Differential Geometry

MAST90143

Volker Schlue
University of Melbourne

Semester 2, 2022
(version: October 21, 2022)

Contents

Preface

These lecture notes are based on a course on Differential Geometry given by Demetrios Christodoulou at ETH Zurich in 2006/07.

Introduction

1828 C.F. Gauss “General Investigations of Curved Surfaces”; Smooth surface in Euclidean 3-dimensional space considered in itself; intrinsic geometry

1854 B. Riemann “On the hypotheses which lie at the foundation of geometry”; Geometry of n-dimensional manifolds; *Note*: From Riemann onwards manifolds were considered in and by themselves, not embedded in Euclidean space.

1930’s Modern notion of a differential manifold is established

1956 Nash demonstrated that any n-dimensional Riemannian manifold can be isometrically embedded in N-dimensional Euclidean space for some suitably large N

Remarks on the development from 1854 to the 1930’s. Riemann did not discuss parallel transport. He introduced the notion of Riemannian metric. He also introduced *curvature*, as the obstruction to the existence of a system of local coordinates where the metric takes the standard Euclidean form in rectangular coordinates. Covariant differentiation began to be investigated by Christofel, further by Ricci, who introduced the concept of tensorfield. 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 connection to general vector bundles. This found application in the gauge theories developed in physics since the 1950’s.

Part I.

Manifolds and Tangent Space

Lecture 1.

Submanifolds

Two basic notions:

- *Continuity*
- *Differentiability*

Correspondingly there are two notions of a manifold:

- *Topological manifold*
- *Differentiable manifold*

We begin with the notion of an m-dimensional *submanifold* of Euclidean n-dimensional space ($n > m$).

Simplest example: Curve on a plane. Define by graph $y = f(x)$ of a function; f continuous for *continuous curve*, f differentiable for a *differentiable curve*. (See Fig. 1.1.)

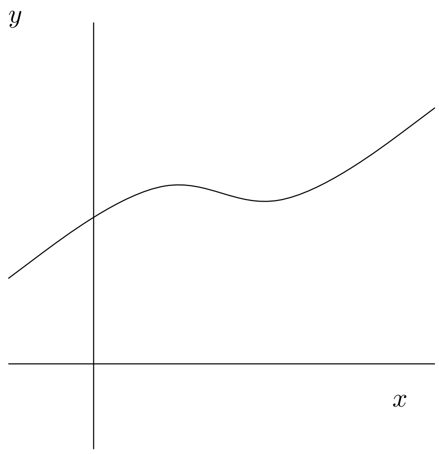


Figure 1.1.: Curve on a plane.

The above definition is too restrictive. It does not even allow for a circle.

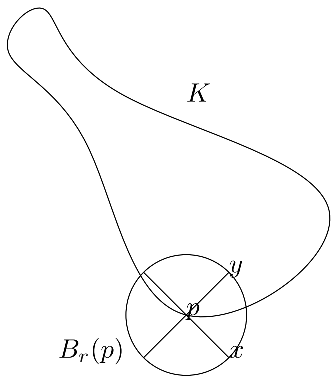


Figure 1.2.: Curve on the plane.

Let K be a closed curve in the plane. (See Fig. 1.2.) $B_r(p)$: open disk with centre at the point $p \in K$ and of radius r ($r > 0$). We consider the part of K which lies in $B_r(p)$: $K \cap B_r(p)$. Choosing r suitably small there should be a system of rectangular axes, with origin p, such that $K \cap B_r(p)$ is the graph of such a function (continuous and differentiable).

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

These definitions can be immediately generalized to the case of an m-dimensional submanifold of n-dimensional Euclidean space ($n > m$).

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

$$x^i = f^i(x^1, \ldots, x^m)\,, \quad i = m+1, \ldots, n \tag{1.1}$$

(the $n - m$ remaining coordinates as functions of the first m coordinates) where the functions $f^i : i = m+1, \ldots, n$ are all continuous (differentiable).

Lecture 2.

Topological Manifolds

We begin with a review of basic notions of topology.

Definition 2.1. A *topological space* is a set X together with a set $\mathcal{T}$ of subsets of X, such that

(1) $\emptyset, X \ \in \mathcal{T}$

(2) arbitrary unions of members of $\mathcal{T}$ are themselves members of $\mathcal{T}$

(3) finite intersections of members of $\mathcal{T}$ are themselves members of $\mathcal{T}$.

$\mathcal{T}$ is called the *topology* of X. The members of $\mathcal{T}$ 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 2.2. A *distance function* d on a set X is a real valued function on $X \times X$ such that

(1) it is symmetric $\quad d(y,x) = d(x,y) \quad \forall\, x, y \in X$

(2) $d(x,y) \geq 0$ with equality if and only if $x = y$

(3) the triangle inequality holds

$$d(x,z) \leq d(x,y) + d(y,z) \qquad \text{for any triplet } x, y, z \in X\,.$$

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

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

Given a distance function d on a set X, we define a topology $\mathcal{T}$ as follows:

$$\mathcal{U} \in \mathcal{T} \text{ accordingly as to whether} \qquad p \in \mathcal{U} \Longrightarrow \exists r > 0 \ : B_r(p) \subset \mathcal{U}\,.$$

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

Examples of metric spaces.

1. Euclidean n-dimensional space: $\mathbb{R}^n$ (a point $x \in \mathbb{R}^n$ is an n-tuplet $x = (x_1, \ldots, x_n)$ with the distance function

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \,. \tag{2.1}$$

2. On any set X define the distance function:

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

   Then

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

   It follows that any subset of X is open. The topology $\mathcal{T}$ induced by this distance function is the power set of X, namely the set of all subsets of X.

Let us recall the notion of *continuity*. Let X and Y be topological spaces and $f : X \to Y$ a mapping of X into Y. f is called continuous if the preimage of each open set in Y is an open set in X:

$$\mathcal{V} \subset Y \,, \mathcal{V} \text{ open} \Longrightarrow f^{-1}[\mathcal{V}] \subset X \text{ open} \,.$$

Proposition 2.1. *If X and Y are metric spaces f is continuous if and only if it is* continuous at x *(as defined below) for each $x \in X$.*

A function $f : X \to Y$ is continuous at x if

$$\forall \varepsilon > 0 \; \exists \delta > 0 \text{ such that } f[B_\delta(x)] \subset B_\varepsilon\big(f(x)\big) \,.$$

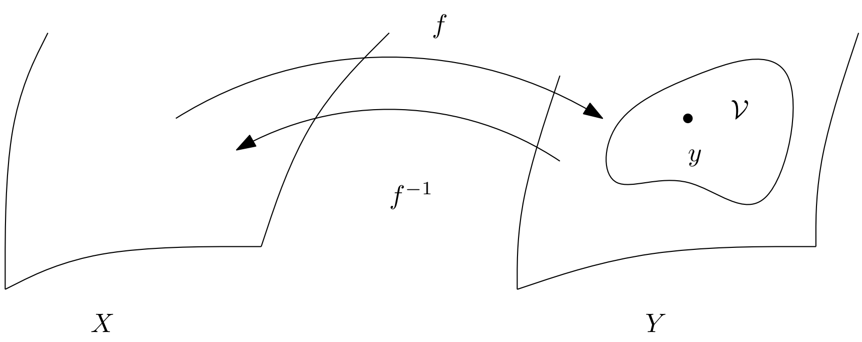


Figure 2.1.: A mapping f of metric spaces X into Y.

Proof. Let $\mathcal{U} = f^{-1}[\mathcal{V}]$. We show $\mathcal{V}$ open $\Longrightarrow \mathcal{U}$ open.

Take $x \in \mathcal{U}$. Then $f(x) = y$ for some $y \in \mathcal{V}$. Since $\mathcal{V}$ is open, $B_\varepsilon(y) \subset \mathcal{V}$ for some suitably small ε. By definition there is a $\delta > 0$ such that $f[B_\delta(x)] \subset B_\varepsilon(y)$. It follows that $B_\delta(x)$ is contained in the preimage of $\mathcal{V}$, namely $\mathcal{U}$. Thus $B_\delta(x) \subset \mathcal{U}$, hence $\mathcal{U}$ is open. □

Exercise 2.1. Prove the other direction.

Homeomorphism. 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; thus f^{-1} is a mapping of Y into X,

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

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

Example: The unit ball with center at the origin in $\mathbb{R}^n$, $B_1(0)$, is homeomorphic to $\mathbb{R}^n$ itself. $f : B_1(0) \to \mathbb{R}^n$ may be defined as follows. Let $x \in B_1(0)$. Set $|x| = d(x, 0) = \sqrt{\sum_{i=1}^n x_i^2}$, then $|x| < 1$. Set

$$f(x) = \frac{x}{\sqrt{1 - |x|^2}}\,, \tag{2.2}$$

then if $y \in \mathbb{R}^n$, we have:

$$f^{-1}(y) = \frac{y}{\sqrt{1 + |y|^2}} \tag{2.3}$$

Definition 2.3. A *topological manifold* $\mathcal{M}$ is a metric topological space for which there is a positive integer n such that every point $p \in \mathcal{M}$ has a neighborhood $\mathcal{U}$ homeomorphic to $\mathbb{R}^n$. (Thus there is a homeomorphism φ of $\mathcal{U}$ onto $\mathbb{R}^n$.) n is called the *dimension* of $\mathcal{M}$.

Problems

1. Let $S \subset \mathbb{R}^2$ be the standard circle: $S^1 = \{x \in \mathbb{R}^2 : |x| = 1\}$. Show that S^1 is a *compact* submanifold of $\mathbb{R}^2$.

2. Let $M \subset \mathbb{R}^2$ be the union of two intervals as follows:

$$M = \{(x, 0) : -1 < x < 1\} \cup \{(0, y) : -1 < y < 1\}\,.$$

   Show that M is not a submanifold of $\mathbb{R}^2$, while $M \backslash \{0\}$ is a submanifold which is not connected.

3. Exhibit a submanifold M of $\mathbb{R}^2$ which is bounded but not compact. (A submanifold M of $\mathbb{R}^n$ is *bounded* if it is contained in the ball $B_R(0)$ for R sufficiently large.)

4. Proof that there is no homeomorphism $\varphi : \mathbb{R} \to \mathbb{R}^2$.

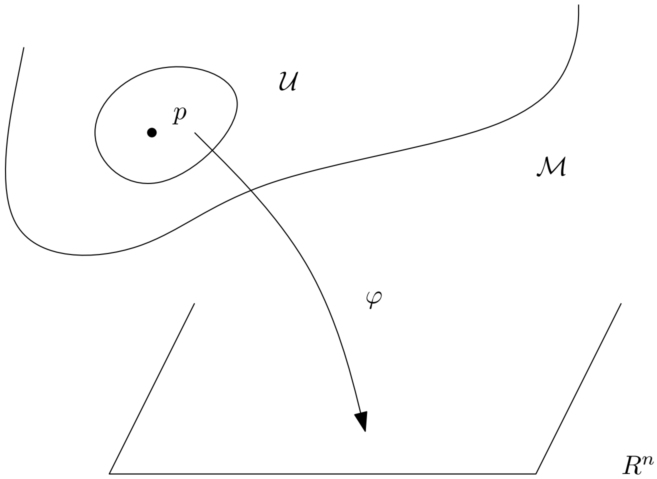


Figure 2.2.: Definition of a topological manifold $\mathcal{M}$.

Lecture 3.
Differentiable Manifolds

3.1. Differentiable Manifolds.

From the definition of a topological manifold we have a collection of open sets

$$\mathcal{A} = \Big\{\mathcal{U}_\alpha \; : \; \alpha \in I\Big\} \qquad I : \text{indexing set}$$

which cover $\mathcal{M}$. That is, for each $p \in \mathcal{M}$ there is some $\mathcal{U}_\alpha \in \mathcal{A}$ such that $p \in \mathcal{U}_\alpha$. Each $\mathcal{U}_\alpha$ is paired a homeomorphism φ_α of $\mathcal{U}_\alpha$ onto $\mathbb{R}^n$. We therefore write

$$\mathcal{A} = \Big\{(\mathcal{U}_\alpha, \varphi_\alpha) \; : \; \alpha \in I\Big\}.$$

Consider a pair of indices α, β such that $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \neq \emptyset$. Since the restriction of a homeomorphism is itself a homeomorphism (onto its image), $\varphi_\alpha|_{\mathcal{U}_\alpha \cap \mathcal{U}_\beta}$ is a homeomorphism of the open set $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ in $\mathcal{M}$ onto its image $\varphi_\alpha(\mathcal{U}_\alpha \cap \mathcal{U}_\beta) \doteq \mathcal{V}_{\alpha,\beta}$, an open set in $\mathbb{R}^n$. Similarly, $\varphi_\beta|_{\mathcal{U}_\alpha \cap \mathcal{U}_\beta}$ is a homeomorphism of $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ onto its image $\varphi_\beta(\mathcal{U}_\alpha \cap \mathcal{U}_\beta) \doteq \mathcal{V}_{\beta,\alpha}$, an open set in $\mathbb{R}^n$. Therefore the mapping

$$\varphi_{\alpha,\beta} = \varphi_\beta \circ \varphi_\alpha^{-1} : \mathcal{V}_{\alpha,\beta} \longrightarrow \mathcal{V}_{\beta,\alpha}$$

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

$$\varphi_{\beta,\alpha} = \varphi_\alpha \circ \varphi_\beta^{-1} : \mathcal{V}_{\beta,\alpha} \longrightarrow \mathcal{V}_{\alpha,\beta}$$

to be also continuously differentiable, we have $\varphi_{\beta,\alpha} = \varphi_{\alpha,\beta}^{-1}$ is continuously differentiable as well.

Definition 3.1. Let $\mathcal{M}$ be a metric topological space. The collection

$$\mathcal{A} = \Big\{(\mathcal{U}_\alpha, \varphi_\alpha) : \alpha \in I\Big\} \tag{3.1}$$

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

$$\varphi_{\alpha,\beta} : \mathcal{V}_{\alpha,\beta} \text{ open } \subset \mathbb{R}^n \longrightarrow \mathcal{V}_{\beta,\alpha} \text{ open } \subset \mathbb{R}^n \tag{3.2}$$

is continuously differentiable in the standard sense of $\mathbb{R}^n$. $\mathcal{M}$ together with such an atlas $\mathcal{A}$ is a *differentiable manifold* $\mathcal{M}_\mathcal{A}$. Each pair $(\mathcal{U}_\alpha, \varphi_\alpha)$ is called a *chart*. The open set $\mathcal{U}_\alpha$ is called the *domain* of the chart.

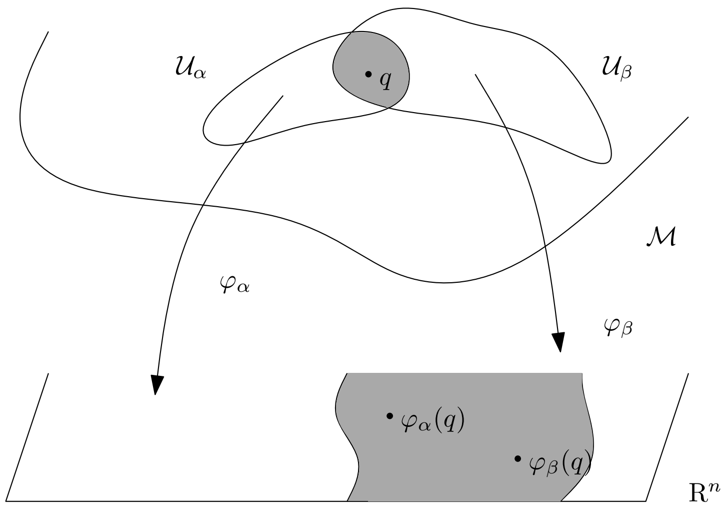


Figure 3.1.: Definition of an atlas of a differentiable manifold.

Differentiable Functions. Let $\mathcal{M}$ be a differentiable manifold. A continuous function f on $\mathcal{M}$ 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 $\mathcal{M}$ and $\mathcal{N}$ be differentiable manifolds with $\dim \mathcal{M} = m$, $\dim \mathcal{N} = n$. Let

$$\mathcal{A} = \Big\{ (\mathcal{U}_\alpha, \varphi_\alpha) : \alpha \in I \Big\}$$

be the atlas of $\mathcal{M}$,

$$\mathcal{B} = \Big\{ (\mathcal{V}_\beta, \psi_\beta) : \beta \in J \Big\}$$

be the atlas of $\mathcal{N}$. Consider a continuous mapping

$$f : \mathcal{M} \longrightarrow \mathcal{N} \, .$$

Let $f(\mathcal{U}_\alpha) \cap \mathcal{V}_\beta \neq \emptyset$ for some pair (α, β), $\alpha \in I$, $\beta \in J$. Since f is a continuous mapping and $\mathcal{V}_\beta$ being an open set in $\mathcal{N}$, $f^{-1}(\mathcal{V}_\beta)$ is an open set in $\mathcal{M}$. Thus $W_{\alpha,\beta} = \mathcal{U}_\alpha \cap f^{-1}(\mathcal{V}_\beta)$ is a non-empty open set in $\mathcal{M}$. $\varphi_\alpha|_{W_{\alpha,\beta}}$ is a homeomorphism of $W_{\alpha,\beta}$ onto its image, an open set in $\mathbb{R}^n$. We say that the mapping f is *continuously differentiable* if for every such pair of indices (α, β) $\psi_\beta \circ f \circ \varphi_\alpha^{-1}$ 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 3.2. A *diffeomorphism* of a differentiable manifold $\mathcal{M}$ onto a differentiable manifold $\mathcal{N}$ is a homeomorphism f of $\mathcal{M}$ onto $\mathcal{N}$ such that both f, f^{-1} are continuously differentiable mappings. Two differentiable manifolds are called *diffeomorphic* if there exists a diffeomorphism of $\mathcal{M}$ onto $\mathcal{N}$.

Let $\mathcal{M}$ be a topological manifold and let

$$\mathcal{A} = \Big\{ (\mathcal{U}_\alpha, \varphi_\alpha) : \alpha \in I \Big\} , \qquad \mathcal{B} = \Big\{ (\mathcal{V}_\beta, \psi_\beta) : \beta \in J \Big\} ,$$

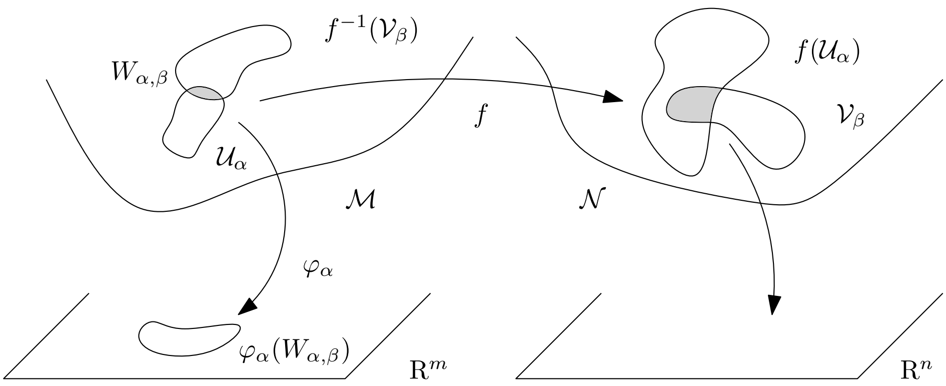


Figure 3.2.: Continuously differentiable mappings.

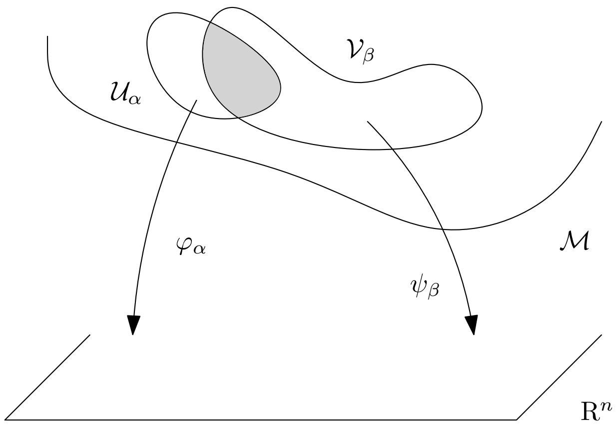


Figure 3.3.: Compatible atlases.

be two atlases for $\mathcal{M}$. We write $\mathcal{M}_\mathcal{A}$, $\mathcal{M}_\mathcal{B}$ for the corresponding differential manifolds. We say the two atlases are *compatible* if the identity mapping

$$\mathrm{id} : \mathcal{M}_\mathcal{A} \longrightarrow \mathcal{M}_\mathcal{B}\,, \quad \mathrm{id}(p) = p \ : \forall p \in \mathcal{M}$$

is a diffeomorphism. Obviously id is a homeomorphism. Since also the inverse $\mathrm{id}^{-1} = \mathrm{id}$, 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 consider all pairs of indices $\alpha \in I$, $\beta \in J$ such that $\mathcal{U}_\alpha \cap \mathcal{V}_\beta \neq \emptyset$. With $W_{\alpha,\beta} = \mathcal{U}_\alpha \cap \mathcal{V}_\beta$ the mapping $\psi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(W_{\alpha,\beta})$ open in $\mathbb{R}^n \to \mathbb{R}^n$ must be continuously differentiable. In other words $\mathcal{A}$ is compatible with $\mathcal{B}$ if and only if $\mathcal{A} \cup \mathcal{B}$ is also an atlas for $\mathcal{M}$. ($\mathcal{A} \cup \mathcal{B}$ consists of all the charts of $\mathcal{A}$ together with all the charts of $\mathcal{B}$.) Compatibility is an equivalence relation.

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

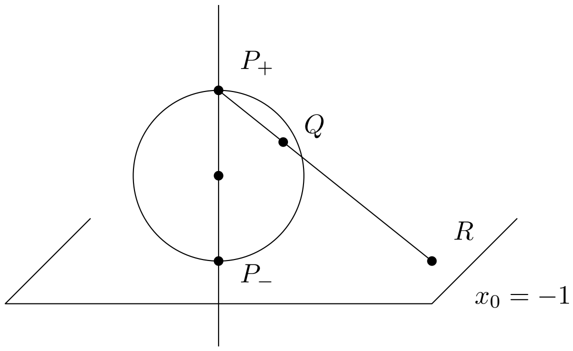


Figure 3.4.: Stereographic projection from the North Pole.

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

$$\mathbb{S}^n = \Big\{(x_0, x) : x_0^2 + |x|^2 = 1\Big\} \quad \text{where } |x| = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

We distinguish the two points

$$P_+ = (1, 0) : \quad \text{North pole of } \mathbb{S}^n$$
$$P_- = (-1, 0) : \quad \text{South pole of } \mathbb{S}^n\,.$$

The stereographic projection from P_+ maps $\mathbb{S}^n \backslash P_+$ onto the plane $x_0 = -1$, which may be identified with $\mathbb{R}^n$. Consider

$$Q = (x_0, x) \in \mathbb{S}^n\,, \quad R = (-1, y)\,,$$

as in Figure 3.4. The condition that P_+, Q, and R are colinear reads:

$$\overleftarrow{RP_+} = \lambda\, \overleftarrow{QP_+}$$
$$\overleftarrow{QP_+} = (x_0, x) - (1, 0) = (x_0 - 1, x)$$
$$\overleftarrow{RP_+} = (-1, y) - (1, 0) = (-2, y)$$

The condition is

$$-2 = \lambda(x_0 - 1)\,, \quad y = \lambda x\,,$$

that is

$$y = \frac{2}{1 - x_0} x\,.$$

This defines a homeomorphism φ_+, $y = \varphi_+(Q)$, of $\mathbb{S}^n \backslash P_+ = \mathcal{U}_+$ onto $\mathbb{R}^n$. Note that the image of P_- is $0 \in \mathbb{R}^n$.

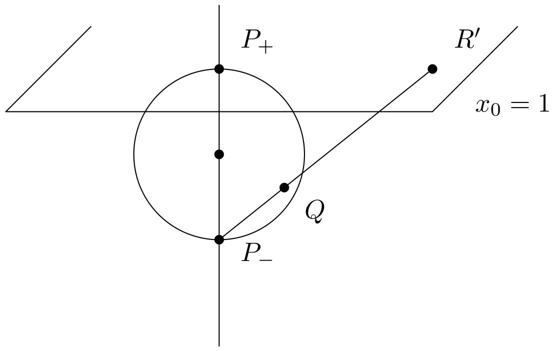


Figure 3.5.: Stereographic projection from P_-.

Similarly, for the stereographic projection from P_- (see figure 3.5), the condition that P_-, Q, and R' are colinear reads:

$$\overleftarrow{R'P_-} = \lambda' \, \overleftarrow{QP_-}$$
$$\overleftarrow{QP_-} = (x_0, x) - (-1, 0) = (x_0 + 1, x)$$
$$\overleftarrow{R'P_-} = (1, y') - (-1, 0) = (2, y')$$

The condition is here

$$2 = \lambda'(x_0 + 1)\,, \quad y' = \lambda' x\,,$$

that is

$$y' = \frac{2}{1 + x_0} x\,, \quad y' = \varphi_-(Q)\,.$$

This defines a homeomorphism φ_- of $\mathbb{S}^n \backslash P_- = \mathcal{U}_-$ onto $\mathbb{R}^n$. The image of P_+ is $0 \in \mathbb{R}^n$.

We have an atlas with two charts

$$\mathcal{A} = \Big\{ (\mathcal{U}_+, \varphi_+), (\mathcal{U}_-, \varphi_-) \Big\}\,.$$

In fact, $\mathcal{U}_+ \cap \mathcal{U}_- = \mathbb{S}^n \backslash \{P_+, P_-\}$ and

$$\varphi_+(\mathcal{U}_+ \cap \mathcal{U}_-) = \mathbb{R}^n \backslash 0 = \varphi_-(\mathcal{U}_+ \cap \mathcal{U}_-)\,.$$

The mappings $\varphi_{+,-} = \varphi_- \circ \varphi_+^{-1}$, $\varphi_{-,+} = \varphi_+ \circ \varphi_-^{-1}$ (where $\varphi_{+,-}^{-1} = \varphi_{-,+}$) are given by ($Q = \varphi_+^{-1}(y)$, $\varphi_-(Q) = y'$)

$$y' = \varphi_{+,-}(y)\,, \quad y = \varphi_{-,+}(y')\,.$$

These mappings are homeomorphisms of $\mathbb{R}^n \backslash 0$ onto itself. We see from the above formulas that

$$y' = \alpha y\,, \quad \alpha > 0\,.$$

To find α consider

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

and thus we obtain simply

$$y' = \frac{4y}{|y|^2}, \quad y = \frac{4y'}{|y'|^2}.$$

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

We conclude this lecture by showing show that the differentiable submanifolds introduced in Lecture 1 are actually differentiable manifolds as defined here.

Proposition 3.1. *Let $\mathcal{M}$ be an m-dimensional* submanifold $\mathcal{M}$ *of* E^n *($m < n$) in sense of Definition 1.2. Then $\mathcal{M}$ is in fact a differentiable manifold.*

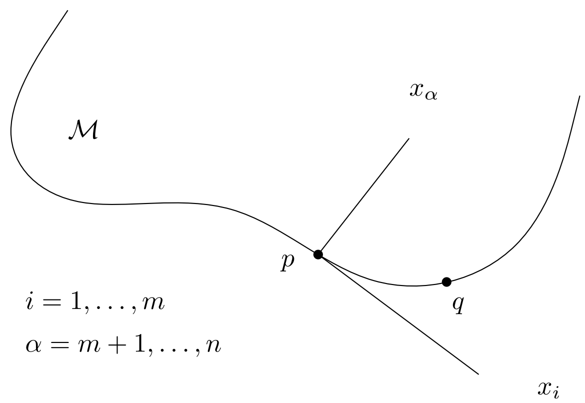
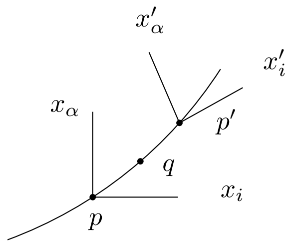


Figure 3.6.: Charts of a differentiable submanifold of Euclidean space.

Proof. At each point $p \in \mathcal{M}$, $\mathcal{M}$ is locally (in some neighborhood $\mathcal{U}_p$) represented as a graph $x_\alpha = f_\alpha(x_1, \ldots, x_m)$ w.r.t. a system of rectangular axes with origin at p. We consider the collection $\{(\mathcal{U}_p, \varphi_p) : p \in \mathcal{M}\} = \mathcal{A}$, where

$$\varphi_p(q) = (x_1, \ldots, x_m) \;\text{ : the first } m \text{ coordinates of } q \in \mathcal{U}_p \text{ w.r.t. the axes based on } p,$$

is a homeomorphism of $\mathcal{U}_p$ onto an open set in $\mathbb{R}^n$ which is homeomorphic to $\mathbb{R}^n$ itself. We must show that the above collection constitutes an atlas for $\mathcal{M}$. Consider then a pair $p, p' \in \mathcal{M}$ such that $\mathcal{U}_p \cap \mathcal{U}_{p'} \neq \emptyset$.



Let $g = \varphi_{p'} \circ \varphi_p^{-1}$, defined on $\varphi_p(\mathcal{U}_p \cap \mathcal{U}_{p'})$: an open set in $\mathbb{R}^m$. Let $q \in \mathcal{U}_p \cap \mathcal{U}_{p'}$. Then q has a representation in both the primed and unprimed axes. But the coordinates of any point, in particular q, w.r.t. two systems of (right handed) axes differ by a translation and rotation. Thus for q:

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

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

$$\begin{aligned} x'_i &= \sum_{j=1}^{n} O_{ij} x_k + c_i = \sum_{j=1}^{m} O_{ij} x_j + \sum_{\alpha=m+1}^{n} O_{i\alpha} x_\alpha + c_i \\ &= \sum_{j=1}^{m} O_{ij} x_j + \sum_{\alpha=m+1}^{n} O_{i\alpha} f_\alpha(x_1, \ldots, x_m) + c_i \,, \qquad (i = 1, \ldots, m). \end{aligned}$$

Since the functions $f_\alpha : \alpha = m+1, \ldots, n$ are continuously differentiable, the mapping g is itself continuously differentiable. ($g = g_{p,p'}$).

Thus the collection $\mathcal{A}$ is in fact an atlas for $\mathcal{M}$ and $\mathcal{M}_{\mathcal{A}}$ is indeed a differentiable manifold.

□

Problems

1. Let us consider $\mathbb{R}$ with its usual differentiable structure given by the chart $\phi : \mathbb{R} \to \mathbb{R}\,, x \mapsto x$, and let us also define the differentiable structure given by the chart $\psi : \mathbb{R} \to \mathbb{R}, x \mapsto x^3$. Show that the two differentiable structures are different (the charts are not compatible), but nevertheless $(\mathbb{R}, \phi)$ and $(\mathbb{R}, \psi)$ are diffeomorphic.

2. We have seen how to define a differentiable structure on $\mathbb{S}^2$ using the stereographic projections. Now we shall do that in another way. Consider the open sets

$$\mathcal{U}_i^{\pm} \doteq \{(x^1, x^2, x^3) \in \mathbb{S}^2 : \pm x^i > 0\}.$$

It is clear that $\{\mathcal{U}_i^{\pm}\}$ is an open covering of $\mathbb{S}^2$, and we define the charts $\phi_i^{\pm} : \mathcal{U}_i^{\pm} \to \mathbb{R}^2$ as follows:

$$\phi_1^{\pm}(x^1, x^2, x^3) = (x^2, x^3)\,, \quad \phi_2^{\pm}(x^1, x^2, x^3) = (x^1, x^3), \quad \phi_3^{\pm}(x^1, x^2, x^3) = (x^1, x^2)\,.$$

Prove that this induces a differentiable structure, and prove that it is equivalent to the one given by the stereographic projection.

Supplement: Differentiability

The concept of differentiabiliy requires a linear structure.

For example, for curves in the plane the derivative at a point p gives the tangent line, which is the closest approximation of a curve by a straight line in a neighborhood of p.

Vector spaces.

Let us first recall the notion of a *linear* or *vector space*.

Definition 3.1. A set V is a *vector space* over the reals, and its elements are called vectors, if it is endowed with the following:

1. *(vector addition)* $v_1 + v_2 \in V$ for every pair $v_1, v_2 \in V$.
   (i) which is associative: $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$
   (ii) and commutative: $v_1 + v_2 = v_2 + v_1$
   (iii) there is a distinguished element $0 \in V$ called the 0 vector such that: $v + 0 = v \quad \forall v \in V$
2. *(scalar multiplication)* $\alpha v \in V$ for every $\alpha \in \mathbb{R}$, $v \in V$.
   (i) $0\ v = 0 \quad \forall v \in V$
   (ii) $1\ v = v \quad \forall v \in V$
   (iii) $(\alpha\beta)v = \alpha(\beta v) \quad \forall \alpha, \beta \in \mathbb{R},\ \forall v \in V$
3. Moreover we have the distributivity properties:
   (i) $\alpha(v_1 + v_2) = \alpha v_1 + \alpha v_2$
   (ii) $(\alpha + \beta)v = \alpha v + \beta v$.

If in 3.(ii) we take $\alpha = 1$ and $\beta = -1$ we have $\alpha + \beta = 0$ so by 2.(i) the left hand side of is 0, the zero vector. The vector $(-1)v$ we denote simply by $-v$, so 3.(ii) becomes in this case, in view of 2.(ii), $0 = v + (-v)$. Thus there is an additive inverse to every vector v, the vector $-v$.

Example: The space of continuous real valued functions on E^n.

$$\text{Vector addition: } (f + g)(x) = f(x) + g(x)$$

satisfies the requirements 1. with the zero vector the function f defined by $f(x) = 0 \quad \forall x \in \mathrm{E}^n$.

$$\text{Scalar multiplication: } (\alpha f)(x) = \alpha f(x) \quad \forall x \in \mathrm{E}^n$$

satisfies the requirements 2. and 3. as well.

Definition 3.2. A vector space V is said to be *of dimension* n, and thus finite dimensional, if there is a set $(e_1, \ldots, e_n)$ of n linearly independent vectors, and every set of $n+1$ vectors is linearly dependent.

Such a set $(e_1, \ldots, e_n)$ is called a basis for V. Given such a basis, any vector $v \in V$ can be expanded as

$$v = \sum_{i=1}^{n} v^i \, e_i \,, \tag{3.1}$$

where $v^i \in \mathbb{R} : i = 1, \ldots, n$ are uniquely determined in the given basis.

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

$$v \in V \longmapsto (v^1, \ldots, v^n) \in \mathbb{R}^n \,.$$

The n-tuplet $(v^1, \ldots, v^n)$ are the components of v in the given basis. On a vector space V we can define a norm $\|\cdot\|$, $\|v\|$ 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. $\|\alpha v\| = |\alpha| \|v\|$ ($|\alpha|$ the absolute value of the real number α)
3. (Minkowski inequality) $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\|$

Remark 3.1. Given a norm $\|\cdot\|$ in a vector space, there is a distance function d defined by this norm:

$$d(u, v) = \|u - v\| \,. \tag{3.2}$$

The Minkowski inequality for $\|\cdot\|$ then translates to the triangle inequality for d.

A normed linear space is called a *Banach space* if it is *complete* with respect to the distance function induced by its norm. We say that a metric space is complete if every Cauchy sequence converges (to some point in the space).

Definition 3.3. A sequence $(x_n : n = 1, 2, \ldots)$ is called a *Cauchy sequence* if

$$\forall \varepsilon > 0 \; \exists N \text{ such that } n, m \geq N \Longrightarrow d(x_n, x_m) < \varepsilon \,.$$

We say that $x_n \to x$ if $d(x_n, x) \to 0$ as $n \to \infty$.

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 a finite dimensional linear space are equivalent.

Let V be an n-dimensional linear space. Let $(e_1, \ldots, e_n)$ be a basis for V. We can assume that the basis is normalized, that is $\|e_i\| = 1 : i = 1, \ldots, n$ (otherwise we set $\hat{e}_i = e_i/\|e_i\| : i = 1, \ldots, n$ to obtain a normalized basis). Any vector $v \in V$ is expanded as

$$v = \sum_{i=1}^{n} v_i e_i \qquad (v_1, \ldots, v_n) \in \mathbb{R}^n \tag{3.3}$$

in a unique way. Thus we have an isomorphism

$$v \in V \longmapsto (v_1, \ldots, v_n) \in \mathbb{R}^n .$$

In $\mathbb{R}^n$ we have the Euclidean norm

$$|x| = \sqrt{\sum_{i=1}^{n} x_i^2}\,, \qquad x = (x_1, \ldots, x_n) \in \mathbb{R}^n . \tag{3.4}$$

We shall show that all norms in a finite dimensional linear space V are equivalent. It will then follow that there are positive constants C_m and C_M such that we have

$$C_m|v| \leq \|v\| \leq C_M|v| . \tag{3.5}$$

Moreover, it will then follow that given a Cauchy sequence in V, the corresponding sequence of components (relative to the basis $(e_1, \ldots, e_n)$) 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.

Recall the Minkowski inequality:

$$\|v + u\| \leq \|v\| + \|u\| \tag{3.6}$$

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

$$\|\sum_{\alpha=1}^{m} \overset{(\alpha)}{v}\| \leq \sum_{\alpha=1}^{m} \|\overset{(\alpha)}{v}\| . \tag{3.7}$$

Let us apply this to the n vectors $v_1 e_1, \ldots, v_n e_n$:

$$\|v\| \leq \sum_{i=1}^{n} \|v_i e_i\| = \sum_{i=1}^{n} |v_i| \tag{3.8}$$

Claim 3.1. *There is a positive constant M such that*

$$\sum_{i=1}^{n} |v_i| \leq M\|v\| . \tag{3.9}$$

Proof. By contradiction. Suppose that no such constant M exists. That is suppose that there is no upper bound for the ratio

$$\frac{1}{\|v\|}\sum_{i=1}^{n}|v_i| \quad \text{in } V\,.$$

Then there is a sequence $(\overset{(m)}{v} : m = 1, \ldots)$ of non-zero vectors such that

$$\sum_{i=1}^{n}|\overset{(m)}{v}_i| = 1\,, \quad m = 1, \ldots$$

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

$$\overset{(m_k)}{v}_i \longrightarrow v_i^* \in [-1, 1] \text{ as } k \to \infty \quad : i = 1, \ldots, n.$$

It follows that

$$1 = \sum_{i=1}^{n}|\overset{(m_k)}{v}_i| \longrightarrow \sum_{i=1}^{n}|v_i^*| \text{ as } k \to \infty\,.$$

Hence

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

Set $v^* = \sum_{i=1}^{n} v_i^* e_i$, then

$$v^* - \overset{(m_k)}{v} = \sum_{i=1}^{n}(v_i^* - \overset{(m_k)}{v}_i)e_i\,,$$

and therefore

$$\|v^* - \overset{(m_k)}{v}\| \leq \sum_{i=1}^{n}|v_i^* - \overset{(m_k)}{v}_i| \longrightarrow 0 \text{ as } k \to \infty\,.$$

By the Minkowski inequality

$$\|v^*\| \leq \|v^* - \overset{(m_k)}{v}\| + \|\overset{(m_k)}{v}\| \longrightarrow 0 \text{ as } k \to \infty\,.$$

It follows that $\|v^*\| = 0$, therefore $v^* = 0$, contradicting the fact that

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

This establishes the claim. □

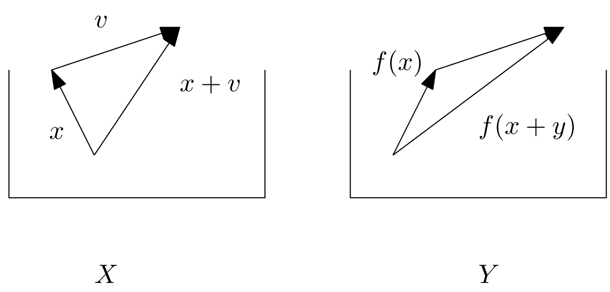


Figure 3.1.: Definition of the derivative of a linear map f.

