim \mathcal{B}_p = \dim T_pM = m$. □

Remark 6.5: The above holds for real vector bundles. In the case of a complex vector bundle $\mathcal{B}$, a metric m on $\mathcal{B}$ is now a continuously differentiable assignment of a Hermitian inner product m_p in $\mathcal{B}_p$ at each $p \in M$. □

Definition 6.6: Let V be a complex vector space. A *Hermitian inner product* on V is a form $h(\cdot, \cdot)$ on V with two arguments such that

1) $h(u, v)$ is anti-linear in the first argument u and linear in the second argument v.

2) $h(v, u) = \overline{h(u, v)}$

3) $h(v, v) \geq 0$ with $h(v, v) = 0 \Leftrightarrow v = 0$.

□

Remark 6.6: Antilinear in the first argument means

$$h(u_1 + u_2, v) \;=\; h(u_1, v) + h(u_2, v) \qquad h(\alpha u, v) \;=\; \overline{\alpha} h(u, v) \quad \forall \alpha \in \mathbb{C}$$

Linear in the second argument means

$$h(u, v_1 + v_2) \;=\; h(u, v_1) + h(u, v_2) \qquad h(u, \alpha v) \;=\; \alpha h(u, v) \quad \forall \alpha \in \mathbb{C}$$

Note that by virtue of 2) $h(v, v)$ is real. □

Remark 6.7: We denote by $\langle \cdot, \cdot \rangle$ the standard Hermitian inner product $\mathbb{C}^n$, which is given by

$$\langle \alpha, \beta \rangle \;:=\; \sum_{\mu=1}^{n} \overline{\alpha}^{\mu} \beta^{\mu} \quad \alpha = (\alpha^1, ..., \alpha^n), \beta = (\beta^1, ..., \beta^n) \in \mathbb{C}^n$$

We also denote by $(e_\mu,\ \mu = 1, ..., n)$ the standard basis of $\mathbb{C}^n$ with

$$e_\mu \;=\; (0_{_1}, ..., 1_{_\mu}, ..., 0_{_n}) \quad \mu = 1, ..., n$$

This is orthonormal relative to $\langle \cdot, \cdot \rangle$, so

$$\langle e_\mu, e_\nu \rangle \;=\; \delta_{\mu\nu}$$

In considering bundle charts (U, ω) we require $\chi_p : \mathcal{B}_p \to \mathbb{C}^n$ to be a linear isometry

$$m_p(u, v) \;=\; \langle \chi_p(u), \chi_p(v) \rangle \quad \forall u, v \in \mathcal{B}_p = \pi^{-1}(p)$$

Then again the local (over U) basis sections σ_μ defined by

$$\sigma_\mu(p) \;=\; \chi_p^{-1} \cdot e_\mu \quad \forall p \in U,\ \mu = 1, ..., n$$

define an orthonormal basis for $\mathcal{B}_p$ at each $p \in U$, that is

$$m_p(\sigma_\mu(p), \sigma_\nu(p)) \;=\; \delta_{\mu\nu} \quad \forall p \in U$$

□

Remark 6.8: Any section ϕ of $\mathcal{B}$ can be locally (over U) expanded as

$$\phi \;=\; \sum_{\mu=1}^{n} \phi^\mu \sigma_\mu$$

The coefficients ϕ^μ are complex-valued functions defined in U. Let ψ be another section of $\mathcal{B}$. Then

$$\psi \;=\; \sum_{\mu=1}^{n} \psi^\mu \sigma_\mu$$

Then the function $m(\phi, \psi)$ is given in U by

$$\begin{aligned} m(\phi, \psi) \;&=\; m\left(\sum_{\mu=1}^{n} \phi^\mu \sigma_\mu, \sum_{\nu=1}^{n} \psi^\nu \sigma_\nu \right) \;=\; \sum_{\mu,\nu=1}^{n} m(\phi^\mu \sigma_\mu, \psi^\nu \sigma_\nu) \\ &=\; \sum_{\mu,\nu=1}^{n} \overline{\phi}^\mu \psi^\nu m(\sigma_\mu, \sigma_\nu) \;=\; \sum_{\mu=1}^{n} \overline{\phi}^\mu \psi^\mu \end{aligned}$$

□

Remark 6.9: In the case $n = 1$ we have a complex line bundle. Then $e_1 = 1$ and σ_1 is defined by

$$\sigma_1(p) \;=\; \chi_p^{-1} \cdot 1 \quad \forall p \in U$$

Thus σ_1 is only subject to the condition

$$m_p(\sigma_1(p), \sigma_1(p)) \;=\; \|e_1\|^2 \;=\; 1$$

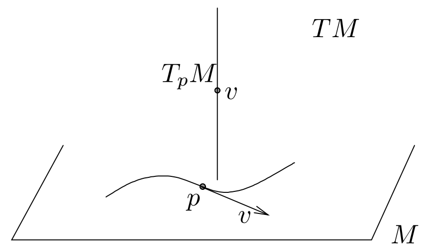
That is, $\sigma_1(p)$ belongs to the unit circle in $\mathcal{B}_p$. Here, $\mathcal{B}_p$ is isomorphis to $\mathbb{C}$. The ambiguity in the choice of σ_1 is reflected in the arbitrariness of phase for the wave function ϕ^1 in Quantum Mechanics

$$\phi \;=\; \phi^1 \sigma_1 \qquad \sigma_1(p) \mapsto e^{i\alpha} \sigma_1(p),\ \phi^1(p) \mapsto e^{-i\alpha} \phi^1(p) \quad \alpha \in \mathbb{R}$$

□

6.2 Arc Length

The concepts of arc length and volume refer exclusively to the case $\mathcal{B} = TM$.



Definition 6.7: The arc *length of a curve* $\gamma : I = [a,b] \to M$ is defined by

$$L(\gamma([a,b])) = \int_a^b \sqrt{g_{\gamma(t)}(\dot{\gamma}(t),\dot{\gamma}(t))}\,\mathrm{d}t$$

where $\dot{\gamma}(t)$ is the tangent vector to γ at $\gamma(t)$. □

Proposition 6.2: This integral is independent of the parametrization of γ, thus it is denoted by $L(\gamma([a,b]))$, $\gamma([a,b])$ being the corresponding path (or arc) in M. □

Proof: Let $\tilde{I} = [\tilde{a},\tilde{b}]$ be another closed interval and $\tilde{\gamma} : \tilde{I} \to M$ be another curve giving the same path in M (a reparametrization of γ). Then the end points coincide

$$\{\tilde{\gamma}(\tilde{a}),\tilde{\gamma}(\tilde{b})\} = \{\gamma(a),\gamma(b)\}$$

Thus either $\tilde{\gamma}(\tilde{a}) = \gamma(a)$ and $\tilde{\gamma}(\tilde{b}) = \gamma(b)$ or $\tilde{\gamma}(\tilde{b}) = \gamma(a)$ and $\tilde{\gamma}(\tilde{a}) = \gamma(b)$. Let us assume that the first alternative holds (the second alternative is handled in a similar manner). Then there is an increasing diffeomorphism $f : \tilde{I} \to I$ such that $\tilde{\gamma} = \gamma \circ f$. Let us set $s = f(t)$, so that $\mathrm{d}s = \dot{f}(t)\mathrm{d}t$. Then

$$\dot{\tilde{\gamma}}(t) = \dot{f}(t)\dot{\gamma}(s)$$

This can be shown using $\phi \in C^\infty(M)$. Then

$$\begin{aligned}
\dot{\tilde{\gamma}}(t)\cdot\phi &= \frac{\mathrm{d}}{\mathrm{d}t'}\,\phi(\tilde{\gamma}(t'))\Big|_{t'=t} = \frac{\mathrm{d}}{\mathrm{d}t'}\,\phi(\gamma(f(t')))\Big|_{t'=t} \\
&= \frac{\mathrm{d}}{\mathrm{d}s'}\,\phi(\gamma(s'))\Big|_{s'=s}\cdot\frac{\mathrm{d}f(t')}{\mathrm{d}t'}\Big|_{t'=t} \\
&= \dot{f}(t)\dot{\gamma}(s)\cdot\phi
\end{aligned}$$

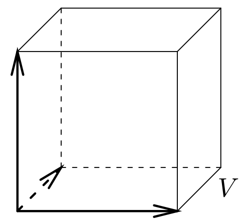
calling $s' = f(t')$. And we have

$$\begin{aligned}
L(\gamma([a,b])) &= \int_a^b \sqrt{g_{\gamma(s)}\big(\dot{\gamma}(s),\dot{\gamma}(s)\big)}\,\mathrm{d}s \\
&= \int_{\tilde{a}}^{\tilde{b}} \sqrt{g_{\tilde{\gamma}(t)}\big(\dot{\tilde{\gamma}}(t),\dot{\tilde{\gamma}}(t)\big)}\cdot\frac{1}{\dot{f}(t)}\cdot\dot{f}(t)\,\mathrm{d}t \\
&= L(\tilde{\gamma}([\tilde{a},\tilde{b}]))
\end{aligned}$$

where we used that $\dot{f}(t)$ is increasing. (QED)

6.3 Orientation

Given a basis $(e_1,...,e_n)$ we want to assign a volume to the corresponding parallelopiped.



Let V be a vector space and $(e_1, ..., e_n)$ a basis for V. Any other basis $(e'_1, ..., e'_n)$ for V can be expressed in terms of the original basis

$$e'_a = \sum_{b=1}^{n} e_b S^b_a \quad a = 1, ..., n$$

The coefficients S^b_a form a non-singular matrix, because we can equally well express the original basis $(e_1, ..., e_n)$ in terms of the new basis $(e'_1, ..., e'_n)$, thus S must be invertible

$$\det S \neq 0 \quad \Rightarrow \quad \text{either } \det S > 0 \text{ or } \det S < 0$$

Definition 6.8: If $\det S > 0$, we say that the new basis has the *same orientation* as the original one and if $\det S < 0$, we say that the new basis has the *opposite orientation*. This partitions the bases for V into two classes. An orientation for V is a choice of class which we call *positive*. The corresponding bases are called *positive*. The other choice would of course be the *negative orientation* with the *negative bases*. □

Definition 6.9: A *volume form* ω on an *oriented vector space* V is a totally antisymmetric n-linear form on V, where $n = \dim V$, such that $\omega(e_1, ..., e_n) > 0$ for any positive basis $(e_1, ..., e_n)$ for V. □

Example 6.1: Let be $n = 2$. Then the change of basis formula takes the form

$$e'_1 = e_1 S^1{}_1 + e_2 S^2{}_1 \qquad e'_2 = e_1 S^1{}_2 + e_2 S^2{}_2$$

and we have

$$\begin{aligned} \omega(e'_1, e'_2) &= \omega(e_1 S^1{}_1 + e_2 S^2{}_1, e_1 S^1{}_2 + e_2 S^2{}_2) \\ &= S^1{}_1 S^2{}_2 \omega(e_1, e_2) + S^2{}_1 S^1{}_2 \omega(e_2, e_1) \\ &= (S^1{}_1 S^2{}_2 - S^1{}_2 S^2{}_1)\omega(e_1, e_2) \\ &= (\det S)\omega(e_1, e_2) \end{aligned}$$

In general, for any n, we have

$$\omega(e'_1, ..., e'_n) = (\det S)\omega(e_1, ..., e_n)$$

This follows from the axioms for a determinant. As a consequence, once ω is positive on one positive basis, it is positive on all positive bases and negative on all negative bases. □

Remark 6.10: Suppose now that the oriented vector space V is endowed with an inner product $\langle \cdot, \cdot \rangle$. Then there is a unique volume form ω corresponding to this inner product. This is defined as follows. Suppose $(e_1, ..., e_n)$ is a positive orthonormal basis for V (with respect to $\langle \cdot, \cdot \rangle$). So $\langle e_a, e_b \rangle = \delta_{ab}$, $a, b = 1, ..., n$. We set $\omega(e_1, ..., e_n) = 1$, because it makes sense. Therefore, if $(e'_1, ..., e'_n)$ is another positive orthonormal basis, then

$$e'_a = \sum_{b=1}^{n} e_b O^b{}_a$$

where O is an orthogonal matrix with $\det O > 0$, hence $\det O = 1$. This is because

$$\begin{aligned} \delta_{ab} &= \langle e'_a, e'_b \rangle = \left\langle \sum_c e_c O^c{}_a, \sum_d e_d O^d{}_b \right\rangle \\ &= \sum_{c,d} O^c{}_a O^d{}_b \langle e_c, e_d \rangle = \sum_c O^c{}_a O^c{}_b = (O^T O)^a{}_b \end{aligned}$$

or simply

$$O^T O = \mathbf{1}$$

Hence

$$(\det O^T)(\det O) \;=\; 1 \quad \text{or} \quad (\det O)^2 \;=\; 1$$

and since $\det O > 0$, it follows $\det O = 1$. Then

$$\omega(e'_1, ..., e'_n) \;=\; (\det O)\omega(e_1, ..., e_n) \;=\; 1$$

□

Definition 6.10: A vector bundle $\mathcal{B}$ over a manifold M is *orientable*, if we can make a continuous choice of orientation for $\mathcal{B}_p$ at each $p \in M$. □

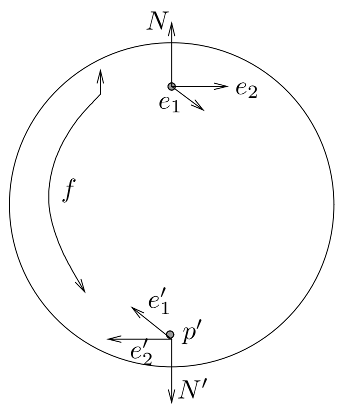
Remark 6.11: More precisely, if $(e_1, ..., e_n)$ is a positive basis for $\mathcal{B}_p$, then if q belongs to a suitably small neighborhood of p in M, and each e'_a, $a = 1, .., n$ is in a suitably small neighborhood of the corresponding e_a, $a = 1, ..., n$ in $\mathcal{B}$, then $(e'_1, ..., e'_n)$ is a positive basis for $\mathcal{B}_q$. □

Definition 6.11: A manifold M is *orientable*, if its tangent bundle TM is orientable in the above sense. □

Example 6.2: The Möbius bundle
The Möbius bundle, a line bundle over S^1, is the classic example of a non-orientable vector bundle. On the other hand, $TS^1 = S^1 \times \mathbb{R}$ is orientable. □

Remark 6.12: In general, if M is an $n-1$ dimensional submanifold of $\mathbb{R}^n$ (a hypersurface in $\mathbb{R}^n$), which is the boundary of a compact set in $\mathbb{R}^n$, we can define at each point $p \in M$ the unit outward normal N_p to M at p. We then define an orientation for M by calling a basis $(e_1, ..., e_{n-1})$ for T_pM positive if and only if $(e_1, ..., e_{n-1}, N_p)$ is a positive basis for $\mathbb{R}^n$. Thus, every such submanifold of $\mathbb{R}^n$ is orientable. Thus $S^{n-1} \subset \mathbb{R}^n$ is orientable. □

Remark 6.13: Consider now $\mathbb{R}P^{n-1}$, the quotient of S^{n-1} by the equivalence relation induced by the antipodal map $f : f(x) = -x$. For now Consider in particular the case $n = 3$, namely the 2-sphere. Then $(\mathrm{d}f)(x) \in \mathscr{L}(T_xS^2, T_{f(x)}S^2)$ ($\mathscr{L}$ denotes linear map) is *orientation-reversing*, so it maps a positive basis to a negative basis.



$e'_i = \mathrm{d}f \cdot e_i$, $i = 1, 2$

(e_1, e_2, N) is a positive basis

(e'_1, e'_2, N') is a negative basis

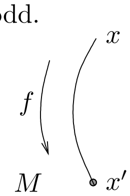
(e_1, e_2, N) is positive while (e'_1, e'_2, N') is negative. □

Remark 6.14: In general, the antipodal map $f : f(x) = -x$ on $S^{n-1} \subset \mathbb{R}^n$ is orientation preserving for n even and orientation reversing for n odd. □

Proposition 6.3: The quotient of an orientable manifold M by the equivalence relation induced by an orientation preserving map is itself orientable, while that by an orientation reversing map is non-orientable. In particular

$$\mathbb{R}P^{n-1} \;=\; S^{n-1}/\sim \quad \text{(antipodal map)}$$

is orientable for n even and non-orientable for n odd.



□

6.4 Volume form of an orientable Rimannian manifold

Let (M, g) be an m-dimensional oriented manifold. Then at each $p \in M$ the inner product g_p in T_pM together with the orientation of T_pM define a volume form ω_p in T_pM, a totally anti-symmetric m-linear ($m = \dim M$) form on T_pM such that $\omega_p(e_1, ..., e_m) = 1$ on any positive orthonormal basis $(e_1, ..., e_m)$. We thus have a section ω of the bundle $\Lambda_m M$ of top degree forms on M (a tensor bundle over M).

