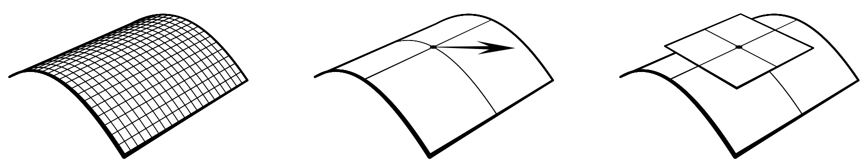
Introduction to differential geometry

Bas Janssens

2023/2024



Contents

Preface

In these notes we first develop the fundamentals of differential geometry, and then specialize to Riemannian geometry. Subsequently we describe Einstein's theory of general relativity, which is built on the foundations of Riemannian geometry.

To a large extent we follow the excellent monograph *Introduction to smooth manifolds* by John M. Lee [L03]. A notable exception is that we introduce tangent vectors using curves rather than derivations. Further we made extensive use of *Riemannian geometry: an introduction to curvature* by Lee [L97] and *Gauge fields, knots and gravity* by John Baez and Javier Muniain [BM94] for the description of Riemannian geometry and general relativity.

There are many topics which one could reasonably expect in an introduction to differential geometry, but which are not covered in these notes. Although it would take too long to list all of them, the most painful omission is undoubtedly the theory of integration on manifolds and Stokes' theorem, for which we refer to Chapter 16 of [L03]. Two further topics on the interface of mathematics and physics that are missing from this introductory text are the appearance of symplectic geometry in hamiltonian mechanics, and the use of principal bundles and connections in gauge theory. For the reader whose interest has been primed by these brief notes, the above sources may provide a useful starting point for further study.

Notation

For two sets X and Y, we denote by $X \setminus Y$ the set of all elements of X that are not in Y. The intersection is denoted by $X \cap Y$, the union by $X \cup Y$, and the disjoint union by $X \sqcup Y$. For a function $F\colon X \to Y$ and subsets $U \subseteq X$ and $V \subseteq Y$, we denote by $F(U) := \{F(x)\,;\, x \in U\}$ the image of U in Y, and by $F^{-1}(V) := \{x \in X\,;\, F(x) \in V\}$ the preimage of V in X. We denote by $\mathbb{C}^\times$ the complex numbers with 0 deleted.

1 Euclidean and Minkowski geometry

It is perhaps fair to say that modern geometry started in 1872, when Felix Klein formulated his *Erlanger Programm*. The central idea of this program is to study *geometry* in terms of its group of *symmetries*. We start by considering two types of geometry: *Euclidean geometry* and *Minkowski geometry*.

In the case of Euclidean geometry, the Erlanger Programm amounts to studying Euclidean space in terms of the Euclidean motion group. This is the group of all transformations that preserve the Euclidean metric.

Minkowski geometry is the type of geometry that captures Einstein's special theory of relativity. The Erlanger Programm then amounts to studying the Poincaré group, the group of all transformations that preserve the Minkowski metric.

Both Euclidean geometry and Minkowski geometry are essentially linear in nature. In Euclidean geometry, the difference between two points is a vector in $\mathbb{R}^n$. The Euclidean metric is adapted to the linear structure, in the sense that the shortest path between two points is a straight line. The central topic in this course is to develop *Riemannian geometry*, a generalization of Euclidean geometry to a certain type of nonlinear spaces called *smooth manifolds*.

In special relativity, the difference between two points in space-time is a vector in $\mathbb{R}^4$, and the speed of light is encoded by the Minkowski metric. Just like Euclidean geometry, Minkowski geometry is linear in the sense that every light ray is a line segment. And just like Euclidean geometry can be generalized to Riemannian geometry, Minkowski geometry has a nonlinear counterpart called *Lorentzian geometry*. The step from Minkowski geometry to Lorentzian geometry is what allowed Einstein to make the step from special to general relativity.

After a brief introduction to Euclidean and Minkowski geometry in Section 1, we will spend most of our time developing the theory of smooth manifolds, in Sections 2 through 8. These are the foundations on which we construct Riemannian geometry in Section 9. Finally, in Section 10, we take the step from Riemannian to Lorentzian geometry, arriving at Einstein's theory of general relativity.