Differentiable Linear Maps.

Let X, Y be normed linear spaces and f a continuous mapping $f : X \to Y$.

Definition 3.4. We say that f is *differentiable* at $x \in X$ if there is a linear map $Df : X \to Y$ such that

$$f(x+v) - f(x) = Df(x) \cdot v + h_x(v) \tag{3.10}$$

with

$$\frac{\|h_x(v)\|_Y}{\|v\|_X} \to 0 \quad \text{as } v \to \ 0\,. \tag{3.11}$$

Then $Df(x)$ is the derivative of f at x.

Remark 3.2. Our notation for linear maps is simply

$$L(v) = L\,v\,.$$

If we fix v with $\|v\| = 1$ and consider the linear span of v,

$$\text{span}\{v\} = \{\alpha v : \alpha \in \mathbb{R}\}\,,$$

then, since

$$f(x+\alpha v) - f(x) = \alpha\, Df \cdot v + h_x(\alpha v)$$

with

$$\frac{\|h_x(\alpha v)\|_Y}{\|\alpha v\|_X} \to 0 \quad \text{as } \|\alpha v\|_X \to 0\,,$$

the condition (3.11) is equivalent to

$$\frac{h_x(\alpha v)}{|\alpha|} \to 0 \quad \text{as } \alpha \to 0\,.$$

Lecture 4.

Quotient spaces

Sets. Let $\sim$ be an equivalence relation on a set X. We write $x \sim y$ to denote $(x, y) \in \sim$, where $\sim$ is viewed as a subset of $X \times X$. The *equivalence class* C_x of x is:

$$\mathcal{C}_x = \Big\{ y \in X : y \sim x \Big\} \tag{4.1}$$

It follows that for any pair $x, y \in X$, $x \sim y$ if and only if $\mathcal{C}_x = \mathcal{C}_y$, and $x \nsim y$ if and only if $\mathcal{C}_x \cap \mathcal{C}_y = \emptyset$. The *quotient* $\mathcal{Q} = X/\sim$ is the set of equivalence classes

$$\mathcal{Q} = \Big\{ \mathcal{C}_x : x \in X \Big\} . \tag{4.2}$$

There is a projection $\pi : X \to \mathcal{Q}$ by

$$\pi(x) = \mathcal{C}_x . \tag{4.3}$$

Topological spaces. If X is a topological space there is a natural topology induced on $\mathcal{Q}$, the *quotient topology*: A subset $\mathcal{U} \subset \mathcal{Q}$ (a set of equivalence classes) is open in Q if and only if $\pi^{-1}(\mathcal{U})$ is an open set in X.

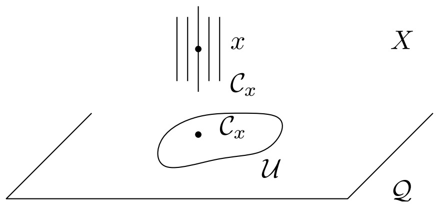


Figure 4.1.: Open sets in the quotient topology.

Consider now the case that X is a topological manifold. Then $X/\sim$ is *not*, in general, a topological manifold.

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

$$\sim = \Delta \cup \{(x', x''), (x'', x')\}$$

where Δ is the diagonal, $\Delta = \{(x, x) : x \in X\}$. Then $\mathcal{Q} = X/\sim$ is not a topological manifold. $\mathcal{Q}$ can be viewed as a folded plane which intersects itself at one point (see Figure 4.2). No neighborhood of the equivalence class $\mathcal{C}_{x'}$ is homeomorphic to $\mathbb{R}^2$.

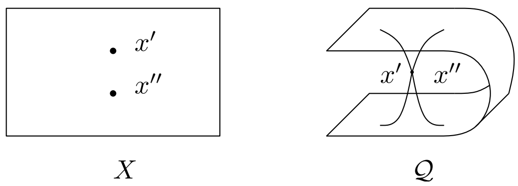


Figure 4.2.: Example of a quotient which is not a topological manifold.

This raises the question which equivalence relations lead to quotients of topological manifolds which themselves are topological manifolds.

The oldest example of a quotient is the *celestial sphere* in ancient astronomy: In $\mathbb{R}^{n+1}$ we introduce the equivalence relation

$$y \sim x \Longleftrightarrow y = \alpha x \quad (\alpha > 0)\,.$$

The equivalence class $\mathcal{C}_x$ is then the ray through x. The quotient $\mathbb{R}^{n+1} \setminus 0/\sim$ is in fact $\mathbb{S}^n$. In each ray there is exactly one representative y, with $\|y\| = 1$.

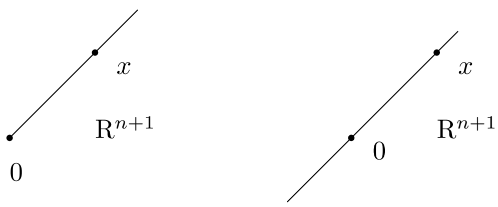


Figure 4.3.: Rays in $\mathbb{R}^{n+1}$.

Another example: In $\mathbb{R}^{n+1} \setminus 0$ we introduce the equivalence relation

$$y \sim x \Longleftrightarrow y = \alpha x \quad (\alpha \in \mathbb{R} \setminus 0)\,.$$

Then $\mathcal{C}_x$ is the line through x. The quotient $\mathbb{R}^{n+1} \setminus 0/\sim$ is called *real projective space* $\mathbb{RP}^n$. In each equivalence class there are exactly two representatives $y, -y$ lying on $\mathbb{S}^n : |y| = 1$. Thus we can also consider $\mathbb{RP}^n$ to be the quotient

$$\mathbb{RP}^n = \mathbb{S}^n/\sim \tag{4.4}$$

where $\sim$ is the equivalence relation $y \sim x$ if and only if either $y = x$ or $y = -x$. An equivalence class is then a pair $\{x, -x\}$ of antipodal points on $\mathbb{S}^n$. So $\mathbb{RP}^n$ is the space of diameters.

The projection π is given by

$$\pi : \mathbb{S}^n \longrightarrow \mathbb{RP}^n\,, \pi(x) = \{x, -x\} \doteq y\,. \tag{4.5}$$

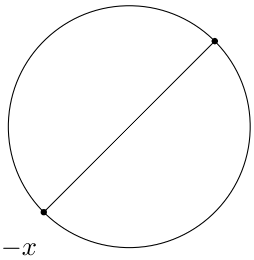


Figure 4.4.: Antipodal points on the sphere.

Let $\mathcal{U}_y$ be a neighborhood of $y \in \mathbb{RP}^n$. Then $\pi^{-1}(\mathcal{U}_y)$ is an open set in $\mathbb{S}^n$, containing x and such that

$$x' \in \pi^{-1}(\mathcal{U}_y) \Rightarrow -x' \in \pi^{-1}(\mathcal{U}_y)\,.$$

In fact, any open set $\mathcal{V}$ in $\mathbb{S}^n$ with these properties ($x \in \mathcal{V}$ and $x' \in \mathcal{V} \Rightarrow -x' \in \mathcal{V}$) projects to a neighborhood of y in $\mathbb{RP}^n$. For any subset $K \subset \mathbb{S}^n$ let us denote $K^* = \{-x : x \in K\}$.

For any $x \in \mathbb{S}^n$ we can find a neighborhood $\mathcal{V}_x$ of x, homeomorphic to $\mathbb{R}^n$, and such that $\mathcal{V}_x^* \cap \mathcal{V}_x = \emptyset$. Let then $\varphi_x : \mathcal{V}_x \longrightarrow \mathbb{R}^n$ be a homeomorphism. The corresponding

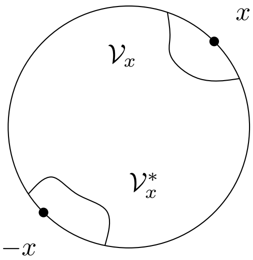


Figure 4.5.: Open sets in projective space.

neighborhood of $y = \{x, -x\}$ in $\mathbb{RP}^n$ is

$$\mathcal{W}_y = \Big\{\{x', -x'\} : x' \in \mathcal{V}_x\Big\} = \pi(\mathcal{V}_x)\,.$$

Given a pair $\{x', -x'\} \in \mathcal{W}_y$ exactly one of the x', $-x'$ belongs to $\mathcal{V}_x$ (and the other to $\mathcal{V}_x^*$). We can then define the mapping $\psi_y : \mathcal{W}_y \longrightarrow \mathbb{R}^n$ by

$$\psi_y(\{x', -x'\}) = \begin{cases} \varphi_x(x') & \text{if } x' \in \mathcal{V}_x \\ \varphi_x(-x') & \text{if } -x' \in \mathcal{V}_x\,. \end{cases}$$

ψ_y is then a homeomorphism and $\mathbb{RP}^n$ is a topological manifold. Moreover, if $\{(\mathcal{V}_x, \varphi_x) : x \in \mathbb{S}^n\}$ is an atlas for $\mathbb{S}^n$ then $\{(W_y, \psi_y) : y \in \mathbb{RP}^n\}$ is an atlas for $\mathbb{RP}^n$ and $\mathbb{RP}^n$ is a differentiable manifold.

The example of $\mathbb{RP}^n$ is an example of a quotient of a differentiable manifold $\mathcal{M}$ by a discrete group G of diffeomorphisms of $\mathcal{M}$. In the case of $\mathbb{RP}^n$, $\mathcal{M} = \mathbb{S}^n$ and $G = \{\mathrm{id}, a\}$ where a is the *antipodal* map $a(x) = -x$, $x \in \mathbb{S}^n$, $(a \circ a = \mathrm{id})$.

Definition 4.1. A *discrete group G* of diffeomorphism is a countable set of diffeomorphisms of $\mathcal{M}$ onto itself which form a group under composition. [Namely, $f_1, f_2 \in G \Rightarrow f_1 \circ f_2 \in G$, $f \in G \Rightarrow f^{-1} \in G$, and $\mathrm{id} \in G$.] A discrete group G of homeomorphisms of a topological manifold $\mathcal{M}$ is defined in a similar way.

There is the following is a *sufficient* condition for $\mathcal{M}/G$ to be a topological (differentiable) manifold: For each $p \in \mathcal{M}$ there is a neighborhood $\mathcal{U}$ of p such that

$$f(\mathcal{U}) \cap \mathcal{U} \neq \emptyset, f \in G \Longrightarrow f = \mathrm{id}\ . \tag{4.6}$$

Here $\mathcal{M}/G$ means the quotient of $\mathcal{M}$ by the equivalence relation

$$q \sim p \Leftrightarrow \exists f \in G \text{ s.t. } q = f(p)\,.$$

Example: The n-dimensional torus $\mathbb{T}^n$ arises as the quotient

$$\mathbb{T}^n = \mathbb{R}^n / Z^n$$

with $Z^n = \{f_k : k \in (k_1, \ldots, k_n) \in \mathbb{Z}^n\}$, where $f_k(x) = x+k = (x_1+k_1, \ldots, x_n+k_n)$. It is easy to check the sufficient condition: For every $x \in \mathbb{R}^n$, $B_{\frac{1}{2}}(x)$ has the required property. For, let $x', x'' \in B_{\frac{1}{2}}(x)$, $x'' = x' + k$, $k \in \mathbb{Z}^n$, $k \neq 0$; then $|k| \geq 1$, that is $|x' - x''| \geq 1$ but this is impossible because the diameter of $B_{\frac{1}{2}}(x)$ is 1. That is, $B_{\frac{1}{2}}(x) \cap f_k(B_{\frac{1}{2}}(x)) = \emptyset$. A continuous (continuously differentiable) function φ on $\mathbb{T}^n$ is induced by a continuous (continuously differentiable) *periodic* function $\tilde{\varphi}$ on $\mathbb{R}^n$, $\tilde{\varphi}(f_k(x)) = \varphi(x) : \forall k \in \mathbb{Z}^n$.

The above sufficient condition (4.6) is *not necessary.*

Example: Quotient of $\mathbb{R}^2$ by a cyclic group G_n of rotations about the origin. Identify $\mathbb{R}^2$ with the complex plane $\mathbb{C}^1$, then $G_n = \{f_k : k = 1, \ldots, n-1\}$ where f_k is the k-fold composition of f, and $f(z) = e^{\frac{2\pi i}{n}} z$, thus $f_k(z) = e^{2\pi i \frac{k}{n}} z$, $f_0 = \mathrm{id}$. Consider the sector $\arg z \in [0, \frac{2\pi}{n}]$. The orbit of z by G_n consists of the points $\{z, e^{\frac{i\pi}{n}}, \ldots, e^{2\pi i \frac{n-1}{n}} z\}$. The quotients are cones; see Figure 4.6. The cone is homeomorphic to $\mathbb{R}^2$ itself. However the origin does not move by any rotation. Therefore if $\mathcal{U}$ is any neighborhood of the origin, then $f_k(\mathcal{U})$ must intersect $\mathcal{U}$ for any $k = 0, \ldots, n-1$. But despite the failure of the sufficient condition the quotient is a topological manifold, although not a differentiable manifold.

Remark 4.1. The n-dim. torus $\mathbb{T}^n$ is diffeomorphic to the product $\mathbb{S}^1 \times \cdots \times \mathbb{S}^1$.

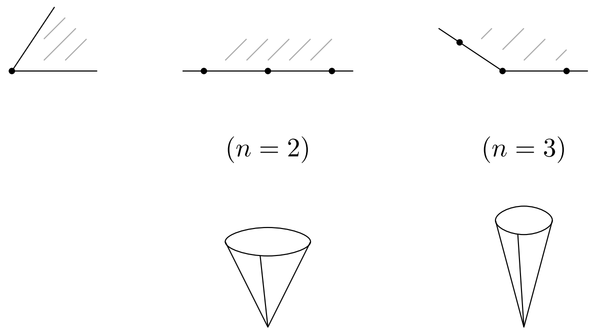


Figure 4.6.: Quotients of the plane by cyclic groups.

Problems

1. Let us consider the group of rotations acting on $\mathbb{R}^3$, $G \doteq \{R_\theta : \theta \in ([0, 2\pi)\}$, where $R_\theta : \mathbb{R}^3 \to \mathbb{R}^3$ is a rotation about the z-axis of angle θ, i.e.

   $$R_\theta(x, y, z) = (x \cos\theta - y \sin\theta, x \sin\theta + y \sin\theta, z)\,.$$

   Now let us consider the subgroup $G_n < G$ given by

   $$G_n \doteq \{R_\theta : \theta = 0, \frac{2\pi}{n}, \frac{4\pi}{n}, \ldots, \frac{(n-1)2\pi}{n}\}\,.$$

   Show that the quotient $(\mathbb{R}^3 \setminus \{(0, 0, z) : z \in \mathbb{R}\})/G_n$ is a manifold of dimension 3.

2. Let G be the group of rotations of the previous exercise. Show that the quotient $\mathbb{R}^3/G$ is homeomorphic to the closed half space $\{(x, z) \in \mathbb{R}^2 : x \geq 0\}$, and observe that, in particular, this quotient is not a manifold in the sense discussed above (it is a *manifold with boundary*).

Lecture 5.

Tangent Vectors

Note. From now on we consider C^∞-differentiable manifolds and C^∞-differentiable functions on them.

For a submanifold $\mathcal{M}$ ($\dim \mathcal{M} = m$) of Euclidean space E^n ($n > m$) the tangent space to $\mathcal{M}$ at a point $p \in \mathcal{M}$, denoted by $T_p\mathcal{M}$ 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 $\mathcal{M}$. (See Fig. 5.1.) That is, X is required to satisfy the conditions:

$$X \cdot g_\alpha = 0 : \quad \alpha = m+1, ..., n, \tag{5.1}$$

where the g_α are the functions

$$g_\alpha = x_\alpha - f_\alpha(x_1, \ldots, x_m) : \quad \alpha = m+1, \ldots, n\,. \tag{5.2}$$

$\mathcal{M}$ being represented in a neighborhood of p as the subset of E^n where these functions vanish.

Remark 5.1. For any continuously differentiable function f on E^n we denote

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

Question: How can one define the tangent space at a point to a manifold intrinsically?

We must define the concept of a tangent vector v at a point p to a manifold $\mathcal{M}$. A tangent vector at p is associated to a continuously differentiable parametrized curve through p.

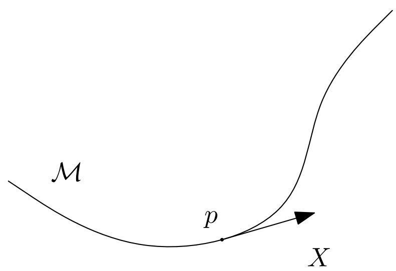


Figure 5.1.: Tangent space to a submanifold of E^n.

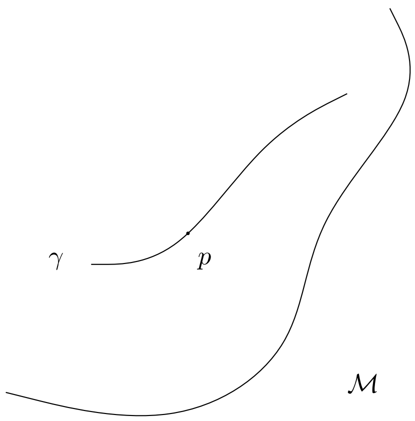


Figure 5.2.: Tangent vector associated to a curve γ.

Definition 5.1. A continuously differentiable parametrized curve γ through $p \in \mathcal{M}$ is a continuoulsy differentiable mapping $\gamma : I \to \mathcal{M}$, where I is an interval of $\mathbb{R}$ containing 0 and $\gamma(0) = p$.

Intuitively, I can be thought of as a time interval and γ 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 the directional derivative. Let f be a continuously differentiable (real-valued) function on $\mathcal{M}$. 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:

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

Consider now the space $\mathrm{C}^\infty(\mathcal{M})$ of all C^∞ functions on $\mathcal{M}$. This is a linear space: For $f, g \in \mathrm{C}^\infty(\mathcal{M})$ we have

$$\begin{aligned}\text{(Addition:)} \quad & (f+g)(q) = f(q) + g(q) : \quad \forall q \in \mathcal{M} \\ \text{(Scalar Multiplication:)} \quad & \alpha \in \mathbb{R}\,,\ (\alpha f)(q) = \alpha f(q) : \quad \forall q \in \mathcal{M}\end{aligned}$$

This is an infinite dimensional linear space. Moreover, $\mathrm{C}^\infty(\mathcal{M})$ is in fact a *commutative ring* w.r.t. the multiplication of functions:

$$\begin{gathered}\text{For any pair } f, g \in \mathrm{C}^\infty(\mathcal{M})\,,\ fg \text{ is defined by} \\ (fg)(q) = f(q)g(q) : \quad \forall q \in \mathcal{M}\end{gathered}$$

We *define* the vector v, the tangent vector v to γ at p to be the linear function on $C^\infty(\mathcal{M})$ given by:

$$v \cdot f = \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \gamma)\big|_{t=0}. \tag{5.3}$$

Note that v is a linear map $\mathrm{C}^\infty(\mathcal{M}) \to \mathbb{R}$. v is linear because

$$\begin{aligned} v \cdot (f_1 + f_2) &= \frac{\mathrm{d}}{\mathrm{d}t}(f_1 \circ \gamma + f_2 \circ \gamma)\big|_{t=0} = v \cdot f_1 + v \cdot f_2, \\ v \cdot (\alpha f) &= \frac{\mathrm{d}}{\mathrm{d}t}(\alpha\, f \circ \gamma)\big|_{t=0} = \alpha\, v \cdot f. \end{aligned} \tag{5.4}$$

Consider now the commutative ring structure of $\mathrm{C}^\infty(\mathcal{M})$:

$$\begin{aligned} v \cdot (fg) &= \frac{\mathrm{d}}{\mathrm{d}t}(fg) \circ \gamma\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\Big((f \circ \gamma)(g \circ \gamma)\Big)\Big|_{t=0} \\ &= (g \circ \gamma)(0)\frac{\mathrm{d}}{\mathrm{d}t}(f \circ \gamma)\Big|_{t=0} + (f \circ \gamma)(0)\frac{\mathrm{d}}{\mathrm{d}t}(g \circ \gamma)\Big|_{t=0} \\ &= g(p)\, v \cdot f + f(p)\, v \cdot g\,, \end{aligned} \tag{5.5}$$

which is the *Leibniz rule*.

Remark 5.2. Two different curves through p may have the same tangent vector at p.

The above motivates the following definition:

Definition 5.2. A *tangent vector* v at $p \in \mathcal{M}$ is a linear (real valued) function on $\mathrm{C}^\infty(\mathcal{M})$ satisfying the Leibniz rule:

$$v \cdot (fg) = g(p)\, v \cdot f + f(p)\, v \cdot g \quad : \text{for every pair } f, g \in \mathrm{C}^\infty(\mathcal{M})\,.$$

Basic properties:

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

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

Combining the two properties we obtain:

$v \cdot f = 0$ if f is constant in a neighborhood of p.

Proof of (1). We can assume that $\alpha \neq 0$. In fact, again by linearity we can assume $\alpha = 1$. Then $f^2 = f$, so by the Leibniz rule: $v \cdot f = v \cdot f^2 = 2f(p)\, v \cdot f = 2\, v \cdot f$ hence $v \cdot f = 0$. □

To prove (2), we begin by constructing a *bump function* on $\mathbb{R}^n$. A bump function on $\mathbb{R}^n$ is a C^∞ function ρ on $\mathbb{R}^n$, which is spherically symmetric, that is

$$\rho(x') = \rho(x) \qquad \text{if } |x'| = |x| \tag{5.6}$$

Also ρ is non-increasing with $|x|$, that is $\rho(x') \leq \rho(x)$ if $|x'| > |x|$. Finally,

$$\begin{aligned} &\rho(0) = 1 \\ &\rho(x) = 0 \text{ if } x \notin B_1(0)\,. \end{aligned} \tag{5.7}$$

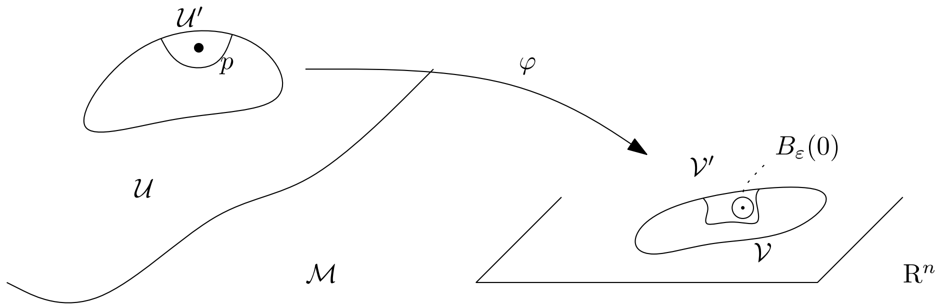


Figure 5.3.: Chart used in the proof of the basic properties of tangent vectors.

(It follows that the range of ρ is $[0,1]$.) An example of a bump function is given by:

$$\rho(x) = \begin{cases} \exp\Big(1 - \frac{1}{1-|x|}\Big) & , \text{ for } |x| < 1 \\ 0 & , \text{ for } |x| \geq 1 \,. \end{cases} \tag{5.8}$$

We now set, for $\varepsilon > 0$,

$$\rho_\varepsilon(x) = \rho(\frac{x}{\varepsilon}) \,, \tag{5.9}$$

then ρ_ε is a bump function with $B_1(0)$ replaced by $B_\varepsilon(0)$.

Proof of (2). Given a C^∞ function f on $\mathcal{M}$ which vanishes in a neighborhood W of p, take a chart $(\mathcal{U},\varphi)$, with $p \in \mathcal{U}$. Then $\mathcal{U}' = \mathcal{U} \cap W$ is a neighborhood of p contained in $\mathcal{U}$ where f vanishes. Denote by $\mathcal{V} = \varphi(\mathcal{U})$, $\mathcal{V}' = \varphi(\mathcal{U}')$. We assume (by translation) that $\varphi(p) = 0$. Since $\mathcal{V}'$ is open, there is an $\varepsilon > 0$ such that $B_\varepsilon(0) \subset \mathcal{V}'$. (C.f. Fig.5.3.) We then define a C^∞ function g on $\mathcal{M}$ by:

$$g(q) = \begin{cases} (\rho_\varepsilon \circ \varphi)(q) & , \text{ if } q \in \mathcal{U}' \\ 0 & , \text{ if } q \notin \mathcal{U}' \,. \end{cases}$$

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

$$0 = v \cdot (fg) = f(p)\, v \cdot g + g(p)\, v \cdot f = v \cdot f \,.$$

□

Problems

1. Let $(\mathbb{R},\varphi)$ and $(\mathbb{R},\psi)$ be $\mathbb{R}$ with two differentiable structures given by the charts $\varphi, \psi : \mathbb{R} \to \mathbb{R}$, $\varphi(x) = x$, $\psi(x) = x^3$. Prove that if a function $f : \mathbb{R} \to \mathbb{R}$ is differentiable with respect to the structure given by ψ then it is also differentiable with respect to the other structure, and for any tangent vector X to $(\mathbb{R},\varphi)$ at 0, we have $X(f) = 0$.

2. Let $\mathcal{M} \subset \mathbb{R}^{n+k}$ be an n-dimensional submanifold. Show that for any abstract tangent vector $X \in \mathrm{T}_p\mathcal{M}$, $p \in \mathcal{M}$, there is a curve $\sigma : (-\epsilon, \epsilon) \to \mathcal{M}$ having the following properties: 1. $\sigma(0) = p$, 2. for every function $f \in \mathrm{C}^\infty$,

$$X(p)(f) = \frac{\mathrm{d}(f \circ \sigma)(t)}{\mathrm{d}t}|_{t=0} \,.$$

Lecture 6.

Tangent Space

Let $\mathcal{M}$ be a differentiable manifold of dimension n.

Definition 6.1. The space of all tangent vectors at $p \in \mathcal{M}$ is called the *tangent space* at p to $\mathcal{M}$, denoted $\mathrm{T}_p\mathcal{M}$.

We shall show that $\mathrm{T}_p\mathcal{M}$ is a n-dimensional vector space. First, $\mathrm{T}_p\mathcal{M}$ is a vector space:

$$\begin{aligned} &\text{(Vector addition:)} \quad (v_1 + v_2) \cdot f = v_1 \cdot f + v_2 \cdot f \quad \forall f \in \mathrm{C}^\infty(\mathcal{M}) \\ &\text{(Scalar multiplication:)} \quad (\alpha v) \cdot f = \alpha\, v \cdot f \quad \forall \alpha \in \mathbb{R}\,, f \in \mathrm{C}^\infty(\mathcal{M}) \end{aligned} \tag{6.1}$$

Let $p \in \mathcal{U}, \mathcal{U}$ the domain of a chart $(\mathcal{U}, \varphi)$. (Refer to Fig. 6.1.) By suitable translation we may assume that $\varphi(p) = 0$: the origin in $\mathbb{R}^n$. Consider the coordinate lines $a_i : i = 1, \ldots, n$ in $\mathbb{R}^n$:

$$a_i : \mathbb{R} \longrightarrow \mathbb{R}, \quad t \mapsto (0, \ldots, \overset{i^{\text{th}}}{t}, \ldots, 0)$$

Consider then the curves $K_i = \varphi^{-1} \circ a_i$ in $\mathcal{M}$, through p. Let l_i be the tangent vector to K_i at p:

$$l_i \cdot f = \frac{\mathrm{d}}{\mathrm{d}t} f \circ K_i \Big|_{t=0} \quad : \forall f \in \mathrm{C}^\infty(\mathcal{M})$$

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

$$f \circ K_i = \tilde{f} \circ a_i = \tilde{f}(0, \ldots, \overset{i^{\text{th}}}{t}, \ldots, 0)\,,$$

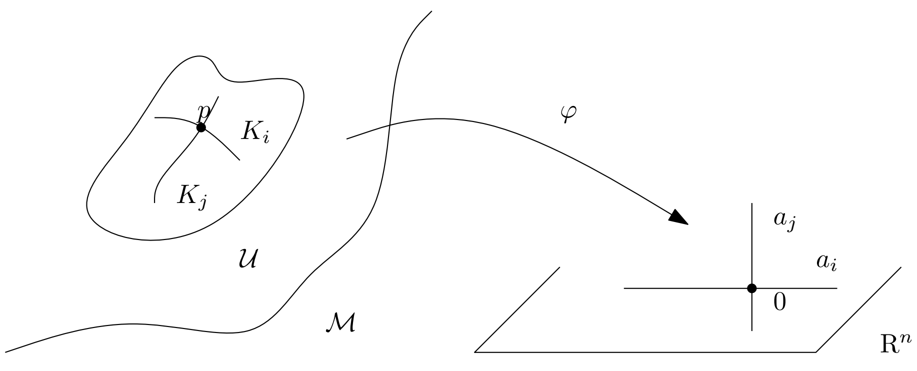


Figure 6.1.: Construction of the coordinate vectorfields.

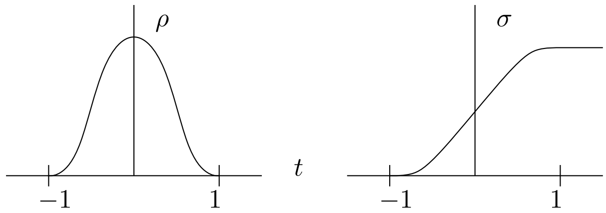


Figure 6.2.: Bump and cutoff functions.

so

$$l_i \cdot f = \frac{\partial \tilde{f}}{\partial x^i}\Big|_0 . \tag{6.2}$$

We denote

$$l_i = \frac{\partial}{\partial x^i}\Big|_p \qquad : i = 1, \ldots, n. \tag{6.3}$$

We shall now show that these vectors form a basis for $\mathrm{T}_p\mathcal{M}$. This implies that $\mathrm{T}_p\mathcal{M}$ is n-dimensional.

Digression on cutoff functions. Let us define on $\mathbb{R}$ the bump function:

$$\rho(t) = \begin{cases} \exp\Big(1 - \frac{1}{1-t^2}\Big) & , \ t \in (-1,1) \\ 0 & , \ t \notin (-1,1) \,. \end{cases} \tag{6.4}$$

ρ is a C^∞ function. We then define the C^∞ function σ by

$$\sigma(t) = \begin{cases} 0 & , \ t \leq 1 \\ \frac{\int_{-1}^{t} \rho(t')\mathrm{d}t'}{\int_{-1}^{1} \rho(t')\mathrm{d}t'} & , \ -1 < t < 1 \\ 1 & , \ t \geq 1 \end{cases} \tag{6.5}$$

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

$$\sigma_{a,b}(t) = \sigma\Big(\frac{t-b}{a}\Big) \tag{6.6}$$

Then $\sigma_{a,b}$ is a non-decreasing C^∞ function with range $[0,1]$,

$$\begin{aligned} \sigma_{a,b}(t) &= 0 \qquad \text{for } t \leq b - a \,, \\ \sigma_{a,b}(t) &= 1 \qquad \text{for } t \geq b + a \,, \end{aligned}$$

Then, given $\varepsilon_1, \varepsilon_2 > 0$, $\varepsilon_2 > \varepsilon_1$, we define on $\mathbb{R}^n$ the C^∞ real function $\eta_{\varepsilon_1,\varepsilon_2}$ by:

$$\eta_{\varepsilon_1,\varepsilon_2}(x) = 1 - \sigma_{a,b}(|x|) \qquad , \ x \in \mathbb{R}^n \tag{6.7}$$

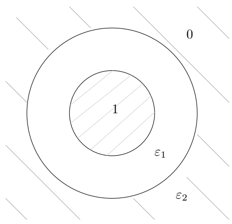


Figure 6.3.: Cutoff function on $\mathbb{R}^n$.

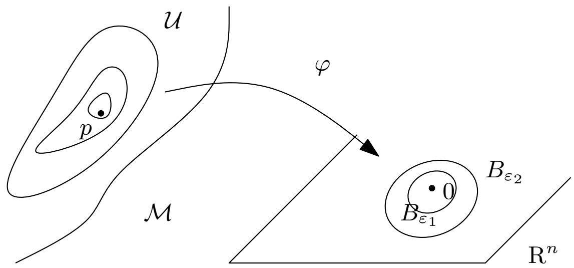


Figure 6.4.: Cutoffs of the coordinate functions.

where a, b are defined by $b - a = \varepsilon_1$, $b + a = \varepsilon_2$. η is spherically symmetric, non-increasing in $|x|$, and 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 $\mathbb{R}^n$:

$$g^i = \eta\, x^i \quad : i = 1, \ldots, n \tag{6.8}$$

These are C^∞ functions with support in $B_{\varepsilon_2}(0)$, agreeing with the coordinate functions in $B_{\varepsilon_1}(0)$. Consider now $g^i \circ \varphi$, functions defined on $\mathcal{U}$, the domain of a chart. $g^i \circ \varphi$ vanishes on $\mathcal{U} \setminus \varphi^{-1}(B_{\varepsilon_2}(0))$ and $g^i \circ \varphi = \varphi^i$ (the i^{th} component of φ, a real function on $\mathcal{U}$) on $\varphi^{-1}(B_{\varepsilon_1}(0))$. We extend the $g^i \circ \varphi$ by 0 outside $\mathcal{U}$, defining in this way the functions $\overline{g}^i$ on $\mathcal{M}$. That is

$$\overline{g}^i = \begin{cases} g^i \circ \varphi & : \text{on } \mathcal{U} \\ 0 & : \text{on } \mathcal{M} \setminus \mathcal{U} \end{cases} \tag{6.9}$$

These are C^∞ functions on $\mathcal{M}$. Since $\overline{g}^i \in C^\infty(\mathcal{M})$ we can apply any given tangent vector to them, in particular a given $v \in T_p\mathcal{M}$, to obtain the real numbers

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

the *components* of v in the chart $(\mathcal{U}, \varphi)$.

Proposition 6.1. *We have*

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

Thus the vectors $\left(\frac{\partial}{\partial x^i}\Big|_p \;: i = 1, \ldots, n\right)$ *form a basis for* $\mathrm{T}_p\mathcal{M}$*, and* $\mathrm{T}_p\mathcal{M}$ *is* n*-dimensional.*

Proof. Consider an arbitrary $f \in \mathrm{C}^\infty(\mathcal{M})$. First, since v vanishes on constant functions, we may substract a suitable constant from f to achieve $f(p) = 0$. Then $\tilde{f} = f \circ \varphi^{-1}$, the representation of f in the chart $(\mathcal{U}, \varphi)$ is a C^∞ function on $\mathbb{R}^n$ vanishing at 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

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

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

$$\tilde{h}_i(x) \longrightarrow \frac{\partial \tilde{f}}{\partial x^i}\Big|_0 \quad , \text{ as } x \to 0\,.$$

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

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

We now define

$$\tilde{h}'_i = \eta\,\tilde{h}_i\,.$$

Then the $\tilde{h}'_i$ agree with the $\tilde{h}_i$ on $B_{\varepsilon_1}(0)$ and vanish in $\mathbb{R}^n \setminus B_{\varepsilon_2}(0)$. We then define C^∞ functions h_i on $\mathcal{M}$ by:

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

We then define a new C^∞ function f' on $\mathcal{M}$ by:

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

f' vanishes on $\mathcal{M} \setminus \mathcal{U}$ and agrees with f on $\varphi^{-1}(B_{\varepsilon_1}(0)) \subset \mathcal{U}$. This is because for $q \in \varphi^{-1}(B_{\varepsilon_1}(0))$ we have $\varphi(q) = x \in B_{\varepsilon_1}(0)$, hence $\eta(x) = 1$, 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 $\varepsilon^{-1}(B_{\varphi_1}(0))$, a neighborhood of q in $\mathcal{M}$. By the basic property (1) of tangent vectors at p, $v \cdot (f - f') = 0$. That is $v \cdot f = v \cdot f'$. Now,

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

By Leibniz rule,

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

But $\overline{g}^i(p) = x^i|_0 = 0$, so the first term in the curly brackets vanishes, while

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

and $v \cdot \overline{g}^i = v^i$: the components of v in the chart φ. We conclude that

$$v \cdot f = v \cdot f' = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i}\big|_p \cdot f$$

and the proposition is proved. □

Problems

1. Given a submanifold $\mathcal{M} \subset \mathbb{R}^{n+k}$ of dimension n, we can express it locally as a the graph of a function $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^k$. Therefore $\mathcal{M}$ is locally the zero set of the function
$$F(x^1, \ldots, x^{n+k}) = (x^{n+1}, \ldots, x^{n+k}) - f(x^1, \ldots, x^n) .$$
Then the tangent space at $p = ((x^1, \ldots, x^n), f(x^1, \ldots, x^n))$ can also be seen as the zero space of $\mathrm{d}F(p)$, i.e.
$$\mathrm{T}_p\mathcal{M} = \big\{ v \in \mathbb{R}^{n+k} \ : \ \mathrm{d}F(p) \cdot v = 0 \big\}.$$
Show that this definition of tangent space is equivalent to the intrinsic one, meaning that the spaces which are obtained are canonically isomorphic (they do not depend on any choice of chart).

2. Given a manifold $\mathcal{M}$ and a tangent vector $v \in \mathrm{T}_p\mathcal{M}$, we can fix a chart $\phi : \mathcal{V} \subset \mathcal{M} \to \mathbb{R}^n$ at p and express v as
$$v = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i} .$$
Show how the coefficients v^i change if we choose another chart $\psi : \mathcal{V} \to \mathbb{R}^n$.

Lecture 7.
The Tangent Bundle

7.1. Definition.

The *tangent bundle* $\mathrm{T}\mathcal{M}$ of a differentiable manifold $\mathcal{M}$ is

$$\mathrm{T}\mathcal{M} = \bigcup_{p\in\mathcal{M}} \mathrm{T}_p\mathcal{M}\,. \tag{7.1}$$

We shall show that $\mathrm{T}\mathcal{M}$ is a differentiable manifold of dimension $2n$. Here $\dim\mathcal{M} = n$. In fact, $\mathrm{T}\mathcal{M}$ is a differentiable *vector bundle* over $\mathcal{M}$.