Definition 6.12: ω is continuously differentiable if g is continuously differentiable. ω is denoted by $\mathrm{d}\mu_g$, the *volume form* of (M, g). □

Remark 6.15: To express things in local coordinates, we express $\frac{\partial}{\partial x^\mu}$, $\mu = 1, ..., m$, the coordinate vectorfields of a chart, in terms of a positive orthonormal frame field (e_μ, $\mu = 1, ..., m$) defined in the domain of the chart

$$\frac{\partial}{\partial x^\mu} = \sum_{\nu=1}^{m} e_\nu S^\nu_\mu$$

Let $g_{\mu\nu} = g\left(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right)$ be the components of g in the chart. Then

$$g_{\mu\nu} = g\left(\sum_k e_k S^k_\mu, \sum_\lambda e_\lambda S^\lambda_\nu\right) = \sum_{k,\lambda} S^k_\mu S^\lambda_\nu g(e_k, e_\lambda) = \sum_k S^k_\mu S^k_\nu$$

In terms of matrices, this means $g = S^T S$. It follows that

$$\det g = (\det S^T)(\det S) = (\det S)^2$$

According to the formula for change of bases, we get

$$\omega\left(\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^m}\right) = (\det S)\omega(e_1, ..., e_m) = \det S$$

If $\left(\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^m}\right)$ is a positive basis, then $\det S > 0$, hence

$$\det S = \sqrt{\det g}$$

We conclude that $\mathrm{d}\mu_g$ is given in the chart by

$$\mathrm{d}\mu_g = \sqrt{\det g}\, \mathrm{d}x^1 \wedge ... \wedge \mathrm{d}x^m$$

The k-fold ($k \leq m$) outer product of covectors $\xi^1, ..., \xi^k$ is given by the full antisymmetrization

$$(\xi^1 \wedge ... \wedge \xi^k) \cdot (X_1, ..., X_k) = \sum_\pi (-1)^{|\pi|} (\xi^1 \cdot X_{\pi(1)}) ... (\xi^k \cdot X_{\pi(k)})$$

The sum is over all permutations π and we denote by $|\pi|$ the length of the permutation (number of pairs exchanged). Thus

$$(-1)^{|\pi|} = \begin{cases} -1, & \text{if } \pi \text{ is odd} \\ 1, & \text{if } \pi \text{ is even} \end{cases}$$

□

Example 6.3: We have

$$\begin{aligned}(\xi^1 \wedge \xi^2 \wedge \xi^3) \cdot (X_1, X_2, X_3) = & (\xi^1 \cdot X_1)(\xi^2 \cdot X_2)(\xi^3 \cdot X_3) + (\xi^1 \cdot X_2)(\xi^2 \cdot X_3)(\xi^3 \cdot X_1) \\ & + (\xi^1 \cdot X_3)(\xi^2 \cdot X_1)(\xi^3 \cdot X_2) - (\xi^1 \cdot X_3)(\xi^2 \cdot X_2)(\xi^3 \cdot X_1) \\ & - (\xi^1 \cdot X_2)(\xi^2 \cdot X_1)(\xi^3 \cdot X_3) - (\xi^1 \cdot X_1)(\xi^2 \cdot X_3)(\xi^3 \cdot X_2)\end{aligned}$$

The number of terms is the order of π for k elements, that is $k!$. □

6.5 Volume of a domain

Let M be an oriented manifold endowed with a volume form ω. Let D be a domain in M contained in the domain of a chart (U, φ), so $D \subset U$. Then $\omega(\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^m})$ is a function in U and $\omega(\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^m}) \circ \varphi^{-1}$ is a function on the open set $\varphi(U) \subset \mathbb{R}^n$ ($\varphi(D) \subset \varphi(U)$).

Definition 6.13: We define the *volume of the domain D* by

$$\mathrm{Vol}(D) \;=\; \int_D \omega \;=\; \int_{\varphi(D)} \omega\left(\frac{\partial}{\partial x^1},\dots,\frac{\partial}{\partial x^m}\right)\circ\varphi^{-1}\,\mathrm{d}x^1\dots\mathrm{d}x^m$$

where $\varphi(D) \subset \mathbb{R}^n$. □

Remark 6.16: To show that this is well defined, let $(\tilde{U},\tilde{\varphi})$ be another chart such that also $D \subset \tilde{U}$. Then with $x^\nu = \varphi^\nu(p)$ and $\tilde{x}^\mu = \tilde{\varphi}^\mu(p)$ where $p \in D$, we have

$$x^\nu \;=\; f^\nu(\tilde{x}^1,\dots,\tilde{x}^m)$$

where $f = \varphi\circ\tilde{\varphi}^{-1}$ and

$$\left.\frac{\partial}{\partial\tilde{x}^\mu}\right|_p \;=\; \sum_{\nu=1}^m \left.\frac{\partial}{\partial x^\nu}\right|_p S^\nu{}_\mu(p) \qquad S^\nu{}_\mu(p) \;=\; \left.\frac{\partial x^\nu}{\partial\tilde{x}^\mu}\right|_p \;=\; \frac{\partial f^\nu}{\partial\tilde{x}^\mu}(\tilde{\varphi}(p))$$

Hence, by the change of bases formula,

$$\omega\left(\frac{\partial}{\partial\tilde{x}^1},\dots,\frac{\partial}{\partial\tilde{x}^m}\right) \;=\; (\det S)\,\omega\left(\frac{\partial}{\partial x^1},\dots,\frac{\partial}{\partial x^m}\right)$$

which are functions on D, and

$$\omega\left(\frac{\partial}{\partial\tilde{x}^1},\dots,\frac{\partial}{\partial\tilde{x}^m}\right)\circ\tilde{\varphi}^{-1}(\tilde{x}) \;=\; \left(\det\frac{\partial f}{\partial\tilde{x}}(\tilde{x})\right)\omega\left(\frac{\partial}{\partial x^1},\dots,\frac{\partial}{\partial x^m}\right)\circ\varphi^{-1}(x)$$

where we used that $\tilde{\varphi}^{-1}(\tilde{x}) = \varphi^{-1}(x) = p$. Now, $\det\frac{\partial f}{\partial\tilde{x}}(\tilde{x})$ is the Jacobian determinant of the transformation $\tilde{x}\mapsto x = f(\tilde{x})$. Recall the change of variables formula $\tilde{x}\mapsto x = f(\tilde{x})$

$$\int_{f(\varepsilon)} \mu(x)\,\mathrm{d}x^1\dots\mathrm{d}x^m \;=\; \int_\varepsilon \mu(f(\tilde{x}))\det\left(\frac{\partial f}{\partial\tilde{x}}(\tilde{x})\right)\mathrm{d}\tilde{x}^1\dots\mathrm{d}\tilde{x}^m$$

Here, we take $\varepsilon = \tilde{\varphi}(D)$ so that $f(\varepsilon) = \varphi(D)$ and

$$\mu \;=\; \omega\left(\frac{\partial}{\partial x^1},\dots,\frac{\partial}{\partial x^m}\right)\circ\varphi^{-1}$$

to obtain

$$\int_{\tilde{\varphi}(D)} \omega\left(\frac{\partial}{\partial\tilde{x}^1},\dots,\frac{\partial}{\partial\tilde{x}^m}\right)\circ\tilde{\varphi}^{-1}\,\mathrm{d}\tilde{x}^1\dots\mathrm{d}\tilde{x}^m \;=\; \int_{\varphi(D)} \omega\left(\frac{\partial}{\partial x^1},\dots,\frac{\partial}{\partial x^m}\right)\circ\varphi^{-1}\,\mathrm{d}x^1\dots\mathrm{d}x^m$$

This is precisely what was needed to be shown. □

6.6 Partitions of Unity

To remove the restriction that D has to be contained in the domain of a chart, we introduce the partitions of unity. Let $\mathcal{A} = \{(U_\alpha,\varphi_\alpha) : \alpha\in I\}$ be an atlas for M. We can assume (by paracompactness) that $\mathcal{A}$ is locally finite, that is, each point $p\in M$ has a neighborhood U_p which intersects only finitely many of the U_α. We can then construct a partition of unity relative to $\mathcal{A}$.

Definition 6.14: A *partition of unity* is a collection of functions $\{f_\alpha : \alpha\in I\}$ such that

1) Each f_α is non-negative, smooth and has support in U_α.

2) We have the sum

$$\sum_\alpha f_\alpha \;=\; 1$$

and because of local finiteness this sum contains only finitely many terms at each point. In fact, by local finiteness for any given $p\in M$, there is a neighborhood U_p such that all, except finitely many of the f_α vanish on U_p.

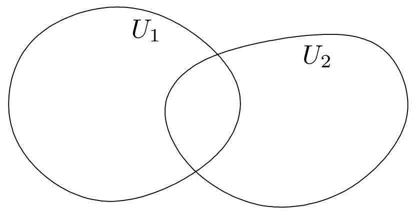
□

Remark 6.17: Given now any domain $D \subset M$ we define

$$\mathrm{Vol}(D) = \sum_\alpha \int_D f_\alpha w = \sum_\alpha \int_{D\cap U_\alpha} f_\alpha w$$

This is either a positive real number or ∞ (if defined so). If $\overline{D}$ is compact, then $\mathrm{Vol}(D)$ is finite. The definition makes sense because it is independent of the particular choice of partition of unity for a given choice of a locally finite atlas for M. □

Example 6.4: Take to intersecting sets



On $U_1 \setminus U_2$, f_2 vanishes, so $f_1 = 1$. Similarly on $U_2 \setminus U_1$ where f_1 vanishes so $f_2 = 1$. But on $U_1 \cap U_2$ we have $f_1 + f_2 = 1$. □

6.7 Volume of a Submanifold

Let N be an n-dimensional submanifold of an m-dimensional *Riemannian manifold* (M, g).

Definition 6.15: We define on N the Riemannian metric h, called the *induced metric* (by g on N),

$$h = g|_{TN}$$

Then (N, h) is itself a Riemannian manifold and the definition of volume applies to any domain on N. □

Remark 6.18: The primary data of geometry are the measurements of arclengths and volumes. □

Example 6.5: Consider $\mathbb{R}^3 \setminus 0$. This is diffeomorphic to $\mathbb{R} \times S^2$. Consider on $\mathbb{R} \times S^2$ the metric

$$g = (\mathrm{d}r)^2 + R^2(r)\gamma^\circ \quad r \in \mathbb{R}$$

where γ° is the standard metric on S^2. In polar coordinates (ϑ, φ) this is

$$\gamma^\circ = (\mathrm{d}\vartheta)^2 + \sin^2\vartheta(\mathrm{d}\varphi)^2$$

where (r, ϑ, φ) are the coordinates on $\mathbb{R} \times S^2$. The vectors tangent to $\{r\} \times S^2 = S_r$, that is to the r-spheres, are linear combinations of $\frac{\partial}{\partial\vartheta}$ and $\frac{\partial}{\partial\varphi}$. So the metric γ_r induced on S_r is

$$\gamma_r = R^2(r)\gamma^\circ$$

The volume of a 2-dimensional manifold is called area. We have

$$\mathrm{Area}(S_r) = \int_{S_r} \mathrm{d}\mu_{\gamma_r} \qquad \mathrm{Area}(S^2) = \int_{S^2} \mathrm{d}\mu_{\gamma^\circ} = 4\pi$$

To check this, we write down in polar coordinates

$$\mathrm{d}\mu_{\gamma^\circ} = \sqrt{\det\gamma^\circ}\,\mathrm{d}\vartheta \wedge \mathrm{d}\varphi$$

where

$$\gamma^\circ = \begin{pmatrix} \gamma^\circ_{\vartheta\vartheta} & \gamma^\circ_{\vartheta\varphi} \\ \gamma^\circ_{\varphi\vartheta} & \gamma^\circ_{\varphi\varphi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\vartheta \end{pmatrix}$$

and therefore

$$\det\gamma^\circ = \sin^2\vartheta \quad \Rightarrow \quad \mathrm{d}\mu_{\gamma^\circ} = \sin\vartheta\,\mathrm{d}\vartheta \wedge \mathrm{d}\varphi$$

So, we get

$$\int_{S^2} \mathrm{d}\mu_{\gamma^\circ} \quad = \quad \int_0^{2\pi}\int_0^{\pi} \sin\vartheta \,\mathrm{d}\vartheta\,\mathrm{d}\varphi \quad = \quad 4\pi$$

Futhermore we have

$$\mathrm{d}\mu_{\gamma_r} \quad = \quad \sqrt{\det\gamma_r}\,\mathrm{d}\vartheta\wedge\mathrm{d}\varphi \quad = \quad R^2(r)\sqrt{\det\gamma^\circ}\,\mathrm{d}\vartheta\wedge\mathrm{d}\varphi \quad = \quad R^2(r)\,\mathrm{d}\mu_{\gamma^\circ}$$

and therefore

$$\mathrm{Area}(S_r) \quad = \quad 4\pi R^2(r)$$

Let now be $B(r_1, r_2) = [r_1, r_2] \times S^2$. Then

$$\mathrm{Vol}(B(r_1,r_2)) \quad = \quad \int_{B(r_1,r_2)} \mathrm{d}\mu_g$$

where

$$g \quad = \quad \begin{pmatrix} 1 & 0 \\ 0 & \gamma_r \end{pmatrix} \quad \Rightarrow \quad \det g \quad = \quad \det\gamma_r$$

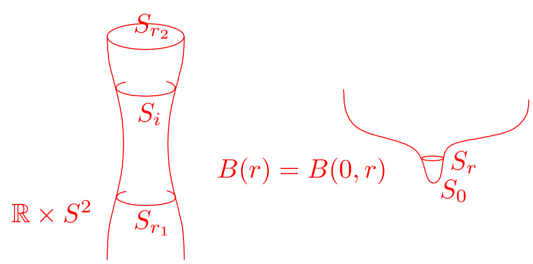
It follows that

$$\mathrm{d}\mu_g \quad = \quad \sqrt{\det\gamma_r}\,\mathrm{d}r\wedge\mathrm{d}\vartheta\wedge\mathrm{d}\varphi \quad = \quad \mathrm{d}r\wedge\mathrm{d}\mu_{\gamma_r}$$

So we get for the volume of $B(r_1, r_2)$

$$\mathrm{Vol}(B(r_1,r_2)) \quad = \quad \int_{r_1}^{r_2}\int_{S_r} \mathrm{d}\mu_{\gamma_r}\,\mathrm{d}r \quad = \quad \int_{r_1}^{r_2} \mathrm{Area}(S_r)\,\mathrm{d}r \quad = \quad 4\pi\int_{r_1}^{r_2} R^2(r)\,\mathrm{d}r$$

Now restrict r to $[0,\infty)$ and set S_0 to be a point by taking $R(0) = 0$. This then is homeomorphic to $\mathbb{R}^3$.



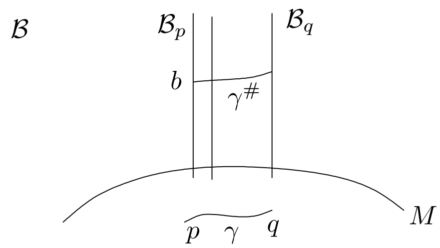
We have

$$\lim_{r\to\infty} \frac{\mathrm{Area}(S_r)}{4\pi r^2} \quad = \quad 1 \qquad\qquad \lim_{r\to\infty} \frac{\mathrm{Vol}(B_r)}{\frac{4\pi r^3}{3}} \quad = \quad 1$$

□

7 Connections

7.1 Parallel transport



Let be $\mathcal{B}$ a vector bundle over M. In general, if q and p are two distinct points in M, we have no way of corresponding elements of $\mathcal{B}_q$ to elements of $\mathcal{B}_p$. So given $b \in \mathcal{B}_p$, we do not know to which element of $\mathcal{B}_q$ this corresponds. There is one exception: 0_q the 0-element of $\mathcal{B}_q$ must correspond to 0_p, the 0-element of $\mathcal{B}_p$. Consider now a curve γ joining p to q

$$\gamma : [0,1] \to M, \;\; \gamma(0) = p, \;\; \gamma(1) = q$$

We then want to find a curve $\gamma^\#$ in $\mathcal{B}$ starting at b, so

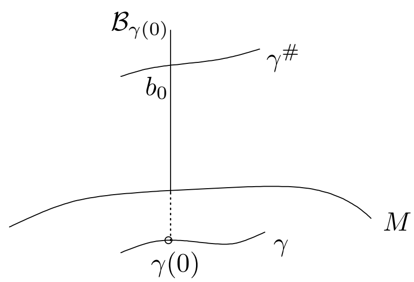
$$\gamma^\# : [0,1] \to \mathcal{B}, \;\; \gamma^\#(0) = b$$

such that $\gamma^\#$ projects to γ, that is $\pi \circ \gamma^\# = \gamma$ or

$$\pi(\gamma^\#(t)) \;\; = \;\; \gamma(t) \quad \forall t \in (0,1]$$

This will be called the horizontal lift of γ to $\mathcal{B}$ initiating at b. It defines the parallel transport of b along γ. Evaluating at the other end point we obtain the corresponding vector in $\mathcal{B}_q$.

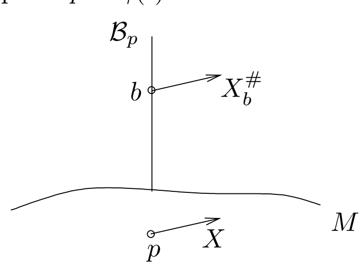
One problem is to define $\gamma^\#$, the horizontal lift to $\mathcal{B}$ of a curve γ in M through a given $b_0 \in \mathcal{B}_{\gamma(0)}$



For this we require an ordinary differential equation for $\gamma^\#$, depending on γ, because $\gamma^\#$ must be uniquely determined by the initial condition $\gamma^\#(0) = b_0$. That is, we must have some rule which specifies the instantaneous velocity $\dot{\gamma}^\#(t)$ of $\gamma^\#$ in terms of the instantaneous position $\gamma^\#(t)$ and the corresponding instantaneous velocity $\dot{\gamma}(t)$ of γ. So we want a rule F of the form

$$\dot{\gamma}^\#(t) \;\; = \;\; F(\gamma^\#(t), \dot{\gamma}(t))$$

Consider the argument of F for a given point $p = \gamma(t) \in M$. Then F is a function of $\mathcal{B}_p \times T_pM$.



Se set $p = \gamma(t)$, $b = \gamma^\#(t)$ and write

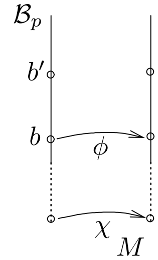
$$F(b, X) \;\; = \;\; X^\#_b \qquad b \in \mathcal{B}_p, \; X \in T_pM$$

This is the horizontal lift of the vector $X \in T_pM$ to $b \in \mathcal{B}_p$, an element of $T_b\mathcal{B}$. The horizontal lift of vectors $X \in T_pM$ to $b \in \mathcal{B}_p$, shall be defined by imposing 4 conditions. But first we need the notion of vertical transformations.

Definition 7.1: Suppose that ϕ is a differentiable mapping of $\mathcal{B}$ into itself. We say that ϕ is *fibre-preserving*, if there is a differentiable mapping χ of M into itself, such that

$$\pi \circ \phi = \chi \circ \pi$$

ϕ being fibre-preserving means in other words, that it maps fibres to fibres.



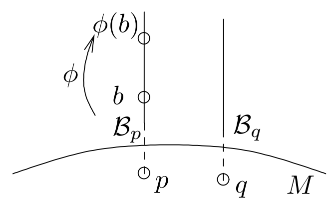
□

Definition 7.2: A special case of fibre-preserving mappings ϕ of $\mathcal{B}$ into itself are *vertical transformations*. In this case

$$\pi \circ \phi = \pi$$

So the mapping χ of M into itself is simply the identity. In other words, $\phi : \mathcal{B} \to \mathcal{B}$ is a vertical transformation if

$$b \in \mathcal{B}_p \quad \Rightarrow \quad \phi(b) \in \mathcal{B}_p \quad \forall b \in \mathcal{B}_p, \forall p \in M$$



□

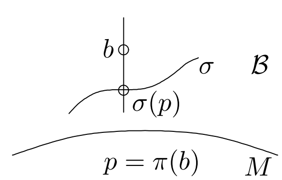
Remark 7.1: The basic vertical transformations are scalar multiplication and vector addition, the basic linear operations.