1.1 Euclidean geometry and the Euclidean motion group

In Euclidean geometry, the central postulate is that

Distance is the same in all orthogonal frames.

Let us make this precise. We model Euclidean space $\mathbb{E}^n$ by the set $\mathbb{R}^n$, equipped with the metric

$$d(p,q) = \sqrt{(p-q, p-q)}, \tag{1}$$

where $(v,w) := v^1w^1 + \ldots + v^nw^n$ denotes the standard inner product.

The point $\mathbb{E}^0$, the line $\mathbb{E}^1$, the plane $\mathbb{E}^2$ and the space $\mathbb{E}^3$ are perhaps the most familiar examples. But once you put them on the same footing by writing

$\mathbb{E}^n$ with $n \in \{0, 1, 2, 3\}$, you soon realise that you may as well allow Euclidean spaces of arbitrary dimension $n \in \mathbb{N}$.

A *coordinate system* is a bijection $x \colon \mathbb{E}^n \to \mathbb{R}^n$, assigning to each point $p \in \mathbb{E}^n$ a unique set $x(p) := (x^1(p), \ldots, x^n(p))$ of coordinates. The metric on $\mathbb{E}^n$ singles out a distinguished set of coordinates on $\mathbb{E}^n$, namely the *orthogonal frames*. An orthogonal frame is a set of coordinates in which the distance is represented correctly,

$$d(p, q) = d(x(p), x(q)) \qquad \text{for all} \qquad p, q \in \mathbb{E}^n.$$

Suppose that a different observer uses a different orthogonal frame $\overline{x}(p) = (\overline{x}^1(p), \ldots, \overline{x}^n(p))$. Then we may express the new coordinates $\overline{x}(p)$ in terms of the old coordinates $x(p)$ by means of a *coordinate transformation*

$$\overline{x}(p) = \kappa(x(p)).$$

Here $\kappa \colon \mathbb{R}^n \to \mathbb{R}^n$ is a transformation that preserves the metric,

$$d(\kappa(x), \kappa(y)) = d(x, y) \qquad \text{for all} \qquad x, y \in \mathbb{R}^n. \tag{2}$$

In other words, κ is an *isometry* with respect to d.

Definition 1.1 (Euclidean motion group)**.** The group of all isometries of the Euclidean metric d is called the *Euclidean motion group $E(n)$*.

This is clearly a group: the identity transformation is an isometry, if two transformations κ and κ' preserve the Euclidean metric, then so does their concatenation $\kappa \circ \kappa'$, and if κ is an isometry, then so is its inverse κ^{-1}.

A simple example of an isometry is the translation $T_v \colon \mathbb{R}^n \to \mathbb{R}^n$, defined by $T_v(x) := x + v$. It shifts every point by $v \in \mathbb{R}^n$. A linear map $R \colon \mathbb{R}^n \to \mathbb{R}^n$ is called an *orthogonal transformation* if it preserves the inner product, $(Rx, Ry) = (x, y)$ for all $x, y \in \mathbb{R}^n$. With respect to an orthonormal basis in $\mathbb{R}^n$, every orthogonal transformation is represented by an orthogonal $n \times n$ matrix. We can therefore identify the group of orthogonal transformations with

$$\mathrm{O}(n) = \{R \in M_n(\mathbb{R})\,;\, R^T R = \mathrm{Id}_n\}.$$

In dimension 2 and 3, orthogonal transformations are either rotations or reflections.

Problem 1.2. Equipped with the matrix multiplication, $\mathrm{O}(n)$ is a group.

Every orthogonal transformation $R \colon \mathbb{R}^n \to \mathbb{R}^n$ is an isometry, since

$$d(Rx, Ry)^2 = (Rx - Ry, Rx - Ry)) = (x - y, x - y) = d(x, y)^2.$$

Since both the translation T_v and the orthogonal transformation R are isometries, so is their product

$$(T_v \circ R)(x) = Rx + v. \tag{3}$$

In fact, it turns out that *every* isometry is of this form.