Let

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

be an atlas for $\mathcal{M}$. At each $p \in \mathcal{U}_\alpha$ the

$$\Big(\frac{\partial}{\partial x^i}\Big|_p : \quad i = 1, \dots, n\Big)\,,$$

defined by φ_α, constitute a basis for $\mathrm{T}_p\mathcal{M}$: For any $v \in \mathrm{T}_p\mathcal{M}$ there is a unique n-tuplet of real numbers $(v^1, \dots, v^n)$, the components of v in the chart, such that

$$v = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i}\Big|_p\,. \tag{7.2}$$

The $(v^i : i = 1, \dots, n)$ are *linear* coordinates in the vector space $\mathrm{T}_p\mathcal{M}$. This means that a straight line

$$t \mapsto a + tb\,, \quad a, b \in \mathrm{T}_p\mathcal{M}$$

in $\mathrm{T}_p\mathcal{M}$ is represented in these coordinates by the linear equations

$$v^i = a^i + tb^i\,, \qquad i = 1, \dots, n\,.$$

Let us write

$$v^i = \chi^i_{\alpha,p} \cdot v\,, \tag{7.3}$$

then $\chi_{\alpha,p} = (\chi^1_{\alpha,p}, \dots, \chi^n_{\alpha,p})$ is a linear isomorphism between $\mathrm{T}_p\mathcal{M}$ and $\mathbb{R}^n$. $\chi_{\alpha,p}$ is associated to φ_α and each $p \in \mathcal{U}_\alpha$:

$$\chi_{\alpha,p} : \mathrm{T}_p\mathcal{M} \longrightarrow \mathbb{R}^n\,. \tag{7.4}$$

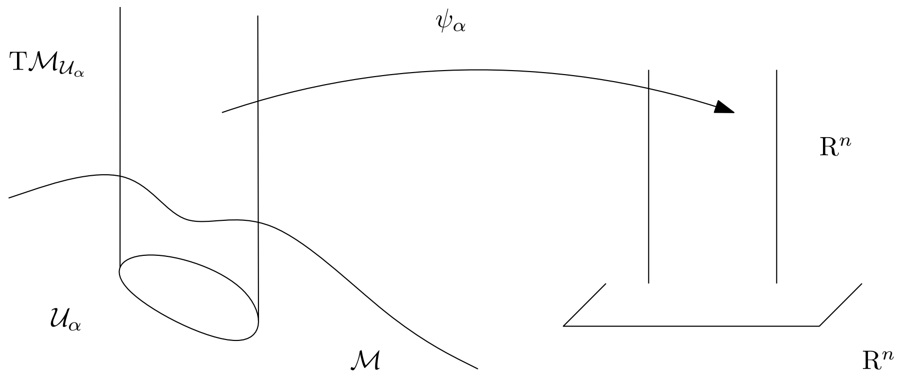


Figure 7.1.: Chart ψ_α for $\mathrm{T}\mathcal{M}|_{\mathcal{U}_\alpha}$.

Consider now

$$\mathrm{T}\mathcal{M}\Big|_{\mathcal{U}_\alpha} = \bigcup_{p\in\mathcal{U}_\alpha} \mathrm{T}_p\mathcal{M}\,.$$

On $\mathrm{T}\mathcal{M}|_{\mathcal{U}_\alpha}$ we define a chart ψ_α by:

$$\psi_\alpha(v) = \Big(\psi_\alpha(p), \chi_{\alpha,p}\cdot v\Big) = \Big((x^1,\ldots,x^n),(v^1,\ldots,v^n)\Big) \;:\; v\in\mathrm{T}_p\mathcal{M}, p\in\mathcal{U}_\alpha \tag{7.5}$$

Then ψ_α maps $\mathrm{T}\mathcal{M}|_{\mathcal{U}_\alpha}$ onto $\mathbb{R}^{2n}$; see Fig. 7.1. The collection

$$\Big\{(\mathrm{T}\mathcal{M}\big|_{\mathcal{U}_\alpha}, \psi_\alpha) \;:\; \alpha\in I\Big\} \tag{7.6}$$

forms an atlas for $\mathrm{T}\mathcal{M}$. To show this, consider $p\in\mathcal{U}_\alpha\cap\mathcal{U}_\beta$. Then a tangent vector $v\in\mathrm{T}_p\mathcal{M}$ has two sets of components:

$$\begin{aligned} v^i_\alpha &= \chi^i_{\alpha,p}\cdot v\,, \quad \text{in the chart } \varphi_\alpha\,,\\ \text{and } v^i_\beta &= \chi^i_{\beta,p}\cdot v\,, \quad \text{in the chart } \varphi_\beta\,. \end{aligned}$$

It holds:

$$v^i_\beta = \sum_{j=1}^n M^i_j v^j_\alpha \;:\; i=1,\ldots,n\,. \tag{7.7}$$

In fact, denoting

$$f = \varphi_\beta\circ\varphi_\alpha^{-1} : \varphi_\alpha(\mathcal{U}_\alpha\cap U_\beta) \longrightarrow \varphi_\beta(\mathcal{U}_\alpha\cap\mathcal{U}_\beta)\,, \tag{7.8}$$

a continuously differentiable mapping from an open set in $\mathbb{R}^n$ to an open set in $\mathbb{R}^n$. Let $x^i_\alpha = \varphi^i_\alpha(q)$ be the coordinates of a point $q\in\mathcal{U}_\alpha\cap\mathcal{U}_\beta$ w.r.t. the chart φ_α, and $x^i_\beta = \varphi^i_\beta(q)$ the coordinates of q w.r.t. φ_β. Then

$$x^i_\beta = f^i(x^i_\alpha,\ldots,x^n_\alpha) \;\; i=1,\ldots,n$$

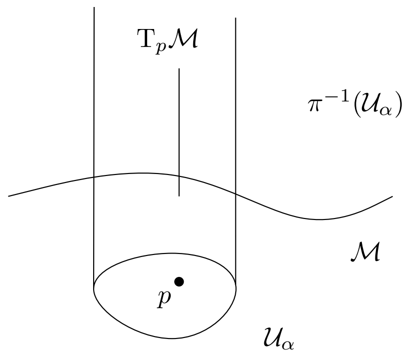


Figure 7.2.: The structure of the tangent bundle.

and we have:

$$M^i_j = \frac{\partial f^i}{\partial x^j_\alpha}\Big|_{\varphi_\alpha(p)} . \tag{7.9}$$

These are continuously differentiable. This shows that (7.6) does satisfy the requirements of an atlas.

7.2. The structure of the tangent bundle.

We describe structural properties of the tangent bundle which are shared more generally by *vector bundles*.

First there is the *projection map*

$$\pi : \mathrm{T}\mathcal{M} \longrightarrow \mathcal{M} , \quad \text{by } \pi(v) = p , \text{ if } v \in \mathrm{T}_p\mathcal{M} ; \tag{7.10}$$

π tells us at which point a tangent vector is attached. Then

$$\mathrm{T}_p\mathcal{M} = \pi^{-1}(p) \tag{7.11}$$

(the set of all vectors attached at p), and

$$\mathrm{T}\mathcal{M}\Big|_{\mathcal{U}_\alpha} = \pi^{-1}(\mathcal{U}_\alpha) ; \tag{7.12}$$

see Fig. 7.2.

The domains $\mathcal{U}_\alpha$ give a covering of $\mathcal{M}$ by open sets $\mathcal{U}_\alpha$,

$$\mathcal{M} = \bigcup_{\alpha \in I} \mathcal{U}_\alpha$$

which have the following property: To $\mathcal{U}_\alpha$ is associated a diffeomorphism

$$\omega_\alpha : \pi^{-1}(\mathcal{U}_\alpha) \longrightarrow \mathcal{U}_\alpha \times \mathbb{R}^n \tag{7.13}$$

such that $\omega_\alpha|_{\mathrm{T}_p\mathcal{M}}$ is a linear isomorphism of $\mathrm{T}_p\mathcal{M}$ on $\mathbb{R}^n$, for each $p \in \mathcal{U}_\alpha$, and we have:

$$\omega_\alpha(v) = \big(p = \pi(v), \omega_{\alpha,p} \cdot v\big) \tag{7.14}$$

where $\omega_{\alpha,p} = \omega_\alpha|_{\mathrm{T}_p\mathcal{M}}$.

We have a basis in $\mathrm{T}_p\mathcal{M}$ given by

$$\frac{\partial}{\partial x^i}\big|_p = \omega_{\alpha,p}^{-1} \cdot e_i \,,$$

where $(e_i \ : \ i = 1, \dots, n)$ the standard basis of $\mathbb{R}^n$, i.e. $e_i = (0, \dots, \overset{i^{\text{th}}}{1}, \dots, 0)$.

This can be generalized to the situation when the open sets $\mathcal{U}_\alpha$ are not necessarily the domains of charts, but only an open cover of $\mathcal{M}$, with the above property, namely that to each $\mathcal{U}_\alpha$ is associated a diffeomorphism

$$\omega_\alpha : \pi^{-1}(\mathcal{U}_\alpha) \longrightarrow \mathcal{U}_\alpha \times \mathbb{R}^n \,, \quad \omega_\alpha(v) = \big(p = \pi(v), \omega_{\alpha,p} \cdot v\big) \,, \tag{7.15}$$

where $\omega_{\alpha,p}$ is a linear isomorphism of $\mathrm{T}_p\mathcal{M}$ on $\mathbb{R}^n$, for each $p \in \mathcal{U}_\alpha$.

In this case for each $p \in \mathcal{U}_\alpha$ the vectors

$$X_{\alpha,i}(p) = \omega_{\alpha,p}^{-1} \cdot e_i \qquad i = 1, \dots, n \tag{7.16}$$

still form a basis for $\mathrm{T}_p\mathcal{M}$.

The tangent bundle $\mathrm{T}\mathcal{M}$ of a manifold $\mathcal{M}$ is called *trivial* if there is a *diffeomorphism*

$$\omega \ : \ \mathrm{T}\mathcal{M} \longrightarrow \mathcal{M} \times \mathbb{R}^n \,, \tag{7.17}$$

such that

$$\omega(v) = (p, \omega_p \cdot v) \,, \qquad v \in \mathrm{T}_p\mathcal{M} \,, \tag{7.18}$$

where ω_p is a linear isomorphism of $\mathrm{T}_p\mathcal{M}$ onto $\mathbb{R}^n$.

Example 7.1. An atlas for $\mathbb{S}^1$ consists of at least two charts. However in the case $\mathcal{M} = \mathbb{S}^1$ we can take as a single domain $\mathcal{U}_\alpha$ the whole of $\mathbb{S}^1$, and find a diffeomorphism

$$\omega \ : \ \mathrm{T}\mathbb{S}^1 \longrightarrow \mathbb{S}^1 \times \mathbb{R} \,. \tag{7.19}$$

This means that the tangent bundle of the circle is *trivial.*

Remark 7.1. This is *not true* for $\mathbb{S}^2$ whose tangent bundle $\mathrm{T}\mathcal{M}$ is *nontrivial.*

Problems

1. Let $\mathbb{S}^1$ be the unit circle. Show that its tangent bundle $\mathrm{T}\mathbb{S}^1$ is trivial, meaning that it is diffeomorphic to $\mathbb{S}^1 \times \mathbb{R}$.

2. Show that the (non-trivial) tangent bundle $\mathrm{T}\mathbb{S}^2$ can be embedded into $\mathbb{R}^6$. More generally one can embed $\mathrm{T}\mathbb{S}^n$ into $\mathbb{R}^{2n+2}$.

Part II.

Vector Bundles

Lecture 8.
Vector bundles

8.1. Definition

Definition 8.1. A differentiable *vector bundle* $\mathcal{B}$ over a differentiable manifold $\mathcal{M}$ is a differentiable manifold with the following structure:

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

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

3. There is an open covering $\{\mathcal{U}_\alpha : \alpha \in I\}$ of $\mathcal{M}$, such that to each $\mathcal{U}_\alpha$ is associated a diffeomorphism
$$\omega_\alpha = \pi^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times V \tag{8.1}$$
where $\omega_\alpha(v) = (p, \omega_{\alpha,p} \cdot v)$, $p = \pi(v)$, and for each $p \in \mathcal{U}_\alpha$, $\omega_{\alpha,p}$ is a linear isomorphism of $\mathcal{B}_p$ onto V.

Remark 8.1. The model vector space V may be real or complex. The vector bundle $\mathcal{B}$ is then accordingly called a *real* or *complex* vector bundle over $\mathcal{M}$. In the real case, if $\dim V = m$, then $\dim \mathcal{B}_p = m$, for each $p \in \mathcal{M}$, and $\dim \mathcal{B} = m + n$, where $\dim M = n$.

Remark 8.2. The case $V = \mathbb{C}$ is realised in quantum mechanics and called the *complex line bundle*.

8.2. Möbius bundle

The Möbius bundle is the simplest *non-trivial* example of a line bundle. Here
$$\mathcal{M} = \mathbb{S}^1 \qquad V = \mathbb{R}\,. \tag{8.2}$$

We begin with the circle $\mathbb{S}^1$ which we view as the quotient
$$\mathbb{S}^1 = (-\tfrac{1}{4}, \tfrac{5}{4})/\sim \tag{8.3}$$
where $\sim$ is the equivalence relation
$$\sim = \Delta \cup \big\{(x, x+1) : x \in (-\tfrac{1}{4}, \tfrac{1}{4})\big\} \cup \big\{(x, x-1) : x \in (\tfrac{3}{4}, \tfrac{5}{4})\big\} \tag{8.4}$$

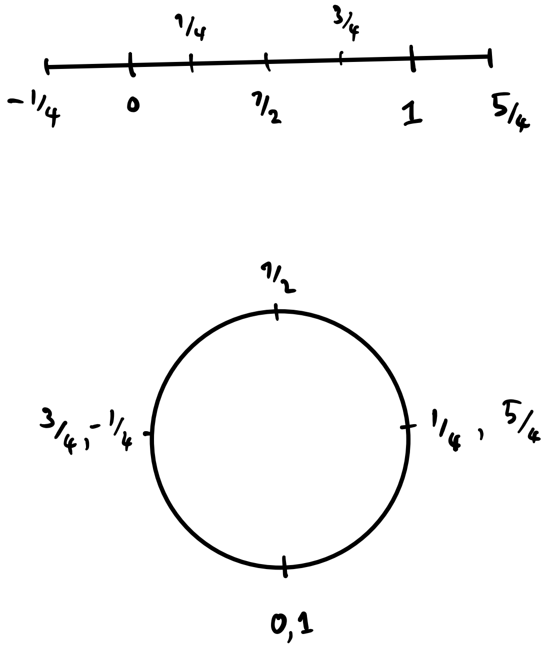


Figure 8.1.: The circle $\mathbb{S}^1$ as a quotient.

The equivalence classes are

$$C_x = \begin{cases} \{x\} & x \in [\frac{1}{4}, \frac{3}{4}] \\ \{(x, x+1)\} & x \in (-\frac{1}{4}, \frac{1}{4}) \\ \{(x, x-1)\} & x \in (\frac{3}{4}, \frac{5}{4}) \end{cases} \tag{8.5}$$

We have a cover $\{\mathcal{U}_-, \mathcal{U}_+\}$ of $\mathbb{S}^1$ where

$$\mathcal{U}_- = \{C_x : x \in (-\tfrac{1}{4}, \tfrac{3}{4})\} \tag{8.6}$$

$$\mathcal{U}_+ = \{C_x : x \in (\tfrac{1}{4}, \tfrac{5}{4})\}\,. \tag{8.7}$$

We define φ_-, a homeomorphism of $\mathcal{U}_-$ onto an open interval in $\mathbb{R}$, similarly φ_+, a homeomorphism of $\mathcal{U}_+$ onto an open interval in $\mathbb{R}$, by

$$\varphi_-(C_x) = x \ : \tag{8.8}$$

if x is the unique representative of C_x lying in the interval $(-\frac{1}{4}, \frac{3}{4})$, and

$$\varphi_+(C_x) = x \ : \tag{8.9}$$

if x is the unique representative of C_x lying in the interval $(\frac{1}{4}, \frac{5}{4})$. Then

$$\varphi_-(\mathcal{U}_-) = (-\tfrac{1}{4}, \tfrac{3}{4}) \tag{8.10}$$

$$\varphi_+(\mathcal{U}_+) = (\tfrac{1}{4}, \tfrac{5}{4})\,. \tag{8.11}$$

The intersection $\mathcal{U}_- \cap \mathcal{U}_+$ consists of two components I_- and I_+, where

$$I_+ = \{C_x : x \in (\tfrac{1}{4}, \tfrac{3}{4})\} = \{\{x\} : x \in (\tfrac{1}{4}, \tfrac{3}{4})\} \tag{8.12}$$
$$I_- = \{C_x : x \in (-\tfrac{1}{4}, \tfrac{1}{4})\} = \{\{x, x+1\} : x \in (-\tfrac{1}{4}, \tfrac{1}{4})\} \tag{8.13}$$

and we have

$$\varphi_-(I_+) = (\tfrac{1}{4}, \tfrac{3}{4}) \tag{8.14}$$
$$\varphi_+(I_+) = (\tfrac{1}{4}, \tfrac{3}{4}) \tag{8.15}$$

and $\varphi_+ \circ \varphi_-^{-1}$ is the identity on $(\frac{1}{4}, \frac{3}{4})$. However,

$$\varphi_-(I_-) = (-\tfrac{1}{4}, \tfrac{1}{4}) \tag{8.16}$$
$$\varphi_+(I_-) = (\tfrac{3}{4}, \tfrac{5}{4}) \tag{8.17}$$

and $\varphi_+ \circ \varphi_-^{-1}$ is the mapping $x \mapsto x + 1$ on $(-\frac{1}{4}, \frac{1}{4})$.

Therefore

$$\mathcal{A} = \{(\mathcal{U}_+, \varphi_+), (\mathcal{U}_+, \varphi_+)\} \tag{8.18}$$

is an atlas for $\mathbb{S}^1$.

We now consider the quotient

$$\mathcal{B} = (-\tfrac{1}{4}, \tfrac{5}{4}) \times \mathbb{R} / \approx \tag{8.19}$$

where $\approx$ is the equivalence relation

$$\begin{aligned} \sim = \Delta &\cup \{\{(x,y),(x+1,-y)\} : x \in (-\tfrac{1}{4}, \tfrac{1}{4}), y \in \mathbb{R}\} \\ &\cup \{\{(x,y),(x-1,-y)\} : x \in (\tfrac{3}{4}, \tfrac{5}{4}), y \in \mathbb{R}\} \end{aligned} \tag{8.20}$$

This is the Möbius bundle. We can picture the Möbius band, which is obtained by replacing $\mathbb{R}$ by $(-1, 1)$, see Figure 8.2.

Let us first show that $\mathcal{B}$ is a differentiable manifold.

Let us denote by $\mathcal{E}_{(x,y)}$ the equivalence class of (x, y). We have

$$\mathcal{E}_{(x,y)} = \begin{cases} \{(x,y)\} & x \in [\tfrac{1}{4}, \tfrac{3}{4}], y \in \mathbb{R} \\ \{(x,y),(x+1,-y)\} & x \in (-\tfrac{1}{4}, \tfrac{1}{4}), y \in \mathbb{R} \\ \{(x,y),(x-1,-y)\} & x \in (\tfrac{3}{4}, \tfrac{5}{4}), y \in \mathbb{R} \end{cases} \tag{8.21}$$

Then

$$\mathcal{V}_- = \{\mathcal{E}_{(x,y)} : x \in (-\tfrac{1}{4}, \tfrac{3}{4}), y \in \mathbb{R}\} \tag{8.22}$$
$$\mathcal{V}_+ = \{\mathcal{E}_{(x,y)} : x \in (\tfrac{1}{4}, \tfrac{5}{4}), y \in \mathbb{R}\}\,. \tag{8.23}$$

provides an open cover for $\mathcal{B}$.

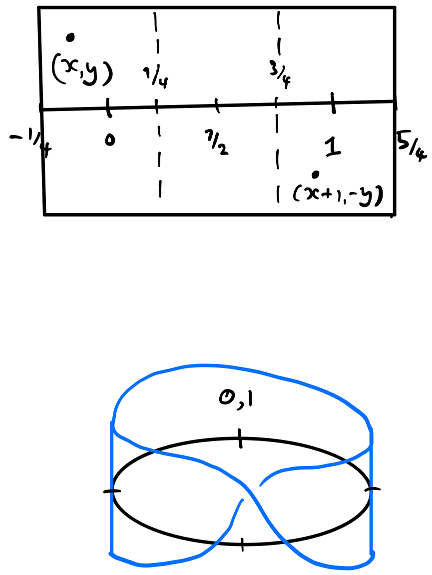


Figure 8.2.: The Möbius band.

Since for $\mathcal{E}_{(x,y)} \in \mathcal{V}_-$ there is a unique representative in $\mathcal{E}_{(x,y)}$ such that its first coordinate belongs to $(-\frac{1}{4}, \frac{3}{4})$, and similarly for $\mathcal{E}_{(x,y)} \in \mathcal{V}_+$ there is a unique representative in $\mathcal{E}_{(x,y)}$ such that its first coordinate belongs to $(\frac{1}{4}, \frac{5}{4})$, we can use these unique representatives to define:

$$\psi_-(\mathcal{E}_{(x,y)}) = (x, y) \qquad x \in (-\tfrac{1}{4}, \tfrac{3}{4}) \tag{8.24}$$

$$\psi_+(\mathcal{E}_{(x,y)}) = (x, y) \qquad x \in (\tfrac{1}{4}, \tfrac{5}{4}) \tag{8.25}$$

Note that

$$\psi_-(\mathcal{V}_-) = (-\tfrac{1}{4}, \tfrac{3}{4}) \times \mathbb{R} \tag{8.26}$$

$$\psi_+(\mathcal{V}_+) = (\tfrac{1}{4}, \tfrac{5}{4}) \times \mathbb{R} \tag{8.27}$$

Moreover, consider $\mathcal{V}_- \cap \mathcal{V}_+$. This consists again of two components J_-, and J_+, where

$$J_+ = \{\mathcal{E}_{(x,y)} : x \in (\tfrac{1}{4}, \tfrac{3}{4}), y \in \mathbb{R}\} = \{\{(x, y)\} : x \in (\tfrac{1}{4}, \tfrac{3}{4}), y \in \mathbb{R}\} \tag{8.28}$$

$$J_- = \{\mathcal{E}_{(x,y)} : x \in (-\tfrac{1}{4}, \tfrac{1}{4}), y \in \mathbb{R}\} = \{\{(x, y), (x + 1, -y)\} : x \in (-\tfrac{1}{4}, \tfrac{1}{4}), y \in \mathbb{R}\} \tag{8.29}$$

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

$$\varphi_-(J_-) = (-\tfrac{1}{4}, \tfrac{1}{4}) \times \mathbb{R} \tag{8.30}$$

$$\varphi_+(J_-) = (\tfrac{3}{4}, \tfrac{5}{4}) \times \mathbb{R} \tag{8.31}$$

and $\psi_+ \circ \psi_-^{-1}$ on $(-\frac{1}{4}, \frac{1}{4}) \times \mathbb{R}$ is given by

$$(x, y) \mapsto (x + 1, -y)\,. \tag{8.32}$$

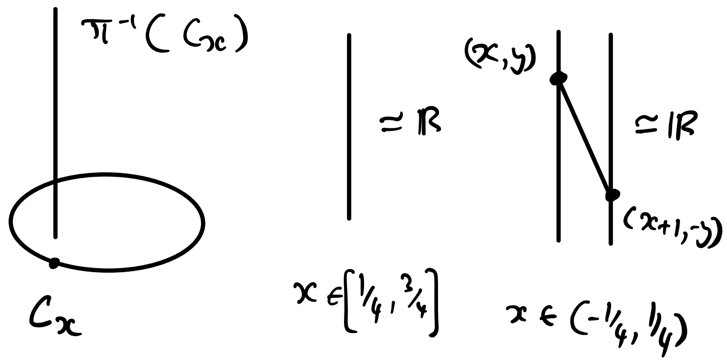


Figure 8.3.: The fibres of the Möbius bundle.

Thus both are differentiable, with differentiable inverses, and

$$\mathcal{A}_{\mathcal{B}} = \{(\mathcal{V}_+, \psi_+), (\mathcal{V}_+, \psi_+)\} \tag{8.33}$$

is an atlas for $\mathcal{B}$.

Next we want to show that $\mathcal{B}$ is in fact a line bundle over $\mathbb{S}^1$. It is nontrivial, meaning that $\mathcal{B}$ is *not* diffeomorphic to $\mathbb{S}^1 \times \mathbb{R}$.

We define the projection $\pi : \mathcal{B} \to \mathcal{M}$ by

$$\pi\big(\mathcal{E}_{(x,y)}\big) = C_x \,. \tag{8.34}$$

Then we have

$$\pi^{-1}(C_x) = \{\mathcal{E}_{(x,y)} : y \in \mathbb{R}\} = \begin{cases} \{\{(x,y)\} : y \in \mathbb{R}\}\,, & x \in [\frac{1}{4}, \frac{3}{4}] \\ \{\{(x,y), (x+1,-y)\} : y \in \mathbb{R}\}\,, & x \in (-\frac{1}{4}, \frac{1}{4}) \\ \{\{(x,y), (x-1,-y)\} : y \in \mathbb{R}\}\,, & x \in (\frac{3}{4}, \frac{5}{4}) \end{cases} \tag{8.35}$$

Each $\mathcal{B}_{C_x} = \pi^{-1}(C_x)$ is isomorphic to $V = \mathbb{R}$; see Figure 8.3.

Note that we have

$$\pi^{-1}(\mathcal{U}_-) = \mathcal{V}_- \,, \qquad \pi^{-1}(\mathcal{U}_+) = \mathcal{V}_+ \,, \tag{8.36}$$

and

$$\pi^{-1}(I_-) = J_- \,, \qquad \pi^{-1}(I_+) = J_+ \,. \tag{8.37}$$

Moreover we have diffeomorphisms

$$\omega_- : \mathcal{V}_- \to \mathcal{U}_- \times \mathbb{R} \,, \qquad \omega_+ : \mathcal{V}_+ \to \mathcal{U}_+ \times \mathbb{R} \,, \tag{8.38}$$

given by

$$\omega_-(\mathcal{E}_{(x,y)}) = (C_x, y) \,, \tag{8.39}$$

where (x, y) is the unique representative of $\mathcal{E}_{(x,y)}$ such that the first component lies in $(-\frac{1}{4}, \frac{3}{4})$, and

$$\omega_+(\mathcal{E}_{(x,y)}) = (C_x, y)\,, \tag{8.40}$$

where (x, y) is the unique representative of $\mathcal{E}_{(x,y)}$ such that the first component lies in $(\frac{1}{4}, \frac{5}{4})$.

Then

$$\omega_{-,C_x} = \omega_-|_{\pi^{-1}(C_x)} \tag{8.41}$$

is the linear isomorphism of $\pi^{-1}(C_x)$ onto $\mathbb{R}$, given by

$$\omega_{-,C_x} \cdot \mathcal{E}_{(x,y)} = y \qquad : x \in (-\tfrac{1}{4}, \tfrac{3}{4}) \tag{8.42}$$

and similarly

$$\omega_{+,C_x} \cdot \mathcal{E}_{(x,y)} = y \qquad : x \in (\tfrac{1}{4}, \tfrac{5}{4})\,. \tag{8.43}$$

While for $C_x \in I_+$,

$$\omega_{+,C_x} \circ \omega_{-,C_x}^{-1} : y \mapsto y \tag{8.44}$$

is the identity on $\mathbb{R}$, we see that for $C_x \in I_-$,

$$\omega_{+,C_x} \circ \omega_{-,C_x}^{-1} : y \mapsto -y \tag{8.45}$$

is a linear isomorphism of $\mathbb{R}$, which is orientation reversing.

Remark 8.3. The bundle $\mathcal{B}$ is *non-orientable.*

Problems

1. Consider a real vector bundle $\mathcal{B}$ over a manifold $\mathcal{M}$. For any $p \in \mathcal{M}$, let $\mathcal{B}_p^*$ be the dual space to $\mathcal{B}_p = \pi^{-1}(p)$. Define

$$\mathcal{B}^* = \bigcup_{p \in \mathcal{M}} \mathcal{B}_p^*\,. \tag{8.46}$$

   Show that $\mathcal{B}^*$ can be given the structure of a vector bundle.

2. Let $\mathcal{M}$ be an n-dimensional manifold. Give a characterisation of the topology of the tangent bundle $\mathrm{T}\mathcal{M}$, i.e. a characterisation of the open sets of $\mathrm{T}\mathcal{M}$.

Lecture 9.

Trivial and orientable bundles

Further Reading

(Spivak, *A Comprehensive Introduction to Differential Geometry*, Chapter 2)

9.1. Trivial bundles

In the previous lecture we have introduced the notion of a (differentiable) vector bundle $\mathcal{B}$ over a (differentiable) manifold $\mathcal{M}$.

The third condition in Definition 8.1 is also referred to as the *local triviality* condition:

(local triviality) For each $p \in \mathcal{M}$, there is a neighbourhood $\mathcal{U}$ of p, and a *homeomorphism*

$$\omega : \pi^{-1}(\mathcal{U}) \to \mathcal{U} \times V\,, \tag{9.1}$$

so that $\omega_q = \omega|_{\pi^{-1}(q)}$ is an isomorphism from $\mathcal{B}_q = \pi^{-1}(q)$ onto V for each $q \in \mathcal{U}$.

Given a topological manifold $\mathcal{M}$, we call $\mathcal{M} \times \mathbb{R}^n$, with the obvious structure of a vector bundle, the **trivial** (n-plane) bundle of $\mathcal{M}$.

We have a notion of *equivalence* of vector bundles over a manifold. Let $\mathcal{B}_1$, and $\mathcal{B}_2$ be vector bundles over $\mathcal{M}$, with projection $\pi_1 : \mathcal{B}_1 \to \mathcal{M}$, and $\pi_2 : \mathcal{B}_2 \to \mathcal{M}$ respectively. We say the two vector bundles are **equivalent** if there is a homeomorphism $h : \mathcal{B}_1 \to \mathcal{B}_2$ which takes each *fibre* $\pi_1^{-1}(p)$ isomorphically onto $\pi_2^{-1}(p)$. The map h is called an **equivalence**.

Note that the local triviality condition says that for each $p \in \mathcal{M}$, there is an equivalence of the vector bundle

$$\mathcal{B}|_{\mathcal{U}} = \bigcup_{q\in\mathcal{U}} \mathcal{B}_q \tag{9.2}$$

and the trivial bundle $\mathcal{U} \times V$, for some neighbourhood $\mathcal{U}$ of p. For this reason, we also call the map ω a *local trivialisation.*

More generally, a vector bundle $\mathcal{B}$ over $\mathcal{M}$ is trivial if it is equivalent to the trivial bundle over $\mathcal{M}$, and in that case there is an equivalence $\omega : \mathcal{B} \to \mathcal{M} \times \mathbb{R}^n$.

Example 9.1. An example of a *non-trival* bundle is the tangent bundle over $\mathbb{S}^2$. For if the tangent bundle of the sphere were trivial, we could use the homeomorphism

$$\omega : \mathrm{T}\mathbb{S}^2 \to \mathbb{S}^2 \times \mathbb{R}^2 \tag{9.3}$$

to construct a collection of tangent vectors, one at each point $p \in \mathbb{S}^2$:

$$v_p = \omega_p^{-1} \cdot (1, 0)\,. \tag{9.4}$$

These vectors are all *non-zero* (because ω_p is an isomorphism), and they vary continuously over the sphere (because ω is a homeomorphism). However, there is no continuous nowhere vanishing vectorfield on the sphere, by a theorem in topology. ("You cannot comb the hair on a sphere.")

9.2. Orientable bundles

Any structure on a vector space leads to a corresponding structure on a vector bundle. We will demonstrate this for the notion of **orientation** in a vector space.

Consider an n-dimensional vector space V. Any two bases $(v_1, \dots, v_n)$ and $(v'_1, \dots, v'_n)$ for V determine an isomorphism $A : V \to V$,

$$v'_i = \sum_{j=1}^{n} a_{ji} v_j\,. \tag{9.5}$$

We call these bases *equally oriented* if $\det A > 0$, and otherwise oppositely oriented if $\det A < 0$. This relation is an equivalence relation dividing the collection of bases into just two equivalence classes. A choice of either of these classes μ is called an *orientation* for V.

We refer to the orientation given by the equivalence class of the standard basis $(e_1, \dots, e_n)$ for $\mathbb{R}^n$ as the *standard orientation.* Also the trivial bundle of a manifold $\mathcal{M}$ has a "standard orientation", if each fibre has the standard orientation.

Let V and W be n-dimensional vector spaces, together with orientations μ, and ν, respectively. An isomorphism $A : V \to W$ is called *orientation preserving* if $(Av_1, Av_2, \dots, Av_n)$ has the orientation ν whenenver $(v_1, \dots, v_n)$ has the orientation μ.

Example 9.2. If $f : \mathcal{M} \times \mathbb{R}^n \to \mathcal{M} \times \mathbb{R}^n$ is an equivalence, and $\mathcal{M}$ is connected, then f is either orientation preserving or reversing, for if we define the functions a_{ij} on $\mathcal{M}$ by

$$f_p \cdot e_i = \sum_{j=1}^{n} a_{ji}(p) e_j\,, \tag{9.6}$$

then $\det(a_{ji}) : \mathcal{M} \to \mathbb{R}$ is continuous and never 0.

Definition 9.1. An *orientation* μ of a vector bundle $\mathcal{B}$ over $\mathcal{M}$ is a choice of orientations μ_p for $\mathcal{B}_p$ which satisfy the following *compatibility condition* for each open set $\mathcal{U} \subset \mathcal{M}$:

If $\omega : \pi^{-1}(\mathcal{U}) \to \mathcal{U} \times V$ is an equivalence, and $\mathcal{U} \times V$ has the standard orientation, then $\omega_q = \omega|_{\mathcal{B}_q}$ is either orientation preserving or orientation reversing for all $q \in \mathcal{U}$.

Remark 9.1. Note that if the compatibility condition is satisfied for one equivalence ω, and $\omega' : \pi^{-1}(\mathcal{U}) \to \mathcal{U} \times V$ is another equivalence, then ω' also satisifes the condition, because $\omega' \circ \omega^{-1} : \mathcal{U} \times V \to \mathcal{U} \times V$ is an equivalence. Therefore the orientations μ_p define an orientation μ of $\mathcal{B}$ if the compatibility condition holds for a collection of open sets $\mathcal{U}$ which cover $\mathcal{B}$.

A vector bundle is called **orientable** if it has an orientation, and **non-orientable** otherwise; and **oriented** bundle is a vector bundle together with an orientation μ.

Remark 9.2. For an oriented vector bundle $\mathcal{B}$, the isomorphisms $\omega_{p,q} : \mathcal{B}_q \to V$ of any local trivialisation ω_p, around $p \in \mathcal{U}_p \subset \mathcal{M}$, are either all orientation preserving or all orientation reversing for $q \in \mathcal{U}_p$.

Definition 9.2. A differentiable manifold $\mathcal{M}$ is called *orientable* (or *non-orientable*) if $\mathrm{T}\mathcal{M}$ is an orientable (non-orientable) bundle. An *oriented manifold* is a manifold $\mathcal{M}$ together with an orientation μ of $\mathrm{T}\mathcal{M}$.

Example 9.3. The sphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ is orientable. To see this we use that $\mathrm{T}_w\mathbb{S}^{n-1}$, for $w = (w^1, \ldots, w^n) \in \mathbb{S}^{n-1}$, is canonically isomorphic to the subspace of all vectors in $\mathbb{R}^n$ orthogonal to w,

$$\mathrm{T}_w\mathbb{S}^{n-1} \simeq \{v \in \mathbb{R}^n : v \cdot w = 0\} \,. \tag{9.7}$$

We can then introduce an orientation μ_w in $\mathrm{T}_w\mathbb{S}^{n-1}$, by defining that a basis $(v_1, \ldots, v_{n-1})$ is in μ_w if $(w, v_1, \ldots, v_n)$ has the standard orientation in $\mathrm{T}_w\mathbb{R}^n$. The orientation $\mu = \{\mu_w : w \in \mathbb{S}^{n-1}\}$ is the standard orientation of the $n-1$-sphere.

9.3. Trivial and non-trivial tangent bundles

All trivial tangent bundles are orientable, but not every orientable manifold has a trivial tangent bundle, as the example of the sphere $\mathbb{S}^2$ shows.

The tangent bundle of $\mathbb{S}^1$ is orientable and trivial. In fact, the same is true for any 3-dimensional orientable manifold.

Problems

1. Show that the Möbius bundle is *non-trivial* and *non-orientable*.
2. Prove that the torus $\mathbb{S}^1 \times \mathbb{S}^1$ is orientable.
3. Decide if the real projective space $\mathbb{RP}^2$ is orientable. What about $\mathbb{RP}^3$?

Lecture 10.

Sections

Definition 10.1. A continuously differentiable *section* of a vector bundle $\mathcal{B}$ over a differentiable manifold $\mathcal{M}$ is a continuously differentiable mapping $\sigma : \mathcal{M} \to \mathcal{B}$ such that

$$\pi \circ \sigma = \mathrm{id}_{\mathcal{M}} . \tag{10.1}$$

In other words, $\sigma(p) \in \mathcal{B}_p$, for each $p \in \mathcal{M}$. See Figure 13.1.

Given a differentiable vector bundle $\mathcal{B}$, recall that we have local trivialisations

$$\omega_\alpha : \pi^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times \mathbb{R}^m , \tag{10.2}$$

if V (the model vector space) is $\mathbb{R}^m$. The **local basis sections** $\sigma_{\alpha,a} : a = 1, \ldots, m$ over $\mathcal{U}_\alpha$ are defined by

$$\sigma_{\alpha,a}(p) = \omega_{\alpha,p}^{-1} \cdot e_a \tag{10.3}$$

where $\omega_{\alpha,p} = \omega_\alpha|_{\mathcal{B}_p}$ is the linear isomorphism of $\mathcal{B}_p$ on V, and $e_a : 1, \ldots, m$ is the standard basis on $\mathbb{R}^m$. These are *basis sections* over $\mathcal{U}_\alpha$ because

$$(\sigma_{\alpha,a}(p) : a = 1, \ldots, m) \tag{10.4}$$

forms a basis for $\mathcal{B}_p$, at each $p \in \mathcal{U}_\alpha$.

Therefore any section ψ of $\mathcal{B}$ over $\mathcal{M}$ can be expanded over $\mathcal{U}_\alpha$ as

$$\psi = \sum_{a=1}^{m} \psi_\alpha^a \sigma_{\alpha,a} \qquad : \text{over } \mathcal{U}_\alpha , \tag{10.5}$$

where the $\psi_\alpha^a : a = 1, \ldots, m$ are real valued functions defined on $\mathcal{U}_\alpha$, which are continuously differentiable if ψ is continuously differentiable.