- *Multiplication:* Given $c \in \mathbb{R}$ we have the vertical transformation

$$M_c(b) = cb \quad \forall b \in \mathcal{B}$$

- *Addition:* Given a section σ of $\mathcal{B}$ we have the vertical transformation

$$S_\sigma(b) = b + \sigma(\pi(b)) \quad \forall b \in \mathcal{B}$$



□

Definition 7.3: The *horizontal lift* $X_b^\#$ of $X \in T_pM$ to $b \in \mathcal{B}_p$, $p \in M$ is defined by 4 conditions:

(1) $\mathrm{d}\pi \cdot X_b^\# = X$

(2) $X_b^\#$ depends linearly on X

(3) It is compatible with scalar multiplication, in the sense that

$$X_{M_c(b)}^\# = \mathrm{d}M_c \cdot X_b^\#$$

(4) It is compatible with vector addition, that is

$$X^{\#}_{S_\sigma(b)} - \mathrm{d}S_\sigma \cdot X^{\#}_b \quad = \quad X^{\#}_{\sigma(p)} - \mathrm{d}\sigma \cdot X$$

for every section σ of $\mathcal{B}$.

The conditions will be explained further in the following remarks. □

Remark 7.2: The first condition is already known. It follows from the condition $\pi \circ \gamma^{\#} = \gamma$ on the horizontal lift of a curve by differentiating both sides with respect to the parameter t and denoting X the vectorfield to γ and $X^{\#}$ the vectorfield to $\gamma^{\#}$ (see Explanation A.10 on page 87). □

Definition 7.4: We define the *horizontal subspace* at b, denoted H_b, by

$$H_b \quad := \quad \{X^{\#}_b : X \in T_pM\} \quad \subset \quad T_b\mathcal{B}$$

This is a linear subspace of $T_b\mathcal{B}$ by condition (2). □

Remark 7.3: By condition (1) the mapping $X \mapsto X^{\#}_b$ is injective. To show this, suppose that $0 = X^{\#}_b \in T_p\mathcal{B}$. Then by (1) $\mathrm{d}\pi \cdot X^{\#}_b = 0 = X \in T_pM$. Hence

$$\dim H_b \quad = \quad \dim T_pM \quad = \quad m$$

and $\mathrm{d}\pi|_{H_b}$ (the inverse of the mapping $X \mapsto X^{\#}_b$) is a linear isomorphism of H_b onto T_pM. □

Definition 7.5: Define

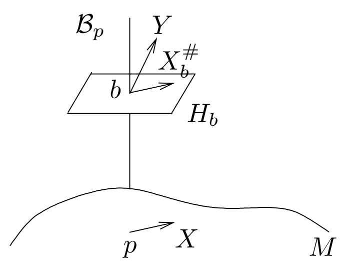
$$V_b \quad := \quad T_b\mathcal{B}_p$$

to be the *vertical subspace* at b. □

Remark 7.4: We have $\dim H_b = m$ and $\mathrm{d}\pi \cdot V = 0$ for every $V \in V_b$, so $V \notin H_b$. It follows

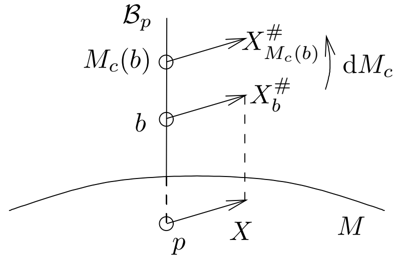
$$T_b\mathcal{B} \quad = \quad H_b \oplus V_b$$

That is, any vector $Y \in T_b\mathcal{B}$ can be uniquely decomposed into a vector $X^{\#}_b \in H_b$, where $X = \mathrm{d}\pi \cdot Y \in T_pM$ with $p = \pi(b)$ and a vector $Y - X^{\#}_b \in V_b$.



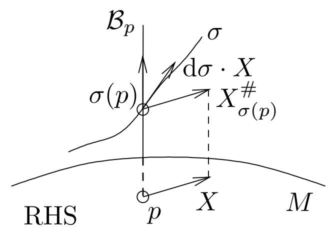
□

Remark 7.5: We could try to picture condition (3) as follows



□

Remark 7.6: The right-hand side in condition (4) takes into account the variation of σ over M.



The vector $\mathrm{d}\sigma \cdot X$ projects to X. That means

$$\mathrm{d}\pi \cdot (\mathrm{d}\sigma \cdot X) \;=\; X$$

which is because $\pi \circ \sigma = \mathrm{id}_M$. So, its horizontal part is $X^{\#}_{\sigma(p)}$. The vertical part is therefore

$$D_X\sigma \;:=\; \mathrm{d}\sigma \cdot X - X^{\#}_{\sigma(p)}$$

what we call *covariant derivative* of σ with respect to X. The fact that $0_q \in \mathcal{B}_q$ corresponds to $0_p \in \mathcal{B}_p$ for every pair of points $p, q \in M$, is reflected in the fact that the covariant derivative of 0-sections σ^0, defined by

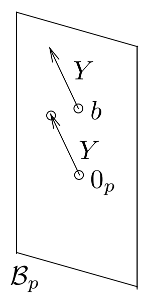
$$\sigma^0(p) \;=\; 0_p \quad \forall p \in M$$

is itself 0_p

$$D_X\sigma^0 \;=\; 0_p \quad \forall X \in T_pM \tag{7.1}$$

(see Explanation A.11 on page 87). □

Remark 7.7: Let Y be a vertical vector at $b \in \mathcal{B}$. That is $Y \in T_b\mathcal{B}_p$ with $p = \pi(b)$.



$\mathcal{B}_p$ being a linear space, a tangent vector at any point can be thought of as an element of the linear space. Thus Y can be thought of as an element of $\mathcal{B}_p$, that is, a vector attached at the origin. □

Remark 7.8: The meaning of condition (4) can be understood as follows. Since $\pi \circ S_\sigma = \pi$ (also $\pi \circ M_c = \pi$, both being vertical transformations), we get that

$$\mathrm{d}\pi \cdot \mathrm{d}S_\sigma \cdot X^{\#}_b \;=\; \mathrm{d}\pi \cdot X^{\#}_b \;=\; \mathrm{d}\pi \cdot X_{S_\sigma(b)} \;=\; X$$

so $\mathrm{d}\pi$ applied to the left-hand side of condition (4) vanishes. Hence the left-hand side is a vertical vector, just like the right-hand side (as shown previously). But the left-hand side is a vertical vector attached at the point $S_\sigma(b) = b + \sigma(p) \in \mathcal{B}_p$, while the right-hand side is a vertical vector attached at the point $\sigma(p) \in \mathcal{B}_p$. But the equality makes sense, because of what we discussed in the basic Remark 7.7. □

7.2 In local Coordinates

Choose a bundle atlas for $\mathcal{B}$ such that the corresponding domains U in M are domains of charts. Consider then any of the domains $U \subset M$ which together cover M. Over U we have the basis sections ι_a, $a = 1, ..., n$. We also have in U the coordinates x^μ, $\mu = 1, ..., m$ ($\dim M = m$, $\dim \mathcal{B} = m + n$). We can then expand any vector $v \in \mathcal{B}_p$ as

$$v \;=\; \sum_{a=1}^{n} v^a \iota_a(p)$$

The $(v^a,\ a = 1, ..., n)$ constitute a system of linear coordinates for $\mathcal{B}_p$. Together with the x^μ, $\mu = 1, ..., m$, we have coordinates for $\mathcal{B}$ in $\pi^{-1}(U)$:

$$(x^1, ..., x^m; v^1, ..., v^n)$$

We may simplify the notation to $(x^\mu; v^a)$. We expand a given $X \in T_pM$ as

$$X = \sum_{\mu=1}^{m} X^\mu \left.\frac{\partial}{\partial x^\mu}\right|_p$$

We may also expand $X_v^\#$ as

$$X_v^\# = \sum_{\mu=1}^{m} Y^\mu \left.\frac{\partial}{\partial x^\mu}\right|_v + \sum_{a=1}^{n} Z^a \left.\frac{\partial}{\partial v^a}\right|_v$$

Here it is $Y^\mu = Y^\mu(v, X)$ and $Z^a = Z^a(v, X)$.

Proposition 7.1: The coefficients Y^μ satisfy

$$Y^\mu(v, X) = X^\mu \quad \mu = 1, ..., m \tag{7.2}$$

which is a consequence of condition (1). □

Proof: Condition one requires $\mathrm{d}\pi \cdot X_v^\# = X$. The projection map π is represented in our coordinates by

$$(x^1, ..., x^m; v^1, ..., v^n) \mapsto (x^1, ..., x^m)$$

$(x^1, ..., x^m)$ being the coordinates of p. Thus

$$\mathrm{d}\pi \cdot \left.\frac{\partial}{\partial x^\mu}\right|_v = \left.\frac{\partial}{\partial x^\mu}\right|_p \qquad \mathrm{d}\pi \cdot \left.\frac{\partial}{\partial v^a}\right|_v = 0$$

Hence, by linearity of the differential, we get

$$\mathrm{d}\pi \cdot X_v^\# = \sum_{\mu=1}^{m} Y^\mu \left.\frac{\partial}{\partial x^\mu}\right|_p = \sum_{\mu=1}^{m} X^\mu \left.\frac{\partial}{\partial x^\mu}\right|_p = X$$

(QED)

Proposition 7.2: The coefficients Z^μ satisfy

$$Z^a(v, X) = \sum_{\mu=1}^{m} Z_\mu^a(v) X^\mu \quad a = 1, ..., n \tag{7.3}$$

which is a consequence of condition (2). □

Proof: Condition (2) requires that $X_v^\#$ is linear in X. In view of (7.2), where linearity is already satisfied for X^μ this reduces to $Z^a(v, X)$ being linear in X. (QED)

Proposition 7.3: The coefficients Z_μ^a satisfy

$$Z_\mu^a(cv) = cZ_\mu^a(v) \quad a = 1, ..., n,\ \mu = 1, ..., m \tag{7.4}$$

which is a consequence of condition (3). □

Proof: We proceed to condition (3), requiring $X^{\#}_{M_c(v)} = \mathrm{d}M_c \cdot X^{\#}_v$. The multiplication M_c is represented in our coordinates by

$$(x^1, ..., x^m; v^1, ..., v^n) \quad \mapsto \quad (x^1, ..., x^m; cv^1, ..., cv^n)$$

So

$$\mathrm{d}M_c \cdot \left.\frac{\partial}{\partial x^\mu}\right|_v \quad = \quad \left.\frac{\partial}{\partial x^\mu}\right|_{M_c(v)} \qquad \mathrm{d}M_c \cdot \left.\frac{\partial}{\partial v^a}\right|_v \quad = \quad c \left.\frac{\partial}{\partial v^a}\right|_{M_c(v)}$$

Hence

$$\mathrm{d}M_c \cdot X^{\#}_v \quad = \quad \sum_{\mu=1}^{m} X^\mu \left.\frac{\partial}{\partial x^\mu}\right|_{M_c(v)} + \sum_{a=1}^{n}\sum_{\mu=1}^{m} cZ^a_\mu(v)X^\mu \left.\frac{\partial}{\partial v^a}\right|_{M_c(v)}$$

while

$$X^{\#}_{M_c(v)} \quad = \quad \sum_{\mu=1}^{n} X^\mu \left.\frac{\partial}{\partial x^\mu}\right|_{M_c(v)} + \sum_{a=1}^{n}\sum_{\mu=1}^{m} Z^a_\mu(cv)X^\mu \left.\frac{\partial}{\partial v^a}\right|_{M_c(v)}$$

(QED)

Proposition 7.4: The coefficients Z^a_μ also satisfy

$$Z^a_\mu(v + \sigma(p)) \quad = \quad Z^a_\mu(v) + Z^a_\mu(\sigma(p)) \tag{7.5}$$

for any section σ of $\mathcal{B}$, which is a consequence of condition (4). □

Proof: We want to express condition (4) in our coordinates. S_σ is represented by

$$(x^1, ..., x^m; v^1, ..., v^n) \quad \mapsto \quad (x^1, ..., x^m; v^1 + \sigma^1(p), ..., v^n + \sigma^n(p))$$

where $(x^1, ..., x^m)$ are the coordinates of p. Consider the x^μ coordinate line

$$t \mapsto (x^1, ..., x^\mu + t, ..., x^m; v^1, ..., v^n)$$

This is mapped by S_σ to the curve

$$t \mapsto \big(x^1, ..., x^\mu + t, ...x^m; v^1 + \sigma^1(\gamma(t)), ..., v^n + \sigma^n(\gamma(t))\big)$$

where $\gamma(t)$ is the curve in M represented by the x^μ coordinate line $t \mapsto (x^1, ..., x^\mu + t, ..., x^m)$, which are the coordinates of $\gamma(t)$ with $\gamma(0) = p$. It follows that

$$\mathrm{d}S_\sigma \cdot \left.\frac{\partial}{\partial x^\mu}\right|_v \quad = \quad \left.\frac{\partial}{\partial x^\mu}\right|_{S_\sigma(v)} + \sum_{a=1}^{n} \left.\frac{\partial \sigma^a}{\partial x^\mu}\right|_p \left.\frac{\partial}{\partial v^a}\right|_{S_\sigma(v)} \tag{7.6}$$

and

$$\mathrm{d}S_\sigma \cdot \left.\frac{\partial}{\partial v^a}\right|_v \quad = \quad \left.\frac{\partial}{\partial v^a}\right|_{S_\sigma(v)} \tag{7.7}$$

(see Explanation A.12 on page 87). Hence

$$\mathrm{d}S_\sigma \cdot X^{\#}_v \quad = \quad \sum_{\mu=1}^{m} X^\mu \left.\frac{\partial}{\partial x^\mu}\right|_{S_\sigma(v)} + \sum_{a=1}^{n}\sum_{\mu=1}^{m} \left(Z^a_\mu(v) + \left.\frac{\partial \sigma^a}{\partial x^\mu}\right|_p \right) X^\mu \left.\frac{\partial}{\partial v^a}\right|_{S_\sigma(v)}$$

Moreover, the section σ is represented by

$$(x^1, ..., x^m) \quad \mapsto \quad (x^1, ..., x^m; \sigma^1(p), ..., \sigma^n(p))$$

$(x^1, ..., x^m)$ being the coordinates of p. Here, p is considered as variable. Then

$$X^{\#}_{\sigma(p)} \quad = \quad \sum_{\mu=1}^{m} X^\mu \left.\frac{\partial}{\partial x^\mu}\right|_{\sigma(p)} + \sum_{a=1}^{n}\sum_{\mu=1}^{m} Z^a_\mu(\sigma(p))X^\mu \left.\frac{\partial}{\partial v^a}\right|_{\sigma(p)}$$

while

$$d\sigma \cdot X \;=\; \sum_{\mu=1}^{m} X^\mu \left.\frac{\partial}{\partial x^\mu}\right|_{\sigma(p)} + \sum_{\mu=1}^{m}\sum_{a=1}^{n} \left.\frac{\partial \sigma^a}{\partial x^\mu}\right|_p X^\mu \left.\frac{\partial}{\partial v^a}\right|_{\sigma(p)}$$

Note that this is because

$$d\sigma \cdot \left.\frac{\partial}{\partial x^\mu}\right|_p \;=\; \left.\frac{\partial}{\partial x^\mu}\right|_{\sigma(p)} + \sum_{a=1}^{n} \left.\frac{\partial \sigma^a}{\partial x^\mu}\right|_p \left.\frac{\partial}{\partial v^a}\right|_{\sigma(p)}$$

which can be shown in the same way as equation 7.6. Therefore

$$D_X\sigma \;:=\; d\sigma \cdot X - X^\#_{\sigma(p)} \;=\; \sum_{a=1}^{n}\sum_{\mu=1}^{m} \left(\left.\frac{\partial \sigma^a}{\partial x^\mu}\right|_p - Z^a_\mu(\sigma(p)) \right) X^\mu \left.\frac{\partial}{\partial v^a}\right|_{\sigma(p)}$$

On the other hand

$$X^\#_{S_\sigma(v)} \;=\; \sum_{\mu=1}^{m} X^\mu \left.\frac{\partial}{\partial x^\mu}\right|_{S_\sigma(v)} + \sum_{a=1}^{n}\sum_{\mu=1}^{m} Z^a_\mu(v+\sigma(p)) X^\mu \left.\frac{\partial}{\partial v^a}\right|_{S_\sigma(v)}$$

hence

$$X^\#_{S_\sigma(v)} - dS_\sigma \cdot X^\#_v \;=\; \sum_{a=1}^{n}\sum_{\mu=1}^{m} \left(Z^a_\mu(v+\sigma(p)) - Z^a_\mu(v) - \left.\frac{\partial \sigma^a}{\partial x^\mu}\right|_p \right) X^\mu \left.\frac{\partial}{\partial v^a}\right|_{S_\sigma(v)}$$

Now, condition (4) can be written as

$$X^\#_{S_\sigma(v)} - dS_\sigma \cdot X^\#_v + D_X\sigma \;=\; 0$$

Substituting, we see that the condition reduces to

$$\sum_{a=1}^{n}\sum_{\mu=1}^{m} \left(Z^a_\mu(v+\sigma(p)) - Z^a_\mu(v) - Z^a_\mu(\sigma(p)) \right) X^\mu \left.\frac{\partial}{\partial v^a}\right|_{S_\sigma(v)} \;=\; 0$$

(QED)

Remark 7.9: Propositions 7.3 and 7.4 together are equivalent to linearity of $Z^a_\mu(v)$ in v. Therefore, there are coefficients $A^a_{\mu b}(p)$, depending only on p, such that

$$Z^a_\mu(v) \;=\; -\sum_{b=1}^{n} A^a_{\mu b}(p) v^b$$

Putting everything together, we have arrived at the following result

$$X^\#_v \;=\; \sum_{\mu=1}^{m} X^\mu \left(\left.\frac{\partial}{\partial x^\mu}\right|_v - \sum_{a,b=1}^{n} A^a_{\mu b}(p) v^b \left.\frac{\partial}{\partial v^a}\right|_v \right) \qquad \forall X \in T_pM,\ \forall v \in \mathcal{B}_p$$

The coefficients $A^a_{\mu b}$ are functions in U, the domain of the chart in M. They are called the *connection coefficients* and they depend not only on the choice of chart for M, but also on the choice of bundle chart through the choice of basis sections ι_a, $a = 1, ..., n$. □