Theorem 1.3. *Every element of $E(n)$ can be decomposed into an orthogonal transformation and a translation,*

$$E(n) = \{T_v \circ R \,;\, v \in \mathbb{R}^n, R \in \mathrm{O}(n)\}.$$

To prove Theorem 1.3, we will need a number of lemmas. Let $\kappa \colon \mathbb{R}^n \to \mathbb{R}^n$ be an isometry. If we set $v := \kappa(0)$, then the map $R := T_{-v} \circ \kappa$ is an isometry that fixes the origin, $R(0) = 0$. To prove the theorem, it suffices to prove that $R \in \mathrm{O}(n)$. We start by showing that R preserves midpoints.

In the Euclidean space $\mathbb{E}^n$, the *midpoint* $m(x, y) := \frac{1}{2}(x + y)$ between x and y is the unique point m that satisfies

$$d(m, x) = d(m, y) = \tfrac{1}{2} d(x, y).$$

Lemma 1.4. *Every isometry κ preserves midpoints:*

$$\kappa(m(x, y)) = m(\kappa(x), \kappa(y)).$$

Proof. If m satisfies $d(m, x) = d(m, y) = \frac{1}{2} d(x, y)$, then since κ is an isometry, $\kappa(m)$ satisfies $d(\kappa(m), \kappa(x)) = d(\kappa(m), \kappa(y)) = \frac{1}{2} d(\kappa(x), \kappa(y))$. It follows that $\kappa(m)$ is a midpoint between $\kappa(x)$ and $\kappa(y)$. Since midpoints are unique in Euclidean space, we have $m(\kappa(x), \kappa(y)) = \kappa(m(x, y))$. □

Next, we show that every isometry $R \colon \mathbb{R}^n \to \mathbb{R}^n$ with $R(0) = 0$ preserves the inner product, $(R(x), R(y)) = (x, y)$ for all $x, y \in \mathbb{R}^n$. For this we use the following *polarisation identity*, which expresses the inner product in terms of the metric.

Lemma 1.5 (Polarisation identity)**.** *For $x, y \in \mathbb{R}^n$, we have*

$$(x, y) = d(m(x, y), 0)^2 - \tfrac{1}{4} d(x, y)^2.$$

Proof. Since $d(x, y)^2 = (x - y, x - y)$ and $d(x + y, 0)^2 = (x + y, x + y)$, we have

$$d(x, y)^2 = (x - y, x - y) = (x, x) - 2(x, y) + (y, y) \quad \text{and} \tag{4}$$

$$d(x + y, 0)^2 = (x + y, x + y) = (x, x) + 2(x, y) + (y, y). \tag{5}$$

Since $d(x{+}y, 0)^2 = 4d(m(x, y), 0)^2$, the result follows by subtracting (4) from (5) and dividing the result by four. □

Since the isometry R preserves midpoints and the origin, it preserves the inner product as well:

$$\begin{aligned}
(R(x), R(y)) &= d\big(m(R(x), R(y)), 0\big)^2 - \tfrac{1}{4} d\big(R(x), R(y)\big)^2 \\
&= d\big(R(m(x, y)), R(0)\big)^2 - \tfrac{1}{4} d(x, y)^2 \\
&= d\big(m(x, y), 0\big)^2 - \tfrac{1}{4} d(x, y)^2 \\
&= (x, y).
\end{aligned}$$

It remains to show that R is linear. Let $e_1, \ldots, e_n$ be an orthonormal basis of $\mathbb{R}^n$. Since R preserves inner products as well as the origin, the vectors $e'_1 := R(e_1)$, $\ldots$, $e'_n := R(e_n)$ are orthonormal. In particular the n vectors $e'_1, \ldots, e'_n$ are independent, and they form a basis of $\mathbb{R}^n$.

The coefficients of $x \in \mathbb{R}^n$ with respect to the orthonormal basis $e_1, \ldots, e_n$ are $x_i = (x, e_i)$. Similarly, the coefficients of Rx with respect to e'_i are $x'_i := (Rx, e'_i)$. But since R preserves the inner product, these coefficients are the same:

$$x'_i = (Rx, e'_i) = (Rx, Re_i) = (x, e_i) = x_i.$$

It follows that

$$R\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i e'_i.$$

In particular, R is a *linear* map. And since R preserves inner products, it is an orthogonal linear transformation. This concludes the proof of Theorem 1.3.

Problem 1.6. Show that every element of $\mathrm{O}(2)$ is of the form

$$R = S\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

where S is one of the two matrices

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Give an explicit description of the Euclidean motion group $E(2)$.