We can proceed similary if $V = \mathbb{C}^m$. In particular, if $V = \mathbb{C}$, there is one local basis section over $\mathcal{U}_\alpha$, given by

$$\sigma_\alpha(p) = \omega_{\alpha,p}^{-1} \cdot 1 . \tag{10.6}$$

Then any section ψ of this complex line bundle can be expanded over $\mathcal{U}_\alpha$ in the form:

$$\psi = \psi_\alpha \sigma_\alpha \qquad : \text{over } \mathcal{U}_\alpha , \tag{10.7}$$

where ψ_α is a complex valued function defined on $\mathcal{U}_\alpha$, which is continuously differentiable if ψ is continuously differentiable.

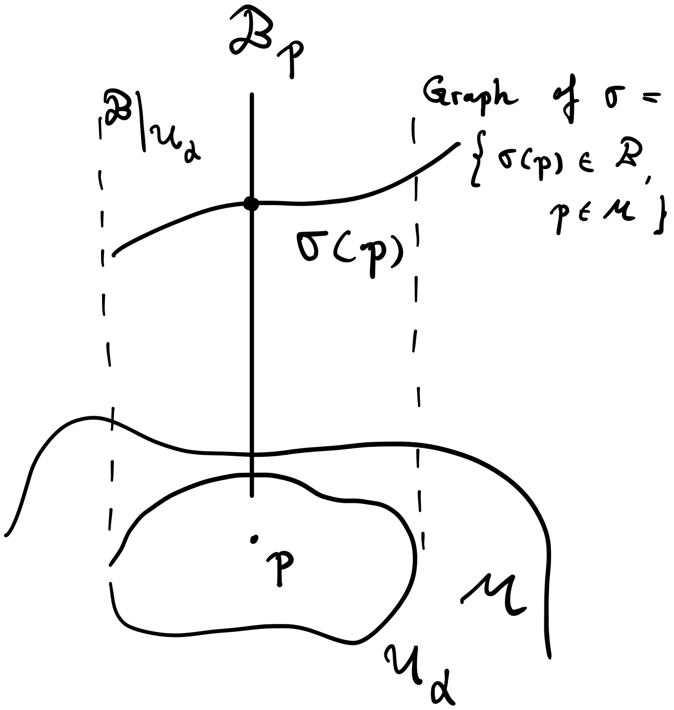


Figure 10.1.: Section of a vector bundle.

10.1. Sections of the tangent bundle

Definition 10.2. A continuously differentiable section X of $\mathrm{T}\mathcal{M}$ is called a continuously differentiable *vectorfield* on $\mathcal{M}$.

Thus $X(p) \in \mathrm{T}_p\mathcal{M}$, for each $p \in \mathcal{M}$.

We may also define a C^∞ vectorfield on $\mathcal{M}$ as a *linear operator* on $\mathrm{C}^\infty(\mathcal{M})$, the space of smooth functions on $\mathcal{M}$. If $f \in \mathrm{C}^\infty(\mathcal{M})$ then the smooth function Xf is defined by

$$(Xf)(p) = X(p) \cdot f \,, \qquad p \in \mathcal{M} \,, \tag{10.8}$$

and so $Xf \in \mathrm{C}^\infty(\mathcal{M})$. *However, not all linear operators on* $\mathrm{C}^\infty(\mathcal{M})$ *correspond to vectorfields.* Vectorfields are distinguished by the fact that they satisfy the Leibniz rule

$$X(fg) = f(Xg) + g(Xf) \,, \tag{10.9}$$

for every pair $f, g \in \mathrm{C}^\infty(\mathcal{M})$.

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

$$(XY)f = X(Yf) \,, \qquad (YX)f = Y(Xf) \,, \qquad f \in \mathrm{C}^\infty(\mathcal{M}) \tag{10.10}$$

but *not* vectorfields. However, the *commutator*

$$[X, Y] = XY - YX \tag{10.11}$$

is indeed a vectorfield:

$$\begin{aligned}[X,Y](fg) =&(XY-YX)(fg) = \\ =&X\big(f(Yg)+g(Yf)\big) - Y\big(f(Xg)+g(Xf)\big) \\ =&fXYg + (Yg)(Xf) + gXYf + (Yf)(Xg) \\ &-fYXg - (Xg)(Yf) - gYXf - (Xf)(Yg) \\ =&f(XY-YX)g + g(XY-YX)f\,,\end{aligned} \tag{10.12}$$

so the Leibniz rule holds.

10.2. Integral curves

Suppose that $\mathcal{M}$ is a differentiable manifold and V a continuously differentiable vectorfield on $\mathcal{M}$.

Definition 10.3. An *integral curve* of V through a point $p \in \mathcal{M}$ is a continuously differentiable curve $\gamma : I \to \mathcal{M}$, where I is an open interval in $\mathbb{R}$ 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 $V(\gamma(t))$,

$$\dot\gamma(t) = V(\gamma(t)) \qquad t \in I \tag{10.13}$$

where $\dot\gamma(t)\cdot f = \frac{\mathrm{d}}{\mathrm{d}t'}(f\circ\gamma)(t')|_{t'=t}$.

Let $(\mathcal{U}, \varphi)$ be a local chart at p. Then γ is represented by the curve

$$\tilde\gamma = \varphi\circ\gamma : J \to \mathbb{R}^n\,, \tag{10.14}$$

a curve in $\mathbb{R}^n$ where $J = \gamma^{-1}(\mathcal{U})$ is an open subinterval of I containing 0. We have

$$\tilde\gamma = (\tilde\gamma^1, \dots, \tilde\gamma^n) \tag{10.15}$$

where $\tilde\gamma^i$ are real functions defined on J. Morever, the vector $V(q)$ for $q \in \mathcal{U}$ is expressed in terms of the chart $(\mathcal{U}, \varphi)$ by

$$V(q) = \sum_{i=1}^n V^i(q)\frac{\partial}{\partial x^i}|_q \tag{10.16}$$

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

$$\tilde V^i = V^i \circ \varphi^{-1}\,, \tag{10.17}$$

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

$$\tilde V(x) = \sum_{i=1}^n \tilde V^i(x)\frac{\partial}{\partial x^i}|_x\,, \quad x \in \mathcal{V}\,. \tag{10.18}$$

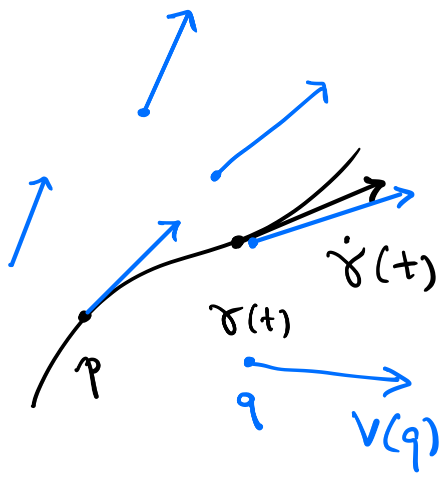


Figure 10.2.: Section of a vector bundle.

Here $\frac{\partial}{\partial x^i}|_x$ is the standard partial derivative with respect to x^i.

Then $\tilde{\gamma}$ is an integral curve of $\tilde{V}$:

$$\dot{\tilde{\gamma}}(t) = \tilde{V}(\tilde{\gamma}(t)) \qquad (t \in J) \tag{10.19}$$

Since

$$\dot{\tilde{\gamma}}(t) = \sum_{i=1}^{n} \frac{\mathrm{d}\tilde{\gamma}^i(t)}{\mathrm{d}t} \frac{\partial}{\partial x^i}|_{\tilde{\gamma}(t)}\,, \tag{10.20}$$

we can express the above in components as

$$\frac{\mathrm{d}\tilde{\gamma}^i}{\mathrm{d}t}(t) = \tilde{V}^i(\tilde{\gamma}(t)) \qquad : i = 1, \ldots, n\,. \tag{10.21}$$

This is a system of *ordinary differential equations* (ODEs), with initial conditions

$$\tilde{\gamma}(0) = \varphi(\gamma(0)) = \varphi(p)\,. \tag{10.22}$$

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.

Problems

1. Show that (10.13) and (10.21) are indeed equivalent.

Digression into ODEs

Local existence theorem.

Let $\mathcal{V}$ be an open set in $\mathbb{R}^n$ and let V be a C^1 vectorfield on $\mathcal{V}$.

Consider the problem

$$\frac{dx}{dt} = V(x) \tag{10.1}$$

$$x(0) = y \tag{10.2}$$

of finding the integral curve of V through $y \in \mathcal{V}$.

The we have the following statement:

Local well-posedness. For any $x_0 \in \mathcal{V}$, there is a ball $B(x_0) \in \mathcal{V}$ centered at x_0, and an interval $I(x_0)$ (centered at 0) such that for any $y \in B(x_0)$ 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.

Global uniqueness theorem

For every $y \in \mathcal{V}$ the problem has a unique *maximal* solution, defined on a maximal open interval $J(y)$, $0 \in J(y)$.

$J(y)$ may be of the form $(-a, \infty)$, or $(-\infty, a)$, for some $a > 0$, or $\mathbb{R}$ itself. In the last case we say we have a *global* solution for the initial condition y.

In particular, if $y \in B(x_0)$ as in the local existence theorem then $I(x_0) \subset J(y)$.

Global solutions

Theorem 10.1 (K)**.** *Set $T_+(y) = \sup J(y)$. Then the following alternative holds:*

> ***Either*** *$T_+(y) = \infty$,*
>
> ***or*** *there is no compact subset $K \subset \mathcal{V}$ such that the motion in $[0, T_+(y))$ is contained in K.*

A similar statement holds for $T_-(y) = \inf J(y)$.

Example 10.1. Let $n = 1$, $\mathcal{V} = \mathbb{R}$, $V(x) = x^2$. The problem

$$\frac{dx}{dt} = x^2 , \qquad x(0) = y \tag{10.3}$$

has the following solution:

1. If $y = 0$, then $x(t) = 0$ for all t, and we have a global solution on $J(0) = \mathbb{R}$.

2. If $y \neq 0$, then
$$x(t) = \frac{1}{\frac{1}{y} - t} . \tag{10.4}$$
For $y > 0$, we have $J(y) = (-\infty, \frac{1}{y})$, and $T_+(y) = \frac{1}{y}$.

   For $y < 0$, we have $J(y) = (\frac{1}{y}, \infty)$, and $T_-(y) = \frac{1}{y}$.

   Note that here the image of $J(y)$ is *not compact.*

Proposition 10.2. *Let V be a C^1 vectorfield on $\mathbb{R}^n$ such that there are positive constants A, B, so that*
$$|V(x)| \leq A|x| + B . \tag{10.5}$$
Then we have a global solution for all initial conditions $y \in \mathbb{R}^n$.

Proof. According to the K theorem either $T_+(y) = \infty$ or it is finite in which case the motion in $[0, T_+(y))$ is not contained in any compact set $K \subset \mathbb{R}^n$. Consider $|x|^2 = \sum_{i=1}^n (x^i)^2$. Differentiating both sides, we have
$$\begin{aligned} 2|x|\frac{\mathrm{d}}{\mathrm{d}t}|x| =& 2\sum_{i=1}^n 2x^i \frac{\mathrm{d}x^i}{\mathrm{d}t} = \sum_{i=1}^n 2x^i V^i(x) \\ \leq& 2|x||V| \leq 2|x|(A|x| + B) , \end{aligned} \tag{10.6}$$
hence
$$\frac{\mathrm{d}|x|}{\mathrm{d}t} \leq A|x| + B , \tag{10.7}$$
or
$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(e^{-At}|x|\Big) \leq Be^{-At} . \tag{10.8}$$
Integrating on $[0, t]$, for $t > 0$, gives
$$e^{-At}|x(t)| - |y| \leq \frac{B}{A}\Big(1 - e^{-At}\Big) \tag{10.9}$$
that is
$$|x(t)| \leq |y|e^{At} + \frac{B}{A}\Big(e^{At} - 1\Big) \leq R \tag{10.10}$$
where
$$R = |y|e^{AT_+(y)} + \frac{B}{A}\Big(e^{AT_+(y)} - 1\Big) . \tag{10.11}$$

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 $T_+(y) = \infty$ is possible. Similarly we establish that $T_-(y) = -\infty$. □

Example from classical mechanics

Consider N particles in Euclidean space $\mathbb{R}^3$.

A *configuration* is a point $(\vec{x}_1, \dots, \vec{x}_N) \in \mathbb{R}^{3N}$, where $\vec{x}_\alpha \in \mathbb{R}^3$ is the *position* of the particle $\alpha = 1, \dots, N$.

Let $m_\alpha > 0$ be the *mass* of the particle α, and let us denote by $\vec{y}_\alpha \in \mathbb{R}^3$ the *momentum* of particle α. Then also $(\vec{y}_1, \dots, \vec{y}_N) \in \mathbb{R}^{3N}$.

The *Hamiltonian* is a function on *phase space* $\mathbb{R}^{3N} \times \mathbb{R}^{3N} = \mathbb{R}^{6N}$:

$$H = \sum_{\alpha=1}^{N} \frac{|\vec{y}_\alpha|^2}{2m_\alpha} + U(\vec{x}_1, \dots, \vec{x}_N) \tag{10.12}$$

where U is the *potential*, a function on *configuration space.*

A class of examples are potentials of the form

$$U(\vec{x}_1, \dots, \vec{x}_N) = \sum_{\alpha<\beta} q_\alpha q_\beta f(|\vec{x}_\alpha - \vec{x}_\beta|) \tag{10.13}$$

where q_α are *generalised charges.*

1. For the Newtonian problem of N gravitating point masses take $q_\alpha = m_\alpha$, and $f(r) = -\frac{1}{r}$.
2. For the problem of N charged particles, take q_α to be the electric charges, and $f(r) = \frac{1}{r}$.
3. The case $q_\alpha = 1$, $m_\alpha = m > 0$, and $f(r) \sim -\frac{1}{r}$ for large r, but bounded from below for $r > 0$ describes a noble gas.

In all these examples the *configuration space* is

$$C = \mathbb{R}^{3N} \setminus \Delta\,, \qquad \Delta = \{(\vec{x}_1, \dots, \vec{x}_N) \in \mathbb{R}^{3N} : \vec{x}_\alpha = \vec{x}_\beta \text{ for some } \alpha \neq \beta\}\,. \tag{10.14}$$

The *canonical equations* are

$$\frac{\mathrm{d}\vec{x}_\alpha}{\mathrm{d}t} = \frac{\partial H}{\partial \vec{y}_\alpha} \tag{10.15}$$

$$\frac{\mathrm{d}\vec{y}_\alpha}{\mathrm{d}t} = -\frac{\partial H}{\partial \vec{x}_\alpha} \tag{10.16}$$

We are considering the integral curves of the *Hamiltonian vectorfield*

$$V = \sum_\alpha \Big\{ \frac{\partial H}{\partial \vec{y}_\alpha} \frac{\partial}{\partial \vec{x}_\alpha} - \frac{\partial H}{\partial \vec{x}_\alpha} \frac{\partial}{\partial \vec{y}_\alpha} \Big\}\,. \tag{10.17}$$

Here

$$\frac{\partial H}{\partial \vec{y}_\alpha} = \frac{\vec{y}_\alpha}{m_\alpha}\,, \qquad \frac{\partial H}{\partial \vec{x}_\alpha} = \frac{\partial U}{\partial \vec{x}_\alpha}\,. \tag{10.18}$$

and thus the equations are

$$\frac{\mathrm{d}\vec{x}_\alpha}{\mathrm{d}t} = \frac{\vec{y}_\alpha}{m_\alpha} \tag{10.19}$$

$$\frac{\mathrm{d}\vec{y}_\alpha}{\mathrm{d}t} = -\frac{\partial U}{\partial \vec{x}_\alpha}\,. \tag{10.20}$$

Theorem 10.3. *Let U be a C^2 function on the configuration space C, which is bounded below: there is a positive constant B such that*

$$U(\vec{x}_1, \ldots, \vec{x}_N) \geq -B\,. \tag{10.21}$$

Then we have global existence for all initial data.

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

$$\frac{\mathrm{d}H}{\mathrm{d}t} = VH = \sum_\alpha \Big\{ \frac{\partial H}{\partial \vec{y}_\alpha}\frac{\partial H}{\partial \vec{x}_\alpha} - \frac{\partial H}{\partial \vec{x}_\alpha}\frac{\partial H}{\partial \vec{y}_\alpha} \Big\} = 0\,. \tag{10.22}$$

Therefore $H(t) = H(0)$: a real number depending only on the initial conditions. Hence

$$\sum_\alpha \frac{|\vec{y}_\alpha|^2}{2m_\alpha} = H(t) - U \leq H(0) + B\,. \tag{10.23}$$

Let $m_M = \max\{m_\alpha : \alpha = 1, \ldots, N\}$, then

$$\sum_\alpha |\vec{y}_\alpha|^2 \leq 2m_M(H(0) + B)\,. \tag{10.24}$$

If we denote by $\vec{y} = (\vec{y}_1, \ldots, \vec{y}_N) \in \mathbb{R}^{3N}$, then

$$|y| \leq \sqrt{2m_M(H(0) + B)}\,. \tag{10.25}$$

To bound the velocities, let $m_m = \min\{m_\alpha : \alpha = 1, \ldots, N\}$, then

$$\Big|\frac{\mathrm{d}\vec{x}_\alpha}{\mathrm{d}t}\Big| = \Big|\frac{\vec{y}_\alpha}{m_\alpha}\Big| \leq \frac{|\vec{y}_\alpha|}{m_m}\,. \tag{10.26}$$

Similarly as above, if we denote $\vec{x} = (\vec{x}_1, \ldots, \vec{x}_N) \in \mathbb{R}^{3N}$, and differentiate

$$|x|^2 = \sum_\alpha |\vec{x}_\alpha|^2\,, \tag{10.27}$$

we have

$$2|x|\frac{\mathrm{d}}{\mathrm{d}t}|x| = \sum_\alpha 2\vec{x}_\alpha \cdot \frac{\vec{x}_\alpha}{\mathrm{d}t} \leq 2\sqrt{\sum_\alpha |\vec{x}_\alpha|^2}\sqrt{\sum_\alpha \Big|\frac{\vec{x}_\alpha}{\mathrm{d}t}\Big|^2}\,. \tag{10.28}$$

It follows that

$$\frac{\mathrm{d}|x|}{\mathrm{d}t} \leq \sqrt{\sum_\alpha |\frac{\vec{x}_\alpha}{\mathrm{d}t}|^2} \leq \frac{1}{m_m}\sqrt{\sum_\alpha |\vec{y}_\alpha|^2} = \frac{|y|}{m_m} \leq \frac{\sqrt{2m_M(H(0)+B)}}{m_m} \tag{10.29}$$

Integrate on $[0,t]$, $t \in (0,T_+)$ to obtain

$$|x(t)| \leq |x(0)| + \frac{\sqrt{2m_M(H(0)+B)}}{m_m}t \leq |x(0)| + cT_+ \tag{10.30}$$

where c is a constant. This shows that on $[0,T_+)$ the motion is contained in a compact set $K = \overline{B_R(0)}$, where $R = |x(0)| + cT_+$.

This contradicts the K theorem.

□

Problems

1. Let us define on $\mathbb{R}^n$ the following *quadratic vectorfield*: $V = (V^1, \ldots, V^n)$, where

$$V^i(x) = \sum_{j,k=1}^n a^i_{jk} x^j x^k \,, \tag{10.31}$$

   where the coefficients a^i_{jk} are constant and satisfy the symmetry condition $a^i_{jk} = a^i_{kj}$ and the *cocycle condition*

$$a^i_{jk} + a^j_{ki} + a^k_{ij} = 0 \,. \tag{10.32}$$

   Prove that for any initial data $x_0 \in \mathbb{R}^n$, the differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = V(x) \tag{10.33}$$

   has a global solution, i.e. a solution $x : \mathbb{R} \to \mathbb{R}^n$, with initial data $x(0) = x_0$.

2. The cycycle condition in the previous exercise is necessary. Show that if it is not satisfied, then there exists initial data $x_0 \in \mathbb{R}^n$ for which the solution with $x(0) = x_0$ is *not* globally defined.

Lecture 11.

Flows

Definition 11.1. A vectorfield V on a manifold $\mathcal{M}$ is *complete* if each integral curve of V is defined for all time.

Now, given $p \in \mathcal{M}$ let us denote by $I(p)$ the maximal interval of existence with initial condition p. Let then V be a C^1 vectorfield on a differentiable manifold $\mathcal{M}$ and let $\mathcal{A} = \{(\mathcal{U}_\alpha, \varphi_\alpha) : \alpha \in I\}$ be an atlas for $\mathcal{M}$. If $p \in \mathcal{U}_\alpha$, to obtain the integral curve of V through p we consider the integral curve of the vectorfield $\tilde{V}$ representing V in the chart, a vectorfield defined on $\mathcal{V}_\alpha = \varphi_\alpha(\mathcal{U}_\alpha)$ open in $\mathbb{R}^n$, through $y = \varphi_\alpha(p) \in \mathcal{V}_\alpha$. This integral curve, $\tilde{\gamma}$, is defined and contained in $\mathcal{V}_\alpha$, for $t \in I(y)$, the maximal interval of existence with initial condition y. Then $\gamma = \varphi_\alpha^{-1} \circ \tilde{\gamma}$ is an integral curve in V through p, defined in $I(y)$ and contained in $\mathcal{U}_\alpha$. Therefore $I(y) \subset I(p)$.

By the local existence theorem, for each $y \in \mathcal{V}_\alpha$ there is a ball $B(y) \subset \mathcal{V}_\alpha$ such that for all $y' \in B(y)$:

$$I(y') \supset [-\varepsilon(y), \varepsilon(y)] \tag{11.1}$$

where $\varepsilon(y) > 0$ and only depends on y, and not y'. Now any point $p \in \mathcal{M}$ belongs to some $\mathcal{U}_\alpha$. Consider

$$\mathcal{W}_p = \varphi_\alpha^{-1}(B(y)), \qquad y = \varphi_\alpha(p). \tag{11.2}$$

Then $\mathcal{W}_p$ is a neighborhood of $p \in \mathcal{M}$, and for every $q \in \mathcal{W}_p$ the motion with initial condition q is defined for all $t \in [-\varepsilon_p, \varepsilon_p]$, where $\varepsilon_p = \varepsilon(y)$. Here $y' = \varphi_\alpha(q)$.

Suppose now that $\mathcal{M}$ is *compact* (for every open cover there is a finite subcover). The collection of open sets $\{W_p : p \in \mathcal{M}\}$ covers $\mathcal{M}$. Thus if $\mathcal{M}$ is compact, a finite subcollection sufficies to cover $\mathcal{M}$:

$$\{\mathcal{W}_{p_i} : i = 1, \ldots, N\} \tag{11.3}$$

Set $\varepsilon = \min\{\varepsilon_{p_i} : i = 1, \ldots, N\} > 0$. Then for each $q \in \mathcal{M}$, the motion with initial condition q is defined for all $t \in [-\varepsilon, \varepsilon]$. Let us set

$$\phi_t(q) = \gamma_q(t) \qquad : t \in [-\varepsilon, \varepsilon] \tag{11.4}$$

where γ_q is the integral curve through q, $\gamma_q(0) = q$. Then from the local existence theorem ϕ_t is a C^1 mapping of $\mathcal{M}$ onto itself. Also $\phi_{-t} = \phi_t^{-1}$, therefore ϕ_t is a *diffeomorphism* of $\mathcal{M}$ onto itself. More generally, if s, t are such that $s + t \in [-\varepsilon, \varepsilon]$, then

$$\phi_s \circ \phi_t = \phi_{s+t}, \tag{11.5}$$

because

$$\phi_s(\phi_t(q)) = \gamma_{\phi_t(q)}(s) = \gamma_{\gamma_q(t)}(s) = \gamma_q(s+t) = \phi_{s+t}(q)\,. \tag{11.6}$$

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

$$t = k\varepsilon + s \tag{11.7}$$

where $k = [t/\varepsilon]$ denotes the integral part of the real number t/ε, and $s \in [0, \varepsilon)$. If $k > 0$ we set $\phi_{k\varepsilon} = \phi_\varepsilon \circ \cdots \circ \phi_\varepsilon$ (k-fold composition), and if $k < 0$ we set $\phi_{k\varepsilon} = \phi_{-\varepsilon} \circ \cdots \circ \phi_{-\varepsilon}$ and finally we set

$$\phi_t = \phi_{k\varepsilon} \circ \phi_s \tag{11.8}$$

where ϕ_s is already defined because $s \in [0, \varepsilon)$. We have defined ϕ_t for all $t \in \mathbb{R}$, and so defined ϕ_t satisfies

$$\phi_0 = \mathrm{id}_{\mathcal{M}}\,, \qquad \phi_s \circ \phi_t = \phi_{s+t} : \qquad s, t \in \mathbb{R}\,. \tag{11.9}$$

Definition 11.2. A 1-parameter family $\{\phi_t : t \in \mathbb{R}\}$ of diffeomorphisms of $\mathcal{M}$ having the property (11.9) is called a *1-parameter group of diffeomorphisms of* $\mathcal{M}$.

For any point $p \in \mathcal{M}$, the integral curve starting at p is given by

$$\gamma_p(t) = \phi_t(p) : t \in \mathbb{R} \tag{11.10}$$

The path $\{\phi_t(p) : t \in \mathbb{R}\} \subset \mathcal{M}$ is called the *orbit* of the group through p. Thus the motion is defined for all time and for any initial condition. Hence the vectorfield V is complete.

Theorem 11.1. *If $\mathcal{M}$ is a compact manifold then any vectorfield V on $\mathcal{M}$ is complete.*

This is *not* true for a non-compact manifold. We have seen the example $\mathcal{M} = \mathbb{R}$, $V(x) = x^2$. However, a complete vectorfield V on a manifold $\mathcal{M}$ (compact or not) generates a 1-parameter group $\{\phi_t\}$ of diffeomorphisms of $\mathcal{M}$ called the *flow* of V.

Problems

1. We have proved that any smooth vectorfield on a compact manifold is complete. We want to give a different proof of this result in the case of a submanifold $M \subset \mathbb{R}^n$. To make things simpler, we shall actually consider only the case of $\mathbb{S}^2 \subset \mathbb{R}^3$. However, the procedure can be applied to any submanifold of Euclidean space.

   Let X be a smooth vectorfield on

   $$\mathbb{S}^2 = \{x :\in \mathbb{R}^3 : |x| = 1\}\,. \tag{11.11}$$

   a) Extend X to a vectorfield $\tilde{X}$ on $\mathbb{R}^3$ using radial projections and a cut-off function with compact support.
   b) Show that the given extension is complete.
   c) Show that if $p \in \mathbb{S}^2$, then any integral curve of $\tilde{X}$ passing through p lies in $\mathbb{S}^2$ for all times.

Lecture 12.

Cotangent bundle

12.1. Differentials

Let $\mathcal{M}$ and $\mathcal{N}$ be differentiable manifolds (not necessarily of the same dimension) and

$$\psi : \mathcal{M} \to \mathcal{N} \tag{12.1}$$

a differentiable mapping. We assume for simplicity that ψ is C^∞.

Definition 12.1. The *differential* of ψ at $p \in \mathcal{M}$, denoted by $\mathrm{d}\psi(p)$, is the linear mapping of $\mathrm{T}_p\mathcal{M}$ into $\mathrm{T}_q\mathcal{N}$, where $q = \psi(p)$, given by:

If $u \in \mathrm{T}_p\mathcal{M}$, then $v = \mathrm{d}\psi(p) \cdot u$ is defined by:

$$v \cdot f = u \cdot (f \circ \psi) \qquad f \in \mathrm{C}^\infty(\mathcal{N}) \,. \tag{12.2}$$

The picture to have in mind here is the following:

Let u be the tangent vector at p to a curve γ in $\mathcal{M}$ through $p \in \mathcal{M}$. Thus for all $g \in \mathrm{C}^\infty(\mathcal{M})$ we have

$$u \cdot g = \frac{\mathrm{d}}{\mathrm{d}t}(g \circ \gamma)(t)|_{t=0} \,. \tag{12.3}$$

Setting $g = f \circ \psi$ in the definition above, we obtain for all $f \in \mathrm{C}^\infty(\mathcal{N})$:

$$v \cdot f = u \cdot g = \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \psi \circ \gamma)(t)|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \kappa)(t)|_{t=0} \tag{12.4}$$

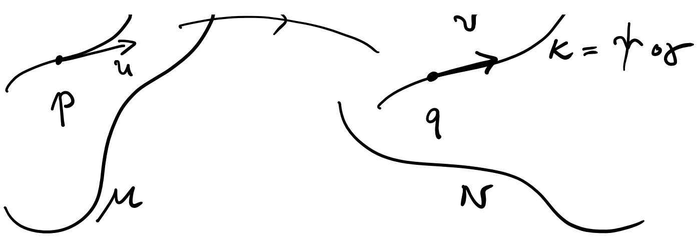


Figure 12.1.: Definition of the differential.

where $\kappa = \psi \circ \gamma$ is a curve in $\mathcal{N}$ through $q = \psi(p)$, the image by ψ of the curve γ. Thus v is the tangent vector at q to κ.

In the special case $\mathcal{N} = \mathbb{R}$ the differential of a C^∞ function ϕ on $\mathcal{M}$ at $p \in \mathcal{M}$, denoted by $\mathrm{d}\phi(p)$, is a linear function on $\mathrm{T}_p\mathcal{M}$ defined by

$$\mathrm{d}\phi(p) \cdot u = u \cdot \phi \qquad u \in \mathrm{T}_p\mathcal{M}\,. \tag{12.5}$$

In fact, in this case $\mathrm{d}\phi$ is a differentiable section of $\mathrm{T}^*\mathcal{M}$, the cotangent bundle of $\mathcal{M}$.

12.2. Cotangent bundle

Definition 12.2. The *cotangent bundle* of $\mathcal{M}$ is

$$\mathrm{T}^*\mathcal{M} = \bigcup_{p\in\mathcal{M}} \mathrm{T}^*_p\mathcal{M} \tag{12.6}$$

where $\mathrm{T}^*_p\mathcal{M}$, the *cotangent space* at p, is the dual space of $\mathrm{T}_p\mathcal{M}$:

$$\mathrm{T}^*_p\mathcal{M} = \left(\mathrm{T}_p\mathcal{M}\right)^*\,. \tag{12.7}$$

Remark 12.1. In general, if V is a real finite dimensional vector space, the dual space V^* is the space of real linear functionals on V. Similarly in the complex.

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

$$\mathcal{B}^* = \bigcup_{p\in\mathcal{M}} \mathcal{B}^*_p \tag{12.8}$$

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

12.3. Sections of dual bundles

Let $(\iota_a : a = 1, 2, \ldots, n)$ $(n = \dim \mathcal{B}_p)$ be the local basis sections of $\mathcal{B}$ over $\mathcal{U} \in \mathcal{M}$. Then the *dual local basis secions* $(\iota^{*a} : a = 1, \ldots, n)$ of $\mathcal{B}^*$ over $\mathcal{U}$ are defined by the conditions that

$$\iota^{*a} \cdot \iota_b = \delta^a_b \qquad : \text{in } \mathcal{U}\,. \tag{12.9}$$

Definition 12.3. A continuously differentiable *1-form* on $\mathcal{M}$ is a continuously differentiable section of $\mathrm{T}^*\mathcal{M}$.

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

Some notation:

$\mathrm{C}^\infty(\mathcal{M})$: the space of smooth functions on $\mathcal{M}$.

$\mathscr{X}^\infty(\mathcal{M})$**:** the space of smooth vectorfields on $\mathcal{M}$.

$\Omega_1^\infty(\mathcal{M})$: the space of smooth 1-forms on $\mathcal{M}$.

Given $f \in \mathrm{C}^\infty(\mathcal{M})$, $X \in \mathscr{X}^\infty(\mathcal{M})$, then fX (defined pointwise) is in $\mathscr{X}^\infty(\mathcal{M})$. And given $\theta \in \Omega_1^\infty(\mathcal{M})$ then $f\theta$ (defined pointwise) $\in \Omega_1^\infty(\mathcal{M})$.

If $\mathcal{U}$ is the domain of a chart $(\mathcal{U}, \varphi)$ then the local basis sections of $\mathrm{T}\mathcal{M}$ over $\mathcal{U}$ are the local coordinate vectorfields $(\iota_\mu = \frac{\partial}{\partial x^\mu} : \mu = 1, \ldots, m)$ $(m = \dim \mathcal{M})$. The dual basis sections of $\mathrm{T}^*\mathcal{M}$ are $(\iota^{*\nu} : \nu = 1, 2, \ldots, m)$ where $\iota^{*\nu} = \mathrm{d}x^\nu$, and $x^\nu = \varphi^\nu$ are the components of the chart of φ. For, we have

$$\mathrm{d}x^\nu \cdot \frac{\partial}{\partial x^\mu} = \frac{\partial x^\nu}{\partial x^\mu} = \delta^\nu_\mu \,. \tag{12.10}$$

Every 1-form θ can be locally (in $\mathcal{U}$) expanded as

$$\theta = \sum_{\nu=1}^{m} \theta_\nu \mathrm{d}x^\nu \,, \tag{12.11}$$

where the coefficients θ_ν are functions defined in $\mathcal{U}$. Recalling that given a vectorfield X we have the expansion

$$X = \sum_{\mu=1}^{m} X^\mu \frac{\partial}{\partial x^\mu} : \qquad \text{in } \mathcal{U} \tag{12.12}$$

then the function $\theta \cdot X$ is expressed in $\mathcal{U}$ as

$$\theta \cdot X = \sum_{\mu=1}^{m} \theta_\mu X^\mu \,. \tag{12.13}$$

Lecture 13.
Pull backs and push forwards

13.1. Pull-back.

This notion applies to *covariant objects*.

Let $\mathcal{M}$ and $\mathcal{N}$ be differentiable manifolds, $\dim \mathcal{M} = n$, and $\dim \mathcal{N} = n$, and let $\phi : \mathcal{M} \to \mathcal{N}$ be a continuously differentiable mapping.

The simplest covariant object is a function. Consider a continuously differentiable function on $\mathcal{N}$ (the target). Then $f \circ \phi$ is a continuously differentiable function on $\mathcal{M}$ (the domain). We call $f \circ \phi$ the **pull-back** by ϕ of f and write

$$\phi^* f = f \circ \phi \,. \tag{13.1}$$

A 1-form is the next simplest covariant structure.

Definition 13.1. Let θ be a 1-form on $\mathcal{N}$ (the target), and $\phi : \mathcal{M} \to \mathcal{N}$ a continuously differentiable mapping. The *pull-back* of θ by ϕ, denoted by $\phi^*\theta$, is the 1-form on $\mathcal{M}$ (the domain) defined by:

$$(\phi^*\theta) \cdot v = \theta \cdot (\mathrm{d}\phi \cdot v)\,, \qquad v \in \mathrm{T}_p\mathcal{M},\ p \in \mathcal{M}\,. \tag{13.2}$$

(Recall that $\mathrm{d}\phi(p)$ is a linear map of $\mathrm{T}_p\mathcal{M}$ into $\mathrm{T}_{\phi(p)}\mathcal{N}$.)

In terms of local coordinates $x^\mu : \mu = 1, 2, \ldots, m$ for $\mathcal{M}$, and $y^a : a = 1, \ldots, n$ for $\mathcal{N}$,

$$\theta = \sum_{a=1}^{n} \theta_a \mathrm{d}y^a \tag{13.3}$$

$$\phi^*\theta = \sum_{\mu=1}^{m} (\phi^*\theta)_\mu \mathrm{d}x^\mu \tag{13.4}$$

where, with ϕ given by $y^a = \phi^a(x_1, \ldots, x^m)$, we have

$$(\phi^*\theta)_\mu(p) = \sum_{a=1}^{n} \theta_a(\phi(p)) \frac{\partial y^a}{\partial x^\mu}|_p\,. \tag{13.5}$$

Exercise 13.1. Verify this formula!

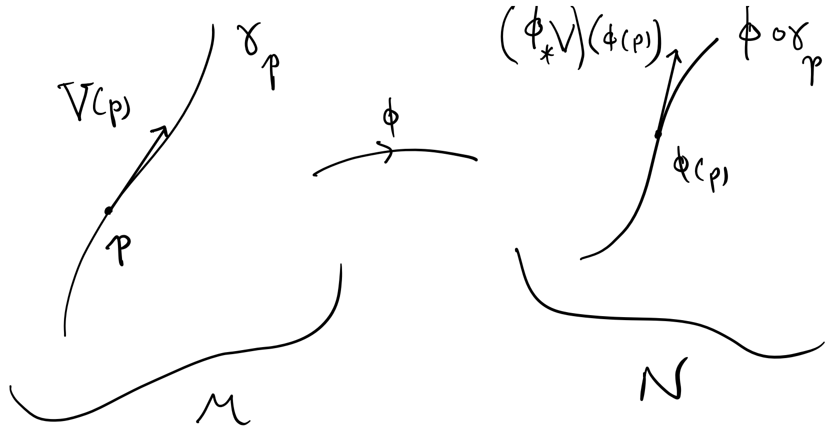


Figure 13.1.: Push-forward of a vectorfield and its integral curves.

13.2. Push-forward.

This applies to *contravariant* objects.

The simplest contravariant object is a vectorfield.

As above we have a continuously differentiable mapping $\phi : \mathcal{M} \to \mathcal{N}$, but *we now require* ϕ *to be one-to-one*. Then for each $q \in \phi(\mathcal{M}) \subset \mathcal{N}$ there is a unique $p \in \mathcal{M}$ such that $\phi(p) = q$.

Definition 13.2. Let V be a vectorfield on $\mathcal{M}$ (the domain), and $\phi : \mathcal{M} \to \mathcal{N}$ one-to-one. The *push-forward* $\phi_* V$ of V by ϕ is the vectorfield defined along the image of $\phi(\mathcal{M}) \subset \mathcal{N}$ (the target) by

$$(\phi_* V)(\phi(p)) = \mathrm{d}\phi(p) \cdot V(p) \,, \qquad p \in \mathcal{M} \,, \tag{13.6}$$

namely

$$(\phi_* V)(q) = (\mathrm{d}\phi \cdot V)(\phi^{-1}(q)) \,, \qquad q \in \phi(\mathcal{M}) \subset \mathcal{N} \,. \tag{13.7}$$

Remark 13.1. Let us now consider the case that $\mathcal{M} = \mathcal{N}$, and ϕ being a diffeomorphism of $\mathcal{M}$ onto itself. Then for a smooth vectorfield V on $\mathcal{M}$, $\phi_* V$ is also a smooth vectorfield on $\mathcal{M}$.

Remark 13.2. The analogue of a pull-back for a vectorfield (more generally, for any contravariant object) is the push-forward by ϕ^{-1}:

$$(\phi_*^{-1} V)(p) = (\mathrm{d}\phi^{-1} \cdot V)(\phi(p)) \,. \tag{13.8}$$

13.3. Integral curves.

If γ_p is the integral curve of V through p, then $\phi \circ \gamma_p$ is the integral curve of $\phi_* V$ through $\phi(p)$.

Exercise 13.2. Verify this!