Remark 7.10: To elliminate the dependence on the choice of chart in M, we consider the *connection* 1*-form*

$$A^a_b \;=\; \sum_{\mu=1}^{m} A^a_{\mu b}\, dx^\mu$$

a matrix valued 1-form on M. This depends on the choice of basis sections ι_a, $a = 1, ..., n$, but not on the local coordinates x^μ, $\mu = 1, ..., m$ in M. □

Remark 7.11: Substituting for Z^a_μ in the formula for $D_X\sigma$ we obtain

$$D_X\sigma \;=\; \sum_{\mu=1}^m\sum_{a=1}^n D_\mu\, \sigma^a\big|_p X^\mu \iota_a(p) \;\in\; \mathcal{B}_p \quad \forall X \in T_pM$$

where

$$D_\mu\, \sigma^a\big|_p \;=\; \left.\frac{\partial \sigma^a}{\partial x^\mu}\right|_p + \sum_{b=1}^n A^a_{\mu b}(p)\sigma^b(p)$$

Note that here

$$D_\mu \sigma^a \;=\; \left(\mathrm{d}\sigma^a + \sum_{b=1}^n A^a_b \sigma^b\right)\cdot \frac{\partial}{\partial x^\mu}$$

If now X is a vectorfield in M, then $D_X\sigma$ is another section of $\mathcal{B}$ defined by

$$(D_X\sigma)(p) \;=\; D_{X(p)}\sigma \;\in\; \mathcal{B}_p \quad \forall p \in M$$

$D_X\sigma$ is called the *covariant derivative* of σ with respect to X. We have

$$D_{\partial/\partial x^\mu}\, \iota_a \;=\; \sum_{b=1}^n A^b_{\mu a}\iota_b$$

This may be thought of as the definition of the connection coefficients. □

Remark 7.12: Covariant differentiation satisfies the Leibnitz rule, because if f is a function on M and σ a section of $\mathcal{B}$, then $f\sigma$ is a section of $\mathcal{B}$ and we have

$$D_X(f\sigma) \;=\; fD_X\sigma + (Xf)\sigma$$

In particular, since we can write

$$\sigma \;=\; \sum_{a=1}^n \sigma^a \iota_a$$

where the σ^a are functions on M, and because $D_X\sigma$ is linear in σ, we have

$$\begin{aligned} D_X\sigma &= D_X\left(\sum_{a=1}^n \sigma^a\iota_a\right) \;=\; \sum_{a=1}^n D_X(\sigma^a\iota_a) \\ &= \sum_{a=1}^n \left(\sigma^a D_X\iota_a + (X\sigma^a)\iota_a\right) \\ &= \sum_{a=1}^n \left(\sigma^a \sum_{b=1}^n \left(A^b_a\cdot X\right)\iota_b + (X\sigma^a)\iota_a\right) \end{aligned}$$

Where we used linearity of $D_X\sigma$ with repsect to $X \in T_pM$ to write

$$D_X\iota_a \;=\; \sum_{\mu=1}^m X^\mu D_{\partial/\partial x^\mu}\iota_a \;=\; \sum_{\mu=1}^n\sum_{b=1}^n X^\mu A^b_{\mu a}\iota_b \;=\; \sum_{b=1}^n (A^b_a\cdot X)\iota_b$$

Writing $X\sigma^a = \mathrm{d}\sigma^a\cdot X$, we obtain

$$D_X\sigma \;=\; \sum_{a=1}^n \left(\mathrm{d}\sigma^a + \sum_{b=1}^n A^a_b\sigma^b\right)\cdot X\iota_a$$

□

Remark 7.13: For $A^a_{\mu b}$, you take coordinates for everything, for A^a_b you take no coordinates for the position, but for the basis sections. For A_μ you take only the coordinates for the position and for A, you take no coordinates. □

7.3 Changes of Basis Sections - Gauge Transformations

Let $(\tilde{\iota}_a,\ a = 1, ..., n)$ be another set of basis sections over the domain $\tilde{U}$ in M. Then, over $U \cap \tilde{U}$ we have two sets of bases and we can express

$$\tilde{\iota}_a \;=\; \sum_{b=1}^{n} \iota_b S_a^b$$

From now on we use the summation convention, so for any pair of upper and lower indices which are the same index, a summation over all possible values of this index is understood. Here S is a non-singular matrix function on $U \cap \tilde{U}$. Given any section ϕ of $\mathcal{B}$, then over $U \cap \tilde{U}$, we have the expressions

$$\phi \;=\; \phi^a \iota_a \;=\; \tilde{\phi}^a \tilde{\iota}_a$$

Substituting for $\tilde{\iota}_a$ in terms of $(\iota_1, ..., \iota_n)$, we obtain

$$\tilde{\phi}^a S_a^b \;=\; \phi^b \qquad \tilde{\phi}^a \;=\; (S^{-1})_b^a \phi^b$$

Proposition 7.5: The connection coefficients transform according to

$$S_c^b \tilde{A}_{\mu a}^c \;=\; A_{\mu c}^b S_a^c + \frac{\partial}{\partial x^\mu} \cdot S_a^b$$

or in matrix notation

$$S \tilde{A}_\mu \;=\; A_\mu S + \frac{\partial}{\partial x^\mu} \cdot S$$

□

Proof: Over U we have

$$D_{\partial/\partial x^\mu} \iota_a \;=\; A_{\mu a}^b \iota_b$$

and over $\tilde{U}$, we have

$$D_{\partial/\partial x^\mu} \tilde{\iota}_a \;=\; \tilde{A}_{\mu a}^b \tilde{\iota}_b$$

Over $U \cap \tilde{U}$, substituting $\tilde{\iota}_a = \iota_b S_a^b$, we obtain, using the Leibnitz rule

$$\begin{aligned} D_{\partial/\partial x^\mu}(\iota_b S_a^b) \;&=\; S_a^b D_{\partial/\partial x^\mu} \iota_b + \frac{\partial}{\partial x^\mu} \cdot S_a^b \, \iota_b \\ &=\; S_a^b A_{\mu b}^c \iota_c + \frac{\partial}{\partial x^\mu} \cdot S_a^b \, \iota_b \\ &=\; \left(S_a^c A_{\mu c}^b + \frac{\partial}{\partial x^\mu} \cdot S_a^b \right) \iota_b \end{aligned}$$

where

$$D_{\partial/\partial x^\mu} \tilde{\iota}_a \;=\; \tilde{A}_{\mu a}^b \tilde{\iota}_b \;=\; \tilde{A}_{\mu a}^b S_b^c \iota_c \;=\; \tilde{A}_{\mu a}^c S_c^b \iota_b$$

We conclude that

$$S_c^b \tilde{A}_{\mu a}^c \;=\; A_{\mu c}^b S_a^c + \frac{\partial}{\partial x^\mu} \cdot S_a^b$$

(QED)

Remark 7.14: Using matrix-column notation, we now denote by ϕ not the section itself, but the column

$$\phi \;=\; \begin{pmatrix} \phi^1 \\ \vdots \\ \phi^n \end{pmatrix}$$

representing the section in the original basis, and by $\tilde{\phi}$ the column representing the section in the new basis. The above transformation formulas then take the form

$$\tilde{\phi} \;=\; S^{-1} \phi$$

with

$$S \tilde{A}_\mu \;=\; A_\mu S + \frac{\partial S}{\partial x^\mu} \quad \rightarrow \quad \tilde{A}_\mu \;=\; S^{-1} A_\mu S + S^{-1} \frac{\partial S}{\partial x^\mu}$$

Consider now

$$D_{\partial/\partial x^\mu}\phi \;=\; (D_\mu\phi^a)\iota_a$$

Here, ϕ is the section itself and

$$D_\mu\phi^a \;=\; \frac{\partial\phi^a}{\partial x^\mu} + A^a_{\mu b}\phi^b \quad\rightarrow\quad D_\mu\phi \;=\; \frac{\partial\phi}{\partial x^\mu} + A_\mu\phi$$

Then, relative to the new basis

$$\begin{aligned}\widetilde{D_\mu\phi} &= \frac{\partial\tilde{\phi}}{\partial x^\mu} + \tilde{A}_\mu\tilde{\phi} \;=\; \frac{\partial}{\partial x^\mu}\left(S^{-1}\phi\right) + S^{-1}\left(A_\mu S + \frac{\partial S}{\partial x^\mu}\right)S^{-1}\phi \\ &= S^{-1}\frac{\partial\phi}{\partial x^\mu} - S^{-1}\frac{\partial S}{\partial x^\mu}S^{-1}\phi + S^{-1}A_\mu\phi + S^{-1}\frac{\partial S}{\partial x^\mu}S^{-1}\phi \\ &= S^{-1}D_\mu\phi\end{aligned}$$

where we used that it follows by the Leibniz rule, that

$$\frac{\partial}{\partial x^\mu}\cdot SS^{-1} \;=\; \frac{\partial S}{\partial x^\mu}S^{-1} + S\frac{\partial S^{-1}}{\partial x^\mu} \quad\Rightarrow\quad \frac{\partial S^{-1}}{\partial x^\mu} \;=\; -S^{-1}\frac{\partial S}{\partial x^\mu}S^{-1}$$

So $D_\mu\phi$ transforms like ϕ under a change of basis

$$\phi \;\mapsto\; S^{-1}\phi \qquad D_\mu\phi \;\mapsto\; S^{-1}D_\mu\phi$$

□

7.4 Curvature

If X is a vectorfield on the base manifold M and $\mathcal{B}$ is a vector bundle over M endowed with a connection A, then we can define $X^\#$, the horizontal lift of X to $\mathcal{B}$, a vectorfield on $\mathcal{B}$, by

$$X^\#(v) \;=\; X(p)^\#_v \quad \forall p\in M,\, \forall v\in\mathcal{B}_p$$

Let X, Y be two vectorfields on M. We compute $[X^\#, Y^\#]$.

Definition 7.6: We introduce the *curvature 2-form*

$$F^a_b \;=\; \mathrm{d}A^a_b + A^a_c\wedge A^c_b$$

where we use local coordinates $(x^1,...,x^m;v^1,...,v^n)$ in $\phi^{-1}(U)$, U the domain of a chart in M. □

Proposition 7.6: The commutator of $X^\#$ and $Y^\#$ satisfies the relation

$$[X^\#,Y^\#]-[X,Y]^\# \;=\; -F^a_b(X,Y)v^b\frac{\partial}{\partial v^a}$$

where again we use local coordinates $(x^1,...,x^m;v^1,...,v^n)$ in $\phi^{-1}(U)$. □

Proof: In local coordinates, we have

$$X^\# \;=\; X^\mu\left(\frac{\partial}{\partial x^\mu} - A^b_{\mu a}v^a\frac{\partial}{\partial v^b}\right) \qquad Y^\# \;=\; Y^\mu\left(\frac{\partial}{\partial x^\mu} - A^b_{\mu a}v^a\frac{\partial}{\partial v^b}\right)$$

Let f be a function on $\mathcal{B}$. It is represented by

$$(x^1,...,x^m;v^1,...,v^n) \;\mapsto\; f(x^1,...,x^m;v^1,...,v^n) \;\in\; \mathbb{R}$$

Then

$$X^\#(Y^\# f) \;=\; X^\# g$$

where

$$g \;=\; Y^\# f \;=\; Y^\nu\left(\frac{\partial f}{\partial x^\nu} - A^d_{\nu c}v^c\frac{\partial f}{\partial v^d}\right) \qquad X^\# g \;=\; X^\mu\left(\frac{\partial g}{\partial x^\mu} - A^b_{\mu a}v^a\frac{\partial g}{\partial v^b}\right)$$

Then we get that

$$\begin{aligned} X^{\#}(Y^{\#}f) &= X^{\mu}\Bigg[\frac{\partial Y^{\nu}}{\partial x^{\mu}}\left(\frac{\partial f}{\partial x^{\nu}} - A^{d}_{\nu c}v^{c}\frac{\partial f}{\partial v^{d}}\right) + Y^{\nu}\left(\frac{\partial^2 f}{\partial x^{\mu}\partial x^{\nu}} - \frac{\partial A^{d}_{\nu c}}{\partial x^{\mu}}v^{c}\frac{\partial f}{\partial v^{d}} - A^{d}_{\nu c}v^{c}\frac{\partial^2 f}{\partial x^{\mu}\partial v^{d}}\right) \\ &\quad - A^{b}_{\mu a}v^{a}Y^{\nu}\left(\frac{\partial^2 f}{\partial v^{b}\partial x^{\nu}} - A^{d}_{\nu c}\delta^{c}_{b}\frac{\partial f}{\partial v^{d}} - A^{d}_{\nu c}v^{c}\frac{\partial^2 f}{\partial v^{b}\partial v^{d}}\right)\Bigg] \end{aligned}$$

Note that

$$\frac{\partial^2 f}{\partial x^{\mu}\partial x^{\nu}}, \qquad -A^{d}_{\nu c}v^{c}\frac{\partial^2 f}{\partial x^{\mu}\partial v^{d}} - A^{b}_{\mu a}v^{a}\frac{\partial^2 f}{\partial v^{b}\partial x^{\nu}}, \qquad A^{b}_{\mu a}v^{a}A^{d}_{\nu c}v^{c}\frac{\partial^2 f}{\partial v^{b}\partial v^{d}}$$

are symmetric in μ, ν. Being multiplied by $X^{\mu}Y^{\nu}$ they are going to cancel when we interchange the roles of X and Y, and subtract. So we only have to consider

$$\begin{aligned} X^{\#}(Y^{\#}f) &\simeq X^{\mu}\left[\frac{\partial Y^{\nu}}{\partial x^{\mu}}\left(\frac{\partial f}{\partial x^{\nu}} - A^{d}_{\nu c}v^{c}\frac{\partial f}{\partial v^{d}}\right) - Y^{\nu}\frac{\partial A^{d}_{\nu c}}{\partial x^{\mu}}v^{c}\frac{\partial f}{\partial v^{d}} + Y^{\nu}A^{b}_{\mu a}v^{a}A^{d}_{\nu b}\frac{\partial f}{\partial v^{d}}\right] \\ &= (XY^{\nu})\left(\frac{\partial f}{\partial x^{\nu}} - A^{d}_{\nu c}v^{c}\frac{\partial f}{\partial v^{d}}\right) - X^{\mu}Y^{\nu}\frac{\partial A^{d}_{\nu c}}{\partial x^{\mu}}v^{c}\frac{\partial f}{\partial v^{d}} - X^{\mu}Y^{\nu}\left(-A^{b}_{\mu a}A^{d}_{\nu b}v^{a}\frac{\partial f}{\partial v^{d}}\right) \end{aligned}$$

We thus obtain

$$\begin{aligned} X^{\#}(Y^{\#}f) - Y^{\#}(X^{\#}f) &= [X,Y]^{\nu}\left(\frac{\partial f}{\partial x^{\nu}} - A^{d}_{\nu c}v^{c}\frac{\partial f}{\partial v^{d}}\right) \\ &\quad - X^{\mu}Y^{\nu}\left[\left(\frac{\partial A^{d}_{\nu c}}{\partial x^{\mu}} - \frac{\partial A^{d}_{\mu c}}{\partial x^{\nu}}\right)v^{c}\frac{\partial f}{\partial v^{d}} - \left(A^{b}_{\mu a}A^{d}_{\nu b} - A^{b}_{\nu a}A^{d}_{\mu b}\right)v^{a}\frac{\partial f}{\partial v^{d}}\right] \end{aligned}$$

where we used that

$$[X,Y]^{\nu} = X^{\mu}\frac{\partial Y^{\nu}}{\partial x^{\mu}} - Y^{\mu}\frac{\partial X^{\nu}}{\partial x^{\mu}}$$

Setting

$$F^{a}_{\mu\nu b} = \frac{\partial A^{a}_{\nu b}}{\partial x^{\mu}} - \frac{\partial A^{a}_{\mu b}}{\partial x^{\nu}} + A^{a}_{\mu c}A^{c}_{\nu b} - A^{a}_{\nu c}A^{c}_{\mu b}$$

or, in matrix notation

$$F_{\mu\nu} = \frac{\partial A_{\nu}}{\partial x^{\mu}} - \frac{\partial A_{\mu}}{\partial x^{\nu}} + A_{\mu}A_{\nu} - A_{\nu}A_{\mu}$$

the formula reduces to the form

$$[X^{\#},Y^{\#}]f - [X,Y]^{\#}f = -X^{\mu}Y^{\nu}F^{a}_{\mu\nu b}v^{b}\frac{\partial f}{\partial v^{a}}$$

Thus we can write

$$[X^{\#},Y^{\#}] - [X,Y]^{\#} = -F^{a}_{b}(X,Y)v^{b}\frac{\partial}{\partial v^{a}}$$

introducing the matrix-valued 2-form

$$F^{a}_{b} = \mathrm{d}A^{a}_{b} + A^{a}_{c}\wedge A^{c}_{b}$$

which we call curvature 2-form. (QED)

Remark 7.15: Note that in general

$$\begin{aligned} (a\wedge b)_{\mu\nu} &= (a\wedge b)\cdot\left(\frac{\partial}{\partial x^{\mu}},\frac{\partial}{\partial x^{\nu}}\right) \\ &= a\left(\frac{\partial}{\partial x^{\mu}}\right)b\left(\frac{\partial}{\partial x^{\nu}}\right) - b\left(\frac{\partial}{\partial x^{\mu}}\right)a\left(\frac{\partial}{\partial x^{\nu}}\right) = a_{\mu}b_{\nu} - b_{\mu}a_{\nu} \end{aligned}$$

$\mathrm{d}A^{a}_{b}$ denotes the exterior derivative, that is

$$\mathrm{d}A^{a}_{b}\cdot(X,Y) = X(A^{a}_{b}\cdot Y) - Y(A^{a}_{b}\cdot X) - A^{a}_{b}\cdot[X,Y]$$

and therefore

$$(\mathrm{d}A^{a}_{b})_{\mu\nu} = \mathrm{d}A^{a}_{b}\cdot\left(\frac{\partial}{\partial x^{\mu}},\frac{\partial}{\partial x^{\nu}}\right) = \frac{\partial A^{a}_{\nu b}}{\partial x^{\mu}} - \frac{\partial A^{a}_{\mu b}}{\partial x^{\nu}}$$

□

Remark 7.16: In terms of matrices

$$F = \frac{1}{2}\sum_{\mu,\nu} F_{\mu\nu}\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \qquad F_{\mu\nu} = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} + A_\mu A_\nu - A_\nu A_\mu$$