The central theme of Klein's Erlanger programm is to study geometry in terms of its group of symmetries. In the following problem, we express the geometric notion of *length* in group theoretic terms.

Problem 1.7. Consider the diagonal action of the Euclidean motion group $E(2)$ on $\mathbb{R}^2 \times \mathbb{R}^2$, defined by $\kappa \cdot (x, y) = (\kappa(x), \kappa(y))$. Then two line segments AB and $A'B'$ are of equal length if and only if their endpoints $(A, B), (A', B') \in \mathbb{R}^2 \times \mathbb{R}^2$ are in the same orbit of $E(2)$.

Similarly, the geometric problem of *congruence* of triangles can be expressed in terms of orbits for the Euclidean motion group.

Problem 1.8. Consider the diagonal action of the Euclidean motion group $E(2)$ on $\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2$, defined by $\kappa \cdot (x, y, z) = (\kappa(x), \kappa(y), \kappa(z))$. Then two triangles ABC and $A'B'C'$ are congruent if and only if their vertices $(A, B, C), (A', B', C') \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2$ are in the same orbit of $E(2)$.

1.2 Special Relativity and the Lorentz group

In 1905, Einstein introduced his special theory of relativity. It features a special type of coordinate systems called *inertial frames*. The central postulate in the special theory of relativity is that

The speed of light has the same value c in all inertial frames.

If you have not seen this before, you may wish to pause here for a second, and reflect on how strange this statement is.

Suppose that observer A uses coordinates (t, x, y, z) to describe space–time. A ray of light comes by, travelling in the x-direction. Observer A measures its coordinates to be $x(t) = ct$, and concludes that it has speed $\Delta x/\Delta t = c$. Now imagine that observer B is moving at speed v in the x-direction relative to observer A. If observer B is at rest with respect to the coordinates $(\bar{t}, \bar{x}, \bar{y}, \bar{z})$, then it seems reasonable to relate these coordinates by the *Galilei transformation*

$$\bar{t} = t, \quad \bar{x} = x - vt, \quad \bar{y} = y, \quad \bar{z} = z. \tag{6}$$

Observer B will describe the ray of light by $\bar{x}(\bar{t}) = (c - v)\bar{t}$, and conclude that it moves at speed $\Delta\bar{x}/\Delta\bar{t} = c - v$. This means that observer A and B come to different conclusions as to the speed of light, violating the central postulate in the special theory of relativity.

Apparently, then, inertial frames are *not* related by Galilei transformations of the type (6). This leaves us with two burning questions: If inertial frames are not related by Galilei transformations, how *should* they be related? And, for that matter, what *are* these inertial frames precisely?

In order to answer these questions, Hermann Minkowski introduced a type of geometry that is radically different from Euclidean geometry. The *Minkowski space* $\mathbb{M}^4$ is the set $\mathbb{R}^4$, equipped with the *Minkowski metric*

$$S^2(p, q) = \eta(p - q, p - q),$$

where $\eta \colon \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$ is the nondegenerate bilinear form

$$\eta(v, w) := -v_0 w_0 + v_1 w_1 + v_2 w_2 + v_3 w_3\,. \tag{7}$$

The word 'metric' is misleading, since the 'Minkowski metric' is not a metric in the usual sense of the word. The minus sign in (7) makes that $\eta(v, v)$ can be negative or even zero, and the same holds for $S^2(p, q)$.

An *inertial frame* for $\mathbb{M}^4$ is a coordinate system $p \mapsto (t(p), x(p), y(p), z(p))$ that respects the Minkowski metric in the following sense:

$$S^2(p, q) = S^2\Big(\big(ct(p), x(p), y(p), z(p)\big), \big(ct(q), x(q), y(q), z(q)\big)\Big). \tag{8}$$

Just like an orthogonal frame is a linear coordinate system in which the inner product on Euclidean space takes the standard form, an inertial system is a linear coordinate system on $\mathbb{M}^4$ in which the Minkowski metric takes a standard form.

The significance of these coordinate systems is that *the speed of light has the same value c in every inertial frame*. Indeed, suppose that a beam of light is emitted at a space-time point $p \in \mathbb{M}^4$, and absorbed at a space–time point $q \in \mathbb{M}^4$. Two observers, using two different inertial frames t, x, y, z and $\bar{t}, \bar{x}, \bar{y}, \bar{z}$, will then calculate the speed of light as follows.