Suppose now that V is a *complete* vectorfield. Then V generates a 1-parameter group ψ_t of diffeomorphisms of $\mathcal{M}$. Then

$$\gamma_p(t) = \psi_t(p)\,. \tag{13.9}$$

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

$$\tilde{\psi}_t(\phi(p)) = (\phi \circ \gamma_p)(t) = \phi(\psi_t(p)) \qquad p \in \mathcal{M}\,, \quad t \in \mathbb{R}\,. \tag{13.10}$$

Thus

$$\tilde{\psi}_t \circ \phi = \phi \circ \psi_t \qquad t \in \mathbb{R}\,. \tag{13.11}$$

or

$$\tilde{\psi}_t = \phi \circ \psi_t \circ \phi^{-1} \qquad t \in \mathbb{R}\,. \tag{13.12}$$

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

$$\phi_* V = V \iff \psi_t \circ \phi = \phi \circ \psi_t \qquad t \in \mathbb{R}. \tag{13.13}$$

Problems

1. Let $\mathcal{M}$ and $\mathcal{N}$ be differentiable manifolds, and $u : \mathcal{M} \to \mathcal{N}$ a differentiable map. We want to define the **pull-back tangent bundle** as follows:

$$u^*\mathrm{T}\mathcal{N} = \bigcup_{p \in \mathcal{M}} \{p\} \times T_{u(p)}\mathcal{N} \tag{13.14}$$

   Show that this has the structure of a vectorbundle over $\mathcal{M}$ by exhibiting suitable local trivialisations.

2. Given a diffeomorphism $\varphi : \mathcal{M} \to \mathcal{M}$, from a differentiable manifold $\mathcal{M}$ onto itself, we can view the push-forward φ_* of vectorfields as a linear map

$$\phi_* : \mathscr{X}^\infty(\mathcal{M}) \to \mathscr{X}^\infty(\mathcal{M})\,. \tag{13.15}$$

   Show that if ϕ, and ψ are any two diffeomorphisms of $\mathcal{M}$ onto itself, then

$$(\phi \circ \psi)_* X = \phi_* \psi_* X \qquad X \in \mathscr{X}^\infty(M)\,. \tag{13.16}$$

Lecture 14.
Lie derivatives

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

$$\pounds_X f = \lim_{t\to 0} \frac{\phi_t^* f - f}{t} = \frac{\mathrm{d}}{\mathrm{d}t}\phi^* f|_{t=0} \tag{14.1}$$

$$\pounds_X \theta = \lim_{t\to 0} \frac{\phi_t^* \theta - \theta}{t} = \frac{\mathrm{d}}{\mathrm{d}t}\phi^* \theta|_{t=0} \tag{14.2}$$

Now, since $\phi_t^* f = f \circ \phi_t$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_t^* f|_{t=0}(p) = \frac{\mathrm{d}}{\mathrm{d}t}(f\circ\phi_t(p))|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}(f\circ\gamma_p(t))|_{t=0} = X(p)\cdot f \qquad p\in\mathcal{M} \tag{14.3}$$

because the tangent vector to γ_p at p is simply $X(p)$.

We conclude that:

$$\pounds_X f = Xf \tag{14.4}$$

Let Y be another vectorfield on $\mathcal{M}$. We define the **Lie derivative** of Y with respect to X by:

$$\pounds_X Y = \lim_{t\to 0} \frac{\phi_{-t*} Y - Y}{t} = \frac{\mathrm{d}}{\mathrm{d}t}\phi_{-t*} Y|_{t=0}\,. \tag{14.5}$$

Proposition 14.1. *We have*

$$\pounds_X Y = [X,Y] \tag{14.6}$$

Proof. Given any function $f\in \mathrm{C}^\infty(\mathcal{M})$ we have

$$(\pounds_X Y)f = \lim_{t\to 0}(t^{-1}E_t)f \tag{14.7}$$

where

$$E_t = \phi_{-t*}Y - Y\,; \tag{14.8}$$

so:

$$\begin{aligned} E_t(p)\cdot f =& \Big((\mathrm{d}\phi_{-t}\cdot Y)(\phi_t(p))\Big)\cdot f - Y(p)\cdot f \\ =& Y(\phi_t(p))\cdot(f\circ\phi_{-t}) - Y(p)\cdot f \qquad (p\in\mathcal{M})\,. \end{aligned} \tag{14.9}$$

Now we can write

$$f \circ \phi_{-t} - f = t g_t \tag{14.10}$$

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

$$g_0 = \frac{\mathrm{d}}{\mathrm{d}t} f \circ \phi_{-t}|_{t=0} = -Xf\,. \tag{14.11}$$

We can then write

$$E_t(p) \cdot f = Y(\phi_t(p)) \cdot f - Y(p) \cdot f + tY(\phi_t(p)) \cdot g_t \tag{14.12}$$

and we have

$$\lim_{t\to 0} \frac{Y(\phi_t(p)) \cdot f - Y(p) \cdot f}{t} = \lim_{t\to 0} \frac{(Yf)(\phi_t(p)) - (Yf)(p)}{t} = (XYf)(p)\,, \tag{14.13}$$

while

$$\lim_{t\to 0} Y(\phi_t(p)) \cdot g_t = Y(p) \cdot g_0 = (Yg_0)(p) = -(YXf)(p)\,. \tag{14.14}$$

□

Let now θ be a 1-form on $\mathcal{M}$, Y a vectorfield on $\mathcal{M}$, and ϕ a diffeomorphism of $\mathcal{M}$ onto itself. Then by the definition of the pull-back and push-forward,

$$\phi^*\theta \cdot Y = (\theta \cdot \phi_* Y) \circ \phi\,. \tag{14.15}$$

Exercise 14.1. Verify this identity!

Since Y is here an arbitrary vectorfield, we may replace Y by $\phi_*^{-1} Y$ to obtain

$$\phi^*\theta \cdot \phi_*^{-1} Y = (\theta \cdot Y) \circ \phi\,. \tag{14.16}$$

Let now X be a complete vectorfield on $\mathcal{M}$ generating the 1-par group ϕ_t. Then we have

$$\phi_t^*\theta \cdot \phi_{-t*} Y = \phi_t^*(\theta \cdot Y)\,. \tag{14.17}$$

Taking the derivative in t yields *Leibniz rule for Lie derivatives*:

$$(\pounds_X\, \theta) \cdot Y + \theta \cdot \pounds_X\, Y = \pounds\,(\theta \cdot X)\,. \tag{14.18}$$

Substituting $\pounds_X\,(\theta \cdot Y) = X(\theta \cdot Y)$, and $\pounds_X\, Y = [X,Y]$, we obtain the following formula for the Lie derivative of a 1-form:

$$(\pounds_X\, \theta) \cdot Y = X(\theta \cdot Y) - \theta \cdot [X,Y]\,. \tag{14.19}$$

Remark 14.1. A 1-form $\alpha \in \Omega_1^\infty(\mathcal{M})$ can be thought of as an assignment of a function to each vectorfield $X \in \mathscr{X}^\infty(\mathcal{M})$ which is linear with respect to the ring $\mathrm{C}^\infty(\mathcal{M})$ of $\mathrm{C}^\infty(\mathcal{M})$ functions on $\mathcal{M}$. In other words,

$$\alpha \cdot (fX) = f\alpha \cdot X \qquad f \in \mathrm{C}^\infty(\mathcal{M})\,, X \in \mathscr{X}^\infty(\mathcal{M})\,. \tag{14.20}$$

Consider then,

$$\begin{aligned}
(\pounds_X\,\theta)\cdot(fY) =& X(\theta\cdot(fY)) - \theta\cdot[X,fY] \\
=& X(f\theta\cdot Y) - \theta\cdot(f[X,Y] + (Xf)Y) \\
=& (Xf)(\theta\cdot Y) + fX(\theta\cdot Y) - f\theta\cdot[X,Y] - (Xf)(\theta\cdot Y) \\
=& f(\pounds_X\,\theta)\cdot Y\,.
\end{aligned} \tag{14.21}$$

Lecture 15.

Flows and commutation

In this lecture we want to prove the following:

Proposition 15.1. *Let X and Y be complete vectorfields on $\mathcal{M}$, generating the 1-parameter groups ϕ_t and ψ_s, respectively. Suppose that X and Y commute:*

$$[X, Y] = 0\,. \tag{15.1}$$

Then we have:

$$\phi_t \circ \psi_s = \psi_s \circ \phi_t \qquad (s, t \in \mathbb{R})\,. \tag{15.2}$$

Proof. Consider

$$\frac{\mathrm{d}}{\mathrm{d}t'} \phi_{-t'*} Y|_{t'=t} = \lim_{r\to 0} \frac{\phi_{-t-r*}Y - \phi_{-t*}Y}{r}\,. \tag{15.3}$$

Now

$$\phi_{-t-r*}Y = \phi_{-r*}(\phi_{-t*}Y)\,; \tag{15.4}$$

in fact, in general, if X is a vectorfield on $\mathcal{M}$, and ϕ, ψ are any two diffeomorphisms of $\mathcal{M}$ onto itself then

$$(\phi \circ \psi)_* X = \phi_* \psi_* X\,. \tag{15.5}$$

Thus

$$\frac{\mathrm{d}}{\mathrm{d}t'} \phi_{-t'*} Y|_{t'=t} = \mathcal{L}_X\,(\phi_{-t*}Y) = [X, \phi_{-t*}Y]\,. \tag{15.6}$$

We now use the following two facts:

1. If X generates ϕ_s then $\phi_{s*}X = X$ for all $s \in \mathbb{R}$.

2. If X, Y are any two vectorfields and ϕ any diffeomorphism, then

$$\phi_*[X, Y] = [\phi_* X, \phi_* Y]\,. \tag{15.7}$$

Proof of 1): If γ_p is the integral curve of X through p then the integral curve of $\phi_{s*}X$ (for s fixed) through $\phi_s(p)$ is $\phi_s \circ \gamma_p$. But $(\phi_s \circ \gamma_p)(t) = \phi_s(\phi_t(p)) = \gamma_p(t+s) = \phi_t(\gamma_p(s)) = \gamma_{\gamma_p(s)}(t)$: this is the integral curve of X itself through the point $\phi_s(p)$. Therefore the integral curves of the vectorfields ϕ_{s*} and X coincide. Hence the vectorfields themselves coincide.

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

$$(\phi_* V)f = \big(V(f \circ \phi)\big) \circ \phi^{-1} \tag{15.8}$$

for all $f \in C^\infty(\mathcal{M})$. Then

$$\begin{aligned}[\phi_* X, \phi_* Y]f &= (\phi_* X)\big((\phi_* Y)f\big) - (\phi_* Y)\big((\phi_* X)f\big) \\ &= \big(X((\phi_* Y)f \circ \phi)\big) \circ \phi^{-1} - \big(Y((\phi_* X)f \circ \phi)\big) \circ \phi^{-1} \\ &= \big(XY(f \circ \phi)\big) \circ \phi^{-1} - \big(YX(f \circ \phi)\big) \circ \phi^{-1} \\ &= \big([X,Y](f \circ \phi)\big) \circ \phi^{-1} \\ &= \big(\phi_*[X,Y]\big)f\,.\end{aligned} \tag{15.9}$$

We return to (15.6). By the two facts above, and the assumption (15.1),

$$[X, \phi_{-t*}Y] = [\phi_{-t*}X, \phi_{-t*}Y] = \phi_{-t*}[X,Y] = 0\,. \tag{15.10}$$

We thus obtain

$$\frac{\mathrm{d}}{\mathrm{d}t'}\phi_{-t'*}Y|_{t'=t} = 0\,, \tag{15.11}$$

hence $\phi_{-t*}Y = \phi_{-0*}Y = 0$ for all $t \in \mathbb{R}$, since ϕ_0 is the identity. We conclude that

$$\phi_{t*}Y = Y \qquad (t \in \mathbb{R})\,. \tag{15.12}$$

Taking any fixed t we recall that if ψ_s is the group generated by Y then the group generated by $\phi_{t*}Y$ is obtained by conjugation with ϕ_t: $\phi_t \circ \psi_s \circ \phi_t^{-1}$ and this must equal ψ_s. Hence $\phi_t \circ \psi_s = \psi_s \circ \phi_t$, for all $t, s \in \mathbb{R}$.

□

Proposition 15.2. *Let X, Y be complete vectorfields on $\mathcal{M}$. 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 $\mathcal{O}(t^3)$ with the 1-parameter group χ_r, generated by $[X,Y]$ for $r = t^2$. More generally,*

$$\psi_{-s} \circ \phi_{-t} \circ \psi_s \circ \phi_t = \chi_{st} + \mathcal{O}\big((s+t)^3\big)\,. \tag{15.13}$$

Proof. Since this is a local statement, we may confine our attention to a coordinate neighborhood of p. We choose coordinates so that

$$x^\mu|_p = 0\,. \tag{15.14}$$

The integral curves of X are represented by the solutions of

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = X^\mu(x(t))\,. \tag{15.15}$$

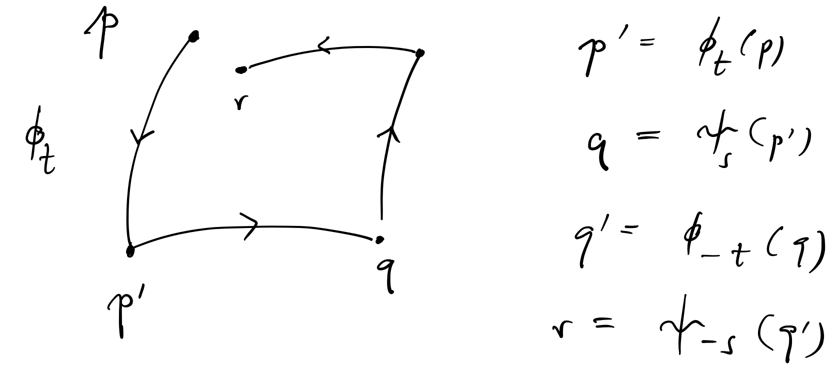


Figure 15.1.: Flow generated by commutator vectorfield.

We have

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}t^2} = \Big(\sum_\nu X^\nu \frac{\partial X^\mu}{\partial x^\nu}\Big)(x(t)) \tag{15.16}$$

Thus, using the Taylor expansion of the functions $x^\mu(t)$ at $t = 0$, along the integral curves of X,

$$\begin{aligned} x^\mu(t) =& x^\mu(0) + t\frac{\mathrm{d}x^\mu}{\mathrm{d}t}(0) + \frac{1}{2}t^2\frac{\mathrm{d}x^\mu}{\mathrm{d}t^2}(0) + \mathcal{O}(t^3) \\ =& x^\mu(0) + tX^\mu(x(0)) + \frac{1}{2}t^2\Big(\sum_\nu X^\nu \frac{\partial X^\mu}{\partial x^\nu}\Big)(x(0)) + \mathcal{O}(t^3) \end{aligned} \tag{15.17}$$

Similarly, the integral curves of Y are represented by the solutions of

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}s} = Y^\mu(x(s))\,. \tag{15.18}$$

We have

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}s^2} = \sum_\nu \frac{\partial Y^\mu}{\partial x^\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}s} = \Big(\sum_\nu Y^\nu \frac{\partial Y^\mu}{\partial x^\nu}\Big)(x(s)) \tag{15.19}$$

Again, using the Taylor expansion of the functions $x^\mu(s)$ at $s = 0$, along the integral curves of Y,

$$\begin{aligned} x^\mu(s) =& x^\mu(0) + s\frac{\mathrm{d}x^\mu}{\mathrm{d}s}(0) + \frac{1}{2}s^2\frac{\mathrm{d}x^\mu}{\mathrm{d}s^2}(0) + \mathcal{O}(s^3) \\ =& x^\mu(0) + sY^\mu(x(0)) + \frac{1}{2}s^2\Big(\sum_\nu Y^\nu \frac{\partial Y^\mu}{\partial x^\nu}\Big)(x(0)) + \mathcal{O}(s^3)\,. \end{aligned} \tag{15.20}$$

Consider the arc from p to p'; see Figure 15.1. This is an integral curve of X. Applying (15.17) with $x(0) = x|_p = 0$, we obtain

$$x^\mu|_{p'} = x^\mu(t) = tX^\mu|_p + \frac{1}{2}t^2\sum_\nu X^\nu|_p \frac{\partial X^\mu}{\partial x^\nu}|_p + \mathcal{O}(t^3) \tag{15.21}$$

Next we have the arc from p' to q; see Figure 15.1. This is an integral curve of Y. Applying (15.20) with $x(0) = x|_{p'}$, we obtain:

$$x^\mu|_q = x^\mu|_{p'} + sY^\mu|_{p'} + \frac{1}{2}s^2 \sum Y^\nu|_{p'} \frac{\partial Y^\mu}{\partial x^\nu}|_{p'} + \mathcal{O}(s^3) \tag{15.22}$$

and we have

$$\begin{aligned} Y^\mu|_{p'} &= Y^\mu|_p + \sum_\nu \frac{\partial Y^\mu}{\partial x^\nu}|_p(x^\nu|_{p'} - x^\nu|_p) + \mathcal{O}(|x^\nu|_{p'} - x^\nu|_p|^2) \\ &= Y^\mu|_p + t\sum_\nu X^\nu|_p \frac{\partial Y^\mu}{\partial x^\nu}|_p + \mathcal{O}(t^2)\,, \end{aligned} \tag{15.23}$$

where we also replaced the last terms at p' by the corresponding terms at p.

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

$$\begin{aligned} x^\mu|_q = tX^\mu|_p + sY^\mu|_p + \frac{1}{2}t^2 \sum_\nu X^\nu|_p \frac{\partial X^\mu}{\partial x^\nu}|_p \\ + st\sum_\nu X^\nu|_p \frac{\partial Y^\mu}{\partial x^\nu}|_p + \frac{1}{2}s^2 \sum Y^\nu|_p \frac{\partial Y^\mu}{\partial x^\nu}|_p + \mathcal{O}((s+t)^3) \end{aligned} \tag{15.24}$$

Next we have the arc from q to q'. This is an integral curve of X. Applying (15.17) with $x(0) = x|_q$ and t replaced by $-t$ we obtain:

$$x^\mu|_{q'} = x^\mu|_q - tX^\mu|_q + \frac{1}{2}t^2 \sum_\nu X^\nu|_q \frac{\partial X^\mu}{\partial x^\nu}|_q + \mathcal{O}(t^3) \tag{15.25}$$

and we have

$$\begin{aligned} X^\mu|_q &= X^\mu|_p + \sum_\nu \frac{\partial X^\mu}{\partial x^\nu}|_p(x^\nu|_q - x^\nu|_p) + \mathcal{O}(|x^\nu|_q - x^\nu|_p|^2) \\ &= X^\mu|_p + \sum_\nu (tX^\nu|_p + sY^\nu|_p) \frac{\partial X^\mu}{\partial x^\nu}|_p + \mathcal{O}((s+t)^2)\,. \end{aligned} \tag{15.26}$$

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

$$\begin{aligned} x^\mu|_{q'} = sY^\mu|_p + \frac{1}{2}t^2 \sum_\nu X^\nu|_p \frac{\partial X^\mu}{\partial x^\nu}|_p + st\sum_\nu X^\nu|_p \frac{\partial Y^\mu}{\partial x^\nu}|_p + \frac{1}{2}s^2 \sum Y^\nu|_p \frac{\partial Y^\mu}{\partial x^\nu}|_p \\ - t\sum_\nu (tX^\nu|_p + sY^\nu|_p) \frac{\partial X^\mu}{\partial x^\nu}|_p + \frac{1}{2}t^2 \sum_\nu X^\nu|_q \frac{\partial X^\mu}{\partial x^\nu}|_q + \mathcal{O}((s+t)^3) \\ = sY^\mu|_p + st\sum_\nu \Big(X^\nu|_p \frac{\partial Y^\mu}{\partial x^\nu}|_p - Y^\nu|_p \frac{\partial X^\mu}{\partial x^\nu}|_p\Big) + \frac{1}{2}s^2 \sum Y^\nu|_p \frac{\partial Y^\mu}{\partial x^\nu}|_p + \mathcal{O}((s+t)^3)\,. \end{aligned} \tag{15.27}$$

Finally we have the arc from q' to r. This is an integral curve of Y. Applying (15.20) with $x(0) = x|_{q'}$ and s replaced by $-s$ we obtain:

$$x^\mu|_r = x^\mu|_{q'} - sY^\mu|_{q'} + \frac{1}{2}s^2 \sum_\nu Y^\nu|_{q'} \frac{\partial Y^\mu}{\partial x^\nu}|_{q'} + \mathcal{O}(s^3)\,. \tag{15.28}$$

and we have

$$\begin{aligned} Y^\mu|_{q'} &= Y^\mu|_p + \sum_\nu \frac{\partial Y^\mu}{\partial x^\nu}|_p (x^\nu|_{q'} - x^\nu|_p) + \mathcal{O}(|x^\nu|_{q'} - x^\nu|_p|^2) \\ &= Y^\mu|_p + s\sum_\nu Y^\nu|_p \frac{\partial Y^\mu}{\partial x^\nu}|_p + \mathcal{O}((s+t)^2)\,. \end{aligned} \tag{15.29}$$

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

$$\begin{aligned} x^\mu|_r = sY^\mu|_p - sY^\mu|_p - s^2 \sum_\nu Y^\nu|_p \frac{\partial Y^\mu}{\partial x^\nu}|_p \\ + st\sum_\nu \Big(X^\nu|_p \frac{\partial Y^\mu}{\partial x^\nu}|_p - Y^\nu|_p \frac{\partial X^\mu}{\partial x^\nu}|_p\Big) + \frac{1}{2}s^2 \sum Y^\nu|_p \frac{\partial Y^\mu}{\partial x^\nu}|_p \\ + \frac{1}{2}s^2 \sum_\nu Y^\nu|_{q'} \frac{\partial Y^\mu}{\partial x^\nu}|_{q'} + \mathcal{O}((s+t)^3) \\ = st\sum_\nu \Big(X^\nu|_p \frac{\partial Y^\mu}{\partial x^\nu}|_p - Y^\nu|_p \frac{\partial X^\mu}{\partial x^\nu}|_p\Big) + \mathcal{O}((s+t)^3)\,. \end{aligned} \tag{15.30}$$

This proves the stated formula, because in local coordinates

$$X = \sum_\nu X^\nu \frac{\partial}{\partial x^\nu} \qquad Y = \sum_\nu Y^\nu \frac{\partial}{\partial x^\nu} \tag{15.31}$$

$$Xf = \sum_\nu X^\nu \frac{\partial f}{\partial x^\nu} \qquad Y = \sum_\nu Y^\nu \frac{\partial f}{\partial x^\nu} \tag{15.32}$$

and so

$$\begin{aligned} XYf =& \sum_\nu X^\nu \frac{\partial Yf}{\partial x^\nu} = \sum_{\nu,\mu} X^\nu \frac{\partial}{\partial x^\nu}\Big(Y^\mu \frac{\partial f}{\partial x^\mu}\Big) \\ =& \sum_{\nu,\mu}\Big(X^\nu Y^\mu \frac{\partial^2 f}{\partial x^\nu \partial x^\mu} + X^\nu \frac{\partial Y^\mu}{\partial x^\nu}\frac{\partial f}{\partial x^\mu}\Big) \end{aligned} \tag{15.33}$$

$$YXf = \sum_{\nu,\mu}\Big(Y^\nu X^\mu \frac{\partial^2 f}{\partial x^\nu \partial x^\mu} + Y^\nu \frac{\partial X^\mu}{\partial x^\nu}\frac{\partial f}{\partial x^\mu}\Big)\,. \tag{15.34}$$

Hence

$$[X,Y]f = \sum_{\nu,\mu}\Big(X^\nu \frac{\partial Y^\mu}{\partial x^\nu} - Y^\nu \frac{\partial X^\mu}{\partial x^\nu}\Big)\frac{\partial f}{\partial x^\mu}\,. \tag{15.35}$$

In conclusion,

$$x^\mu|_r = st[X,Y]^\mu|_p + \mathcal{O}((s+t)^3)\,. \tag{15.36}$$

□

Lecture 16.

2-forms and exterior derivatives

16.1. 2-forms

Let us define the bundle

$$\Lambda_2\mathcal{M} = \bigcup_{p\in\mathcal{M}} \Lambda_2(\mathrm{T}_p\mathcal{M}) \tag{16.1}$$

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

Definition 16.1. A 2-form ω on $\mathcal{M}$ is a continuously differentiable section of $\Lambda_2(\mathcal{M})$.

Outer product. If α and β are two elements of V^* we define $\alpha \wedge \beta \in \Lambda_2(V)$ by

$$(\alpha \wedge \beta) \cdot (u, v) = (\alpha \cdot u)(\beta \cdot v) - (\alpha \cdot v)(\beta \cdot u) \qquad u, v \in V\,. \tag{16.2}$$

Now if $(e_1, \dots, e_n)$ is a basis for V and $(e^{*1}, \dots, e^{*n})$ is the dual basis for V^* and, then the $\frac{n(n-1)}{2}$ elements of $\Lambda_2(V)$:

$$e^{*i} \wedge e^{*j} \qquad : \ i < j = 1, 2, \dots, n \tag{16.3}$$

form a basis for $\Lambda_2(V)$.

This applies in particular to the case $V = \mathrm{T}_p\mathcal{M}$. Suppose that $(\mathcal{U}, \varphi)$ is a chart of $\mathcal{M}$. Then the vectors

$$\frac{\partial}{\partial x^1}|_p, \dots, \frac{\partial}{\partial x^1}|_p \tag{16.4}$$

form a basis for $\mathrm{T}_p\mathcal{M}$, $p \in \mathcal{U}$ and the corresponding dual basis for $\mathrm{T}_p^*\mathcal{M}$ is

$$\mathrm{d}x^1|_p, \mathrm{d}x^2|_p, \dots, \mathrm{d}x^n|_p\,. \tag{16.5}$$

Therefore any 2-form ω can be locally in $\mathcal{U}$ expanded as

$$\omega = \sum_{\mu<\nu} \omega_{\mu\nu} \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu\,, \tag{16.6}$$

where the coefficients

$$\omega_{\mu\nu} = \omega\Big(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\Big) = -\omega_{\nu\mu} \tag{16.7}$$

form an antisymmetric matrix. We can also write:

$$\omega = \frac{1}{2} \sum_{\mu,\nu=1}^{n} \omega_{\mu\nu} \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu\,, \tag{16.8}$$

16.2. Exterior derivatives

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

$$\mathrm{d}\theta \cdot (X, Y) = X(\theta \cdot Y) - Y(\theta \cdot X) - \theta \cdot [X, Y]\,. \tag{16.9}$$

This defines a 2-form because it is bilinear with respect to multiplication by the ring of C^∞ functions on $\mathcal{M}$.

Exercise 16.1. Verify that for all $f \in \mathrm{C}^\infty(\mathcal{M})$, $\mathrm{d}\theta \cdot (fX, Y) = f\mathrm{d}\theta \cdot (X, Y)$, and $\mathrm{d}\theta \cdot (X, fY) = f\mathrm{d}\theta \cdot (X, Y)$.

Consider a chart $(\mathcal{U}, \varphi)$. Take X, Y to be local (defined in $\mathcal{U}$) vectorfields,

$$X = \frac{\partial}{\partial x^\mu}\,, Y = \frac{\partial}{\partial x^\nu}\,. \tag{16.10}$$

Then

$$\theta \cdot X = \theta_\mu\,, \theta \cdot Y = \theta_\nu\,, \tag{16.11}$$

and the definition reduces to

$$(\mathrm{d}\theta)_{\mu\nu} = \frac{\partial \theta_\nu}{\partial x^\mu} - \frac{\partial \theta_\mu}{\partial x^\nu}\,. \tag{16.12}$$

Lie derivatives. Comparing the formulas for $\pounds_X\, \theta$ and $\mathrm{d}\theta$ we also see that

$$\pounds_X\, \theta = i_X \mathrm{d}\theta + \mathrm{d}\, i_X \theta\,. \tag{16.13}$$

Here i_X denotes the 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

$$i_X\omega \cdot Y = \omega \cdot (X, Y) \qquad Y \in \mathscr{X}^\infty(\mathcal{M})\,. \tag{16.14}$$

Exercise 16.2. Verify this identity.

Part III.

Lecture 17.

Lie Groups

17.1. Lie Groups

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

Group multiplications:

$$G \times G \longrightarrow G\,, \quad (a,b) \mapsto ab \tag{17.1}$$

Taking the inverse:

$$G \longrightarrow G\,, \quad a \mapsto a^{-1} \tag{17.2}$$

are smooth mappings.

We have two distinguished groups of diffeomorphisms of G onto itself:

Right multiplications: $\{r_a : a \in G\}$,

$$r_a(b) = ba \qquad (b \in G)\,. \tag{17.3}$$

Left multiplications: $\{l_a : a \in G\}$,

$$l_a(b) = ab \qquad (b \in G)\,. \tag{17.4}$$

Note that $r_a \circ l_b = l_b \circ r_a$ for all $a, b \in G$.

17.2. Lie Algebra

This is the following distinguished subspace of $\mathscr{X}^\infty(G)$:

$$\mathscr{G} = \Big\{X \in \mathscr{X}^\infty(G) : (l_a)_* X = X\,, a \in G\Big\} \tag{17.5}$$

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

Proposition 17.1. *If $X, Y \in \mathscr{G}$ then $[X, Y] \in \mathscr{G}$.*

Proof. Let $Z = [X, Y] \in \mathscr{X}^\infty(G)$. Recall that we have proven in Lecture 15 that for any diffeomorphism ϕ,

$$\phi_*[X, Y] = [\phi_* X, \phi_* Y]\,. \tag{17.6}$$

We can apply this with $\phi = l_a, a \in G$, then

$$(l_a)_* Z = [(l_a)_* X, (l_a)_* Y] = [X, Y] = Z\,, \quad a \in G\,. \tag{17.7}$$

Therefore $Z \in \mathscr{G}$. □

17.3. Construction of left invariant vectorfields

Let us denote by e the identity element in G. Let $X \in \mathscr{G}$. Then

$$l_{a*}X = X\,, \quad a \in G\,. \tag{17.8}$$

We have

$$(l_{a*}X)(l_a(b)) = \mathrm{d}l_a \cdot X(b)\,, \qquad b \in G\,. \tag{17.9}$$

Set $b = e$ to obtain

$$(l_{a*}X)(a) = \mathrm{d}l_a \cdot X(e)\,, \qquad a \in G\,. \tag{17.10}$$

But $l_{a*}X = X$, so we conclude that

$$X(a) = \mathrm{d}l_a \cdot X(e)\,, \qquad a \in G\,. \tag{17.11}$$

This defines X on G given its value at e.

Consider then the evaluation map ϵ_a at $a \in G$, restricted to $\mathscr{G}$.

Remark 17.1. In general, the **evaluation map** is defined by

$$\epsilon_p : \mathscr{X}^\infty(\mathcal{M}) \to \mathrm{T}_p\mathcal{M}\,, \quad \epsilon_p(X) = X(p)\,. \tag{17.12}$$

Since l_a is a diffeomorphism, $\mathrm{d}l_a$ is an isomorphism of T_eG onto T_aG. It follows that ϵ_a restricted to $\mathscr{G}$ is an isomorphism of $\mathscr{G}$ onto T_aG. Therefore, at each $a \in G$, $\epsilon_a^{-1}(v)$ is some element of $\mathscr{G}$ for each tangent vector $v \in \mathrm{T}_aG$. In other words, given any tangent vector $v \in \mathrm{T}_aG$, a an arbitrary element of G, there is a unique $X \in \mathscr{G}$ such that $X(a) = v$.

Remark 17.2. This implies that any Lie group is *parallelizable*. Given a basis for T_eG, the corresponding left invariant vectorfields define a basis for $\mathscr{G}$, hence (since ϵ_a is an isomorphism at each $a \in G$) a *frame* at each $a \in G$.

Given $X \in \mathscr{G}$, let us denote by ${}^{(X)}\gamma : t \mapsto {}^{(X)}a_t$ the integral curve of X through e.

Proposition 17.2. *$\{{}^{(X)}a_t : t \in \mathbb{R}\}$ is a 1-parameter subgroup of G.*

Proof. Recall that if γ_p is the integral curve of X through p, the $\phi \circ \gamma_p$ is the integral curve of $\phi_* X$ through $\phi(p)$.

Let us fix a number s. Then (taking $\phi = l_{{}^{(X)}a_s}$) the integral curve of

$$l_{{}^{(X)}a_s *}X = X\,, \tag{17.13}$$

through $l_{{}^{(X)}a_s}(e) = {}^{(X)}a_s$ is

$$l_{{}^{(X)}a_s} \circ {}^{(X)}\gamma : t \mapsto l_{{}^{(X)}a_s}({}^{(X)}a_t) = {}^{(X)}a_s {}^{(X)}a_t\,. \tag{17.14}$$

It follows that

$${}^{(X)}a_{s+t} = {}^{(X)}a_s {}^{(X)}a_t\,. \tag{17.15}$$

□

Furthermore, from the proof, we also have, by taking $\phi = l_b$, $b \in G$ that the integral curve of $l_{b*}X = X$ through $l_b(e) = b$ is

$$l_b \circ {}^{(X)}\gamma : t \mapsto l_b({}^{(X)}a_t) = b^{(X)}a_t \,. \tag{17.16}$$

We have thus proved:

Proposition 17.3. *The integral curve of X through $b \in G$ is*

$$t \mapsto b^{(X)}a_t \,, \tag{17.17}$$

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. Thus the 1-parameter group of diffeomorphisms generated by X is:

$$\left\{ r_{{}^{(X)}a_t} : t \in \mathbb{R} \right\} . \tag{17.18}$$

Thus the left invariant vectorfields generate right multiplications.

17.4. Examples

17.4.1. Affine group

The simplest non-trivial example is the affine group of the real line

$$s \mapsto e^x s + y \,, \tag{17.19}$$

namely scaling by e^x, and translating by y.

The group manifold is here $G = \mathbb{R}^2$, and points are pairs (x, y). The group multiplication $(x_1, y_1) \cdot (x_2, y_2)$ is defined as follows: Since

$$(x_2, y_2)s = s' = e^{x_2}s + y_2 \tag{17.20}$$

and

$$(x_1, y_1)(x_2, y_2)s = (x_1, y_1)s' = e^{x_1}s' + y_1 = e^{x_1+x_2}s + y_1 + e^{x_1}y_2 \,, \tag{17.21}$$

we set

$$(x_1, y_1)(x_2, y_2) = (x_1 + x_2, y_1 + e^{x_1}y_2) \,. \tag{17.22}$$

The 1-parameter subgroups are

Scaling $s \mapsto e^t$, generated by

$$\left\{ (t, 0) : t \in \mathbb{R} \right\} \subset G \tag{17.23}$$

Translation $s \mapsto s + t$, generated by

$$\left\{ (0, t) : t \in \mathbb{R} \right\} \subset G \,. \tag{17.24}$$

Right multiplication by the scaling subgroup:

$$(x, y) \mapsto (x, y)(t, 0) = (x + t, y)\,. \tag{17.25}$$

The corresponding left invariant vectorfield is

$$X = \frac{\partial}{\partial x}\,. \tag{17.26}$$

Right multiplication by the translation subgroup:

$$(x, y) \mapsto (x, y)(0, t) = (x, y + e^x t) \tag{17.27}$$

The corresponding left invariant vectorfield is

$$Y = e^x \frac{\partial}{\partial y}\,. \tag{17.28}$$

Then

$$[X, Y] = Y\,. \tag{17.29}$$

17.4.2. Heisenberg group

Here the group manifold is

$$G = \Big\{(x, y, z)\Big\} = \mathbb{R}^3 \tag{17.30}$$

and G acts on 1-dimensional square integrable complex valued functions on $\mathbb{R}$:

$$((x, y, z)\psi)(s) = e^{iys+iz}\psi(s + x) \tag{17.31}$$

See problems below for the description of the left invariant vectorfields X, Y, Z, satisfying

$$[X, Y] = Z\,. \tag{17.32}$$

This appears in 1-dimensional quantum mechanics.

17.4.3. Applications: Crystal dislocations

The affine group and the Heisenberg group also appear in the continuum description of crystal dislocations.

Edge dislocations. Suppose that an extra half-line of atoms has been inserted along the negative 1st axis; see Figure 17.1.

In a 3-dimensional crystal an extra half-plane would have been inserted.

A circuit of translations alternately along the 1st and 2nd axes (say m along the 1st, n along the second, $-m$ along the 1st, and $-n$ along the 2nd) which does not include the origin O closes. However, a circuit of translations alternately along the 1st and 2nd

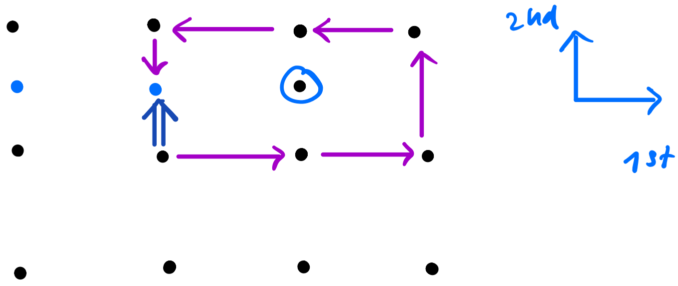


Figure 17.1.: Edge dislocations.

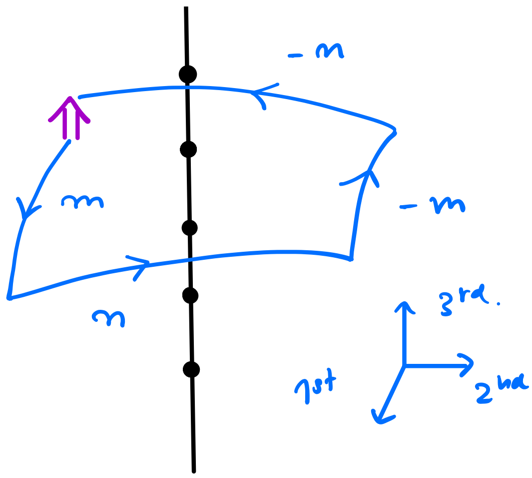


Figure 17.2.: Screw dislocations.

axes does not close if the circuit contains the origin O. We arrive at an atom which can be reached in a single step along the 2nd axis.

The origin is a *dislocation point*. It is actually a dislocation line in 3 dimensions, the edge of the extra half plane. The single translation by which the final point is reached from the initial point is called **Burgers vector**.

A *uniform distribution* of **edge** dislocations is described by

$$[X, Y] = Y \tag{17.33}$$

Screw dislocation. This appears in a 3-dimensional crystal; see Figure 17.2.

Here a curcuit of translations alternately along the 1st and 2nd axes which includes the 3rd axis does not close. In fact we arrive at an atom which can be reached ina single step along the 3rd axis by 11.

Exercise 17.1. Give a similar description of the lattice structure for screw dislocations as we have given for edge dislocations.