Under a change of basis sections, F transforms homogeneously

$$F \mapsto \tilde{F}_{\mu\nu} = S^{-1}F_{\mu\nu}S$$

(see Explanation A.14 on page 88) and $F_{\mu\nu}$ defines a linear transformation of $\mathcal{B}_p$

$$v^a \mapsto F^a_{\mu\nu b}v^b$$

namely the *curvature transformation*. We could also write

$$v \mapsto F_{\mu\nu} \cdot v$$

At $p \in M$, F itself is an antisymmetric bilinear form in T_pM $(F(X,Y))$ with values (not in $\mathbb{R}$ but) in $\mathcal{L}(\mathcal{B}_p, \mathcal{B}_p)$, the space of linear transformations of $\mathcal{B}_p$. In particular, it is

$$F_{\mu\nu} = F\left(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right)$$

□

Remark 7.17: We have

$$\left([X^\#, Y^\#] - [X,Y]^\#\right)(v) = -F(X,Y) \cdot v \in \mathcal{B}_p \quad \forall v \in \mathcal{B}_p, \forall p \in M$$

The right-hand side is considered to be a vertical vector at $v \in \mathcal{B}_p$. Now, suppose that $[X,Y] = 0$ and X, Y are complete. Suppose that X generates the group χ_t and Y generates the group ϕ_s. Then

$$\chi_t \circ \phi_s = \phi_s \circ \chi_t \quad \forall t, s$$

It follows that

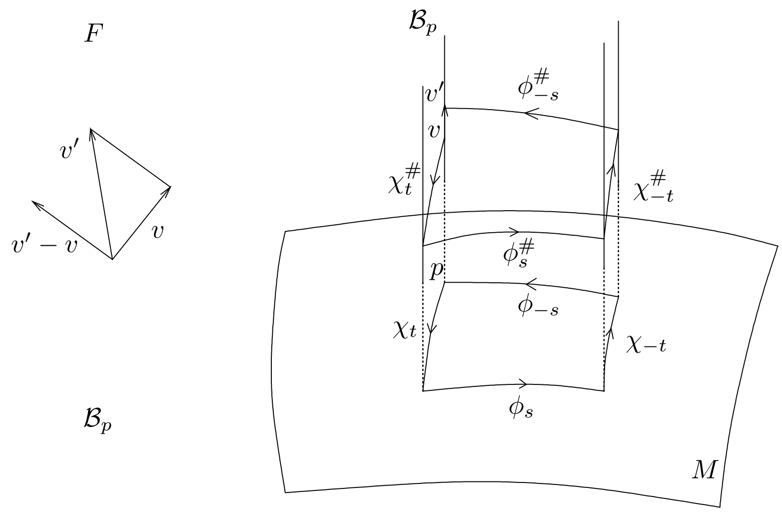
$$(\phi_{-s} \circ \chi_{-t} \circ \phi_s \circ \chi_t)(p) = p \quad \forall p \in M$$

Now $X^\#$ generates the group $\chi^\#_t$, $Y^\#$ generates the group $\phi^\#_s$, with the orbit $t \mapsto \chi^\#_t(v)$ being the horizontal lift of the orbit $t \mapsto \chi_t(p)$ and similarly for $\phi^\#_s$ and ϕ_s. Here, $p \in M$ and $v \in \mathcal{B}_p$. Then

$$\left(\phi^\#_{-s} \circ \chi^\#_{-t} \circ \phi^\#_s \circ \chi^\#_t\right)(v) = v' \in \mathcal{B}_p \quad v \in \mathcal{B}_p$$

and from Proposition 5.4 follows, that up to order $\mathcal{O}((s+t^3))$

$$v' - v = -ts[X^\#, Y^\#] = -ts\, F(X,Y)(p) \cdot v$$



□

Proposition 7.7: Suppose that X, Y are vectorfields on M and ϕ a section of $\mathcal{B}$. Then

$$D_X D_Y \phi - D_Y D_X \phi - D_{[X,Y]}\phi \quad = \quad F(X,Y)\phi$$

□

Proof: Consider the case $X = \frac{\partial}{\partial x^\mu}$, $Y = \frac{\partial}{\partial x^\nu}$. Then we have

$$D_{\partial/\partial x^\mu}\phi \quad = \quad (D_\mu \phi^a)\iota_a$$

and $D_X\phi$ as well as $D_Y\phi$ are represented by the columns

$$D_\mu\phi \quad = \quad \frac{\partial\phi}{\partial x^\mu} + A_\mu\phi \qquad D_Y\phi \quad = \quad \frac{\partial\phi}{\partial x^\nu} + A_\nu\phi$$

Hence $D_X D_Y\phi$ is represented by the column

$$\begin{aligned} D_\mu D_\nu\phi \quad &= \quad D_\mu\left(\frac{\partial\phi}{\partial x^\nu} + A_\nu\phi\right) \quad = \quad \frac{\partial}{\partial x^\mu}\left(\frac{\partial\phi}{\partial x^\nu} + A_\nu\phi\right) + A_\mu\left(\frac{\partial\phi}{\partial x^\nu} + A_\nu\phi\right) \\ &= \quad \frac{\partial^2\phi}{\partial x^\mu\partial x^\nu} + \left(\frac{\partial A_\nu}{\partial x^\mu} + A_\mu A_\nu\right)\phi + A_\nu\frac{\partial\phi}{\partial x^\mu} + A_\mu\frac{\partial\phi}{\partial x^\nu} \end{aligned}$$

where the first and the last two terms together are symmetric in μ, ν while $[X,Y] = [\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}] = 0$. Interchanging μ, ν and subtracting, we get

$$D_\mu D_\nu\phi - D_\nu D_\mu\phi \quad = \quad F_{\mu\nu}\phi$$

which shows the proposition (see Explanation A.15 on page 88). (QED)

Remark 7.18: One can consider a connection on a vector bundle $\mathcal{B}$ as an assignment

$$(X, \sigma) \quad \mapsto \quad D_X\sigma$$

to a vectorfield X on M and a differentiable section σ of $\mathcal{B}$ of a new section $D_X\sigma$ of $\mathcal{B}$, called the *covariant derivative* of σ with respect to X satisfying the following three conditions

(1) $D_X\sigma$ is linear in σ, in the sense that

$$D_X(\sigma_1 + \sigma_2) \quad = \quad D_X\sigma_1 + D_X\sigma_2$$

(2) $D_X\sigma$ is linear in X with respect to multiplication by the ring of functions, that is if f is a function on M, then

$$D_{fX}\sigma \quad = \quad fD_X\sigma$$

(3) $D_X\sigma$ satisfies the Leibnitz rule

$$D_X(f\sigma) \quad = \quad fD_X\sigma + (Xf)\sigma$$

if we multiply σ by a differentiable function f.

Conversely any assignment of this type defines a connection A on $\mathcal{B}$. Because taking local basis sections ι_a, $a = 1, ..., n$ over U, the domain of a chart in M, we define the *connection coefficients* $A^a_{\mu b}$ by

$$D_{\partial/\partial x^\mu}\iota_a \quad = \quad A^b_{\mu a}\iota_b$$

Then if $\sigma = \sigma^a\iota_a$, we know by (1) and (3)

$$\begin{aligned} D_{\partial/\partial x^\mu}\sigma \quad &= \quad \sum_a D_{\partial/\partial x^\mu}(\sigma^a\iota_a) \quad = \quad \sum_a\left(\sigma^a D_{\partial/\partial x^\mu}\iota_a + \left(\frac{\partial\sigma^a}{\partial x^\mu}\right)\iota_a\right) \\ &= \quad \sum_a\left(\sum_b \sigma^a A^b_{\mu a}\iota_b + \left(\frac{\partial\sigma^a}{\partial x^\mu}\right)\iota_a\right) \quad = \quad \sum_a\left(\frac{\partial\sigma^a}{\partial x^\mu} + \sum_b A^a_{\mu b}\sigma^b\right)\iota_a \\ &= \quad \sum_a (D_\mu\sigma^a)\iota_a \end{aligned}$$

and by (2) with $X = X^\mu\frac{\partial}{\partial x^\mu}$

$$D_X\sigma \quad = \quad \sum_\mu X^\mu D_{\partial/\partial x^\mu}\sigma \quad = \quad \sum_{\mu,a} X^\mu(D_\mu\sigma^a)\iota_a$$

□

7.5 Connections in Vector Bundles with Metric

Let $(\mathcal{B}, h)$ be a vector bundle with metric. We introduce the *compatibility condition* of a connection A on $\mathcal{B}$ with the metric h. This is the condition, that the magnitude of a vector is invariant under parallel transport. That is, if γ is a curve in M through p and $\gamma^{\#}$ is its horizontal lift to $\mathcal{B}$ through some $v \in \mathcal{B}_p$, then

$$\|\gamma^{\#}(t)\|^2 = h(\gamma^{\#}(t), \gamma^{\#}(t))$$

must be independent of t.

Proposition 7.8: The compatibility condition is equivalent to

$$\sum_{a,b=1}^{n} v^a v^b A^a_{\mu b}(p) = 0 \quad \forall p \in U$$

where we chose a set of orthonormal basis sections ι_a, $a = 1, ..., n$ over $U \subset M$ with $h(\iota_a, \iota_b) = \delta_{ab}$ and the corresponding local coordinates in $\pi^{-1}(U) \subset \mathcal{B}$. □

Proof: As before, with the corresponding local coordinates, we express

$$X^{\#}_v = X^\mu \left(\frac{\partial}{\partial x^\mu} - A^a_{\mu b} v^b \frac{\partial}{\partial v^a} \right)$$

So, $\gamma^{\#}$ is represented by $t \mapsto (x^1(t), ..., x^m(t), v^1(t), ..., v^n(t))$, where

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = \dot{\gamma}^\mu(t) \qquad \frac{\mathrm{d}v^a}{\mathrm{d}t} = -A^a_{\mu b}(\gamma(t))\dot{\gamma}^\mu(t) v^b$$

The $v^a(t)$ are the components of the parallel transported vector

$$v(t) = \gamma^{\#}(t) \qquad v(t) = v^a(t)\iota_a(\gamma(t))$$

Then

$$\begin{aligned}
\|v(t)\|^2 &= h(v(t), v(t)) = h(v^a(t)\iota_a(\gamma(t)), v^b(t)\iota_b(\gamma(t))) \\
&= v^a(t)v^b(t)h(\iota_a(\gamma(t)), \iota_b(\gamma(t))) \\
&= v^a(t)v^b(t)\delta_{ab} \\
&= \sum_a (v^a(t))^2
\end{aligned}$$

Thus the compatibility condition $\frac{\mathrm{d}}{\mathrm{d}t}\|v(t)\|^2 = 0$ reads

$$\sum_{a=1}^{n} v^a \frac{\mathrm{d}v^a}{\mathrm{d}t} = 0 \quad \Leftrightarrow \quad \sum_{a,b=1}^{n} v^a v^b A^a_{\mu b}(\gamma(t))\dot{\gamma}^\mu(t) = 0$$

This must hold for all possible parallel transports, hence for all possible tangent vectors $\dot{\gamma}$. It follows that the connection A must satisfy the condition

$$\sum_{a,b=1}^{n} v^a v^b A^a_{\mu b}(p) = 0 \quad \forall p \in U$$

(QED)

Proposition 7.9: Metric compatibility is equivalent to antisymmetry of the matrix A_μ. That is

$$A_\mu = -A_\mu$$

□

Proof: Decomposing the matrix A_μ into its symmetric and antisymmetric part, under transposition, that is

$$A_\mu = \frac{1}{2}\left(A_\mu + A_\mu^T\right) + \frac{1}{2}\left(A_\mu - A_\mu^T\right)$$

we obtain that the quadratic form

$$Q_\mu = \frac{1}{2}\left(A_\mu + A_\mu^T\right) \qquad Q_\mu(v,v) = \sum_{a,b=1} Q^a_{\mu b} v^a v^b$$

must satisfies

$$Q_\mu(v,v) = 0 \quad \forall v$$

which follows with Proposition 7.8. By polarization, we then get

$$Q_\mu(u,v) = 0 \qquad \forall u,v$$

that is

$$Q_\mu = 0 \quad \Rightarrow \quad A_\mu + A_\mu^T = 0$$

We conclude that metric compatibility is equivalent to the antisymmetry of the matrix A_μ. (QED)

Remark 7.19: A_μ is antisymmetric, that is $A_\mu^T = -A_\mu$ or $A^b_{\mu a} = -A^a_{\mu b}$. It follows that also $F_{\mu\nu}$ is antisymmetric, so $F^T_{\mu\nu} = -F_{\mu\nu}$, because

$$(A_\mu A_\nu)^T = A_\nu^T A_\mu^T = A_\nu A_\mu$$

□

Remark 7.20: Since we chose a set of orthonormal basis sections for our derivations in chapter, our Gauge Transformations should only change from one orthonormal basis to another orthonormal basis, meaning that the transformation matrix S must be orthogonal. So we write O instead of S with

$$\tilde{\iota}_a = \sum_{b=1}^n \iota_b O^b_a \quad O(p) \in O(n), \forall p \in U$$

In terms of columns and matrices, we have

$$\tilde{\phi} = O^{-1}\phi \qquad \widetilde{D_\mu \phi} = O^{-1} D_\mu \phi \qquad \tilde{A}_\mu = O^{-1} A_\mu O + O^{-1}\frac{\partial O}{\partial x^\mu} \qquad \tilde{F}_{\mu\nu} = O^{-1} F_{\mu\nu} O$$

Note that the antisymmetric matrices, like A_μ and $F_{\mu\nu}$, constitute the Lie Algebra of $O(n)$. Futhermore, if A is an antisymmetric matrix and O an orthogonal matrix, then

$$A \mapsto O^{-1} A O$$

is the *adjoint action* of the Lie group $O(n)$ where

$$(O^{-1}AO)^T = O^T A^T (O^{-1})^T = -O^{-1}AO \qquad O^{-1} = O^T$$

For an arbitary Lie group G, we write

$$\mathrm{Ad}_a : G \to G \;\; : \;\; (a,g) \mapsto a^{-1} g a$$

where $\mathrm{Ad}_a(e) = e$. The adjoint action then is $\mathrm{d}\,\mathrm{Ad}_a(e) : T_e G \to T_e G$ and we identify the Lie algebra $\mathcal{Y}$ with $T_e G$. □

Proposition 7.10: In the case of a complex vector bundle $\mathcal{B}$ with a Hermitian metric h, the metric compatibility condition is equivalent to

$$\overline{A}^b_{\mu a} + A^a_{\mu b} = 0 \quad \Rightarrow \quad A^\dagger_\mu = -A_\mu$$

□

Remark 7.21: Here, if M is any complex matrix, $M^\dagger$, the adjoint is $M^\dagger = \overline{M}^T$. So A_μ and also $F_{\mu\nu}$ are anti-Hermitian matrices. The gauge transformations are now restricted to those with values in the unitary group $U(n)$

$$\tilde{\iota}_a \;=\; \sum_{b=1}^{n} \iota_b U_a^b \qquad U(p) \in U(n),\, \forall p \in U$$

and we have, in terms of columns and matrices.

$$\tilde{\phi} \;=\; U^{-1}\phi \qquad \widetilde{D_\mu \phi} \;=\; U^{-1} D_\mu \phi \qquad \tilde{A}_\mu \;=\; U^{-1} A_\mu U + U^{-1}\frac{\partial U}{\partial x^\mu} \qquad \tilde{F}_{\mu\nu} \;=\; U^{-1} F_{\mu\nu} U$$

□

Remark 7.22: We compare our notation with the one, which is more common in Physics

Our Notation	*Physics Notation*
A_μ and $F_{\mu\nu}$ (anti-Hermitian matrices)	iA_μ and $iF_{\mu\nu}$ (Hermitian matrices)
$D_\mu\phi = \partial_\mu\phi + A_\mu\phi$	$D_\mu\phi = \partial_\mu\phi + iA_\mu\phi$
$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + A_\mu A_\nu - A_\nu A_\mu$	$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + i(A_\mu A_\nu - A_\nu A_\mu)$

□

Proof of Proposition 7.10: The metric compatibility reads

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{a=1}^{n} |v^a|^2 \;=\; 0 \quad \text{or} \quad \sum_{a=1}^{n}\left(\overline{v}^a \frac{\partial v^a}{\partial t} + v^a \frac{\mathrm{d}\overline{v}^a}{\mathrm{d}t}\right) \;=\; 0$$

which is equivalent to

$$\sum_{a,b=1}^{n}\left(\overline{v}^a v^b A^a_{\mu b} + v^a \overline{v}^b \overline{A}^a_{\mu b}\right) \;=\; 0$$

Again, we write

$$A_\mu \;=\; H_\mu + G_\mu \;=\; \frac{1}{2}\left(A_\mu + A_\mu^\dagger\right) + \frac{1}{2}\left(A_\mu - A_\mu^\dagger\right)$$

with the Hermitian part H_μ and the anti-Hermitian part G_μ. The contribution of the anti-Hermitian part is

$$\sum_{a,b=1}^{n}\left(\overline{v}^a v^b G^a_{\mu b} + v^a \overline{v}^b \overline{G}^a_{\mu b}\right) \;=\; \frac{1}{2}\sum_{a,b=1}^{n}\left(\overline{v}^a v^b \left(A^a_{\mu b} - \overline{A}^b_{\mu a}\right) + v^a \overline{v}^b \left(\overline{A}^a_{\mu b} - A^b_{\mu a}\right)\right) \;=\; 0$$

Our condition thus reduces to

$$\sum_{a,b=1}^{n}\left(\overline{v}^a v^b H^a_{\mu b} + v^a \overline{v}^b \overline{H}^a_{\mu b}\right) \;=\; 2\sum_{a,b=1}^{n} \overline{v}^a v^b H^a_{\mu b} \;=\; 0$$

Thus, we can write

$$\langle v, H_\mu v\rangle \;=\; 0 \quad \forall v \in \mathbb{C}^n$$

where $\langle\cdot,\cdot\rangle$ is the standard Hermitian product on $\mathbb{C}^n$. It remains to be shown that this implies $H_\mu = 0$. We show this by *Polarization*. We have

$$0 \;=\; \langle v+u, H_\mu(v+u)\rangle - \langle v-u, H_\mu(v-u)\rangle \;=\; 2\left(\langle u, H_\mu v\rangle + \langle v, H_\mu u\rangle\right) \quad \forall v, u$$

and it follows that

$$\langle u, H_\mu v\rangle \;=\; -\langle v, H_\mu u\rangle$$

Now replace u by iu to get

$$\langle iu, H_\mu v\rangle + \langle v, H_\mu iu\rangle \;=\; -i\langle u, H_\mu v\rangle + i\langle v, H_\mu u\rangle$$