The first observer, using coordinates t, x, y, z, will assign space–time coordinates $t(p), x(p), y(p), z(p)$ and $t(q), x(q), y(q), z(q)$ to the emission and absorption of the light ray. Calculating the differences $\Delta t = t(q) - t(p)$, $\Delta x = x(q) - x(p)$, $\Delta y = y(q) - y(p)$ and $\Delta z = z(q) - z(p)$, this observer concludes that the speed of light is

$$c = \frac{\sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}}{\Delta t},$$

or, equivalently, that

$$0 = -c^2(\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2. \tag{9}$$

Since this is equivalent to $S^2(p, q) = 0$, the other observer, using inertial coordinates $\overline{t}, \overline{x}, \overline{y}, \overline{z}$, will conclude that

$$0 = -c^2(\Delta\overline{t})^2 + (\Delta\overline{x})^2 + (\Delta\overline{y})^2 + (\Delta\overline{z})^2, \tag{10}$$

and hence that

$$\frac{\sqrt{(\Delta\overline{x})^2 + (\Delta\overline{y})^2 + (\Delta\overline{z})^2}}{\Delta\overline{t}} = c$$

for the *same* constant c. In other words, the two observers, using different inertial frames t, x, y, z and $\overline{t}, \overline{x}, \overline{y}, \overline{z}$, will measure the *same* speed c.

To express the inertial frame $\overline{t}, \overline{x}, \overline{y}, \overline{z}$ of one observer in terms of the inertial frame t, x, y, z of the other observer, we need a transformation $\kappa\colon \mathbb{R}^4 \to \mathbb{R}^4$ that respects the Minkowski metric.

Definition 1.9 (Poincaré group)**.** The Poincaré group P is the group of all bijections $\kappa\colon \mathbb{R}^4 \to \mathbb{R}^4$ such that $S^2(\kappa(p), \kappa(q)) = S^2(p, q)$.

Just like the Euclidean motion group is the group of symmetries of $(\mathbb{E}^n, d)$, the Poincaré group is the group of symmetries of $(\mathbb{M}^4, S^2)$. Clearly, every translation $T_v(\xi) = \xi + v$ is an element of the Poincaré group. A *Lorentz transformation* is a linear map $\Lambda\colon \mathbb{R}^4 \to \mathbb{R}^4$ that respects the bilinear form η,

$$\eta(\Lambda v, \Lambda w) = \eta(v, w) \quad \text{for all} \quad v, w \in \mathbb{R}^4. \tag{11}$$

The group of Lorentz transformations is denoted $\mathrm{O}(3, 1)$. Every Lorentz transformation is an element of the Poincaré group, since

$$S^2(\Lambda v, \Lambda w) = \eta\Big(\Lambda(v - w), \Lambda(v - w)\Big) = \eta(v - w, v - w) = S^2(v, w).$$

In fact, *every* element of the Poincaré group is of the form $\kappa(\xi) = \Lambda\xi + v$.

Theorem 1.10. *Every element of the Poincaré group P can be decomposed in a translation and a Lorentz transformation,*

$$P = \{T_v \circ \Lambda\,;\, v \in \mathbb{R}^4, \Lambda \in \mathrm{O}(3, 1)\}.$$

To describe the Lorentz group more explicitly, we express η in terms of the usual inner product

$$(v, w) := v_0 w_0 + v_1 w_1 + v_2 w_2 + v_3 w_3$$

and the diagonal matrix

$$H := \operatorname{diag}(-1, 1, 1, 1)$$

as $\eta(v, w) = (v, Hw)$. From (11), we then find that

$$(\Lambda v, H\Lambda w) = (v, Hw)$$

for all v, w, so that $\Lambda^T H \Lambda = H$. We find that the Lorentz group is given by

$$\mathrm{O}(3,1) = \{\Lambda \in M(4, \mathbb{R})\,;\, \Lambda^T H \Lambda = H\}.$$

Problem 1.11. Prove that $\mathrm{O}(3,1)$ is a group.