A *uniform distribution* of **screw** dislocations is described by

$$[X, Y] = Z \tag{17.34}$$

Problems

1. Let the Heisenberg group be defined as follows: Every triple $(x, y, z) \in G = \mathbb{R}^3$ acts on the Lebesgue space of square summable complex valued functions $\mathrm{L}^2(\mathbb{R}; \mathbb{C})$ as follows:

$$((x, y, z)\psi)(s) = e^{iys+iz}\psi(s + x)\,, \qquad s \in \mathbb{R}\,, \qquad \psi \in \mathrm{L}^2(\mathbb{R}; \mathbb{C})\,. \tag{17.35}$$

   a) Show that the definition is well-posed, meaning that $(x, y, z)\psi \in \mathrm{L}^2(\mathbb{R}; \mathbb{C})$.

   b) Find the group multiplication law, namely given $(x_i, y_i, z_i) : i = 1, 2$, define the product so that

$$(x_1, y_1, z_1)\big((x_2, y_2, z_2)\psi\big) = \big((x_1, y_1, z_1)(x_2, y_2, z_2)\big)\psi\,. \tag{17.36}$$

   c) Show that

$$G_1 = \big\{(x, 0, 0) : x \in \mathbb{R}\big\} \tag{17.37}$$
$$G_2 = \big\{(0, y, 0) : y \in \mathbb{R}\big\} \tag{17.38}$$
$$G_3 = \big\{(0, 0, z) : z \in \mathbb{R}\big\} \tag{17.39}$$

   are Lie subgroups of G.

   d) Find the 3 left invariant vectorfields X, Y, Z on G generating G_1, G_2, and G_3, respectively.

   e) Show that

$$[Y, Z] = [Z, X] = 0\,, \qquad [X, Y] = Z\,. \tag{17.40}$$

2. Given a Lie group G, recall that the evaluation map $\epsilon_a : \mathscr{G} \to T_aG$ given by $\epsilon_a(X) = X(a)$ is an isomorphism.

   Define the Lie algebra valued 1-form ν by

$$\nu(v) = \epsilon_a^{-1}(v) \in \mathscr{G} \qquad (v \in \mathrm{T}_aG)\,. \tag{17.41}$$

   Show that

$$\mathrm{d}\nu = -\lambda \tag{17.42}$$

   where λ is the $\mathscr{G}$-valued 2-form given by

$$\lambda(v_1, v_2) = [\epsilon_a^{-1}(v_1), \epsilon_a^{-1}(v_2)] \in \mathscr{G}\,, \tag{17.43}$$

   for every $v_1, v_2 \in \mathrm{T}_aG$.

Lecture 18.

Metrics

Quadratic forms. Given a *real* vector space V we denote by $S_2(V)$ the space of symmetric bilinear forms on V (also called **quadratic forms** on V). If $(e_1, \ldots, e_n)$ is a basis for V, then given $h \in S_2(V)$,

$$h_{ij} = h(e_i, e_j) = h_{ji} \tag{18.1}$$

are the components of h in this basis.

Tensor products. If $\alpha, \beta \in V^*$ we define the **tensor product** $\alpha \otimes \beta$, a bilinear form on V, by

$$(\alpha \otimes \beta) \cdot (u, v) = \alpha(u)\beta(v) \qquad u, v \in V\,. \tag{18.2}$$

If $(e^{*1}, \ldots, e^{*n})$ is the basis for V^* which is dual to the basis $(e_1, \ldots, e_n)$ for V, we have the expansion

$$h = \sum_{i,j=1}^{n} h_{ij} e^{*i} \otimes e^{*j} = \frac{1}{2} \sum_{i,j=1}^{n} h_{ij} \Big(e^{*i} \otimes e^{*j} + e^{*j} \otimes e^{*i} \Big) \tag{18.3}$$

The bilinear form $(1/2)(\alpha \otimes \beta + \beta \otimes \alpha) \in S_2(V)$ is the *symmetric tensor product* of α and β.

If now $\mathcal{B} = \bigcup_{p \in \mathcal{M}} \mathcal{B}_p$ is a vector bundle over a manifold $\mathcal{M}$, then

$$S_2(\mathcal{B}, \mathcal{M}) = \bigcup_{p \in \mathcal{M}} S_2(\mathcal{B}_p) \tag{18.4}$$

is also a vector bundle over $\mathcal{M}$.

Inner products. Going back to the definition of $S_2(V)$, we denote by $S_2^+(V)$ the subset of $S_2(V)$ consisting of those quadratic forms which are positive definite:

Definition 18.1. We say $h \in S_2^+(V)$ if $h \in S_2(V)$ and $h(v, v) \geq 0$ with equality if and only if $v = 0$.

Thus $S_2^+(V)$ is the space of inner products on V.

Proposition 18.1. *$S_2^+(V)$ is an open positive convex cone in $S_2(V)$.*

That $S_2^+(V)$ is a positive cone means: $h \in S_2^+(V)$ and $\lambda > 0$ implies $\lambda h \in S_2^+(V)$. Whereas convex means: 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 $S_2^+(V)$:

$$(1-\lambda)h_1 + \lambda h_2 \in S_2^+(V) \qquad \lambda \in [0,1]\,. \tag{18.5}$$

Definition 18.2 (Metric)**.** A *metric* m on a real vector bundle $\mathcal{B}$ is a continuously differentiable section of

$$S_2^+(\mathcal{B}, \mathcal{M}) = \bigcup_{p \in \mathcal{M}} S_2^+(\mathcal{B}_p)\,, \tag{18.6}$$

and open subbundle of $S_2(\mathcal{B}, \mathcal{M})$.

Thus m is a continuously differentiable assignment of an inner product, m_p, in $\mathcal{B}_p$, at each $p \in \mathcal{M}$.

Given two continuously differentiable sections σ, τ of $\mathcal{B}$, $m(\sigma, \tau)$ defined pointwise,

$$m(\sigma,\tau)(p) = m_p(\sigma(p), \tau(p)) \tag{18.7}$$

is a continuously differentiable function on $\mathcal{M}$.

Riemannian metric. In the case $\mathcal{B} = \mathrm{T}\mathcal{M}$, $m = g$ is called a **Riemannian metric** on $\mathcal{M}$. So g is a continuously differentiable section of

$$S_2^+\mathcal{M} = \bigcup_{p \in \mathcal{M}} S_2^+(\mathrm{T}_p\mathcal{M})\,, \tag{18.8}$$

a continuously differentiable assignment of an inner product, g_p, in $\mathrm{T}_p\mathcal{M}$, at each $p \in \mathcal{M}$.

If $(\mathcal{U}, \varphi)$ is a chart on $\mathcal{M}$, and

$$g_{ij} = g\Big(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^i}\Big) \tag{18.9}$$

are the components of g in the coordinates basis, then we can expand g as follows:

$$g = \sum_{i,j=1}^{n} g_{ij} \mathrm{d}x^i \otimes \mathrm{d}x^j\,. \tag{18.10}$$

Remark 18.1. The tensor product $\otimes$ is often dropped from the notation, and we simply write:

$$g = \sum_{i,j=1}^{n} g_{ij} \mathrm{d}x^i \mathrm{d}x^j\,. \tag{18.11}$$

Remark 18.2. The bundle

$$S_2\mathcal{M} = \bigcup_{p \in \mathcal{M}} S_2(\mathrm{T}_p\mathcal{M}) \tag{18.12}$$

is a **tensor bundle** over $\mathcal{M}$, and its sections are called *symmetric 2-covariant tensors* on $\mathcal{M}$. Similarly $\Lambda_1\mathcal{M} = \mathrm{T}^*\mathcal{M}$ is a tensor bundle, its sections being 1-forms, or *1-covariant tensors* on $\mathcal{M}$. $\Lambda_2(\mathcal{M})$ is also tensor bundle, whose sections are 2-forms on $\mathcal{M}$, or *antisymmetric 2-covariant tensors.*

Local trivialisation and linear isometry. Recall that for any real vector bundle $\mathcal{B}$ over $\mathcal{M}$, we have local trivialisations $(\mathcal{U}, \omega)$, where $\mathcal{U}$ is an open domain in $\mathcal{M}$ and ω is a diffeomorphism

$$\omega : \pi^{-1}(\mathcal{U}) \longrightarrow \mathcal{U} \times \mathbb{R}^n \tag{18.13}$$

where π is the projection of the bundle, and for $b \in \mathcal{B}_p = \pi^{-1}(p)$, $\omega(b) = (p, \chi_p(b))$, and for each $p \in \mathcal{U}$, $\chi_p : \mathcal{B}_p \to \mathbb{R}^n$ is a linear isomorphism.

Suppose now that $\mathcal{B}$ is equipped with a metric m. Since $\mathbb{R}^n$ is endowed with the standard inner product $\langle \cdot, \cdot \rangle$, we then require χ_p to *preserve norms*:

$$\sqrt{m_p(b, b)} = \|\chi_p(b)\| \tag{18.14}$$

Here, for $\alpha = (\alpha^1, \ldots, \alpha^n) \in \mathbb{R}^n$, and $\beta = (\beta^1, \ldots, \beta^n) \in \mathbb{R}^n$,

$$\langle \alpha, \beta \rangle = \sum_{\mu=1}^{n} \alpha^\mu \beta^\mu \, , \tag{18.15}$$

and $\| \cdot \|$ denotes the corresponding norm: $\|\alpha\| = \sqrt{\langle \alpha, \alpha \rangle}$.

We say that χ_p is a *linear isometry*, for each $p \in \mathcal{U}$. It follows that

$$m_p(u, v) = \langle \chi_p(u), \chi_p(v) \rangle \, , \qquad u, v \in \pi^{-1}(p) = \mathcal{B}_p \, . \tag{18.16}$$

Orthonormal basis sections. Let e_μ be the standard basis for $\mathbb{R}^n$. This is an orthonormal basis relative to the standard inner product $\langle \cdot, \cdot \rangle$: $\langle e_\mu, e_\nu \rangle = \delta_{\mu\nu}$. We then define the local basis sections σ_μ of $\mathcal{B}$ by:

$$\sigma_\mu(p) = \chi_p^{-1} \cdot e_\mu : \qquad p \in \mathcal{U} \, . \tag{18.17}$$

We then have

$$m_p(\sigma_\mu(p), \sigma_\nu(p)) = \langle e_\mu, e_\nu \rangle = \delta_{\mu\nu} : \qquad p \in \mathcal{U} \tag{18.18}$$

Thus the sections $\sigma_\mu : \mu = 1, 2, \ldots, n$ form an orthonormal basis for $\mathcal{B}_p$, at each $p \in \mathcal{U}$.

In the case $\mathcal{B} = \mathrm{T}\mathcal{M}$ a section σ of $\mathcal{B}$ is a vectorfield X, and the set $(\sigma_\mu : \mu = 1, \ldots, n)$ is an *orthonormal framefield* $(X_1, \ldots, X_n)$. Here $m = \dim \mathcal{M}$, and $n = \dim \mathcal{B}_p = \dim \mathrm{T}_p\mathcal{M} = m$.

Complex vector bundles. The above considerations hold for real vector bundles. In the case of a *complex* vector bundle $\mathcal{B}$, a metric m on $\mathcal{B}$ is a continuously differentiable assignment of a *Hermitian* inner product m_p in $\mathcal{B}_p$, at each $p \in \mathcal{M}$.

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 antilinear in u and linear in v, namely $h(u_1 + u_2, v) = h(u_1, v) + h(u_2, v)$ but $h(\alpha u, v) = \overline{\alpha} h(u, v)$ for all $\alpha \in \mathbb{C}$.

2. $h(v, u) = \overline{h(u, v)}$. In particular $h(v, v)$ is real.

3. $h(v, v) \geq 0$ with equality if and only if $v = 0$.

We denote by $\langle \cdot, \cdot \rangle$ the standard Hermitian inner product on $\mathbb{C}^n$,

$$\langle \alpha, \beta \rangle = \sum_{\mu=1}^{n} \overline{\alpha}^{\mu} \beta^{\mu} \tag{18.19}$$

for $\alpha = (\alpha^1, \ldots, \alpha^n) \in \mathbb{C}^n$, and $\beta = (\beta^1, \ldots, \beta^n) \in \mathbb{C}^n$. We also denote by $(e_\mu : \mu = 1, 2, \ldots, n)$ the standard basis of $\mathbb{C}^n$:

$$e_\mu = (0, \ldots, 1, \ldots, 0) : \mu = 1, \ldots, n\,. \tag{18.20}$$

This is orthonormal relative to the standard inner product: $\langle e_\mu, e_\nu \rangle = \delta_{\mu\nu}$.

When considering bundle charts $(\mathcal{U}, \omega)$ we require $\chi_p : \mathcal{B}_p \to \mathbb{C}^n$ to be a linear isometry

$$m_p(u, v) = \langle \chi_p \cdot u, \chi_p \cdot v \rangle \;\; : \qquad u, v \in \mathcal{B}_p = \pi^{-1}(p)\,. \tag{18.21}$$

Then again the local basis sections σ_μ defined by

$$\sigma_\mu(p) = \chi_p^{-1} \cdot e_\mu \;\; : \quad p \in \mathcal{U}, \mu = 1, \ldots, n \tag{18.22}$$

define an orthonormal basis for $\mathcal{B}_p$ at each $p \in \mathcal{U}$:

$$m_p(\sigma_\mu(p), \sigma_\nu(p)) = \delta_{\mu\nu} \;\; : \qquad p \in \mathcal{U}\,. \tag{18.23}$$

Sections of complex vector bundles. Any section ϕ of $\mathcal{B}$ can be locally — over $\mathcal{U}$ — expanded as

$$\phi = \sum_{\mu=1}^{n} \phi^\mu \sigma_\mu \tag{18.24}$$

The coefficients ϕ^μ are complex-valued functions defined in $\mathcal{U}$. Let ψ be another section of $\mathcal{B}$, then

$$\psi = \sum_{\mu=1}^{n} \psi^\mu \sigma_\mu \tag{18.25}$$

Then the function $m(\phi, \psi)$ is given in $\mathcal{U}$ by:

$$\begin{aligned} m(\phi, \psi) =& m(\sum_{\mu=1}^{n} \phi^\mu \sigma_\mu, \sum_{\nu=1}^{n} \psi^\nu \sigma_\nu) \\ =& \sum_{\mu,\nu=1}^{n} \overline{\phi}^\mu \psi^\nu m(\sigma_\mu, \sigma_\nu) = \sum_{\mu=1}^{n} \overline{\phi}^\mu \psi^\mu \end{aligned} \tag{18.26}$$

Line bundles. In the case $n = 1$, we have a complex line bundle. Then $e_1 = 1$ and σ_1 is defined by:

$$\sigma_1(p) = \chi_p^{-1} \cdot 1 \tag{18.27}$$

Thus σ_1 is only subject to the condition

$$m_p(\sigma_1(p), \sigma_1(p)) = \|e_1\|^2 = 1\,. \tag{18.28}$$

That means $\sigma_1(p)$ belongs to the unit circle in $\mathcal{B}_p$; here $\mathcal{B}_p$ is isomorphic to $\mathbb{C}$.

The ambiguity in the choice of σ_1 is reflected in the arbitrariness of phase of the function ϕ^1,

$$\phi = \phi^1 \sigma_1\,, \tag{18.29}$$

where $\sigma_1(p) \mapsto e^{i\alpha}\sigma_1(p)$, for $\alpha \in \mathbb{R}$, corresponds to $\phi^1(p) \mapsto e^{-i\alpha}\phi^1(p)$.

Problems

1. The purpose of this exercise is to define a left invariant metric on a Lie group G. We know that the Lie algebra $\mathscr{G}$ of G is a finite dimensional vector space. Let us then consider any inner product on $\mathscr{G}$ and define the inner product g_a on T_aG as

$$g_a(u, v) = \gamma\big(\epsilon_a^{-1}(u), \epsilon_a^{-1}(v)\big)\,, \qquad u, v \in \mathrm{T}_aG\,, \tag{18.30}$$

where $\epsilon_a : \mathscr{G} \to \mathrm{T}_aG$ is the evaluation map

$$\epsilon_a(X) = X(a)\,. \tag{18.31}$$

Show that the 2-covariant tensor on G thus defined is invariant under left multiplication, namely

$$(l_a)^* g = g \qquad (a \in G) \tag{18.32}$$

where $l_a(b) = ab$ for every $b \in G$.

Note: The pull-back of a r-covariant tensor W on a manifold $\mathcal{M}$ by a diffeomorphism ϕ is defined as follows:

$$(\phi^* W)(p)(v_1, \ldots, v_n) = W(\phi(p))(\mathrm{d}\phi(p) \cdot v_1, \ldots, \mathrm{d}\phi(p) \cdot v_r)\,. \tag{18.33}$$

Lecture 19.

Arc length and volume

The concepts of *arc length* and *volume* refer exclusively to the case $\mathcal{B} = \mathrm{T}\mathcal{M}$.

19.1. Arc length

The arc length of a curve $\gamma : I = [a, b] \to \mathcal{M}$ is defined by

$$L(\gamma[a,b]) = \int_a^b \sqrt{g_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))}\mathrm{d}t \tag{19.1}$$

where $\dot\gamma(t)$ is the tangent vector to γ at $\gamma(t)$.

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 $\mathcal{M}$.

For let $\tilde{I} = [\tilde{a}, \tilde{b}]$ be another closed interval and $\tilde\gamma : \tilde{I} \to \mathcal{M}$ be another curve giving the same path in $\mathcal{M}$: a reparametrization of γ. Then the end points coincide

$$\{\tilde\gamma(\tilde a), \tilde\gamma(\tilde b)\} = \{\gamma(a), \gamma(b)\}\,. \tag{19.2}$$

Thus either $\tilde\gamma(\tilde a) = \gamma(a)$ and $\tilde\gamma(\tilde b) = \gamma(b)$ or $\tilde\gamma(\tilde a) = \gamma(b)$ and $\tilde\gamma(\tilde b) = \gamma(a)$. Let us assume that the first alternative holds. Then there is an increasing diffeomorphism

$$f : \tilde I \to I \quad \text{s.t.} \quad \tilde\gamma = \gamma \circ f\,. \tag{19.3}$$

Let us set $s = f(t)$, then

$$\dot{\tilde\gamma}(t) = f'(t)\dot\gamma(s)\,. \tag{19.4}$$

We verify this as follows: Let $\phi \in \mathrm{C}^\infty(\mathcal{M})$, then

$$\begin{aligned}\dot{\tilde\gamma}(t)\cdot\phi &= \frac{\mathrm{d}}{\mathrm{d}t'}\phi(\tilde\gamma(t'))|_{t'=t} \\ &= \frac{\mathrm{d}}{\mathrm{d}t'}\phi(\gamma(f(t')))|_{t'=t} = \frac{\mathrm{d}}{\mathrm{d}s'}\phi(\gamma(s'))|_{s'=f(t)}\frac{\mathrm{d}f(t')}{\mathrm{d}t'}|_{t'=t} \\ &= f'(t)\dot\gamma(s)\cdot\phi\,.\end{aligned} \tag{19.5}$$

Thus

$$\begin{aligned} L(\gamma[a,b]) &= \int_a^b \sqrt{g_{\gamma(s)}(\dot\gamma(s),\dot\gamma(s))}\mathrm{d}s \\ &= \int_{\tilde a}^{\tilde b} \sqrt{g_{\tilde\gamma(t)}(\dot\gamma(f(t)),\dot\gamma(f(t)))} f'(t)\mathrm{d}t \\ &= \int_{\tilde a}^{\tilde b} \sqrt{g_{\tilde\gamma(t)}(\dot{\tilde\gamma}(t),\dot{\tilde\gamma}(t))}\mathrm{d}t = L(\tilde\gamma[\tilde a,\tilde b])\,. \end{aligned} \tag{19.6}$$

19.2. Volume

Let V be a vector space. Given a basis $(e_1,\ldots,e_n)$ we want to assign a volume to the corresponding parallelepiped.

Orientation. Let V be a vector space, and $(e_1,\ldots,e_n)$ a basis for V. Any other basis $(e'_1,\ldots,e'_n)$ can be expressed in terms of the original basis:

$$e'_a = \sum_{b=1}^n e_b S^b_a \qquad a = 1,\ldots,n \tag{19.7}$$

The coefficients S^b_a form a non-singular matrix, because we can equally well express the original basis $(e_1,\ldots,e_n)$ in terms of the new basis $(e'_1,\ldots,e'_n)$, thus S must be invertible:

$$\det S \neq 0\,. \tag{19.8}$$

Recall that if $\det(S) > 0$ we say that the new basis has the same *orientation* as the old. If $\det(S) < 0$ we say that the new basis has the opposite orientation. This partitions the bases for V into two classes. As we have seen and *orientation* for V is a choice of a class which we call *positive*. The corresponding bases are called *positive*.

Definition 19.1. A *volume form* ω on an *oriented* vector space V is a totally antisymmetric n-linear form ($n = \dim V$) on V such that

$$\omega(e_1,\ldots,e_n) > 0 \tag{19.9}$$

for any positive basis $(e_1,\ldots,e_n)$ for V.

Example 19.1. Let $n = 2$. Then the change of bases formula takes the form

$$e'_1 = e_1 S^1_1 + e_2 S^2_1 \tag{19.10}$$

$$e'_1 = e_1 S^1_2 + e_2 S^2_2 \tag{19.11}$$

and we have

$$\begin{aligned} \omega(e'_1,e'_2) &= \omega(e_1 S^1_1 + e_2 S^2_1, e_1 S^1_2 + e_2 S^2_2) \\ &= S^1_1 S^2_2\ \omega(e_1,e_2) + S^2_1 S^1_2\ \omega(e_2,e_1) \\ &= (S^1_1 S^2_2 - S^1_2 S^2_1)\ \omega(e_1,e_2) \\ &= \det(S)\ \omega(e_1,e_2)\,. \end{aligned} \tag{19.12}$$

In general, for any n, we have

$$\omega(e'_1, \dots, e'_n) = \det(S)\omega(e_1, \dots, e_n) \tag{19.13}$$

This follows from the properties of the determinant.

As a consequence, once ω is positive on one positive basis it is positive on all positive bases.

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 $(e_1, \dots, e_n)$ is a positive *orthonormal* basis w.r.t. $\langle \cdot, \cdot \rangle$ for V. So $\langle e_a, e_b \rangle = \delta_{ab} : a = 1, 2, \dots, n$. We set

$$\omega(e_1, \dots, e_n) = 1\,. \tag{19.14}$$

This makes sense, because if $(e'_1, \dots, e'_n)$ is another positive orthonormal basis then

$$e'_a = \sum_{b=1}^{n} e_b O^b_a \tag{19.15}$$

where O is an orthogonal matrix with $\det O > 0$, hence $\det O = 1$. Indeed,

$$\begin{aligned} \delta_{ab} =& \langle e'_a, e'_b \rangle = \sum_{c,d=1}^{n} O^c_a O^d_a \langle e_c, e_d \rangle \\ =& \sum_{c=1}^{n} O^c_a O^c_a = (O^T O)^a_b \end{aligned} \tag{19.16}$$

so $O^T O = \mathbb{1}$, hence $\det O^T \det O = 1$, or $(\det O)^2 = 1$, so $\det O = 1$ since $\det O > 0$.

Volume form of an orientable Riemannian manifold. Let $(\mathcal{M}, g)$ be an m-dimensional oriented manifold. Then at each $p \in \mathcal{M}$ the inner product g_p in $\mathrm{T}_p\mathcal{M}$ together with the orientation of $\mathrm{T}_p\mathcal{M}$, define a volume form ω_p in $\mathrm{T}_p\mathcal{M}$, a totally antisymmetric m-linear ($m = \dim \mathcal{M}$) form on $\mathrm{T}_p\mathcal{M}$ such that $\omega_p(e_1, \dots, e_n) = 1$ on any positive orthonormal basis $(e_1, \dots, e_m)$.

We thus have a section ω of the bundle $\Lambda_m\mathcal{M}$ of top degree forms on $\mathcal{M}$ (a tensor bundle over $\mathcal{M}$). ω is continuously differentiable if g is continuously differentiable, and ω is the **volume form of** $(\mathcal{M}, g)$, denoted by

$$\mathrm{d}\mu_g\,. \tag{19.17}$$

In local coordinates, we express $\frac{\partial}{\partial x^\mu} : \mu = 1, \dots, m$, the coordinate vectorfields of a chart, in terms of a positive orthonormal frame field $(e_\mu : \mu = 1, \dots, m)$ defined in the domain of the chart:

$$\frac{\partial}{\partial x^\mu} = \sum_{\nu=1}^{m} e_\nu S^\nu_\mu\,. \tag{19.18}$$

Let

$$g_{\mu\nu} = g\Big(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\Big) \tag{19.19}$$

be the components of g in the chart. Then

$$g_{\mu\nu} = g\Big(\sum_{\kappa=1}^{m} e_\nu S_\mu^\kappa, \sum_{\lambda=1}^{m} e_\lambda S_\nu^\lambda\Big) = \sum_{\kappa,\lambda=1}^{n} S_\mu^\kappa S_\nu^\lambda g(e_\kappa, e_\lambda) = \sum_{\kappa=1}^{n} S_\mu^\kappa S_\nu^\kappa\,. \tag{19.20}$$

In terms of matrices, $g = S^T S$, and it follows that

$$\det g = \det S^T \det S = (\det S)^2\,. \tag{19.21}$$

According to the formula for the change of bases:

$$\omega\Big(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\Big) = \det(S)\omega(e_1, \dots, e_n) = \det S\,. \tag{19.22}$$

If $(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n})$ is a positive basis, then $\det(S) > 0$, hence $\det(S) = \sqrt{\det g}$.
In conclusion:

Proposition 19.1. *In any chart* $(\mathcal{U}, \varphi)$ *the volume form* $\mathrm{d}\mu_g$ *is given by*

$$\mathrm{d}\mu_g = \sqrt{\det g}\ \mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^m\,. \tag{19.23}$$

The k-fold ($k \leq m$) outer product of covectors $\xi^1, \dots, \xi^k$ is given by the full antisymmetrization:

$$(\xi^1 \wedge \ldots \wedge \xi^k) \cdot (X_1, \dots, X_k) = \sum_\pi (-1)^{|\pi|} (\xi^1 \cdot X_{\pi(1)}) \cdots (\xi^k \cdot X_{\pi(k)}) \tag{19.24}$$

where the sum is taken over all permutations π and we denote by $|\pi|$ the length of the permutation (number of pairs exchanged). Thus

$$(-1)^{|\pi|} = \begin{cases} -1 & \text{if } \pi \text{ is odd} \\ 1 & \text{if } \pi \text{ is even} \end{cases} \tag{19.25}$$

Example 19.2.

$$\begin{aligned}(\xi^1 \wedge \xi^2 \wedge \xi^3) \cdot (X_1, X_2, X_3) =& (\xi^1 \cdot X_1)(\xi^2 \cdot X_2)(\xi^3 \cdot X_3) + (\xi^1 \cdot X_2)(\xi^2 \cdot X_3)(\xi^3 \cdot X_1) \\ &+ (\xi^1 \cdot X_3)(\xi^2 \cdot X_1)(\xi^3 \cdot X_2) - (\xi^1 \cdot X_3)(\xi^2 \cdot X_2)(\xi^3 \cdot X_1) \\ &- (\xi^1 \cdot X_2)(\xi^2 \cdot X_1)(\xi^3 \cdot X_3) - (\xi^1 \cdot X_1)(\xi^2 \cdot X_3)(\xi^3 \cdot X_2)\end{aligned} \tag{19.26}$$

Volume of a domain. Let $\mathcal{M}$ be an oriented manifold endowed with a volume form ω. Let $\mathcal{D}$ be a domain of a chart $(\mathcal{U}, \varphi)$, $\mathcal{D} \subset \mathcal{U}$. Then $\omega(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^m})$ is a function in $\mathcal{U}$, and

$$\omega\Big(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^m}\Big) \circ \varphi^{-1} \tag{19.27}$$

is a function on $\varphi(\mathcal{U})$, and open set in $\mathbb{R}^m$, and $\varphi(\mathcal{D}) \subset \varphi(\mathcal{U})$. We then define

$$\begin{aligned} \mathrm{Vol}(\mathcal{D}) &= \int_{\mathcal{D}} \omega \\ &= \int_{\varphi(\mathcal{D})} \omega\Big(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^m}\Big) \circ \varphi^{-1} \mathrm{d}x^1 \dots \mathrm{d}x^m \,. \end{aligned} \tag{19.28}$$

To show that this is well-defined, let $(\tilde{\mathcal{U}}, \tilde{\varphi})$ be another chart such that also $\mathcal{D} \subset \tilde{\mathcal{U}}$. Then with $x^\nu = \varphi^\nu(p)$, $\tilde{x}^\mu = \tilde{\varphi}^\mu(p)$, $p \in \mathcal{D}$, we have

$$x^\nu = f^\nu(\tilde{x}^1, \dots, \tilde{x}^m)\,, \qquad f = \varphi \circ \tilde{\varphi}^{-1}\,, \quad \nu = 1, \dots, m\,. \tag{19.29}$$

and

$$\frac{\partial}{\partial \tilde{x}^\mu}|_p = \sum_{\nu=1}^m \frac{\partial}{\partial x^\nu}|_p S^\nu_\mu(p) \tag{19.30}$$

where

$$S^\nu_\mu(p) = \frac{\partial x^\nu}{\partial \tilde{x}^\mu}|_p = \frac{\partial f^\nu}{\partial \tilde{x}^\mu}(\tilde{\varphi}(p))\,. \tag{19.31}$$

Hence, by the change of bases formula, as functions on $\mathcal{D}$,

$$\omega\Big(\frac{\partial}{\partial \tilde{x}^1}, \dots, \frac{\partial}{\partial \tilde{x}^n}\Big) = \det(S)\omega\Big(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\Big) \tag{19.32}$$

and

$$\omega\Big(\frac{\partial}{\partial \tilde{x}^1}, \dots, \frac{\partial}{\partial \tilde{x}^n}\Big) \circ \tilde{\varphi}^{-1}(\tilde{x}) = \Big(\det \frac{\partial f}{\partial \tilde{x}}(\tilde{x})\Big)\omega\Big(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\Big) \circ \tilde{\varphi}^{-1}(\tilde{x})\,. \tag{19.33}$$

Now $\det \frac{\partial f}{\partial \tilde{x}}(\tilde{x})$ is the Jacobian determinant of the transformation $\tilde{x} \to x = f(\tilde{x})$. Recall that under the change of variables $\tilde{x} \mapsto x = f(\tilde{x})$,

$$\int_{f(\mathcal{E})} \mu(x) \mathrm{d}x^1 \dots \mathrm{d}x^n = \int_{\mathcal{E}} \mu(f(\tilde{x})) \det \frac{\partial f}{\partial \tilde{x}}(\tilde{x}) \mathrm{d}\tilde{x}^1 \dots \mathrm{d}\tilde{x}^n\,. \tag{19.34}$$

Here we take $\mathcal{E} = \tilde{\varphi}(\mathcal{D})$ so that $f(\mathcal{E}) = \varphi(\mathcal{D})$ and

$$\mu = \omega\Big(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\Big) \circ \varphi^{-1} \tag{19.35}$$

to obtain

$$\int_{\tilde{\varphi}(\mathcal{D})} \omega\Big(\frac{\partial}{\partial \tilde{x}^1}, \dots, \frac{\partial}{\partial \tilde{x}^n}\Big) \circ \tilde{\varphi}^{-1}(\tilde{x}) \mathrm{d}\tilde{x}^1 \dots \mathrm{d}\tilde{x}^n = \int_{\varphi(\mathcal{D})} \omega\Big(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n}\Big) \circ \varphi^{-1} \mathrm{d}x^1 \dots \mathrm{d}x^n \tag{19.36}$$

This is precisely what we needed to show.

Partition of unity. To remove the restriction that $\mathcal{D}$ be contained in the domain of a chart, we use a partition of unity.

Let

$$\mathcal{A} = \Big\{(\mathcal{U}_\alpha, \varphi_\beta) : \alpha \in I\Big\} \tag{19.37}$$

be an atlas for $\mathcal{M}$. We will assume (paracompactness) that $\mathcal{A}$ is locally finite, namely each point $p \in \mathcal{M}$ has a neighbourhood $\mathcal{V}_p$ which intersects only finitely many of the $\mathcal{U}_\alpha$. We can then construct a **partition of unity** relative to $\mathcal{A}$: 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 is supported in $\mathcal{U}_\alpha$.
2. $\sum_\alpha f_\alpha = 1$

By local finiteness this sum contains only finitely many terms at each point. In fact, by local finiteness for any given $p \in \mathcal{M}$ there is a neighbourhood $\mathcal{V}_p$ such that all except finitely many of the f_α vanish on $\mathcal{V}_p$.

Given now *any* domain $\mathcal{D} \subset \mathcal{M}$ we define

$$\mathrm{Vol}(\mathcal{D}) = \sum_\alpha \int_{\mathcal{D}} f_\alpha \omega = \sum_\alpha \int_{\mathcal{D}\cap\mathcal{U}_\alpha} f_\alpha \omega \tag{19.38}$$

This is either a positive real number of ∞. (If $\overline{\mathcal{D}}$ is compact then $\mathrm{Vol}(\mathcal{D})$ is finite.)

Volume of a submanifold. Let $\mathcal{N}$ be a n-dimensional submanifold of a m-dimensional Riemannian manifold $(\mathcal{M}, g)$. We define on $\mathcal{N}$ the Riemannian metric h, called the *induced metric* (by g on $\mathcal{N}$),

$$h = g|_{\mathrm{T}\mathcal{N}} \, . \tag{19.39}$$

Then $(\mathcal{N}, h)$ is itself a Riemannian manifold and the definition of volume applies to any domain in $\mathcal{N}$.

19.3. Examples

Consider $\mathbb{R}^3 \setminus 0$. This is diffeomorphic to $\mathbb{R} \times \mathbb{S}^2$. Consider on $\mathbb{R} \times \mathbb{S}^2$ the metric

$$g = \mathrm{d}r^2 + R^2(r) \overset{\circ}{\gamma} \tag{19.40}$$

where $\overset{\circ}{\gamma}$ is the standard metric on $\mathbb{S}^2$. In polar coordinates,

$$\overset{\circ}{\gamma} = (\mathrm{d}\vartheta)^2 + \sin^2(\vartheta)(\mathrm{d}\varphi)^2 \, . \tag{19.41}$$

Here (r, ϑ, φ) are coordinates on $\mathbb{R} \times \mathbb{S}^2$. The vectors tangent to $\{r\} \times \mathbb{S}^2 = S_r$ are linear combinations of $\frac{\partial}{\partial\vartheta}$ and $\frac{\partial}{\partial\varphi}$. So the metric γ_r induced on S_r is

$$\gamma_r = R^2(r) \overset{\circ}{\gamma} \; . \tag{19.42}$$

The volume of a 2-dimensional manifold is called **area**. We have

$$\text{Area}(S_r) = \int_{S_r} \mathrm{d}\mu_{\gamma_r} \tag{19.43}$$

$$\text{Area}(\mathbb{S}^2) = \int_{\mathbb{S}^2} \mathrm{d}\mu_{\overset{\circ}{\gamma}} = 4\pi \tag{19.44}$$

To check (19.44): In polar coordinates we have

$$\mathrm{d}\mu_{\overset{\circ}{\gamma}} = \sqrt{\det \overset{\circ}{\gamma}}\ \mathrm{d}\vartheta \wedge \mathrm{d}\varphi \tag{19.45}$$

Now as a matrix

$$\overset{\circ}{\gamma} = \begin{pmatrix} \overset{\circ}{\gamma}_{\vartheta\vartheta} & \overset{\circ}{\gamma}_{\vartheta\varphi} \\ \overset{\circ}{\gamma}_{\varphi\vartheta} & \overset{\circ}{\gamma}_{\varphi\varphi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2(\vartheta) \end{pmatrix} \tag{19.46}$$

so

$$\det(\overset{\circ}{\gamma}) = \sin^2(\vartheta) \tag{19.47}$$

$$\mathrm{d}\mu_{\overset{\circ}{\gamma}} = \sin\vartheta\ \mathrm{d}\vartheta \wedge \mathrm{d}\varphi \tag{19.48}$$

Hence

$$\int_{\mathbb{S}^2} \mathrm{d}\mu_{\overset{\circ}{\gamma}} = \int_0^{2\pi} \Big\{ \int_0^{\pi} \sin\vartheta \mathrm{d}\vartheta \Big\} \mathrm{d}\varphi = 4\pi \tag{19.49}$$

To check (19.43): We have

$$\mathrm{d}\mu_{\gamma_r} = \sqrt{\det \gamma_r} \mathrm{d}\vartheta \wedge \mathrm{d}\varphi = R^2(r) \sqrt{\det \overset{\circ}{\gamma}} \mathrm{d}\vartheta \wedge \varphi = R^2(r) \mathrm{d}\mu_{\overset{\circ}{\gamma}}\,. \tag{19.50}$$

Therefore

$$\text{Area}(S_r) = 4\pi R^2(r)\,. \tag{19.51}$$

Let now $B(r_1, r_2) = [r_1, r_2] \times \mathbb{S}^2$. Then

$$\text{Vol}(B(r_1, r_2)) = \int_{B(r_1,r_2)} \mathrm{d}\mu_g\,. \tag{19.52}$$

Since

$$g = \begin{pmatrix} 1 & 0 \\ 0 & \gamma_r \end{pmatrix}, \qquad \det g = \det \gamma_r \tag{19.53}$$

and so

$$\mathrm{d}\mu_g = \sqrt{\det \gamma_r}\ \mathrm{d}r \wedge \mathrm{d}\vartheta \wedge \mathrm{d}\varphi = \mathrm{d}r \wedge \mathrm{d}\mu_{\gamma_r} \tag{19.54}$$

we obtain

$$\begin{aligned} \text{Vol}(B(r_1, r_2)) &= \int_{r_1}^{r_2} \Big\{ \int_{S_r} \mathrm{d}\mu_{\gamma_r} \Big\} \mathrm{d}r \\ &= \int_{r_1}^{r_2} \text{Area}(S_r) \mathrm{d}r = 4\pi \int_{r_1}^{r_2} R^2(r) \mathrm{d}r\,. \end{aligned} \tag{19.55}$$

19.4. Problems

1. In this problem we will define a metric on $\mathbb{R}^n \setminus \{0\}$. Let $R : (0, \infty) \to (0, \infty)$ be a smooth function, and $\mathring{\gamma}$ be the standard metric on $\mathbb{S}^{n-1}$.
   a) Show that $\mathbb{R}^n \setminus \{0\}$ is diffeomorphic to $\mathbb{R}_+ \times \mathbb{S}^{n-1}$, where $\mathbb{R}_+ = (0, \infty)$. In fact, show that
   $$\varphi(x) = \Big(|x|, \frac{x}{|x|}\Big) \tag{19.56}$$
   is a diffeomorphism.
   b) Let us define a metric g on $\mathbb{R}_+ \times \mathbb{S}^{n-1}$ by
   $$g = \mathrm{d}r^2 + R^2(r)\,\mathring{\gamma}\ . \tag{19.57}$$
   We can pull this metric back by φ to a metric on $\mathbb{R}^n \setminus \{0\}$.
   Then we want to extend $\varphi^* g$ to the origin by assuming that
   $$\lim_{r \to 0} R(r) = 0\,, \tag{19.58}$$
   and setting $R(0) = 0$.
   Under which assumptions is the extended metric *continuous* as a metric on $\mathbb{R}^n$?