So we obtain

$$\langle u, H_\mu v\rangle \;=\; \langle v, H_\mu u\rangle \;=\; -\langle v, H_\mu u\rangle$$

It follows that

$$\langle v, H_\mu u\rangle \;=\; 0 \quad \forall v, u$$

Therefore, $H_\mu = 0$. (QED)

Proposition 7.11: For a general vector bundle $\mathcal{B}$ with metric h, the condition of metric compatibility of a connection A on $\mathcal{B}$ may be expressed in the form

$$X \cdot h(\sigma, \tau) \;=\; h(D_X\sigma, \tau) + h(\sigma, D_X\tau)$$

for every pair σ, τ of sections of $\mathcal{B}$ and every vectorfield X on M. □

Proof: It is enough to check the case $X = \frac{\partial}{\partial x^\mu}$, $\sigma = \iota_a$, $\tau = \iota_b$ (see Explanation A.16 on page 89). Then for a complex vector bundle $\mathcal{B}$, $h(\sigma, \tau) = h(\iota_a, \iota_b) = \delta_{ab}$, so

$$X \cdot h(\sigma, \tau) \;=\; \frac{\partial \delta_{ab}}{\partial x^\mu} \;=\; 0$$

while

$$D_X\sigma \;=\; D_{\partial/\partial x^\mu}\iota_a \;=\; A^c_{\mu a}\iota_c \qquad D_X\tau \;=\; D_{\partial/\partial x^\mu}\iota_b \;=\; A^c_{\mu b}\iota_c$$

so

$$h(D_X\sigma, \tau) \;=\; h(A^c_{\mu a}\iota_c, \iota_b) \;=\; \overline{A}^c_{\mu a}\delta_{cb} \;=\; \overline{A}^b_{\mu a}$$

and

$$D(\sigma, D_X\tau) \;=\; h(\iota_a, A^c_{\mu b}\iota_c) \;=\; A^c_{\mu b}\delta_{ac} \;=\; A^a_{\mu b}$$

and the condition reduces to

$$\overline{A}^b_{\mu a} + A^a_{\mu b} \;=\; 0$$

(QED)

7.6 Connections on the Tangent Bundle

We now restrict ourselves to the case $\mathcal{B} = TM$. Then a section σ is a vectorfield on M. So in $D_X\sigma$, X and σ are of the same kind. We are going to require in this case a certain symmetry condition. In the case $\mathcal{B} = TM$, we denote A by Γ, F by R (Riemann) and D by ∇. The *Symmetry condition* then is

$$\nabla_X Y - \nabla_Y X \;=\; [X, Y]$$

It is due to this condition, that the General Relativity differs from the other Gauche theories, it's not due to the metric. The above is independent of whether or not $\mathcal{B}$ is endowed with a metric.

Suppose that U is the domain of a chart of M. Then in U we have the coordinate vectorfields $\frac{\partial}{\partial x^\mu}$, $\mu = 1, ..., m$, forming a set of basis sections.

Proposition 7.12: In the coordinate basis, the symmetry condition without metric reduces to

$$\Gamma^\lambda_{\mu\nu} - \Gamma^\lambda_{\nu\mu} \;=\; 0$$

□

Proof: The connection coefficients are defined by

$$\nabla_{\partial/\partial x^\mu}\frac{\partial}{\partial x^\nu} \;=\; \Gamma^\lambda_{\mu\nu}\frac{\partial}{\partial x^\lambda}$$

Γ_μ is a metric with entries $\Gamma^\lambda_{\mu\nu}$. The curvature is then given by

$$T_{\mu\nu} \;=\; R\left(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right) \;=\; \frac{\partial \Gamma_\nu}{\partial x^\mu} - \frac{\partial \Gamma_\mu}{\partial x^\nu} + \Gamma_\mu\Gamma_\nu - \Gamma_\nu\Gamma_\mu$$

We write

$$T^\kappa_{\lambda\mu\nu} \;=\; (R_{\mu\nu})^\kappa_\lambda \;=\; \frac{\partial \Gamma^\kappa_{\nu\lambda}}{\partial x^\mu} - \frac{\partial \Gamma^\kappa_{\mu\lambda}}{\partial x^\nu} + \Gamma^\kappa_{\mu\alpha}\Gamma^\alpha_{\nu\lambda} - \Gamma^\kappa_{\nu\alpha}\Gamma^\alpha_{\mu\lambda}$$

In terms of the connection 1-form, we have

$$\Gamma^\kappa_\lambda \;=\; \Gamma^\kappa_{\mu\lambda}\,\mathrm{d}x^\mu$$

and the curvature 2-form R^κ_λ is given by

$$R^\kappa_\lambda \;=\; \mathrm{d}\Gamma^\kappa_\lambda + \Gamma^\kappa_\alpha \wedge \Gamma^\alpha_\lambda$$

In the coordinate basis we have $\left[\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right] = 0$, so the symmetry condition reduces to

$$\Gamma^\lambda_{\mu\nu} - \Gamma^\lambda_{\nu\mu} \;=\; 0$$

(QED)

Remark 7.23: With metric, that is in the case of a Riemannian manifold (M, g), the connection coefficients in an orthonormal frame field E_μ, $\mu = 1, ..., m$ with $g(E_\mu, E_\nu) = \delta_{\mu\nu}$, are defined by

$$\nabla_{E_\mu} E_\alpha \;=\; \Gamma^\beta_{\mu\alpha} E_\beta$$

The connection 1-form is then

$$\Gamma^\alpha_\beta \;=\; \Gamma^\alpha_{\mu\beta}\theta^\mu$$

where θ^μ, $\mu = 1, ..., m$ is the dual basis of 1-forms defined in U. As we have seen for vector bundles with metric, the metric compatibility is equivalent to the condition that Γ_μ is an antisymmetric matrix (and therefore an element of the Lie Algebra of $\mathrm{O}(m)$). So we have

$$(\Gamma_\mu)^T \;=\; -\Gamma_\mu \quad\Leftrightarrow\quad \Gamma^\alpha_{\mu\beta} \;=\; -\Gamma^\beta_{\mu\alpha}$$

The *structure funcions* $\Lambda^\mu_{\alpha\beta}$ of the frame field $(E_1, ..., E_m)$ are defined by

$$[E_\alpha, E_\beta] \;=\; \Lambda^\mu_{\alpha\beta} E_\mu$$

Consider now the dual basis 1-forms $\theta^1, ..., \theta^m$. We have $\theta^\mu \cdot E_\nu = \delta^\mu_\nu$, so

$$\begin{aligned}(\mathrm{d}\theta^\mu)(E_\alpha, E_\beta) &= E_\alpha(\theta^\mu \cdot E_\beta) - E_\beta(\theta^\mu \cdot E_\alpha) - \theta^\mu \cdot [E_\alpha, E_\beta] \\ &= -\theta^\mu \cdot (\Lambda^\nu_{\alpha\beta} E_\nu) \;=\; -\Lambda^\mu_{\alpha\beta}\end{aligned}$$

Hence

$$\mathrm{d}\theta^\mu \;=\; -\frac{1}{2}\Lambda^\mu_{\alpha\beta}\theta^\alpha \wedge \theta^\beta$$

because $\Lambda^\mu_{\beta\alpha} = -\Lambda^\mu_{\alpha\beta}$. □

Proposition 7.13: There is a unique connection Γ on M which is symmetric and compatible with the metric g. It is called the *Levi-Civita connection.* □

Proof: Using an orthonormal frame field
The symmetry condition

$$\nabla_{E_\alpha} E_\beta - \nabla_{E_\beta} E_\alpha \;=\; [E_\alpha, E_\beta]$$

reads

$$\Gamma^\mu_{\alpha\beta} - \Gamma^\mu_{\beta\alpha} \;=\; \Lambda^\mu_{\alpha\beta}$$

and metric compatibility requires

$$\Gamma^\mu_{\alpha\beta} \;=\; -\Gamma^\beta_{\alpha\mu}$$

Thus we get

$$\begin{aligned}\Gamma^\mu_{\alpha\beta} &= -\Gamma^\beta_{\alpha\mu} \;=\; -\Gamma^\beta_{\mu\alpha} - \Lambda^\beta_{\alpha\mu} \;=\; \Gamma^\alpha_{\mu\beta} - \Lambda^\beta_{\alpha\mu} \;=\; \Gamma^\alpha_{\beta\mu} + \Lambda^\alpha_{\mu\beta} - \Lambda^\beta_{\alpha\mu} \\ &= -\Gamma^\mu_{\beta\alpha} + \Lambda^\alpha_{\mu\beta} - \Lambda^\beta_{\alpha\mu} \;=\; -\Gamma^\mu_{\alpha\beta} - \Lambda^\mu_{\beta\alpha} + \Lambda^\alpha_{\mu\beta} - \Lambda^\beta_{\alpha\mu}\end{aligned}$$

Hence

$$\Gamma^\mu_{\alpha\beta} \;=\; \frac{1}{2}\left(-\Lambda^\mu_{\beta\alpha} + \Lambda^\alpha_{\mu\beta} - \Lambda^\beta_{\alpha\mu}\right)$$

or simply

$$\Gamma^\mu_{\alpha\beta} \;=\; \frac{1}{2}\left(\Lambda^\mu_{\alpha\beta} + \Lambda^\alpha_{\mu\beta} + \Lambda^\beta_{\mu\alpha}\right)$$

(QED)

Proof: Using a coordinate frame field
As shown in Proposition 7.12, the symmetry condition is $\Gamma^\mu_{\alpha\beta} = \Gamma^\mu_{\beta\alpha}$ and the metric compatibility with

$$g_{\alpha\beta} = g\left(\frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial x^\beta}\right)$$

reads with Proposition 7.11

$$\begin{aligned} \frac{\partial g_{\alpha\beta}}{\partial x^\mu} &= g\left(\nabla_{\frac{\partial}{\partial x^\mu}} \frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial x^\beta}\right) + g\left(\frac{\partial}{\partial x^\alpha}, \nabla_{\frac{\partial}{\partial x^\mu}} \frac{\partial}{\partial x^\beta}\right) \\ &= \Gamma^\nu_{\mu\alpha} g_{\nu\beta} + \Gamma^\nu_{\mu\beta} g_{\alpha\nu} \end{aligned} \tag{7.8}$$

We interchange μ and α to write

$$\frac{\partial g_{\mu\beta}}{\partial x^\alpha} = \Gamma^\nu_{\alpha\mu} g_{\nu\beta} + \Gamma^\nu_{\alpha\beta} g_{\nu\mu} \tag{7.9}$$

and then interchange α and β to get

$$\frac{\partial g_{\mu\alpha}}{\partial x^\beta} = \Gamma^\nu_{\beta\mu} g_{\nu\alpha} + \Gamma^\nu_{\beta\alpha} g_{\nu\mu} \tag{7.10}$$

Equation (7.9) plus equation (7.10) minus equation (7.8) gives, in view of the symmetry condition

$$2\Gamma^\nu_{\alpha\beta} g_{\nu\mu} = \frac{\partial g_{\mu\beta}}{\partial x^\alpha} + \frac{\partial g_{\mu\alpha}}{\partial x^\beta} - \frac{\partial g_{\alpha\beta}}{\partial x^\mu}$$

where we used that

$$\begin{aligned} &\Gamma^\nu_{\alpha\mu} g_{\nu\beta} + \Gamma^\nu_{\alpha\beta} g_{\nu\mu} + \Gamma^\nu_{\beta\mu} g_{\nu\alpha} + \Gamma^\nu_{\beta\alpha} g_{\nu\mu} - \Gamma^\nu_{\mu\alpha} g_{\nu\beta} - \Gamma^\nu_{\mu\beta} g_{\alpha\nu} \\ &= (\Gamma^\nu_{\alpha\mu} g_{\nu\beta} - \Gamma^\nu_{\mu\alpha} g_{\nu\beta}) + (\Gamma^\nu_{\alpha\beta} g_{\nu\mu} + \Gamma^\nu_{\beta\alpha} g_{\nu\mu}) + (\Gamma^\nu_{\beta\mu} g_{\nu\alpha} - \Gamma^\nu_{\mu\beta} g_{\alpha\nu}) \\ &= 2\Gamma^\nu_{\alpha\beta} g_{\nu\mu} \end{aligned}$$

We thus obtain

$$\Gamma^\mu_{\alpha\beta} = \frac{1}{2}(g^{-1})^{\mu\nu}\left(\frac{\partial g_{\nu\beta}}{\partial x^\alpha} + \frac{\partial g_{\nu\alpha}}{\partial x^\beta} - \frac{\partial g_{\alpha\beta}}{\partial x^\nu}\right)$$

(QED)

Definition 7.7: A *geodesic* is a curve whose tangent vector is parallel transported along the curve, that is

$$\nabla_{\dot\gamma}\dot\gamma = 0$$

□

Remark 7.24: The definition does not require a metric on M. □

Remark 7.25: Denoting $v := \dot\gamma$ the condition $\nabla_v v = 0$ reads, in local coordinates

$$\frac{\mathrm{d}v^\mu}{\mathrm{d}t} + \Gamma^\mu_{\alpha\beta} v^\alpha v^\beta = 0$$

Here γ is represented by $x^\mu = x^\mu(t)$ and v is represented by $v^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}t}(t)$. The argument of $\Gamma^\mu_{\alpha\beta}$ is $x(t)$. In the case that the connection Γ comes from a metric g, a geodesic has constant speed

$$\frac{\mathrm{d}}{\mathrm{d}t}\|v\| = 0$$

Since the arc length is given by

$$s = \int \|v\|\,\mathrm{d}t$$

for a geodesic $\frac{\mathrm{d}s}{\mathrm{d}t}$ is constant, therefore the parameter t is proportional to the arc length. For the components of the curvature 2-form in a possibly orthonormal basis E_μ, $\mu = 1, ..., m$, we have

$$R^\alpha_{\beta\mu\nu} = R(E_\mu, E_\nu)^\alpha_\beta = (\mathrm{d}\Gamma)(E_\mu, E_\nu)^\alpha_\beta + \Gamma(E_\mu)^\alpha_\gamma \Gamma(E_\nu)^\gamma_\beta - \Gamma(E_\nu)^\alpha_\gamma \Gamma(E_\mu)^\gamma_\beta$$

We write $\Gamma(E_\mu)^\alpha_\beta = \Gamma^\alpha_{\mu\beta}$ and put in

$$\begin{aligned}(\mathrm{d}\Gamma)(E_\mu, E_\nu)^\alpha_\beta &= E_\mu(\Gamma(E_\nu)^\alpha_\beta) - E_\nu(\Gamma(E_\mu)^\alpha_\beta) - \Gamma([E_\mu, E_\nu])^\alpha_\beta \\ &= E_\mu(\Gamma^\alpha_{\nu\beta}) - E_\nu(\Gamma^\alpha_{\mu\beta}) - \Lambda^\gamma_{\mu\nu}\Gamma^\alpha_{\gamma\beta}\end{aligned}$$

where $[E_\mu, E_\nu] = \Lambda^\gamma_{\mu\nu}E_\gamma$. We thus obtain

$$R^\alpha_{\beta\mu\nu} = E_\mu(\Gamma^\alpha_{\nu\beta}) - E_\nu(\Gamma^\alpha_{\mu\beta}) - \Lambda^\gamma_{\mu\nu}\Gamma^\alpha_{\gamma\beta} + \Gamma^\alpha_{\mu\gamma}\Gamma^\gamma_{\nu\beta} - \Gamma^\alpha_{\nu\gamma}\Gamma^\gamma_{\mu\beta}$$

□

Theorem 7.1: Let $\mathcal{B}$ be a vector bundle (without metric). The vanishing of the curvature R is necessary and sufficient for the existence locally of basis sections $(\iota_1, ..., \iota_n)$ relative to which the connection coefficients

$$\Gamma^a_{\mu b} = 0$$

□

Proof: Recall the change of basis formula

$$\tilde{\iota}_a = \sum_{b=1}^{n} \iota_b S^b_a$$

where S is a non-singular matrix at each point. We have

$$S\tilde{\Gamma}_\mu = \Gamma_\mu S + \frac{\partial S}{\partial x^\mu}$$

Thus, to make $\tilde{\Gamma}_\mu = 0$ we must solve

$$\frac{\partial S}{\partial x^\mu} = -\Gamma_\mu S \quad \leftrightarrow \quad \frac{\partial S^b_a}{\partial x^\mu} = -\Gamma^b_{\mu c}S^c_a$$

So we look for an S with $\mathrm{d}S = -\Gamma S$. This locally has a solution S, if and only if

$$\frac{\partial^2 S}{\partial x^\mu \partial x^\nu}$$

is symmetric in μ, ν or $\mathrm{d}^2 S = 0$. We have

$$\frac{\partial^2 S}{\partial x^\mu \partial x^\nu} = -\frac{\partial \Gamma_\mu}{\partial x^\nu} S - \Gamma_\mu \frac{\partial S}{\partial x^\nu} = \left(-\frac{\partial \Gamma_\mu}{\partial x^\nu} + \Gamma_\mu \Gamma_\nu\right) S$$

so interchanging μ, ν and subtracting yields

$$0 = R_{\mu\nu} S$$

(QED)

Theorem 7.2: Let $\mathcal{B}$ be a vector bundle with metric h. The vanishing of the curvature R is necessary and sufficient for the existence locally of orthonormal basis sections $(\iota_1, ..., \iota_n)$ relative to which the connection coefficients vanish. □

Proof: Here

$$\tilde{\iota}_a = \sum_{b=1}^{n} \iota_b O^b_a$$

where O is a $O(m)$ matrix. Let be $R_{\mu\nu} = 0$. Then the integrability condition of the previous theorem holds. So we obtain a non-singular matrix S satisfying

$$\mathrm{d}S = -\Gamma S$$

Now we want to show that S can be taken to be an $O(n)$ matrix. We have

$$\mathrm{d}S^T = -S^T\Gamma^T = S^T\Gamma$$

where we used that $\Gamma^T = -\Gamma$, and

$$\mathrm{d}(SS^T) \;=\; (\mathrm{d}S)S^T + S(\mathrm{d}S^T) \;=\; -\Gamma SS^T + SS^T\Gamma$$

or, setting $M = SS^T$

$$\mathrm{d}M \;=\; -\Gamma M + M\Gamma$$

This equation for M has a unique solution for a given value of M at a fixed point p_0. Since $M = \mathbf{1}$ is a solution, it is the only solution such that $M(p_0) = \mathbf{1}$. It follows that $SS^T = \mathbf{1}$, hence S can be chosen to be an $O(n)$-matrix. (QED)