Problem 1.12. For two-dimensional Minkowski space, the corresponding symmetry group is

$$\mathrm{O}(1,1) = \{\Lambda \in M(2, \mathbb{R})\,;\, \Lambda^T H \Lambda = \eta\},$$

where $H = \operatorname{diag}(-1, 1)$ is the diagonal matrix with entries -1 and 1. Show that every element $\Lambda \in \mathrm{O}(1,1)$ is of the form

$$\Lambda = R \begin{pmatrix} \cosh(\chi) & \sinh(\chi) \\ \sinh(\chi) & \cosh(\chi) \end{pmatrix},$$

with $\chi \in \mathbb{R}$ and where R is one of the four matrices

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad T = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad PT = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Problem 1.13 (Time dilation and length contraction)**.** Observer B moves with speed v in the x-direction relative to observer A. Observer B is at rest with respect to the inertial frame (t, x, y, z), and observer A is at rest with respect to the inertial frame $(\overline{t}, \overline{x}, \overline{y}, \overline{z})$. There is one point $\mathcal{O}$ in space–time at which both observers coincide, and they agree to put their space–time origin there. Since their relative motion is in the x-direction, both observers agree on the y and z-coordinates. The relation between their inertial frames is given by the Lorentz transformation

$$\begin{pmatrix} c\overline{t} \\ \overline{x} \\ \overline{y} \\ \overline{z} \end{pmatrix} = \begin{pmatrix} \cosh(\chi) & \sinh(\chi) & 0 & 0 \\ \sinh(\chi) & \cosh(\chi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}.$$

Since observer B is at rest with respect to the inertial frame (t, x, y, z), her space–time coordinates in this frame are described by the curve

$$(t(\tau), x(\tau), y(\tau), z(\tau)) = (\tau, 0, 0, 0).$$

a) Express the space–time position of observer B in the inertial frame $(\bar{t}, \bar{x}, \bar{y}, \bar{z})$, and determine the speed $v = \Delta\bar{x}/\Delta\bar{t}$ in terms of χ.

b) Determine $\cosh(\chi)$ and $\sinh(\chi)$ in terms of v.

c) According to observer B, a time $\Delta t = \tau$ has elapsed between the space–time points with coordinates $(t, x, y, z) = (0, 0, 0, 0)$ and $(t, x, y, z) = (\tau, 0, 0, 0)$. Determine the amount of time $\Delta\bar{t}$ that has elapsed according to observer A, and derive the *time dilation* formula

$$\Delta\bar{t}/\Delta t = \frac{1}{\sqrt{1 - v^2/c^2}}.$$

In particular, a moving clock will tick more slowly than one that is standing still.

c) Observer B is holding a rod of length L, and she points it in the x-direction. In her inertial system (t, x, y, z), the two endpoints of the rod then trace out the curves $\tau \mapsto (\tau, 0, 0, 0)$ and $s \mapsto (s, L, 0, 0)$. Transform these two lines to the inertial system $(\bar{t}, \bar{x}, \bar{y}, \bar{z})$, and determine the difference $\Delta\bar{x}$ between the $\bar{x}$-coordinates of these two curves at $\bar{t} = 0$. The length of the rod as measured by observer A is then $\overline{L} = |\Delta\bar{x}|$. Derive the *length contraction* formula

$$\overline{L}/L = \sqrt{1 - v^2/c^2}.$$

In particular, a moving rod will be shorter than one that is standing still.

Problem 1.14 (The orthochronous Lorentz group)**.** Let $H = \{v \in \mathbb{R}^4\,;\, \eta(v, v) = -1\}$ be the hyperboloid in $\mathbb{R}^4$.

a) Show that it consists of two sheets $H^{\pm} = \{v \in H\,;\, \pm v_0 \geq 1\}$.

b) Show that every $g \in \mathrm{O}(3, 1)$ maps H to itself, $g(v) \in H$ for all $v \in H$.

c) The *orthochronous Lorentz group* $\mathrm{O}^{\uparrow}(3, 1)$ is the set of all $g \in \mathrm{O}(3, 1)$ that map H^+ to itself, $g(v) \in H^+$ for all $v \in H^+$. Show that $\mathrm{O}^{\uparrow}(3, 1)$ is a normal subgroup of index 2.

d) Show that $\mathrm{O}^{\uparrow}(3, 1)$ is both open and closed in $\mathrm{O}(3, 1)$ with respect to the subspace topology of $\mathrm{O}(3, 1) \subseteq M_4(\mathbb{R})$.