2. Let us endow $\mathbb{R}^3$ with a metric like in the previous problem:
   $$g = \mathrm{d}r^2 + R^2(r)\,\mathring{\gamma} \tag{19.59}$$
   Let (r, ϑ, φ) be spherical coordinates. Prove that the curve
   $$\sigma : [a, b] \to \mathbb{R}^3 \setminus \{0\}\,, \qquad \sigma(t) = (t, \vartheta_0, \varphi_0), \tag{19.60}$$
   minimizes the arclength among all curves from $\sigma(a)$ to $\sigma(b)$.

Part IV.

Connections

Lecture 20.

Connections in a vector bundle

20.1. Horizontal lift

Consider a vector bundle $\mathcal{B}$ over $\mathcal{M}$. In general, if p and q are two distinct points in $\mathcal{M}$ we have no way of associating elements of $\mathcal{B}_q$ with elements of $\mathcal{B}_p$. There is one exception: The zero element, 0_q should correspond to 0_p, the zero element in $\mathcal{B}_p$.

Consider a curve γ joining p to q, see Figure 20.1:

$$\gamma : [0,1] \to \mathcal{M}\,, \qquad \gamma(0) = p\,, \quad \gamma(1) = q\,. \tag{20.1}$$

We want to find a curve $\gamma^\sharp$ in $\mathcal{B}$ starting at $b \in \mathcal{B}_p$,

$$\gamma^\sharp : [0,1] \to \mathcal{B}\,, \qquad \gamma^\sharp(0) = b\,, \tag{20.2}$$

such that $\gamma^\sharp$ projects to γ:

$$\pi \circ \gamma^\sharp = \gamma\,, \quad \pi\Big(\gamma^\sharp(t)\Big) = \gamma(t)\ (t \in [0,1])\,. \tag{20.3}$$

This will be called the **horizontal lift** of γ to $\mathcal{B}$, starting at b. It defines the **parallel transport** of b along γ, giving us the corresponding vector to $b \in \mathcal{B}_p$ in $\mathcal{B}_q$ by evaluating $\gamma^\sharp$ at the other end point, $\gamma^\sharp(1) \in \mathcal{B}_q$.

Question: How can we define $\gamma^\sharp$, the horizontal lift to $\mathcal{B}$ of a curve γ in $\mathcal{M}$ through a given $b_0 \in \mathcal{B}_{\gamma(0)}$?

The idea is to impose a first order ODE for $\gamma^\sharp$ (for a given γ). Then $\gamma^\sharp$ will be uniquely determined by the initial condition $\gamma^\sharp(0) = b_0$.

In other words, we must have some rule which specifies the instantaneous velocity $\dot{\gamma}^\sharp(t)$ of $\gamma^\sharp$ in terms of the instantaneous position $\gamma^\sharp(t)$ and the corresponding instantaneous velocity $\dot{\gamma}(t)$ of γ.

So we want a rule F such that

$$\dot{\gamma}^\sharp(t) = F\Big(\gamma^\sharp(t), \dot{\gamma}(t)\Big)\,. \tag{20.4}$$

Consider the arguments of F for a given point $p = \gamma(t) \in \mathcal{M}$. Then F is a function on $\mathcal{B}_p \times \mathrm{T}_p\mathcal{M}$, where $p = \gamma(t)$, $b = \gamma^\sharp(t)$, and we write

$$F(b, X) = X^\sharp_b \qquad (b \in \mathcal{B}_p, X \in \mathrm{T}_p\mathcal{M})\,. \tag{20.5}$$

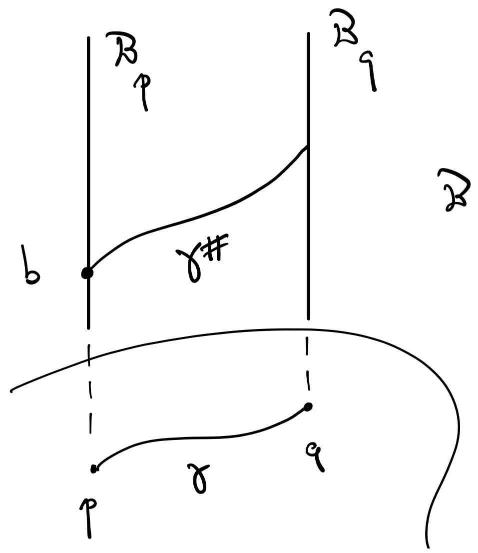


Figure 20.1.: Horizontal lift.

This is the **horizontal lift** of a vector $X \in \mathrm{T}_p\mathcal{M}$ to $b \in \mathcal{B}_p$, and element of $T_b\mathcal{B}$.

The horizontal lift of vectors $X \in \mathrm{T}_p\mathcal{M}$ to $b \in \mathcal{B}_p$ shall be defined by imposing *four* conditions.

The first condition follows from the following condition on the horizontal lift of a curve:

$$\pi \circ \gamma^\sharp = \gamma\,. \tag{20.6}$$

For, differentiating both sides with respect to the parameter t and setting $X = \dot{\gamma}(t)$ we obtain:

$$\boxed{\mathrm{d}\pi \cdot X^\sharp = X\,.} \tag{20.7}$$

Exercise 20.1. Verify that (20.6) indeed implies (20.7).

We also require that

$$\boxed{X^\sharp \text{ depends linearly on } X\,.} \tag{20.8}$$

The other two conditions will be discussed in the next section.

We define the **horizontal subspace** at b, denoted by H_b, by

$$H_b = \left\{X_b^\sharp : \; X \in \mathrm{T}_p\mathcal{M}\right\} \in \mathrm{T}_b\mathcal{B}\,. \tag{20.9}$$

This is a *linear subspace* by the condition (20.8).

Moreover by the condition (20.7) the mapping $X \mapsto X_b^\sharp$ is *injective*. For, if $X_b^\sharp = 0 \in \mathrm{T}_b\mathcal{B}$, we infer from (20.7) that $X = \mathrm{d}\pi \cdot X_b^\sharp = 0 \in \mathrm{T}_p\mathcal{M}$. Hence,

$$\dim H_b = \dim \mathrm{T}_p\mathcal{M} = m \tag{20.10}$$

and $\mathrm{d}\pi|_{H_b}$, the inverse of the mapping $X \to X_b^\sharp$, is a linear isomorphism of H_b onto $\mathrm{T}_p\mathcal{M}$.

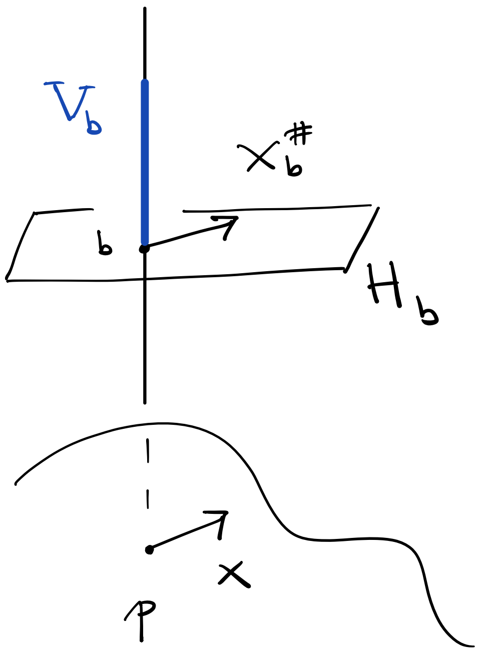


Figure 20.2.: Vertical subspace.

20.2. Vertical transformations

We define the **vertical subspace** at $b \in \mathcal{B}_p$ to be

$$V_b = \mathrm{T}_b\mathcal{B}_p \,. \tag{20.11}$$

We have

$$\mathrm{T}_b\mathcal{B} = H_b \oplus V_b \,. \tag{20.12}$$

Thus any vector $Y \in \mathrm{T}_b\mathcal{B}$ can be uniquely decomposed into a vector $X_b^\sharp \in H_B$, where $X = \mathrm{d}\pi(b) \cdot Y \in \mathrm{T}_p\mathcal{M}$, $p = \pi(b)$, and a vector

$$Y - X_b^\sharp \in V_b \,. \tag{20.13}$$

See Figure 20.2.

Remark 20.1. Let Y be vertical vector at $v \in \mathcal{B}$, namely $Y \in \mathrm{T}_v\mathcal{B}_p$, $p = \pi(v)$. Since $\mathcal{B}_p$ is a linear space, a tangent vector at any point can be thought of as an element of the linear space. Thus Y can be identified with an element in $\mathcal{B}_p$, namely a vector attached at the origin; see Fig. 20.3.

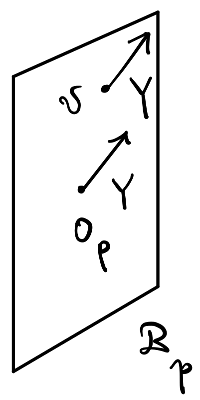


Figure 20.3.: Tangent space to a linear space.

Now suppose 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 $\mathcal{M}$ into itself such that

$$\pi \circ \phi = \chi \circ \pi \,. \tag{20.14}$$

In other words, if ϕ is fibre-preserving then it maps fibres to fibres.

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 $\mathcal{M}$ into itself is simply the identity. So $\phi : \mathcal{B} \to \mathcal{B}$ is a vertical transformation if $v \in \mathcal{B}_p$ implies $\phi(v) \in \mathcal{B}_p$, $p \in \mathcal{M}$.

The basic vertical transformations correspond to scalar multiplication and vector addition — the basic linear operations.

Multiplication: Given $c \in \mathbb{R}$ we have the vertical transformation

$$M_c(v) = cv \qquad (v \in \mathcal{B}) \,. \tag{20.15}$$

Addition: Given a section σ of $\mathcal{B}$ we have the vertical transformation

$$S_\sigma(v) = v + \sigma(\pi(v)) \qquad (v \in \mathcal{B}) \,. \tag{20.16}$$

Recall that we have already given two

conditions on the horizontal lift: (20.7) and (20.8).

We now introduce two further conditions, the *compatibility conditions with the linear structure of* $\mathcal{B}$:

Compatibility with scalar multiplication: For every $c \in \mathbb{R}$, $v \in \mathcal{B}_p$, $p \in \mathcal{M}$, and every $X \in \mathrm{T}_p\mathcal{M}$,

$$\boxed{X^\sharp_{M_c(v)} = \mathrm{d}M_c \cdot X^\sharp_v} \tag{20.17}$$

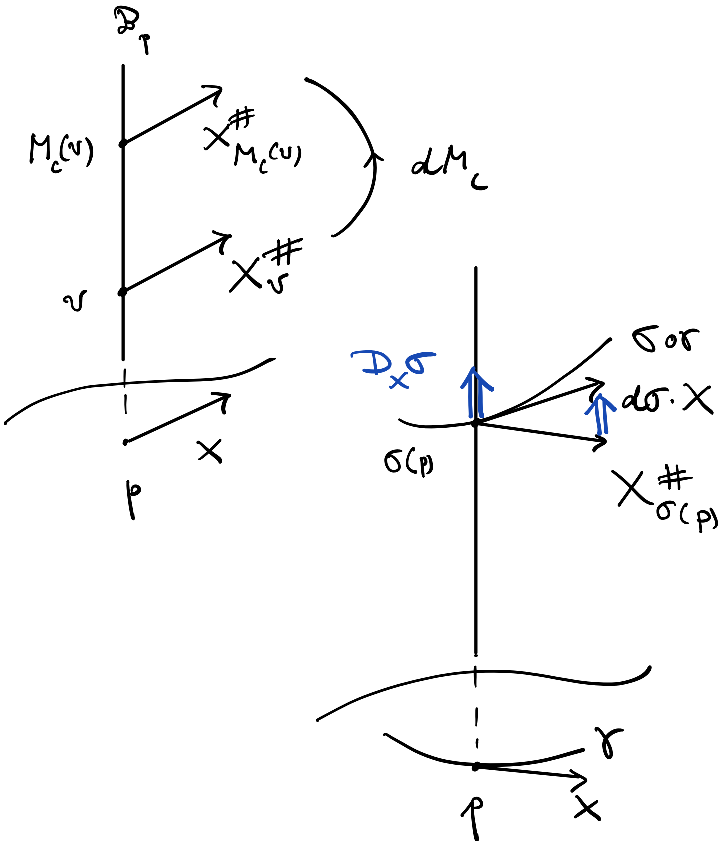


Figure 20.4.: Compatibility with linear structure.

Compatibility with vector addition: For every section σ of $\mathcal{B}$, and all $v \in \mathcal{B}_p$, $p \in \mathcal{M}$, and every $X \in \mathrm{T}_p\mathcal{M}$,

$$\boxed{X^\sharp_{S_\sigma(v)} - \mathrm{d}S_\sigma \cdot X^\sharp_v = X^\sharp_{\sigma(p)} - \mathrm{d}\sigma \cdot X} \tag{20.18}$$

The RHS takes into account the variation of σ over $\mathcal{M}$; see Fig. 20.4. The vector $\mathrm{d}\sigma \cdot X$ projects to X, because $\pi \circ \sigma = \mathrm{id}_{\mathcal{M}}$. So its horizontal part is $X^\sharp_{\sigma(p)}$, and the *vertical* part is therefore

$$\mathrm{d}\sigma \cdot X - X^\sharp_{\sigma(p)}\,. \tag{20.19}$$

Definition 20.1. The **covariant derivative** of σ with respect to X is

$$D_X\sigma = \mathrm{d}\sigma \cdot X - X^\sharp_{\sigma(p)}\,. \tag{20.20}$$

The fact that 0_q corresponds to 0_p for every pair of points $p, q \in \mathcal{M}$, is reflected in the fact that the covariant derivative of the 0-section $\mathring{\sigma}$, defined by

$$\mathring{\sigma}(p) = 0_p \qquad (p \in \mathcal{M}) \tag{20.21}$$

is itself the zero element:

$$D_X\mathring{\sigma} = 0_p \qquad (X \in \mathrm{T}_p\mathcal{M}) \tag{20.22}$$

20.3. Connection coefficients

Local coordinates. Choose an atlas for $\mathcal{B}$ such that the corresponding domains $\mathcal{U}$ in $\mathcal{M}$ are domains of charts. Consider then any of the domains $\mathcal{U} \subset \mathcal{M}$ which together cover $\mathcal{M}$. Over $\mathcal{U}$ we have the basis sections $(\iota_a : a = 1, \ldots, n)$. We also have in $\mathcal{U}$ the coordinates $x^\mu : \mu = 1, \ldots, m$, where $\dim \mathcal{M} = m$, and $\dim \mathcal{B} = m + n$. We can then expand any vector $v \in \mathcal{B}_p$ as

$$v = \sum_{a=1}^{n} v^a \iota_a(p) \tag{20.23}$$

The $(v^a : a = 1, \ldots, n)$ constitute a system of linear coordinates for $\mathcal{B}_p$. Together with the $(x^\mu : \mu = 1, \ldots, m)$ we have coordinates for $\mathcal{B}$ in $\pi^{-1}(\mathcal{U})$,

$$(x^1, \ldots, x^m; v^1, \ldots, v^n) \tag{20.24}$$

Let us simplify the notation to $(x^\mu; v^a)$.

We expand a given $X \in \mathrm{T}_p\mathcal{M}$ as

$$X = \sum_{\mu=1}^{m} X^\mu \frac{\partial}{\partial x^\mu}|_p \,. \tag{20.25}$$

We may also expand $X_v^\sharp$ as

$$X_v^\sharp = \sum_{\mu=1}^{m} Y^\mu \frac{\partial}{\partial x^\mu}|_v + \sum_{a=1}^{n} Z^a \frac{\partial}{\partial v^a}|_v \tag{20.26}$$

Here $Y^\mu = Y^\mu(v, X)$, and $Z^\mu = Z^\mu(v, X)$.

Consider now the condition (20.7),

$$\mathrm{d}\pi \cdot X_v^\sharp = X \,. \tag{20.27}$$

The projection π is represented in our coordinates by:

$$(x^1, \ldots, x^m; v^1, \ldots, v^n) \mapsto (x^1, \ldots, x^m) \,, \tag{20.28}$$

where $(x^1, \ldots, x^m)$ are the coordinates of p. Thus

$$\mathrm{d}\pi \cdot \frac{\partial}{\partial x^\mu}|_v = \frac{\partial}{\partial x^\mu}|_p \,, \tag{20.29}$$

$$\mathrm{d}\pi \cdot \frac{\partial}{\partial v^a}|_v = 0 \,. \tag{20.30}$$

Hence
$$\mathrm{d}\pi \cdot X^\sharp_v = \sum_{\mu=1}^{m} Y^\mu \frac{\partial}{\partial x^\mu}|_p \,, \tag{20.31}$$

Therefore condition (20.7) is equivalent to

$$\boxed{Y^\mu(v,X) = X^\mu \qquad : \ \mu = 1,\ldots,m\,.} \tag{20.32}$$

Consider next the condition (20.8), that $X^\sharp_v$ is linear in X. In view of (20.32) this reduces to $Z^a(v,X)$ being linear in X,

$$\boxed{Z^a(v,X) = \sum_{\mu=1}^{m} Z^a_\mu(v) X^\mu \qquad : \ a = 1,\ldots,n.} \tag{20.33}$$

We proceed to condition (20.17), namely

$$X^\sharp_{M_c(v)} = \mathrm{d}M_c \cdot X^\sharp_v \,. \tag{20.34}$$

The multiplication M_c is represented in our coordinates by :

$$(x^1,\ldots,x^m; v^1,\ldots,v^n) \mapsto (x^1,\ldots,x^m; cv^1,\ldots,cv^n)\,, \tag{20.35}$$

so

$$\mathrm{d}M_c \cdot \frac{\partial}{\partial x^\mu}|_v = \frac{\partial}{\partial x^\mu}|_{M_c(v)}\,, \tag{20.36}$$
$$\mathrm{d}M_c \cdot \frac{\partial}{\partial v^a}|_v = c\frac{\partial}{\partial v^a}|_{M_c(v)}\,. \tag{20.37}$$

Hence

$$\mathrm{d}M_c \cdot X^\sharp_v = \sum_{\mu=1}^{m} X^\mu \frac{\partial}{\partial x^\mu}|_{M_c(v)} + \sum_{a=1}^{n}\sum_{\mu=1}^{m} cZ^a_\mu(v) X^\mu \frac{\partial}{\partial v^a}|_{M_c(v)}\,, \tag{20.38}$$

while
$$X^\sharp_{M_c(v)} = \sum_{\mu=1}^{m} X^\mu \frac{\partial}{\partial x^\mu}|_{M_c(v)} + \sum_{a=1}^{n}\sum_{\mu=1}^{m} Z^a_\mu(cv) X^\mu \frac{\partial}{\partial v^a}|_{M_c(v)} \tag{20.39}$$

Therefore the condition (20.17) reduces to

$$\boxed{Z^a_\mu(cv) = cZ^a_\mu(v) \qquad a = 1,\ldots,m\,.} \tag{20.40}$$

Finally, consider (20.18), namely

$$X^\sharp_{S_\sigma(v)} - \mathrm{d}S_\sigma \cdot X^\sharp_v = X^\sharp_{\sigma(p)} - \mathrm{d}\sigma \cdot X = -D_X\sigma\,. \tag{20.41}$$

Remark 20.2. Recall that $\pi \circ S_\sigma = \pi$ (and also $\pi \circ M_c = \pi$, both being vertical transformations) and thus

$$\mathrm{d}\pi \cdot \mathrm{d}S_\sigma \cdot X^\sharp_v = \mathrm{d}\pi \cdot X^\sharp_v = X\,, \tag{20.42}$$

and also

$$\mathrm{d}\pi \cdot X^\sharp_{S_\sigma(v)} = X \tag{20.43}$$

so $\mathrm{d}\pi$ applied to the LHS of (20.41) vanishes. Hence the LHS is a vertical vector, just like the RHS as we have shown in (20.21). But the LHS is a vertical vector attached at the point

$$S_\sigma(v) = v + \sigma(p) \in \mathcal{B}_p \tag{20.44}$$

while the RHS is a vertical vector attached at the point $\sigma(p) \in \mathcal{B}_p$. However in view of the general Remark 20.1 this equality makes sense because both vertical vectors can be thought of as elements of $\mathcal{B}_p$. Thus a vertical vector V at v,

$$\sum_{a=1}^{n} V^a \frac{\partial}{\partial v^a}\Big|_v \in \mathrm{T}_v\mathcal{B}_p \qquad (p = \pi(v)) \tag{20.45}$$

can be identified with the following element of $\mathcal{B}_p$:

$$\sum_{a=1}^{n} V^a \iota_a(p) \in \mathcal{B}_p\,. \tag{20.46}$$

We now turn to the task of expressing condition (20.41) in our coordinates. The transformation S_σ is represented by

$$(x^1, \ldots, x^m; v^1, \ldots, v^n) \mapsto (x^1, \ldots, x^m; v^1 + \sigma^1(p), \ldots, v^n + \sigma^n(p))\,, \tag{20.47}$$

the coordinates of p being $(x^1, \ldots, x^n)$. Consider the x^μ coordinate line

$$t \mapsto (x^1, \ldots, x^\mu + t, \ldots, x^m; v^1, \ldots, v^n)\,. \tag{20.48}$$

This is mapped by S_σ to the curve

$$t \mapsto (x^1, \ldots, x^\mu + t, \ldots, x^m; v^1 + \sigma^1(\gamma(t)), \ldots, v^n + \sigma^n(\gamma(t)))\,, \tag{20.49}$$

where $\gamma(t)$ is the curve in $\mathcal{M}$ represented by the x^μ coordinate line

$$t \mapsto (x^1, \ldots, x^\mu + t, \ldots, x^m)\,. \tag{20.50}$$

where $\gamma(0) = p$. It follows that

$$\mathrm{d}S_\sigma \cdot \frac{\partial}{\partial x^\mu}\Big|_v = \frac{\partial}{\partial x^\mu}\Big|_{S_\sigma(v)} + \sum_{a=1}^{n} \frac{\partial \sigma^a}{\partial x^\mu}\Big|_p \frac{\partial}{\partial v^a}\Big|_{S_\sigma(v)} \tag{20.51}$$

and

$$\mathrm{d}S_\sigma \cdot \frac{\partial}{\partial v^a}\Big|_v = \frac{\partial}{\partial v^a}\Big|_{S_\sigma(v)} \tag{20.52}$$

Hence

$$\mathrm{d}S_\sigma \cdot X_v^\sharp = \sum_{\mu=1}^m X^\mu \frac{\partial}{\partial x^\mu}|_{S_\sigma(v)} + \sum_{a=1}^n \sum_{\mu=1}^m \Big(Z_\mu^a(v) + \frac{\partial \sigma^a}{\partial x^\mu}|_p \Big) X^\mu \frac{\partial}{\partial v^a}|_{S_\sigma(v)} \tag{20.53}$$

On the other hand,

$$X_{S_\sigma(v)}^\sharp = \sum_{\mu=1}^m X^\mu \frac{\partial}{\partial x^\mu}|_{S_\sigma(v)} + \sum_{a=1}^n \sum_{\mu=1}^m Z_\mu^a(v+\sigma(p)) X^\mu \frac{\partial}{\partial v^a}|_{S_\sigma(v)} \tag{20.54}$$

hence

$$X_{S_\sigma(v)}^\sharp - \mathrm{d}S_\sigma \cdot X_v^\sharp = \sum_{a=1}^n \sum_{\mu=1}^m \Big(Z_\mu^a(v+\sigma(p)) - Z_\mu^a(v) - \frac{\partial \sigma^a}{\partial x^\mu}|_p \Big) X^\mu \frac{\partial}{\partial v^a}|_{S_\sigma(v)} \tag{20.55}$$

Moreover, the section σ is represented by

$$(x^1, \dots, x^m) \mapsto (x^1, \dots, x^m; \sigma^1(p), \dots, \sigma^n(p)) \tag{20.56}$$

where $(x^1, \dots, x^n)$ are the coordinates of p, and thus

$$\mathrm{d}\sigma \cdot \frac{\partial}{\partial x^\mu}|_p = \frac{\partial}{\partial x^\mu}|_{\sigma(p)} + \sum_{a=1}^n \frac{\partial \sigma^a}{\partial x^\mu}|_p \frac{\partial}{\partial v^a}|_{\sigma(p)} \tag{20.57}$$

Then

$$X_{\sigma(p)}^\sharp = \sum_{\mu=1}^m X^\mu \frac{\partial}{\partial x^\mu}|_{\sigma(p)} + \sum_{a=1}^n \sum_{\mu=1}^m Z_\mu^a(\sigma(p)) X^\mu \frac{\partial}{\partial v^a}|_{\sigma(p)} \tag{20.58}$$

while

$$\mathrm{d}\sigma \cdot X = \sum_{\mu=1}^m X^\mu \frac{\partial}{\partial x^\mu}|_{\sigma(p)} + \sum_{\mu=1}^m \sum_{a=1}^n \frac{\partial \sigma^a}{\partial x^\mu}|_p X^\mu \frac{\partial}{\partial v^a}|_{\sigma(p)} \,. \tag{20.59}$$

Therefore

$$\begin{aligned} D_X\sigma =& \mathrm{d}\sigma \cdot X - X_{\sigma(p)}^\sharp = \\ =& \sum_{\mu=1}^m \sum_{a=1}^n \Big(\frac{\partial \sigma^a}{\partial x^\mu}|_p - Z_\mu^a(\sigma(p)) \Big) X^\mu \frac{\partial}{\partial v^a}|_{\sigma(p)} \,. \end{aligned} \tag{20.60}$$

Now substituting into (20.41),

$$X_{S_\sigma(v)}^\sharp - \mathrm{d}S_\sigma \cdot X_v^\sharp + D_X\sigma = 0 \,. \tag{20.61}$$

we see that the condition reduces to:

$$\boxed{Z_\mu^a(v+\sigma(p)) = Z_\mu^a(v) + Z_\mu^a(\sigma(p))} \tag{20.62}$$

Together with (20.40), namely

$$Z^a_\mu(cv) = cZ^a_\mu(v) \qquad a = 1, \ldots, m\,, \tag{20.63}$$

we see that the conditions (20.17) and (20.18) are equivalent to the linearity of $Z^a_\mu(v)$ in v. *Therefore, there are coefficients $A^a_{\mu b}(p)$, depending only on p, such that*

$$Z^a_\mu(v) = -\sum_{b=1}^{n} A^a_{\mu b}(p)v^b\,. \tag{20.64}$$

Putting everything together, we have arrived at the following result:

$$X^\sharp_v = \sum_{\mu=1}^{m} X^\mu \Big\{ \frac{\partial}{\partial x^\mu}|_v - \sum_{a,b=1}^{n} A^a_{\mu b}(p)v^b \frac{\partial}{\partial v^a}|_v \Big\} \qquad (X \in \mathrm{T}_p\mathcal{M}, v \in \mathcal{B}_p) \tag{20.65}$$

The coefficients $A^a_{\mu b}$ are functions in $\mathcal{U}$, the domain of a chart in $\mathcal{U}$. They are called the **connection coefficients**, and they depend not only on the choice of a chart on $\mathcal{M}$ but also on the choice of a bundle chart through the choice of basis sections $\iota_a : a = 1, \ldots, n$.

We can eliminate the dependence on the choice of chart on $\mathcal{M}$ when considering the **connection 1-form**:

$$A^a{}_b = \sum_{\mu=1}^{m} A^a_{\mu b} \mathrm{d}x^\mu \tag{20.66}$$

which is a matrix-valued 1-form on $\mathcal{M}$. This depends on the choice of basis sections $(\iota_a : a = 1, \ldots, n)$ *but not* on the local coordinates $(x^\mu : \mu = 1, \ldots, m)$.

Covariant derivative. We can substitute for Z^a_μ in the formula for $D_X\sigma$ to obtain

$$D_X\sigma = \sum_{\mu=1}^{m}\sum_{a=1}^{n} D_\mu\sigma^a|_p X^\mu \iota_a(p) \tag{20.67}$$

where

$$D_\mu\sigma^a|_p = \frac{\partial\sigma^a}{\partial x^\mu}|_p + \sum_{b=1}^{n} A^a_{\mu b}(p)\sigma(p)^b\,. \tag{20.68}$$

If now X is a vectorfield in $\mathcal{M}$ then $D_X\sigma$ is another section of $\mathcal{B}$ defined by

$$(D_X\sigma)(p) = D_{X(p)}\sigma \in \mathcal{B}_p \qquad (p \in \mathcal{M}) \tag{20.69}$$

The section $D_X\sigma$ is called the **covariant derivative** of σ with respect to X.

We have

$$D_{\frac{\partial}{\partial x^\mu}}\iota_a = \sum_{b=1}^{n} A^b_{\mu a}\iota_b \tag{20.70}$$

This identity may also be thought of as the *definition* of the connection coefficients.

Exercise 20.2. Prove the formula (20.70) using the above.

Covariant differentiation satisfies the Leibniz rule: if f is a function on $\mathcal{M}$ and σ a section of $\mathcal{B}$ then $f\sigma$ is a section of $\mathcal{B}$ and we have

$$D_X(f\sigma) = fD_X\sigma + (Xf)\sigma \tag{20.71}$$

In particular, since

$$\sigma = \sum_{a=1}^{n} \sigma^a \iota_a \tag{20.72}$$

where σ^a are functions, and $D_X\sigma$ is linear in σ we have

$$\begin{aligned} D_X\sigma = D_X\Big(\sum_{a=1}^{n}\sigma^a\iota_a\Big) = \\ = \sum_{a=1}^{n}\Big\{\sigma^a D_X\iota_a + (X\sigma^a)\iota_a\Big\} \\ = \sum_{a=1}^{n}\Big\{\sigma^a\sum_{b=1}^{n}(A^a{}_b\cdot X)\iota_b + (X\sigma^a)\iota_a\Big\} \end{aligned} \tag{20.73}$$

where we used that, by linearity of $D_X\sigma$ with respect to X, with $X \in \mathrm{T}_p\mathcal{M}$,

$$\begin{aligned} D_X\iota_a &= \sum_{\mu=1}^{m} X^\mu D_{\partial/\partial x^\mu}\iota_a \\ &= \sum_{\mu=1}^{m}\sum_{b=1}^{n} X^\mu A^b_{\mu a}\iota_b = \sum_{b=1}^{n}(A^b_a\cdot X)\iota_b \end{aligned} \tag{20.74}$$

Writing

$$X\cdot\sigma^a = \mathrm{d}\sigma^a\cdot X \tag{20.75}$$

we obtain

$$D_X\sigma = \sum_{a=1}^{n}\Big(\mathrm{d}\sigma^a + \sum_{b=1}^{n}A^a{}_b\sigma^b\Big)\cdot X\ \ \iota_a \tag{20.76}$$

Also note that

$$D_\mu\sigma^a = \big(\mathrm{d}\sigma^a + \sum_{b=1}^{n}A^a{}_b\sigma^b\big)\cdot\frac{\partial}{\partial x^\mu}\,. \tag{20.77}$$

Remark 20.3. With the above notation, we refer to a connection with either of the following notation:

$$A^b_{\mu b}\,,\qquad A^a{}_b\,,\qquad A_\mu\,,\qquad A\,. \tag{20.78}$$

Problems

1. Let $\gamma : [0,1] \to \mathbb{R}^2$ be a closed simple curve. Using coordinates (x^1, x^2) on $\mathbb{R}^2$ we can write

$$\gamma(t) = (x^1(t), x^2(t))\,. \tag{20.79}$$

Let now $y(t) = (y^1(t), y^2(t)$ solve the following system of ODEs:

$$\dot{y}^1(t) = -x^1(t)\, y^2(t)\, \dot{x}^2(t) \tag{20.80}$$

$$\dot{y}^2(t) = x^1(t)\, y^1(t)\, \dot{x}^2(t) \tag{20.81}$$

$$y(0) = y_0 \tag{20.82}$$

for some $y_0 \in \mathbb{R}^2$. Prove that $y(1)$ is the rotation of y_0 by an angle which is proportional to the area enclosed by γ, i.e. if we use polar coordinates

$$y^1 = m \cos\varphi \qquad y^2 = m \sin\varphi \tag{20.83}$$

with m and φ depending on t, then m is constant and $\varphi(1) = \varphi(0) + A(\gamma)$, where $A(\gamma)$ is the *oriented area* enclosed by γ.

2. Let $\mathcal{M} = \mathbb{R}^2$ and consider its tangent bundle $\mathcal{B} = \mathrm{T}\mathbb{R}^2 \simeq \mathbb{R}^2 \times \mathbb{R}^2$. Let

$$A^a{}_b = \sum_\mu A^a_{\mu b} \mathrm{d}x^\mu \tag{20.84}$$

be the connection 1-form given by

$$A = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \tag{20.85}$$

where a is the 1-form

$$a = x^1 \mathrm{d}x^2\,, \tag{20.86}$$

and (x^1, x^2) are the standard coordinates on $\mathbb{R}^2$.

Let $\gamma : [0,1] \to \mathbb{R}^2$ be a simple closed curve with

$$\gamma(0) = \gamma(1) = x_0 = (x_0^1, x_0^2)\,, \tag{20.87}$$

and write in coordinates $\gamma(t) = (x^1(t), x^2(t))$, and consider any vector

$$y_0 = (y_0^1, y_0^2) \in \mathrm{T}_{x_0}\mathbb{R}^2\,. \tag{20.88}$$

Show that, after parallel transport along γ, the vector y_0 is rotated by a factor equal to the oriented area enclosed by γ.

Supplement: Gauge transformations

20.1. Changes of basis sections

Let $(\tilde{\iota}_a : a = 1, \dots, n)$ be another set of basis sections over the domain $\tilde{\mathcal{U}}$ in $\mathcal{M}$. Then over $\mathcal{U} \cap \tilde{\mathcal{U}}$ we have two sets of bases and we can express

$$\tilde{\iota}_a = \sum_{b=1}^{n} \iota_b S^b{}_a \,. \tag{20.1}$$

Remark 20.1. From now we use the *summation convention*: 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 $\mathcal{U} \cap \tilde{\mathcal{U}}$. Given any section ϕ of $\mathcal{B}$, then over $\mathcal{U} \cap \tilde{\mathcal{U}}$ we have expressions

$$\phi = \phi^a \iota_a = \tilde{\phi}^a \tilde{\iota}_a \tag{20.2}$$

Substituting for $\tilde{\iota}_a$ in terms of $(\iota_1, \dots, \iota_n)$ we obtain

$$\tilde{\phi}^a S^b{}_a = \phi^b \qquad \tilde{\phi}^a = (S^{-1})^a{}_b \phi^b \,. \tag{20.3}$$

Over $\mathcal{U}$ we have

$$D_{\frac{\partial}{\partial x^\mu}} \iota_a = A^b_{\mu a} \iota_b \tag{20.4}$$

and over $\tilde{\mathcal{U}}$:

$$D_{\frac{\partial}{\partial x^\mu}} \tilde{\iota}_a = \tilde{A}^b_{\mu a} \tilde{\iota}_b \tag{20.5}$$

Then, over $\mathcal{U} \cap \tilde{\mathcal{U}}$ using (20.1) we obtain

$$\begin{aligned} D_{\frac{\partial}{\partial x^\mu}} \tilde{\iota}_a =& D_{\frac{\partial}{\partial x^\mu}} \left(\iota_b S^b{}_a \right) \\ =& S^b{}_a D_{\frac{\partial}{\partial x^\mu}} \iota_b + \frac{\partial S^b{}_a}{\partial x^\mu} \iota_b \\ =& S^b{}_a A^c_{\mu b} \iota_c + \frac{\partial S^b{}_a}{\partial x^\mu} \iota_b = \left(S^c{}_a A^b_{\mu c} + \frac{\partial S^b{}_a}{\partial x^\mu} \right) \iota_b \end{aligned} \tag{20.6}$$

On the other hand,

$$D_{\frac{\partial}{\partial x^\mu}} \tilde{\iota}_a = \tilde{A}^b_{\mu a} \tilde{\iota}_b = \tilde{A}^b_{\mu a} S^c{}_b \iota_c = \tilde{A}^c_{\mu a} S^b{}_c \iota_b \,. \tag{20.7}$$

We conclude that

$$S^b{}_c \tilde{A}^c_{\mu a} = A^b_{\mu c} S^c{}_a + \frac{\partial S^b{}_a}{\partial x^\mu} \tag{20.8}$$

or in matrix notation

$$S\tilde{A}_\mu = A_\mu S + \frac{\partial S}{\partial x^\mu} \tag{20.9}$$

Let us now use matrix-column notation, and denote by ϕ not the section itself but the column

$$\phi = \begin{pmatrix} \phi^1 \\ \vdots \\ \phi^n \end{pmatrix} \tag{20.10}$$

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

$$\tilde{\phi} = S^{-1}\phi \qquad S\tilde{A}_\mu = A_\mu S + \frac{\partial S}{\partial x^\mu} \tag{20.11}$$

or

$$\tilde{A}_\mu = S^{-1}A_\mu S + S^{-1}\frac{\partial S}{\partial x^\mu}\,. \tag{20.12}$$

Consider now

$$D_{\frac{\partial}{\partial x^\mu}}\phi = (D_\mu \phi^a)\iota_a \tag{20.13}$$

where

$$D_\mu \phi^a = \frac{\partial \phi^a}{\partial x^\mu} + A^a_{\mu b}\phi^b\,. \tag{20.14}$$

In matrix-column notation this reads

$$D_\mu \phi = \frac{\partial \phi}{\partial x^\mu} + A_\mu \phi\,. \tag{20.15}$$

Then, relative to the new basis,

$$\begin{aligned} \widetilde{D_\mu \phi} =& \frac{\partial \tilde{\phi}}{\partial x^\mu} + \tilde{A}_\mu \tilde{\phi} \\ =& \frac{\partial}{\partial x^\mu}(S^{-1}\phi) + S^{-1}\Big(A_\mu S + \frac{\partial S}{\partial x^\mu}\Big)S^{-1}\phi \\ =& S^{-1}\frac{\partial \phi}{\partial x^\mu} - S^{-1}\frac{\partial S}{\partial x^\mu}S^{-1}\phi + S^{-1}A_\mu\phi + S^{-1}\frac{\partial S}{\partial x^\mu}S^{-1}\phi \\ =& S^{-1}D_\mu\phi \end{aligned} \tag{20.16}$$

So $D_\mu\phi$ transforms like ϕ, under a change of basis:

$$\phi \mapsto S^{-1}\phi \qquad D_\mu\phi \mapsto S^{-1}D_\mu\phi\,. \tag{20.17}$$

Lecture 21.