Theorem 7.3: Riemann
Let (M, g) be a Riemannian manifold. The vanishing of the curvature R is necessary and sufficient for the existence of local coordinates $(x^1, ..., x^m)$ in which the metric takes the Euclidean form

$$\sum_{\mu=1}^{n} \mathrm{d}x^\mu \otimes \mathrm{d}x^\mu \qquad g_{\mu\nu} \;=\; \delta_{\mu\nu}$$

□

Proof: We apply Theorems 7.2 to $\mathcal{B} = TM$, $h = g$. We then have that $R = 0$ is necessary and sufficient for the existence locally of an orthonormal frame field $(E_1, ..., E_m)$ relative to which the connection coefficients $\Gamma^\mu_{\alpha\beta} = 0$. Consider the dual basis 1-forms $(\theta^1, ..., \theta^m)$. Recall the general formula

$$\mathrm{d}\theta^\mu \;=\; -\frac{1}{2}\Lambda^\mu_{\alpha\beta}\theta^\alpha \wedge \theta^\beta$$

where the $\Lambda^\mu_{\alpha\beta}$ are the structure functions of $(E_1, ..., E_m)$

$$\Lambda^\mu_{\alpha\beta} \;=\; \Gamma^\mu_{\alpha\beta} - \Gamma^\mu_{\beta\alpha}$$

$\Gamma^\mu_{\alpha\beta} = 0$ implies $\Gamma^\mu_{\beta\alpha} = 0$ and conversely. Thus we have

$$\mathrm{d}\theta^\mu \;=\; 0$$

This implies the existence, locally, of functions x^μ such that $\theta^\mu = \mathrm{d}x^\mu$. It follows that

$$(E_1, ..., E_m) \;=\; \left(\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^m}\right)$$

hence

$$\delta_{\mu\nu} \;=\; g(E_\mu, E_\nu) \;=\; g\left(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right)$$

and

$$g \;=\; \sum_{\mu=1}^{m} \mathrm{d}x^\mu \otimes \mathrm{d}x^\mu$$

(QED)

7.7 The Bianchi Identities

We consider again the case of a general vector bundle. We want to show that the covariant exterior derivative of F satisfies

$$DF \;=\; 0$$

But first, we have to define DF. To do so, we consider $F(X, Y)$ for given vectorfields X, Y on M. At each $p \in M$, $F(X, Y)(p)$ is a linear map of $\mathcal{B}_p$ into itself. Thus $F(X, Y)$ is a section of the tensor bundle

$$\mathrm{End}(\mathcal{B}) \;=\; \bigcup_{p \in M} \mathcal{L}(\mathcal{B}_p, \mathcal{B}_p)$$

where for any vectorspaces U, V, $\mathcal{L}(U, V)$ denotes the space of linear maps $U \to V$. Let now T be any differentiable section of $\mathrm{End}(\mathcal{B})$. Given a vectorfield X on M, we define $D_X T$ by the Leibniz rule

$$D_X(T \cdot \sigma) \;=\; (D_X T)\sigma + T \cdot (D_X \sigma)$$

for any differentiable section σ of $\mathcal{B}$. $T \cdot \sigma$ will then also be a differentiable section of $\mathcal{B}$. So everything except the first term on the right is known. The Leibniz rule then defines $D_X T$ as follows.

Definition 7.8: For a section $T \in \mathrm{End}(\mathcal{B})$, the *covariant derivative* is defined by

$$(D_X T) \cdot \sigma \;=\; D_X(T \cdot \sigma) - T \cdot (D_X \sigma)$$

for any section σ of $\mathcal{B}$. □

Definition 7.9: We now define the *covariant exterior derivative* DF for the curvature F at each $p \in M$, a totally antisymmetric trilinear form in T_pM into values in $\mathcal{L}(\mathcal{B}_p, \mathcal{B}_p)$, by

$$\begin{aligned} &DF \cdot (X,Y,Z) \\ &\quad := \; D_Z(F(X,Y)) + D_X(F(Y,Z)) + D_Y(F(Z,X)) - F([X,Y],Z) - F([Y,Z],X) - F([Z,X],Y) \end{aligned}$$

□

Remark 7.26: Note that if ω is an ordinary 2-form on M, then $\mathrm{d}\omega$ is an ordinary 3-form on M defined by

$$\begin{aligned} &\mathrm{d}\omega(X,Y,Z) \\ &\quad := \; Z(\omega(X,Y)) + X(\omega(Y,Z)) + Y(\omega(Z,X)) - \omega([X,Y],Z) - \omega([Y,Z],X) - \omega([Z,X],Y) \end{aligned}$$

for every triplet X, Y, Z of vectorfields on M. $\mathrm{d}\omega$ is trilinear in X, Y, Z with respect to multiplication by the ring of $C^\infty(M)$-functions. Furthermore, $\mathrm{d}\omega = 0$ if $\omega = \mathrm{d}\theta$ for some 1-form θ. See Explanation A.17 on page 89. □

Proposition 7.14: The covariant exterior derivative of the curvature F satisfies

$$DF \;=\; 0$$

□

Proof: Take three vectorfields X, Y, Z on M. Then, by definition

$$\big(D_Z(F(X,Y))\big) \cdot \sigma \;=\; D_Z(F(X,Y) \cdot \sigma) - F(X,Y) \cdot D_Z \sigma$$

where

$$F(X,Y) \cdot \sigma \;=\; D_X D_Y \sigma - D_Y D_X \sigma - D_{[X,Y]}\sigma \;=\; [D_X, D_Y]\sigma - D_{[X,Y]}\sigma$$

which leads to

$$D_Z(F(X,Y) \cdot \sigma) \;=\; D_Z[D_X, D_Y]\sigma - D_Z D_{[X,Y]}\sigma$$

Using this, we get

$$\begin{aligned} &D_Z(F(X,Y) \cdot \sigma) + D_X(F(Y,Z) \cdot \sigma) + D_Y(F(Z,X) \cdot \sigma) \\ &\quad = \; D_Z[D_X, D_Y]\sigma + D_X[D_Y, D_Z]\sigma + D_Y[D_Z, D_X]\sigma - D_Z D_{[X,Y]}\sigma - D_X D_{[Y,Z]}\sigma - D_Y D_{[Z,X]}\sigma \end{aligned}$$

The first three terms on the right-hand side are

$$\begin{aligned} &(D_Z D_X D_Y - D_Z D_Y D_X + D_X D_Y D_Z - D_X D_Z D_Y + D_Y D_Z D_X - D_Y D_X D_Z)\sigma \\ &\quad = \; [D_Y, D_Z] D_X \sigma + [D_Z, D_X] D_Y \sigma + [D_X, D_Y] D_Z \sigma \\ &\quad = \; F(Y,Z) \cdot D_X\sigma + F(Z,X) \cdot D_Y\sigma + F(X,Y) \cdot D_Z\sigma + D_{[Y,Z]} D_X \sigma + D_{[Z,X]} D_Y \sigma + D_{[X,Y]} D_Z \sigma \end{aligned}$$

Substituting in the previous equatioin, we obtain

$$\begin{aligned} &D_Z(F(X,Y) \cdot \sigma) + D_X(F(Y,Z) \cdot \sigma) + D_Y(F(Z,X) \cdot \sigma) \\ &\quad = \; F(X,Y) \cdot D_Z\sigma + F(Y,Z) \cdot D_X\sigma + F(Z,X) \cdot D_Y\sigma \\ &\qquad + [D_{[X,Y]}, D_Z]\sigma + [D_{[Y,Z]}, D_X]\sigma + [D_{[Z,X]}, D_Y]\sigma \end{aligned}$$

The first three terms in the right-hand side are absorbed into the left-hand side and we obtain

$$\begin{aligned} &\big(D_Z(F(X,Y)) + D_X(F(Y,Z)) + D_Y(F(Z,X))\big) \cdot \sigma \\ &\quad = \; F([X,Y],Z)\sigma + F([Y,Z],X) \cdot \sigma + F([Z,X],Y) \cdot \sigma + D_{[[X,Y],Z]}\sigma + D_{[[Y,Z],X]}\sigma + D_{[[Z,X],Y]}\sigma \end{aligned}$$

where

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

which is the Jacobi identity. So, because of linearity, the sum of the last three terms vanishes. This already proves the proposition. (QED)

Remark 7.27: We obtained $DF = 0$. In the case of a connection in TM, we can define

$$(\nabla_X R)(Y, Z) \;=\; \nabla_X(R(Y, Z)) - R(\nabla_X Y, Z) - R(Y, \nabla_X Z)$$

which is the Leibniz rule. Now, by the symmetry condition

$$\begin{aligned} & R(\nabla_X Y, Z) + R(Y, \nabla_X Z) + R(\nabla_Y Z, X) + R(Z, \nabla_Y X) + R(\nabla_Z X, Y) + R(X, \nabla_Z Y) \\ &= R(\nabla_X Y - \nabla_Y X, Z) + R(\nabla_Y Z - \nabla_Z Y, X) + R(\nabla_Z X - \nabla_X Z, Y) \\ &= R([X, Y], Z) + R([Z, X], Y) + R([Y, Z], X) \end{aligned}$$

Using $DR = 0$, we get the *(second) Bianchi identity*

$$(\nabla_X R)(Y, Z) + (\nabla_Y R)(Z, X) + (\nabla_Z R)(X, Y) \;=\; 0$$

For any symmetric connection in TM we also have the cyclic identity

$$R(X, Y) \cdot Z + R(Y, Z) \cdot X + R(Z, X) \cdot Y \;=\; 0$$

for a triplet of vectorfields X, Y, Z on M. This identity is also called *(first) Bianchi identity* and is a trivial consequence of the symmetry condition together with the Jacobi identity (see Explanation A.18 on page 89). □

A Additional explanations

Explanation A.1: On page 26, we want to find the transformation of components $v^i_\beta(v^1_\alpha, ..., v^n_\alpha)$, $i = 1, ..., n$. For that, we write down

$$v \cdot g \;=\; \sum_{i=1}^n v^i_\alpha \left.\frac{\partial}{\partial x^i_\alpha}\right|_p \cdot g \;=\; \sum_{i=1}^n v^i_\beta \left.\frac{\partial}{\partial x^i_\beta}\right|_p \cdot g$$

We know that

$$\left.\frac{\partial}{\partial x^i_\alpha}\right|_p \cdot g \;=\; \left.\frac{\partial \tilde{g}_\alpha}{\partial x^i_\alpha}\right|_{\varphi_\alpha(p)} \qquad \left.\frac{\partial}{\partial x^i_\beta}\right|_p \cdot g \;=\; \left.\frac{\partial \tilde{g}_\beta}{\partial x^i_\beta}\right|_{\varphi_\beta(p)}$$

where $\tilde{g}_\alpha = g \circ \varphi_\alpha^{-1}$ and $\tilde{g}_\beta = g \circ \varphi_\beta^{-1}$. With $f = \varphi_\beta \circ \varphi_\alpha^{-1}$ we then get that

$$\tilde{g}_\alpha \;=\; \tilde{g}_\beta \circ \varphi_\beta \circ \varphi_\alpha^{-1} \;=\; \tilde{g}_\beta \circ f$$

We use this now, to calculate

$$\left.\frac{\partial \tilde{g}_\alpha}{\partial x^i_\alpha}\right|_{\varphi_\alpha(p)} \;=\; \sum_{j=1}^n \left.\frac{\partial \tilde{g}_\beta}{\partial x^j_\beta}\right|_{\varphi_\beta(p)} \left.\frac{\partial f^j}{\partial x^i_\alpha}\right|_{\varphi_\alpha(p)} \;=:\; \sum_{j=1}^n M^j_i \left.\frac{\partial \tilde{g}_\beta}{\partial x^j_\beta}\right|_{\varphi_\beta(p)}$$

Putting the last result into the initial equation, we obtain

$$\sum_{i=1}^n v^i_\alpha \left.\frac{\partial}{\partial x^i_\alpha}\right|_p \cdot g \;=\; \sum_{i=1}^n v^i_\alpha \sum_{j=1}^n M^j_i \left.\frac{\partial}{\partial x^j_\beta}\right|_p \cdot g \;=\; \sum_{j=1}^n v^j_\beta \left.\frac{\partial}{\partial x^j_\beta}\right|_p \cdot g$$

We conclude that

$$v^j_\beta \;=\; \sum_{i=1}^n M^j_i v^i_\alpha \qquad M^j_i \;=\; \left.\frac{\partial f^j}{\partial x^i_\alpha}\right|_{\varphi_\alpha(p)}$$

□

Explanation A.2: At the end of the proof of Theorem 4.4 on page 34, instead of integrating the terms

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-At}|x(t)|\right) \;\leq\; Be^{-At}$$

over $[0, t)$ with $t > 0$, we do the same over $[0, t)$ with $t > 0$ to obtain

$$e^{At}|x(-t)| - |y| \;\leq\; \frac{B}{A}\left(1 - e^{At}\right)$$

It follows then analogically

$$|x(-t)| \;\leq\; |y|e^{-At} + \frac{B}{A}\left(e^{-At} - 1\right) \;\leq\; |y|e^{-AT_-(y)} + \frac{B}{A}\left(e^{-AT_-(y)} - 1\right) \;=:\; R_-$$

and the obviously the same holds for $|x(t)|$. □

Explanation A.3: Equation (4.1) on page 38 follows from the definition of the differential, where $\forall f \in C^\infty(\mathbb{R})$

$$\begin{aligned}
&(\mathrm{d}\psi(p) \cdot u) \cdot f \;=\; u \cdot (f \circ \psi) \\
&\quad\Rightarrow\; \frac{\mathrm{d}}{\mathrm{d}t} f(k(t))\Big|_{t=0} \;=\; \frac{\mathrm{d}}{\mathrm{d}t} f \circ \psi(\gamma(t))\Big|_{t=0} \\
&\quad\Rightarrow\; f'(k(0))\, \mathrm{d}\psi(p) \cdot u \;=\; f'(\psi(\gamma(0)))\, \frac{\mathrm{d}}{\mathrm{d}t}\psi(\gamma(t))\Big|_{t=0} \;=\; f'(\psi(\gamma(0)))\, u \cdot \psi \\
&\quad\Rightarrow\; \mathrm{d}\psi(p) \cdot u \;=\; u \cdot \psi
\end{aligned}$$

because $f'(k(0)) = f'(\psi(\gamma(0)))$ is just a scalar (or a simple vector) and can be removed. In this case, $\mathrm{d}\psi(p) \cdot u$ or $u \cdot \psi$ can be interpreted as number (directional derivative), or as tangent vector in $T_p\mathbb{R}$. □

Explanation A.4: Equation (4.2) on page 39 can also be calculated setting $q = \varphi(p)$ and calculating

$$\begin{aligned} \mathrm{d}x^\nu \cdot \frac{\partial}{\partial x^\mu} &= \frac{\mathrm{d}}{\mathrm{d}t}\, x^\nu \left(\varphi^{-1}(q^1, ..., q^\mu + t, ..., q^m)\right)\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}(q^1, ..., q^\mu + t, ..., q^m)^\nu\Big|_{t=0} = \delta^\nu_\mu \end{aligned}$$

because by definition, we have $x^\nu = \varphi^\nu$. □

Explanation A.5: We derive equation (5.1) on page 40

$$\begin{aligned} (\phi^*\theta)\cdot v &= \theta\cdot(\mathrm{d}\phi\cdot v) = \theta_\mu \mathrm{d}y^\mu \cdot \left(\frac{\partial y^\xi}{\partial x^\eta}\, v^\eta\right)\frac{\partial}{\partial y^\xi} \\ &= \theta_\mu \frac{\partial y^\xi}{\partial x^\eta}\, v^\eta\, \mathrm{d}y^\mu \frac{\partial}{\partial y^\xi} = \theta_\mu \frac{\partial y^\xi}{\partial x^\eta}\, v^\eta\, \delta^\mu_\xi \\ &= \theta_\mu \frac{\partial y^\mu}{\partial x^\eta}\, v^\eta \end{aligned}$$

for all $v \in T_pM$. It follows equation (5.1). □

Explanation A.6: We show equation (5.2) on page 43

$$\begin{aligned} (\phi\circ\psi)_* X(p)\cdot f &= \mathrm{d}(\phi\circ\psi)\cdot X\left((\phi\circ\psi)^{-1}p\right)\cdot f = X\left((\phi\circ\psi)^{-1}p\right)\cdot(f\circ(\phi\circ\psi)) \\ &= X\left((\phi\circ\psi)^{-1}p\right)\cdot((f\circ\phi)\circ\psi) = \mathrm{d}\psi\cdot X\left((\phi\circ\psi)^{-1}p\right)\cdot(f\circ\phi) \\ &= \mathrm{d}\phi\cdot\mathrm{d}\psi\cdot X\left((\psi^{-1}\circ\phi^{-1})p\right)\cdot f = \mathrm{d}\phi\cdot\psi_* X\left(\phi^{-1}(p)\right)\cdot f \\ &= \phi_*\psi_* X(p)\cdot f \end{aligned}$$

□

Explanation A.7: Equation (5.3) on page 44 follows from

$$\begin{aligned} \left((\phi\circ\psi)^*\theta\right)\cdot v &= \theta\cdot\left(\mathrm{d}(\phi\circ\psi)\cdot v\right) = \theta\cdot\left(\mathrm{d}\phi\cdot\mathrm{d}\psi\cdot v\right) \\ &= (\phi^*\theta)\cdot(\mathrm{d}\psi\cdot v) = (\psi^*\phi^*\theta)\cdot v \end{aligned}$$

for any 1-form θ and any $v \in T_pM$. □

Explanation A.8: We verify equation (5.6) on page 48 $\forall f \in C^\infty(M)$

$$\begin{aligned} \mathrm{d}\theta\cdot(fX, Y) &= fX(\theta\cdot Y) - Y(\theta\cdot fX) - \theta\cdot[fX, Y] \\ &= fX(\theta\cdot Y) - Yf(\theta\cdot X) - \theta\cdot(fXY - YfX) \\ &= fX(\theta\cdot Y) - (Yf)(\theta\cdot X) - fY(\theta\cdot X) - f\theta\cdot XY + \theta\cdot((Yf)X + fYX) \\ &= fX(\theta\cdot Y) - (Yf)(\theta\cdot X) - fY(\theta\cdot X) - f\theta\cdot XY + (Yf)(\theta\cdot X) + f\theta\cdot YX \\ &= f(X(\theta\cdot Y) - Y(\theta\cdot X) - \theta\cdot[X, Y]) \\ &= f\mathrm{d}\theta\cdot(X, Y) \end{aligned}$$

□

Explanation A.9: We want to show that $\mathrm{d}\ell_a(e)$ is an isomorphism of T_eG onto T_aG, as claimed in Remark 5.15 on page 50. It is injective, because for $X, Y \in T_eG$, we get

$$\begin{aligned} \mathrm{d}\ell_a\cdot X\cdot f \neq \mathrm{d}\ell_a\cdot Y\cdot f &\Leftrightarrow X\cdot(f\circ\ell_a) \neq Y\cdot(f\circ\ell_a) \\ &\Leftrightarrow X \neq Y \end{aligned}$$

where the equations always hold for all $f \in C^\infty(M)$. It is surjective, because for $X \in T_aG$ with

$$X\cdot f = \frac{\mathrm{d}}{\mathrm{d}t}\, f\circ\gamma_X(t)\Big|_{t=0}$$

we get that there is a $Y \in T_eG$ with $\gamma_Y(t) = \ell_a^{-1} \circ \gamma_X(t)$, such that

$$\begin{aligned} \mathrm{d}\ell_a \cdot Y \cdot f &= Y \cdot (f \circ \ell_a) = \frac{\mathrm{d}}{\mathrm{d}t} f \circ \ell_a \circ \ell_a^{-1} \circ \gamma_X(t)\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t} f \circ \gamma_X(t)\Big|_{t=0} = X \cdot f \end{aligned}$$

for all $f \in C^\infty(M)$ and therefore $\mathrm{d}\ell_a \cdot Y = X$. □

Explanation A.10: We show that the first condition of the horizontal lift follows from $\pi \circ \gamma^\# = \gamma$, as claimed in Remark 7.2 on page 64

$$\begin{aligned} \pi \circ \gamma^\# = \gamma \quad &\Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}t} f \circ \pi \circ \gamma^\#\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} f \circ \gamma\Big|_{t=0} \\ &\Rightarrow \quad X^\# \cdot (f \circ \pi) = X \cdot f \\ &\Rightarrow \quad \mathrm{d}\pi \cdot X^\# \cdot f = X \cdot f \\ &\Rightarrow \quad \mathrm{d}\pi \cdot X^\# = X \end{aligned}$$

where we used $f \in C^\infty(M)$. □

Explanation A.11: We show equation 7.1 on page 65 for $f \in C^\infty(\mathcal{B})$

$$\begin{aligned} X^\#(0_p) \cdot f &= \frac{\mathrm{d}}{\mathrm{d}t} f \circ \gamma^\#\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} f \circ \sigma^0 \circ \pi \circ \gamma^\#\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t} f \circ \sigma^0 \circ \gamma\Big|_{t=0} = X(p) \cdot (f \circ \sigma^0) \\ &= \mathrm{d}\sigma^0 \cdot X(p) \cdot f \end{aligned}$$

where we used that here $\sigma^0 \circ \pi \circ \gamma^\# = \gamma^\#$. It follows

$$\mathrm{d}\sigma^0 \cdot X(p) = X^\#(0_p) \quad \Rightarrow \quad \mathrm{d}\sigma^0 \cdot X(p) - X^\#(0_p) = 0_p$$

□

Explanation A.12: We want to shaw equation 7.6 on page 67 for $f \in C^\infty(\mathcal{B})$. Let ψ be the chart of $\mathcal{B}$ over U and $\Gamma(t) = \psi^{-1}(x^1, ..., x^\mu + t, ..., x^m; v_1, ..., v^n)$. Then

$$\begin{aligned} \mathrm{d}S_\sigma \cdot \frac{\partial}{\partial x^\mu}\Big|_v \cdot f &= \frac{\partial}{\partial x^\mu}\Big|_v \cdot (f \circ S_\sigma) = \frac{\mathrm{d}}{\mathrm{d}t} f \circ \psi^{-1} \circ \psi \circ S_\sigma \circ \Gamma(t)\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t} \tilde{f}\big(x^1, ..., x^\mu + t, ..., x^m; v^1 + \sigma^1(\gamma(t)), ..., v^n + \sigma^n(\gamma(t))\big)\Big|_{t=0} \\ &= \mathrm{d}\tilde{f}(S_\sigma(v)) \cdot \left(0, ..., 1, ..., 0; \frac{\partial \sigma^1}{\partial x^\mu}, ..., \frac{\partial \sigma^n}{\partial x^\mu}\right) = \frac{\mathrm{d}\tilde{f}}{\mathrm{d}x^\mu}\Big|_{S_\sigma(v)} + \frac{\partial \sigma^a}{\partial x^\mu}\Big|_p \frac{\partial \tilde{f}}{\partial v^a}\Big|_{S_\sigma(v)} \\ &= \frac{\partial}{\partial x^\mu}\Big|_{S_\sigma(v)} \cdot f + \frac{\partial \sigma^a}{\partial x^\mu}\Big|_p \frac{\partial}{\partial v^a}\Big|_{S_\sigma(v)} \cdot f \end{aligned}$$

where we set $\tilde{f} = f \circ \psi^{-1}$. With $\Gamma(t) = \psi^{-1}(x^1, ..., x^m; v^1, ..., v^a + t, ..., v^n)$ we can also show equation 7.7 just in the same way

$$\begin{aligned} \mathrm{d}S_\sigma \cdot \frac{\partial}{\partial v^a}\Big|_v \cdot f &= \frac{\partial}{\partial v^a}\Big|_v \cdot (f \circ S_\sigma) = \frac{\mathrm{d}}{\mathrm{d}t} f \circ \psi^{-1} \circ \psi S_\sigma \circ \Gamma(t)\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t} \tilde{f}\big(x^1, ..., x^m; v^1 + \sigma^1(p), ..., v^a + \sigma^a(p) + t, ..., v^n + \sigma^n(p)\big)\Big|_{t=0} \\ &= \mathrm{d}\tilde{f}(S_\sigma(v)) \cdot (0, ..., 0; 0, ..., 1, ..., 0) = \frac{\mathrm{d}\tilde{f}}{\mathrm{d}x^\mu}\Big|_{S_\sigma(v)} \\ &= \frac{\partial}{\partial v^a}\Big|_{S_\sigma(v)} \cdot f \end{aligned}$$

□

Explanation A.13: To verify the Leibnitz rule for $D_X\sigma$, as mentioned in Remark 7.12 on page 69, we use the formula of 7.11

$$\begin{aligned}
D_X(f\sigma) &= \left(\left.\frac{\partial f\sigma^a}{\partial x^\mu}\right|_p + A^a_{\mu b}(p) f\sigma^b(p)\right) X^\mu \iota_a(p) \\
&= f\left(\left.\frac{\partial \sigma^a}{\partial x^\mu}\right|_p + A^a_{\mu b}(p)\sigma^b(p)\right) X^\mu \iota_a(p) + \sigma^a(p) \left.\frac{\partial f}{\partial x^\mu}\right|_p X^\mu \iota_a(p) \\
&= f D_X\sigma + (Xf)\sigma
\end{aligned}$$

□

Explanation A.14: We want to show that $F_{\mu\nu}$ transforms homogeneously, as said in remark 7.16 on page 73. For this, we use that

$$\tilde{A}_\mu = S^{-1}A_\mu S + S^{-1}\frac{\partial S}{\partial x^\mu}$$

It follows, that

$$\begin{aligned}
\tilde{F}_{\mu\nu} &= \frac{\partial \tilde{A}_\nu}{\partial x^\mu} - \frac{\partial \tilde{A}_\mu}{\partial x^\nu} + \tilde{A}_\mu\tilde{A}_\nu - \tilde{A}_\nu\tilde{A}_\mu \\
&= \frac{\partial}{\partial x_\mu}\left(S^{-1}A_\nu S + S^{-1}\frac{\partial S}{\partial x^\nu}\right) - \frac{\partial}{\partial x_\nu}\left(S^{-1}A_\mu S + S^{-1}\frac{\partial S}{\partial x^\mu}\right) \\
&\quad + \left(S^{-1}A_\mu S + S^{-1}\frac{\partial S}{\partial x^\mu}\right)\left(S^{-1}A_\nu S + S^{-1}\frac{\partial S}{\partial x^\nu}\right) - \left(S^{-1}A_\nu S + S^{-1}\frac{\partial S}{\partial x^\nu}\right)\left(S^{-1}A_\mu S + S^{-1}\frac{\partial S}{\partial x^\mu}\right) \\
&= \left(\frac{\partial S^{-1}}{\partial x^\mu}A_\nu - \frac{\partial S^{-1}}{\partial x^\nu}A_\mu\right)S + S^{-1}\left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}\right)S + S^{-1}\left(A_\nu\frac{\partial S}{\partial x^\mu} - A_\mu\frac{\partial S}{\partial x^\nu}\right) \\
&\quad + \frac{\partial S^{-1}}{\partial x^\mu}\frac{\partial S}{\partial x^\nu} - \frac{\partial S^{-1}}{\partial x^\nu}\frac{\partial S}{\partial x^\mu} + S^{-1}\left(A_\mu A_\nu - A_\nu A_\mu\right)S \\
&\quad + S^{-1}\left(A_\mu\frac{\partial S}{\partial x^\nu} - A_\nu\frac{\partial S}{\partial x^\mu}\right) + \left(S^{-1}\frac{\partial S}{\partial x^\mu}S^{-1}A_\nu - S^{-1}\frac{\partial S}{\partial x^\mu}S^{-1}A_\mu\right)S \\
&\quad + S^{-1}\frac{\partial S}{\partial x^\mu}S^{-1}\frac{\partial S}{\partial x^\nu} - S^{-1}\frac{\partial S}{\partial x^\nu}S^{-1}\frac{\partial S}{\partial x^\mu} \\
&= S^{-1}\left(\frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu}\right)S + S^{-1}\left(A_\mu A_\nu - A_\nu A_\mu\right)S \\
&= S^{-1}F_{\mu\nu}S
\end{aligned}$$

where we used that

$$\frac{\partial}{\partial x^\mu}S^{-1}S = \frac{\partial S^{-1}}{\partial x^\mu}S + S^{-1}\frac{\partial S}{\partial x^\mu} = 0 \quad\Rightarrow\quad S^{-1}\frac{\partial S}{\partial x^\mu}S^{-1} = \frac{\partial S^{-1}}{\partial x^\mu}$$

□

Explanation A.15: We want to show, why in the proof of Proposition 7.7 on page 74 it is enough to show the identity for $X = \frac{\partial}{\partial x^\mu}$ and $Y = \frac{\partial}{\partial x^\nu}$. To do so, we put in general vectorfields $X = X^\mu\partial_\mu$ and $Y = Y^\nu\partial_\nu$, denoting $\frac{\partial}{\partial x^\mu}$ by ∂_μ, and use linearity as well as the Leibnitz rule

$$\begin{aligned}
&D_X D_Y\phi - D_Y D_X\phi - D_{[X,Y]}\phi - F(X,Y)\phi \\
&\quad= X^\mu D_{\partial_\mu}Y^\nu D_{\partial_\nu}\phi + Y^\nu D_{\partial_\nu}X^\mu D_{\partial_\mu}\phi - D_{[X^\mu\partial_\mu, Y^\nu\partial_\nu]}\phi - X^\mu Y^\nu F(\partial_\mu,\partial_\nu)\phi \\
&\quad= X^\mu Y^\nu\left(D_{\partial_\mu}D_{\partial_\nu}\phi - D_{\partial_\nu}D_{\partial_\mu}\phi - D_{[\partial_\mu,\partial_\nu]}\phi - F(\partial_\mu,\partial_\nu)\phi\right) \\
&\qquad\quad + X^\mu(\partial_\mu Y^\nu)D_{\partial_\nu}\phi + Y^\nu(\partial_\nu X^\mu)D_{\partial_\mu}\phi - X^\mu D_{(\partial_\mu Y^\nu)\partial_\nu}\phi + Y^\nu D_{(\partial_\nu X^\mu)\partial_\mu}\phi \\
&\quad= X^\mu Y^\nu\left(D_{\partial_\mu}D_{\partial_\nu}\phi - D_{\partial_\nu}D_{\partial_\mu}\phi - D_{[\partial_\mu,\partial_\nu]}\phi - F(\partial_\mu,\partial_\nu)\phi\right)
\end{aligned}$$

□

Explanation A.16: We want to show that for the proof of Proposition 7.11 on page 78, it is enough to show the equation for $X = \frac{\partial}{\partial x^\mu}$, $\sigma = \iota_a$ and $\tau = \iota_b$. This follows by

$$\begin{aligned}
&X \cdot h(\sigma,\tau) - h(D_X\sigma,\tau) - h(\sigma, D_X\tau) \\
&\quad = X^\mu \frac{\partial}{\partial x^\mu} \cdot h(\sigma^a\iota_a, \tau^b\iota_b) - h(X^\mu D_{\partial/\partial x^\mu}\sigma^a\iota_a, \tau^b\iota_b) - h(\sigma^a\iota_a, X^\mu D_{\partial/\partial x^\mu}\tau^b\iota_b) \\
&\quad = X^\mu \frac{\partial}{\partial x^\mu} \cdot \overline{\sigma}^a\tau^b h(\iota_a,\iota_b) - X^\mu\overline{\sigma}^a\tau^b h(D_{\partial/\partial x^\mu}\iota_a,\iota_b) - X^\mu\overline{\sigma}^a\tau^b h(\iota_a, D_{\partial/\partial x^\mu}\iota_b) \\
&\qquad - X^\mu \frac{\partial\overline{\sigma}^a}{\partial x^\mu}\tau^b h(\iota_a,\iota_b) - X^\mu\sigma^a\frac{\partial\tau^b}{\partial x^\mu} h(\iota_a,\iota_b) \\
&\quad = X^\mu\overline{\sigma}^a\tau^b\left(\frac{\partial}{\partial x^\mu}\cdot h(\iota_a,\iota_b) - h(D_{\partial/\partial x^\mu}\iota_a,\iota_b) - h(\iota_a, D_{\partial/\partial x^\mu}\iota_b)\right)
\end{aligned}$$

where we used that

$$D_{\partial/\partial x^\mu}\sigma^a\iota_a = \sigma^a D_{\partial/\partial x^\mu}\iota_a + \frac{\partial\sigma^a}{\partial x^\mu}\iota$$

Here we used that $X \in TM$, so $\overline{X}^\mu = X^\mu$, because M is real. □

Explanation A.17: We want to show, what's stated in Remark 7.26 on page 83. So, first of all $\mathrm{d}\omega$ is totally antisymmetric, because for example

$$\begin{aligned}
&\mathrm{d}\omega(Y,X,Z) \\
&\quad = Z(\omega(Y,X)) + Y(\omega(X,Z)) + X(\omega(Z,Y)) - \omega([Y,X],Z) - \omega([X,Z],Y) - \omega([Z,Y],X) \\
&\quad = -Z(\omega(X,Y)) - Y(\omega(Z,X)) - X(\omega(Y,Z)) + \omega([X,Y],Z) + \omega([Z,X],Y) + \omega([Y,Z],X) \\
&\quad = -\mathrm{d}\omega(X,Y,Z)
\end{aligned}$$

where we used that ω itself is totally antisymmetric. $\mathrm{d}\omega$ is trilinear in X, Y, Z with respect to multiplication by the ring of $C^\infty(M)$-functions, because with $f \in C^\infty(M)$, we get

$$\begin{aligned}
&\mathrm{d}\omega(fX,Y,Z) \\
&\quad = Z(\omega(fX,Y)) + fX(\omega(Y,Z)) + Y(\omega(Z,fX)) - \omega([fX,Y],Z) - \omega([Y,Z],fX) - \omega([Z,fX],Y) \\
&\quad = f\mathrm{d}\omega(X,Y,Z) + (Zf)\omega(X,Y) + (Yf)\omega(Z,X) + (Yf)\omega(X,Z) - (Zf)\omega(X,Y) \\
&\quad = f\mathrm{d}\omega(X,Y,Z)
\end{aligned}$$

where we used that ω is bilinear with respect to the ring of $C^\infty(M)$-functions and

$$[fX,Y] = fXY - (Yf)X - fYX$$

Finally, if $\omega = \mathrm{d}\theta$ for a 1-form θ, then

$$\begin{aligned}
&\mathrm{d}\omega(X,Y,Z) \\
&\quad = Z(\mathrm{d}\theta(X,Y)) + X(\mathrm{d}\theta(Y,Z)) + Y(\mathrm{d}\theta(Z,X)) - \mathrm{d}\theta([X,Y],Z) - \mathrm{d}\theta([Y,Z],X) - \mathrm{d}\theta([Z,X],Y) \\
&\quad = Z\big(X(\theta\cdot Y) - Y(\theta\cdot X) - \theta\cdot[X,Y]\big) - [X,Y](\theta\cdot Z) + Z(\theta\cdot[X,Y]) + \theta\cdot[[X,Y],Z] \\
&\qquad + X\big(Y(\theta\cdot Z) - Z(\theta\cdot Y) - \theta\cdot[Y,Z]\big) - [Y,Z](\theta\cdot X) + X(\theta\cdot[Y,Z]) + \theta\cdot[[Y,Z],X] \\
&\qquad + Y\big(Z(\theta\cdot X) - X(\theta\cdot Z) - \theta\cdot[Z,X]\big) - [Z,X](\theta\cdot Y) + Y(\theta\cdot[Z,X]) + \theta\cdot[[Z,X],Y] \\
&\quad = \theta\cdot\big([[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y]\big) = 0
\end{aligned}$$

where we used the Jacobi identity. □

Explanation A.18: We want to proof the first Bianchi identity, as stated in Remark 7.27 on page 84. We know that

$$R(X,Y)\cdot Z = [\nabla_X,\nabla_Y]Z - \nabla_{[X,Y]}Z = \nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z - \nabla_{[X,Y]}Z$$

With the symmetry condition, it follows

$$\begin{aligned}
&R(X,Y)\cdot Z + R(Y,Z)\cdot X + R(Z,X)\cdot Y \\
&\quad = \nabla_X[Y,Z] + \nabla_Y[Z,X] + \nabla_Z[X,Y] - \nabla_{[X,Y]}Z - \nabla_{[Y,Z]}X - \nabla_{[Z,X]}Y \\
&\quad = [X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0
\end{aligned}$$

where we used the Jacobi identity. □

Stichwortverzeichnis

Symbols

A

B

C

D

E

F

G

H

I

L