Curvature

If X is a vectorfield on the base manifold $\mathcal{M}$ and $\mathcal{B}$ is a vector bundle over $\mathcal{M}$ endowed with a connection A, then we can define $X^\sharp$, the horizontal lift of X to $\mathcal{B}$, a vectorfield on $\mathcal{B}$, by:

$$X^\sharp(v) = X(p)^\sharp_v\,, \qquad p \in \mathcal{M},\ v \in \mathcal{B}_p\,. \tag{21.1}$$

Let X, and Y be two vectorfields on $\mathcal{M}$. We compute $[X^\sharp, Y^\sharp]$. We shall use the local coordinates $(x^1, \dots, x^m; v^1, \dots, v^n)$ in $\pi^{-1}(\mathcal{U})$, $\mathcal{U}$ the domain of a chart in $\mathcal{M}$. We have

$$X^\sharp = X^\mu\Big(\frac{\partial}{\partial x^\mu} - A^b_{\mu a} v^a \frac{\partial}{\partial v^b}\Big) \tag{21.2}$$

$$Y^\sharp = Y^\mu\Big(\frac{\partial}{\partial x^\mu} - A^b_{\mu a} v^a \frac{\partial}{\partial v^b}\Big) \tag{21.3}$$

Let f be a function on $\mathcal{B}$. It is represented by

$$(x^\mu; v^a) \mapsto f(x^\mu; v^a) \tag{21.4}$$

Then,

$$X^\sharp(Y^\sharp f) = X^\sharp g \qquad g = Y^\sharp f = Y^\nu\Big(\frac{\partial f}{\partial x^\nu} - A^d_{\nu c} v^c \frac{\partial f}{\partial v^d}\Big) \tag{21.5}$$

$$\begin{aligned} X^\sharp(Y^\sharp f) &= X^\mu\Big(\frac{\partial g}{\partial x^\mu} - A^b_{\mu a} v^a \frac{\partial g}{\partial v^b}\Big) \\ &= X^\mu\Big\{\frac{\partial Y^\nu}{\partial x^\mu}\Big(\frac{\partial f}{\partial x^\nu} - A^d_{\nu c} v^c \frac{\partial f}{\partial v^d}\Big) + Y^\nu\Big(\frac{\partial^2 f}{\partial x^\mu \partial x^\nu} - \frac{\partial A^d_{\nu c}}{\partial x^\mu} v^c \frac{\partial f}{\partial v^d} - A^d_{\nu c} v^c \frac{\partial^2 f}{\partial x^\mu \partial v^d}\Big) \\ &\qquad - A^b_{\mu a} v^a Y^\nu\Big(\frac{\partial^2 f}{\partial v^b \partial x^\nu} - A^d_{\nu c}\delta^c_b \frac{\partial f}{\partial v^d} - A^d_{\nu c} v^c \frac{\partial^2 f}{\partial v^b \partial v^d}\Big)\Big\} \end{aligned} \tag{21.6}$$

Note that

$$-A^d_{\nu c} v^c \frac{\partial^2 f}{\partial x^\mu \partial v^d} - A^b_{\mu a} v^a \frac{\partial^2 f}{\partial v^b \partial x^\nu} \tag{21.7}$$

is symmetric in μ, ν, so when multiplied by $X^\mu Y^\nu$ these terms will cancel when we interchange the roles of X and Y, and subtract:

$$\begin{aligned} X^\sharp(Y^\sharp f) - Y^\sharp(X^\sharp f) &= [X, Y]^\nu\Big(\frac{\partial f}{\partial x^\nu} - A^d_{\nu c} v^c \frac{\partial f}{\partial v^d}\Big) \\ &\quad - X^\mu Y^\nu\Big\{\Big(\frac{\partial A^d_{\nu c}}{\partial x^\mu} - \frac{\partial A^d_{\mu c}}{\partial x^\nu}\Big) v^c \frac{\partial f}{\partial v^d} - \Big(A^b_{\mu a} A^d_{\nu b} - A^b_{\nu a} A^d_{\mu b}\Big) v^a \frac{\partial f}{\partial v^d}\Big\} \end{aligned} \tag{21.8}$$

Here we also used that

$$A^b_{\mu a} v^a A^d_{\nu c} v^c \frac{\partial^2 f}{\partial v^b \partial v^d} \tag{21.9}$$

is symmetric in μ, ν, so it likewise cancels.

Setting

$$F_{\mu\nu}{}^a{}_b = \frac{\partial A^a_{\nu b}}{\partial x^\mu} - \frac{\partial A^a_{\mu b}}{\partial x^\nu} + A^a_{\mu c} A^c_{\nu b} - A^a_{\nu c} A^c_{\mu b} \tag{21.10}$$

or in matrix notation,

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\nu} + A_\mu A_\nu - A_\nu A_\mu \tag{21.11}$$

the formula above reduces to

$$[X^\sharp, Y^\sharp] f - [X, Y]^\sharp f = -X^\mu Y^\nu F_{\mu\nu}{}^a{}_b v^b \frac{\partial f}{\partial v^a} \,. \tag{21.12}$$

We can also write this formula as

$$[X^\sharp, Y^\sharp] f - [X, Y]^\sharp f = -F(X, Y)^a{}_b v^b \frac{\partial f}{\partial v^a} \,. \tag{21.13}$$

where F is a matrix-valued 2-form, the **curvature 2-form**:

$$F^a{}_b = \mathrm{d}A^a{}_b + A^a{}_c \wedge A^c{}_b \tag{21.14}$$

This is because for any 1-forms a, b,

$$\begin{aligned}(a \wedge b)_{\mu\nu} &= (a \wedge b)\Big(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\Big) \\ &= a\Big(\frac{\partial}{\partial x^\mu}\Big) b\Big(\frac{\partial}{\partial x^\nu}\Big) - b\Big(\frac{\partial}{\partial x^\mu}\Big) a\Big(\frac{\partial}{\partial x^\nu}\Big) = a_\mu b_\nu - b_\mu a_\nu\end{aligned} \tag{21.15}$$

and similarly for the components of the exterior derivative we have derived (16.12). In terms of matrices,

$$F = \frac{1}{2} \sum_{\mu,\nu} F_{\mu\nu} \mathrm{d}x^\mu \wedge \mathrm{d}x^\nu \tag{21.16}$$

where $F_{\mu\nu}$ is given by (21.11).

Remark 21.1. Under a change of basis sections $F_{\mu\nu}$ transforms homogeneously,

$$F_{\mu\nu} \mapsto \widetilde{F}_{\mu\nu} = S^{-1} F_{\mu\nu} S \,. \tag{21.17}$$

21.1. Curvature transformation

The linear transformation in $\mathcal{B}_p$,

$$v \mapsto F_{\mu\nu} \cdot v \qquad v^a \mapsto F_{\mu\nu}{}^a{}_b v^b \tag{21.18}$$

is called the **curvature transformation**.

At $p \in \mathcal{M}$, F itself is an antisymmetric bilinear form in $\mathrm{T}_p\mathcal{M}$, $F(X,Y) = F_{\mu\nu}X^\mu Y^\nu$, with values in $\mathcal{L}(\mathcal{B}_p, \mathcal{B}_p)$, namely the space of linear transformations on $\mathcal{B}_p$. In particular,

$$F_{\mu\nu}(p) = F\Big(\frac{\partial}{\partial x^\mu}|_p, \frac{\partial}{\partial x^\nu}|_p\Big) \tag{21.19}$$

Remark 21.2. For any vectorspaces U, and V, let us denote by $\mathcal{L}(U,V)$ the space of linear maps $U \to V$.

We have obtained

$$\Big(\big[X^\sharp, Y^\sharp\big] - [X,Y]^\sharp\Big)(v) = -F(X,Y)\cdot v \in \mathcal{B}_p \tag{21.20}$$

The RHS can be viewed as a *vertical vector* at $v \in \mathcal{B}_p$. Now suppose that $[X,Y] = 0$, and that X and Y are complete. Suppose that X generates the group χ_t, and Y generates the group ϕ_s. Then

$$\chi_t \circ \phi_s = \phi_s \circ \chi_t \qquad (t, s \in \mathbb{R}) \tag{21.21}$$

So:

$$(\phi_{-s} \circ \chi_{-t} \circ \phi_s \circ \chi_t)(p) = p \qquad (p \in \mathcal{M}) \tag{21.22}$$

Now $X^\sharp$ generates the group $\chi_t^\sharp$, and $Y^\sharp$ generates the group $\phi_s^\sharp$, where the orbit $t \mapsto \chi_t^\sharp(p)$ is the horizontal lift of the orbit $t \mapsto \chi_t(p)$ and similarly for $\phi_s^\sharp$ and ϕ_s.

Then

$$(\phi_{-s}^\sharp \circ \chi_{-t}^\sharp \circ \phi_s^\sharp \circ \chi_t^\sharp)(v) = v' \in \mathcal{B}_p \qquad (v \in \mathcal{B}_p) \tag{21.23}$$

and to order ts,

$$v' - v = -ts\, F(X,Y)(p)\cdot v\,. \tag{21.24}$$

See Figure 21.1 for the geometric interpretation of the curvature transformation.

21.2. Curvature and covariant differentiation

Proposition 21.1. *Suppose that X, Y are vectorfields on $\mathcal{M}$ and ϕ is a section of $\mathcal{B}$. Then*

$$D_X D_Y \phi - D_Y D_X \phi - D_{[X,Y]}\phi = F(X,Y)\phi \tag{21.25}$$

Proof. Consider the case

$$X = \frac{\partial}{\partial x^\mu} \qquad Y = \frac{\partial}{\partial x^\nu}\,. \tag{21.26}$$

Then $D_X\phi$ is represented by the column

$$D_\mu \phi = \frac{\partial \phi}{\partial x^\mu} + A_\mu \phi \tag{21.27}$$

and $D_Y\phi$ by

$$D_\nu \phi = \frac{\partial \phi}{\partial x^\nu} + A_\nu \phi\,. \tag{21.28}$$

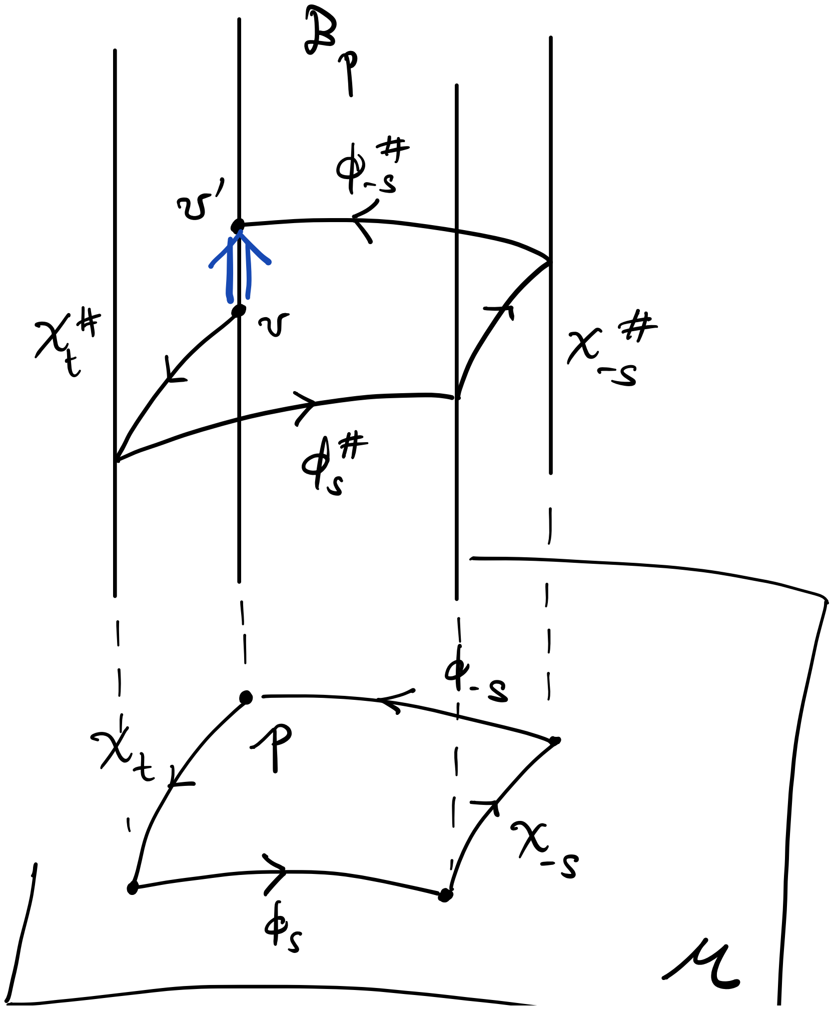


Figure 21.1.: Curvature transformation.

Hence $D_X D_Y \phi$ is represented by

$$\begin{aligned} D_\mu D_\nu \phi =& D_\mu\Big(\frac{\partial \phi}{\partial x^\nu} + A_\nu \phi\Big) = \frac{\partial}{\partial x^\mu}\Big(\frac{\partial \phi}{\partial x^\nu} + A_\nu\phi\Big) + A_\mu\Big(\frac{\partial \phi}{\partial x^\nu} + A_\nu \phi\Big) \\ =& \frac{\partial^2 \phi}{\partial x^\mu \partial x^\nu} + \Big(\frac{\partial A_\nu}{\partial x^\mu} + A_\mu A_\nu\Big)\phi + A_\nu \frac{\partial \phi}{\partial x^\mu} + A_\mu \frac{\partial \phi}{\partial x^\nu} \end{aligned} \tag{21.29}$$

and $D_Y D_X \phi$ is represented by

$$D_\nu D_\mu \phi = \frac{\partial^2 \phi}{\partial x^\nu \partial x^\mu} + \Big(\frac{\partial A_\mu}{\partial x^\nu} + A_\nu A_\mu\Big)\phi + A_\mu \frac{\partial \phi}{\partial x^\nu} + A_\nu \frac{\partial \phi}{\partial x^\mu} \tag{21.30}$$

while

$$[X, Y] = \Big[\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\Big] = 0\,. \tag{21.31}$$

Subtracting the two formulas above it follows that

$$D_\mu D_\nu \phi - D_\nu D_\mu \phi = F_{\mu\nu}\phi\,. \tag{21.32}$$

The general formula now follows easily using Leibniz rule. □

Connections. One can view a connection on a vector bundle $\mathcal{B}$ as an assignment

$$(X, \sigma) \mapsto D_X \sigma \tag{21.33}$$

of a vectorfield X on $\mathcal{M}$ and a differentiable section σ of $\mathcal{B}$, to a new section $D_X\sigma$ of $\mathcal{B}$, called the covariant derivative of σ with respect to X satisfying the following conditions:

1. $D_X\sigma$ is linear in σ: $D_X(\sigma_1 + \sigma_2) = D_X\sigma_1 + D_X\sigma_2$
2. $D_X\sigma$ is linear in X with respect to multiplication by the ring of functions. I.e. if f is a function on $\mathcal{M}$, then
$$D_{fX}\sigma = f D_X \sigma\,. \tag{21.34}$$
3. $D_X\sigma$ satisfies Leibniz rule
$$D_X(f\sigma) = f D_X\sigma + (Xf)\sigma \tag{21.35}$$
if we multiply σ by a differentiable function.

Conversely any assignment of this type defines a connection A on $\mathcal{B}$. For taking local basis sections $\iota_a : a = 1, \ldots, n$ over $\mathcal{U}$, the domain of a chart in $\mathcal{M}$, we define the connection coefficients

$$D_{\frac{\partial}{\partial x^\mu}} \iota_a = A^b_{\mu a}\iota_b \tag{21.36}$$

Thus if $\sigma = \sigma^a \iota_a$, we have by the first and third condition:

$$\begin{aligned} D_{\frac{\partial}{\partial x^\mu}}\sigma =& D_{\frac{\partial}{\partial x^\mu}}(\sigma^a \iota_a) = \sigma^a D_{\frac{\partial}{\partial x^\mu}}\iota_a + \frac{\partial \sigma^a}{\partial x^\mu}\iota_a \\ =& \sigma^a A^b_{\mu a}\iota_b + \frac{\partial \sigma^a}{\partial x^\mu}\iota_a = \Big(\frac{\partial \sigma^a}{\partial x^\mu} + A^b_{\mu a}\sigma^b\Big)\iota_a \\ =& (D_\mu \sigma^a)\iota_a \end{aligned} \tag{21.37}$$

and by the second condition, if $X = X^\mu \frac{\partial}{\partial x^\mu}$, then

$$D_X \sigma = X^\mu D_{\frac{\partial}{\partial x^\mu}} \sigma = X^\mu (D_\mu \sigma^a) \iota_a \,. \tag{21.38}$$

Then finally with our calculation above,

$$D_X D_Y \sigma - D_Y D_X \sigma - D_{[X,Y]} \sigma = F(X,Y) \sigma \,. \tag{21.39}$$

Lecture 22.

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 metric h. This is the condition that *the magnitude of a vector is invariant under parallel transport.*

This means, if γ is a curve in $\mathcal{M}$ through p, and $\gamma^\sharp$ is its horizontal lift to $\mathcal{B}$ through some $v \in \mathcal{B}_p$, then

$$\|\gamma^\sharp(t)\|^2 = h\Big(\gamma^\sharp(t), \gamma^\sharp(t)\Big) \tag{22.1}$$

must be independent of t.

We choose a set of *orthonormal* basis sections ($\iota_a := 1, \ldots, n$) over $\mathcal{U} \in \mathcal{M}$:

$$h(\iota_a, \iota_b) = \delta_{ab} \tag{22.2}$$

As before we use the corresponding local coordinates in $\pi^{-1}(\mathcal{U}) \subset \mathcal{B}$ to express

$$X_v^\sharp = X^\mu \Big(\frac{\partial}{\partial x^\mu} - A^a_{\mu b} v^b \frac{\partial}{\partial v^a}\Big) \tag{22.3}$$

so $\gamma^\sharp$ is represented by

$$t \mapsto \left(x^1(t), \ldots, x^m(t); v^1(t), \ldots, v^n(t)\right) \tag{22.4}$$

where

$$\frac{\mathrm{d}x^\mu}{\mathrm{d}t} = \dot\gamma^\mu(t) \tag{22.5}$$

$$\frac{\mathrm{d}v^a}{\mathrm{d}t} = -A^a_{\mu b}(\gamma(t))\dot\gamma^\mu(t) v^b \tag{22.6}$$

The $v^a(t)$ are the components of the parallel transported vector:

$$v(t) = \gamma^\sharp(t) \qquad v(t) = v^a(t)\iota_a(\gamma(t)) \tag{22.7}$$

Then

$$\|v(t)\|^2 = h(v(t), v(t)) = v^a(t) v^b(t) h(\iota_a(\gamma(t)), \iota_b(\gamma(t))) = \sum_{a=1}^n (v^a(t))^2 \tag{22.8}$$

Thus the compatibility condition becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\|v(t)\|^2 = \sum_{a=1}^{n} v^a \frac{\mathrm{d}v^a}{\mathrm{d}t} = 0 \tag{22.9}$$

which is equivalent to

$$\sum_{a,b=1}^{n} v^a v^b A^a_{\mu b}(\gamma(t))\dot{\gamma}^\mu(t) = 0 \tag{22.10}$$

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

$$\sum_{a,b=1}^{n} v^a v^b A^a_{\mu b}(p) = 0 \qquad p \in \mathcal{M} \tag{22.11}$$

When decomposing the matrix A_μ into its symmetric and antisymmetric part,

$$A_\mu = \frac{1}{2}(A_\mu + A_\mu^T) + \frac{1}{2}(A_\mu - A_\mu^T) \tag{22.12}$$

we obtain that the quadratic form

$$Q_\mu(v,v) = 0 \qquad Q_\mu = \frac{1}{2}(A_\mu + A_\mu^T) \tag{22.13}$$

Here

$$Q_\mu(v,v) = \sum_{a,b=1}^{n} Q^a_{\mu b} v^a v^b = 0 \tag{22.14}$$

and by polarisation it follows that

$$Q_\mu(u,v) = 0 \tag{22.15}$$

namely $Q_\mu = 0$, or

$$A_\mu + A_\mu^T = 0\,. \tag{22.16}$$

We conclude that metric compatibility is equivalent to the antisymmetry of the matrix A_μ:

$$A_\mu^T = -A_\mu \qquad A^b_{\mu a} = -A^a_{\mu b}\,. \tag{22.17}$$

It follows that $F_{\mu\nu}$ is anti-symmetric as well:

$$F_{\mu\nu}^T = -F_{\mu\nu} \tag{22.18}$$

because

$$(A_\mu A_\nu)^T = A_\nu^T A_\mu^T = A_\nu A_\mu\,. \tag{22.19}$$

Remark 22.1. Gauge transformations to be discussed.

Remark 22.2. Metric compatbility in complex vector bundle to be discussed.

Remark 22.3. Gauge transformations in complex vector bundle to be discussed.

Lecture 23.

Connections on the tangent bundle

We now restrict ourselves to the case $\mathcal{B} = \mathrm{T}\mathcal{M}$. Then a section σ is a vectorfield on $\mathcal{M}$. 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 $\mathcal{B} = \mathrm{T}\mathcal{M}$ we denote A by Γ, F by R (Riemann) and D by ∇.

Symmetry condition.

$$\nabla_X Y - \nabla_Y X = [X, Y] \tag{23.1}$$

Without metric. Suppose that $\mathcal{U}$ is the domain of a chart of $\mathcal{M}$. Then in $\mathcal{U}$ we have the coordinate vectorfields

$$\frac{\partial}{\partial x^\mu} : \mu = 1, \ldots, m \tag{23.2}$$

forming a set of basis sections. In such a coordinate basis the connection coefficients are defined by

$$\nabla_{\frac{\partial}{\partial x^\mu}} \frac{\partial}{\partial x^\nu} = \Gamma^\lambda_{\mu\nu} \frac{\partial}{\partial x^\lambda} \tag{23.3}$$

Here Γ_μ is a matrix with entries $\Gamma^\lambda_{\mu\nu}$. The curvature is then given by

$$R_{\mu\nu} = R(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}) = \frac{\partial \Gamma_\nu}{\partial x^\mu} - \frac{\partial \Gamma_\mu}{\partial x^\nu} + \Gamma_\mu \Gamma_\nu - \Gamma_\nu \Gamma_\mu \tag{23.4}$$

We write

$$R^\kappa{}_{\lambda\mu\nu} = (R_{\mu\nu})^\kappa{}_\lambda = \frac{\partial \Gamma^\kappa_{\nu\lambda}}{\partial x^\mu} - \frac{\partial \Gamma^\kappa_{\mu\lambda}}{\partial x^\nu} + \Gamma^\kappa_{\mu\alpha} \Gamma^\alpha_{\nu\lambda} - \Gamma^\kappa_{\nu\alpha} \Gamma^\alpha_{\mu\lambda} \tag{23.5}$$

In terms of the connection 1-form

$$\Gamma^\kappa{}_\lambda = \Gamma^\kappa_{\mu\lambda} \mathrm{d}x^\mu \tag{23.6}$$

the curvature 2-form $R^\kappa{}_\lambda$ is given by

$$R^\kappa{}_\lambda = \mathrm{d}\Gamma^\kappa{}_\lambda + \Gamma^\kappa{}_\alpha \wedge \Gamma^\alpha{}_\lambda \tag{23.7}$$

In the coordinate basis we have $[\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}] = 0$ so the symmetry condition reduces to

$$\Gamma^\lambda_{\mu\nu} - \Gamma^\lambda_{\nu\mu} = 0\,. \tag{23.8}$$

With metric. In this case we have a Riemannian manifold $(\mathcal{M}, g)$. The connection coefficients in an *orthonormal* frame field $(E_\mu : \mu = 1, \ldots, m)$, $g(E_\mu, E_\nu) = \delta_{\mu\nu}$, are defined by

$$\nabla_{E_\mu} E_\alpha = \Gamma^\beta_{\mu\alpha} E_\beta \tag{23.9}$$

The connection 1-form is then

$$\Gamma^\alpha{}_\beta = \Gamma^\alpha_{\mu\beta} \theta^\mu \tag{23.10}$$

where θ^μ is the dual basis of 1-forms defined in $\mathcal{U}$.

Metric compatibility is equivalent, as we have seen, to the condition that Γ_μ is an anti-symmetric matrix

$$\Gamma^T_\mu = -\Gamma_\mu \,. \tag{23.11}$$

The **structure coefficients** $\Lambda^\alpha_{\mu\beta}$ of the frame field $(E_1, \ldots, E_m)$ are defined by

$$[E_\alpha, E_\beta] = \Lambda^\mu_{\alpha\beta} E_\mu \tag{23.12}$$

Consider now the dual basis 1-forms $(\theta^1, \ldots, \theta^m)$. We have

$$\theta^\mu \cdot E_\nu = \delta^\mu{}_\nu \,, \tag{23.13}$$

so

$$(\mathrm{d}\theta^\mu)(E_\alpha, E_\beta) = E_\alpha(\theta^\mu \cdot E_\beta) - E_\beta(\theta^\mu \cdot E_\alpha) - \theta^\mu \cdot [E_\alpha, E_\beta] = -\theta^\mu \cdot (\Lambda^\nu_{\alpha\beta} E_\nu) = -\Lambda^\mu_{\alpha\beta} \,. \tag{23.14}$$

Hence

$$\mathrm{d}\theta^\mu = -\frac{1}{2}\, \Lambda^\mu_{\alpha\beta}\, \theta^\alpha \wedge \theta^\beta \tag{23.15}$$

Proposition 23.1. *There is a unique connection* Γ *on* $\mathrm{T}\mathcal{M}$ *which is symmetric and compatible with the metric* g*. It is called the* Levi-Civita connection.

Proof (using the orthonormal frame field). The symmetry condition

$$\nabla_{E_\alpha} E_\beta - \nabla_{E_\beta} E_\alpha = [E_\alpha, E_\beta] \tag{23.16}$$

reads

$$\Gamma^\mu_{\alpha\beta} - \Gamma^\mu_{\beta\alpha} = \Lambda^\mu_{\alpha\beta} \,. \tag{23.17}$$

Metric compatibility is

$$\Gamma^\mu_{\alpha\beta} = -\Gamma^\beta_{\alpha\mu} \,. \tag{23.18}$$

Thus:

$$\begin{aligned}
\Gamma^\mu_{\alpha\beta} =& -\Gamma^\beta_{\alpha\mu} = -\Gamma^\beta_{\mu\alpha} - \Lambda^\beta_{\alpha\mu} \\
=& \Gamma^\alpha_{\mu\beta} - \Lambda^\beta_{\alpha\mu} = \Gamma^\alpha_{\beta\mu} + \Lambda^\alpha_{\mu\beta} - \Lambda^\beta_{\alpha\mu} \\
=& -\Gamma^\mu_{\beta\alpha} + \Lambda^\alpha_{\mu\beta} - \Lambda^\beta_{\alpha\mu} = -\Gamma^\mu_{\alpha\beta} - \Lambda^\mu_{\beta\alpha} + \Lambda^\alpha_{\mu\beta} - \Lambda^\beta_{\alpha\mu}
\end{aligned} \tag{23.19}$$

hence

$$\Gamma^\mu_{\alpha\beta} = \frac{1}{2}\Big(-\Lambda^\mu_{\beta\alpha} + \Lambda^\alpha_{\mu\beta} - \Lambda^\beta_{\alpha\mu}\Big) \tag{23.20}$$

or

$$\boxed{\Gamma^{\mu}_{\alpha\beta} = \frac{1}{2}\Big(\Lambda^{\mu}_{\alpha\beta} + \Lambda^{\alpha}_{\mu\beta} + \Lambda^{\beta}_{\mu\alpha}\Big)} \tag{23.21}$$

□

Proof (using a coordinate frame field). Then the symmetry condition is

$$\Gamma^{\mu}_{\alpha\beta} = \Gamma^{\mu}_{\beta\alpha} \tag{23.22}$$

and metric compatibility reads, with $g_{\alpha\beta} = g(\frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial x^\beta})$,

$$\begin{aligned}\frac{\partial g_{\alpha\beta}}{\partial x^\mu} =& g\Big(\nabla_{\frac{\partial}{\partial x^\mu}}\frac{\partial}{\partial x^\alpha}, \frac{\partial}{\partial x^\beta}\Big) + g\Big(\frac{\partial}{\partial x^\alpha}, \nabla_{\frac{\partial}{\partial x^\mu}}\frac{\partial}{\partial x^\beta}\Big)\\ =&\Gamma^{\nu}_{\mu\alpha}g_{\nu\beta} + \Gamma^{\nu}_{\mu\beta}g_{\nu\alpha}\end{aligned} \tag{23.23}$$

Interchange μ, α:

$$\frac{\partial g_{\mu\beta}}{\partial x^\alpha} = \Gamma^{\nu}_{\alpha\mu}g_{\nu\beta} + \Gamma^{\nu}_{\alpha\beta}g_{\nu\mu} \tag{23.24}$$

Interchange μ, β:

$$\frac{\partial g_{\alpha\mu}}{\partial x^\beta} = \Gamma^{\nu}_{\beta\alpha}g_{\nu\mu} + \Gamma^{\nu}_{\beta\mu}g_{\nu\alpha} \tag{23.25}$$

Adding the last two equations and subtracting the first gives, in view of the symmetry condition:

$$2\ \Gamma^{\nu}_{\alpha\beta}\ g_{\nu\mu} = \frac{\partial g_{\mu\beta}}{\partial x^\alpha} + \frac{\partial g_{\mu\alpha}}{\partial x^\beta} - \frac{\partial g_{\alpha\beta}}{\partial x^\mu} \tag{23.26}$$

We thus obtain

$$\boxed{\Gamma^{\mu}_{\alpha\beta} = \frac{1}{2}(g^{-1})^{\mu\nu}\Big(\frac{\partial g_{\nu\beta}}{\partial x^\alpha} + \frac{\partial g_{\nu\alpha}}{\partial x^\beta} - \frac{\partial g_{\alpha\beta}}{\partial x^\nu}\Big)\,.} \tag{23.27}$$

□

23.1. Geodesics

Definition 23.1. A *geodesic* is a curve γ whose tangent vector is parallel transported along the curve:

$$\nabla_{\dot\gamma}\dot\gamma = 0 \tag{23.28}$$

Remark 23.1. The definition does not require a metric on $\mathcal{M}$.

In local coordinates, γ is represented by $x^\mu = x^\mu(t)$, and denoting $\dot\gamma = v$, v is represented by $v^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}t}(t)$. The condition $\nabla_v v = 0$ then reads

$$\frac{\mathrm{d}v^\mu}{\mathrm{d}t} + \Gamma^{\mu}_{\alpha\beta}(x(t))v^\alpha v^\beta = 0\,. \tag{23.29}$$

In the case that the connection Γ comes from a metric g, a geodesic has constant speed:

$$\frac{\mathrm{d}}{\mathrm{d}t}\|v\| = 0\,. \tag{23.30}$$

Since the arclength is given by

$$s = \int \|v\| \mathrm{d}t\,, \tag{23.31}$$

for a geodesic $\frac{\mathrm{d}s}{\mathrm{d}t}$ is constant, therefore the parameter t is proportional to the arc length. Such a parameter is called an *affine parameter*.

We may ask if solutions to (23.29) exist for all time? Given that the system of equations (23.29) is non-linear, global existence is not guaranteed. Moreover it may fail for the following reasons:

1. The geodesic moves from one chart to another in finite time. But of course, this is not a real problem: we can always extend the curve to the next adjacent chart.

2. Consider the example when the manifold is an open subset of $\mathbb{R}^n$. The geodesics are straight lines an thus always run out of the manifold in finite time.

Definition 23.2. A Riemannian manifold $(\mathcal{M}, g)$ is geodesically complete if each geodesic may be extended ad infinitum for an affine parameter.

The distance on a Riemannian manifold is

$$d(p,q) = \inf_{\gamma[p,q]} L(\gamma[p,q]) \tag{23.32}$$

where $\gamma[p,q]$ is any C^1-curve from p to q.

As a consequence of the **Hopf-Rinow theorem** a manifold is *geodesically complete* if and only if it is *metrically complete*, namely if $(\mathcal{M}, d)$ is complete as a metric topological space.

Geodesics are curves with minimal arc length:

Theorem 23.2. *If $q \in B_\varepsilon(p)$, then the geodesic γ from p to q has the least arc length of any curve connecting p, q. Moreover it is unique, thus*

$$d(p,q) = L[\gamma]\,. \tag{23.33}$$

We will not prove this theorem in this lecture. However, note that geodesics do not necessarily minimize arc length globally.

Problems

1. Let $u : \mathcal{M} \to \mathcal{N}$ be a smooth map between two differentiable manifolds $\mathcal{M}$ and $\mathcal{N}$. We have already defined the pull-back bundle

$$u^*\mathrm{T}\mathcal{N} = \bigcup_{p\in\mathcal{M}} \{p\} \times \mathrm{T}_{u(p)}\mathcal{N} \tag{23.34}$$

with a suitable vector bundle structure; cf. Lecture 13.

Let now $\mathrm{T}\mathcal{N}$ be endowed with a symmetric connection A. Show that we can define a connection A^* on $u^*\mathrm{T}\mathcal{N}$, called the *induced connection* by requiring that if

$$t \mapsto \gamma(t) \tag{23.35}$$

is a curve in $\mathcal{M}$ through p, then the curve

$$t \mapsto \gamma^\sharp(t)\,, \tag{23.36}$$

the horizontal lift of γ to $u^*\mathrm{T}\mathcal{N}$ through

$$v \in \mathcal{B}_p = \{p\} \times \mathrm{T}_{u(p)}\mathcal{N}\,, \tag{23.37}$$

is simply

$$\gamma^\sharp(t) = (\gamma(t), (u \circ \gamma)^\sharp(t)) \tag{23.38}$$

where $(u \circ \gamma)^\sharp$ is the horizontal lift of the curve

$$t \mapsto (u \circ \gamma)(t) \tag{23.39}$$

in $\mathcal{N}$ to $\mathrm{T}\mathcal{N}$ through $v \in \mathrm{T}_{u(p)}\mathcal{N}$.

Let now $(x^1, \ldots, x^m)$ and $(y^1, \ldots, y^n)$ be local coordinates on $\mathcal{M}$, and $\mathcal{N}$, respectively. Show that in these coordinates

$$A^{*a}_{\mu b}(p) = A^a_{cb}(u(p))\frac{\partial u^c}{\partial x^\mu}|_p\,, \tag{23.40}$$

where $\{u^c : c = 1, \ldots, n\}$ are the components of u in the coordinate system $(y^1, \ldots, y^n)$.

Lecture 24.

Characterisation of curvature

Theorem 24.1. *Let $\mathcal{B}$ be a vector bundle (without metric). The vanishing of the curvature F is necessary and sufficient for the existence locally of basis sections $(\iota_1, \ldots, \iota_n)$ relative to which the connection coefficients vansish: $A^a_{\mu b} = 0$.*

Proof. Recall the change of basis formula

$$\tilde{\iota}_a = \sum_{b=1}^{n} \iota_b S^b_a \,, \tag{24.1}$$

where S is a non-singular matrix at each point.

$$S\tilde{A}_\mu = A_\mu S + \frac{\partial S}{\partial x^\mu} \tag{24.2}$$

Thus, to make $\tilde{A}_\mu = 0$ we must solve

$$\frac{\partial S}{\partial x^\mu} = -A_\mu S \tag{24.3}$$

for S. This equation,

$$\mathrm{d}S = -AS \tag{24.4}$$

has a solution S locally if and only if $\dfrac{\partial S}{\partial x^\mu \partial x^\nu}$ is symmetric in μ, ν, or $\mathrm{d}S^2 = 0$. We have

$$\frac{\partial^2 S}{\partial x^\mu \partial x^\nu} = -\frac{\partial A_\mu}{\partial x^\nu} S - A_\mu \frac{\partial S}{\partial x^\nu} = \Big(-\frac{\partial A_\mu}{\partial x^\nu} + A_\mu A_\nu\Big) S \tag{24.5}$$

so interchanging μ, ν and subtracting yields:

$$0 = F_{\mu\nu} S \,. \tag{24.6}$$

□

Theorem 24.2. *The $\mathcal{B}$ be a vector bundle with metric h. The vanishing of the curvature F is necessary and sufficient for the existence locally of orthonormal basis sections $(\iota_1, \ldots, \iota_n)$ relative to which the connection coefficients vanish.*

Proof. Let $F_{\mu\nu} = 0$. Then the integrability condition of the previous theorem holds. So we obtain a non-singular matrix S satisfying

$$\mathrm{d}S = -AS\,. \tag{24.7}$$

Now we want to show that S can be taken to be an orthogonal matrix O,

$$\tilde{\iota}_a = \iota_b O^b_a\,. \tag{24.8}$$

We have

$$\mathrm{d}S^T = -S^T A^T = S^T A \tag{24.9}$$

because $A^T = -A$, and

$$\mathrm{d}(SS^T) = (\mathrm{d}S)S^T + S\mathrm{d}S^T = -ASS^T = SS^T A \tag{24.10}$$

so setting $M = SS^T$,

$$\mathrm{d}M = -AM + MA\,. \tag{24.11}$$

This equation for M has a unique solution for a given value of M at a fixed point p. Since $M = I$ is a solution it is the only solution that $M(p) = I$. It follows that $SS^T = I$, hence S is an orthogonal matrix. □

Theorem 24.3 (Riemann)**.** *Let $(\mathcal{M}, g)$ be a Riemannian manifold. The vanishing of the curvature R is necessary and sufficient for the existence of local coordinates $(x^1, \ldots, x^m)$ in which the metric g takes the Euclidean form*

$$\sum_{\mu=1}^{n} \mathrm{d}x^\mu \otimes \mathrm{d}x^\nu\,, \qquad g_{\mu\nu} = \delta_{\mu\nu}\,. \tag{24.12}$$

Proof. We apply the previous theorem to $\mathcal{B} = \mathrm{T}\mathcal{M}$, $h = g$. We then have the following: $R = 0$ is necessary and sufficient for the existence locally of an orthonormal frame field $(E_1, \ldots, E_n)$ relative to which the connection coefficients $\Gamma^\mu_{\alpha\beta} = 0$. Consider the dual basis 1-forms $(\theta^1, \ldots, \theta^n)$. Recall the general formula

$$\mathrm{d}\theta^\mu = -\frac{1}{2}\Lambda^\mu_{\alpha\beta}\theta^\alpha \wedge \theta^\beta \tag{24.13}$$

where the $\Lambda^\mu_{\alpha\beta}$ are the structure functions of $(E_1, \ldots, E_n)$.

$$\Lambda^\mu_{\alpha\beta} = \Gamma^\mu_{\alpha\beta} - \Gamma^\mu_{\beta\alpha}. \tag{24.14}$$

Hence $\Gamma^\mu_{\alpha\beta} = 0$ implies $\Lambda^\mu_{\alpha\beta} = 0$, and thus we have

$$\mathrm{d}\theta^\mu = 0\,. \tag{24.15}$$

This implies the existence, locally, of functions x^μ such that $\theta^\mu = \mathrm{d}x^\mu$. It follows that

$$(E_1, \ldots, E_n) = \Big(\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^1}\Big) \tag{24.16}$$

henec

$$\delta_{\mu\nu} = g(E_\mu, E_\nu) = g\Big(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\Big) \tag{24.17}$$

and

$$g = \sum_{\mu=1}^{n} \mathrm{d}x^\mu \otimes \mathrm{d}x^\mu \,. \tag{24.18}$$

□

Bibliography

Spivak, Michael. *A Comprehensive Introduction to Differential Geometry.* Vol. I. Publish or Perish, Inc, 2005.