2 Manifolds

In many applications, one has to deal with (partial) differential equations on spaces M with a geometry which is *not* that of the familiar Euclidean space $\mathbb{R}^n$. To handle such situations, we need to understand what it means for a function to be differentiable. The notion of a *smooth manifold* allows one to handle *differentiability* in a rather general context.

Roughly speaking, a smooth manifold is a topological space that 'locally looks like $\mathbb{R}^n$'. This allows us to transport the entire machinery of calculus to the world of differential geometry. Also, one can formulate partial differential equations in this setting, such as the Einstein equation in general relativity, the Poisson equation in harmonic analysis, and the Cauchy–Riemann equation in complex geometry.

2.1 Definition of a smooth manifold

We want to define a *manifold* M as a topological space that 'locally looks like $\mathbb{R}^n$'. Using the tools from Appendix A, we can make this statement more precise: we require that every point $p \in M$ has a neighbourhood $U \subseteq M$ that is homeomorphic to an open subset of $\mathbb{R}^n$.

Definition 2.1 (Charts)**.** A *chart* on M is a homeomorphism

$$\phi\colon M \supseteq U \to \phi(U) \subseteq \mathbb{R}^n$$

from an open subset $U \subseteq M$ to an open subset $\phi(U) \subseteq \mathbb{R}^n$.

In other words, ϕ identifies an open neighbourhood U in M with an open neighbourhood $\phi(U)$ in $\mathbb{R}^n$. The open subsets $U \subseteq M$ are called *coordinate patches*, and we think of $\phi(p) \in \mathbb{R}^n$ as the coordinates of p with respect to the chart (U, ϕ). To emphasize this point of view, we write

$$\phi(p) = (x^1(p), \ldots, x^n(p)),$$

where $x^\mu(p)$ is the μth coordinate of p with respect to the chart (U, ϕ).

Definition 2.2 (Topological Atlas)**.** Let $\mathcal{A} = \{(U_\alpha, \phi_\alpha)\,;\, \alpha \in A\}$ be a collection of charts, labelled by an index set A. We say that $\mathcal{A}$ is a *topological atlas* for M if M is the union of the coordinate patches, $M = \bigcup_{\alpha\in A} U_\alpha$.

Let (U_α, ϕ_α) and (U_β, ϕ_β) be two charts with nonempty intersection $U_\alpha \cap U_\beta$, and let $p \in U_\alpha \cap U_\beta$. Then the *transition function* $\kappa_{\alpha\beta}$ maps the coordinates $\phi_\alpha(p) = (x^1, \ldots, x^n)$ of p with respect to the chart (U_α, ϕ_α) to the coordinates $\phi_\beta(p) = (x^{\overline{1}}, \ldots, x^{\overline{n}})$ of the same point p with respect to the other chart (U_β, ϕ_β). Note that the the transition function $\kappa_{\alpha\beta}$ is only defined if p is in the intersection $U_\alpha \cap U_\beta$ of the two coordinate patches, and it is completely determined by the two charts ϕ_α and ϕ_β,

$$\kappa_{\alpha\beta}\colon \phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta) \quad \text{is defined by} \quad \kappa_{\alpha\beta} := \phi_\beta \circ \phi_\alpha^{-1}. \tag{12}$$

Since ϕ_α and ϕ_β are homeomorphisms, the transition functions $\kappa_{\alpha\beta} = \phi_\beta \circ \phi_\alpha^{-1}$ and $\kappa_{\beta\alpha} = \phi_\alpha \phi_\beta^{-1}$ are automatically homeomorphisms as well.

Definition 2.3 (Smooth Atlas)**.** Two charts (U_α, ϕ_α) and (U_β, ϕ_β) are called *compatible* if both $\kappa_{\alpha\beta}$ and $\kappa_{\beta\alpha}$ are *smooth*. A topological atlas $\mathcal{A}$ is called *smooth* is all its charts are compatible.

We define a *smooth manifold* as a Hausdorff topological space M, together with a collection $\mathcal{A}$ of coordinate patches that covers M, and for which all transition functions are smooth.

Definition 2.4 (Smooth manifold)**.** A *smooth manifold* is a Hausdorff topological space M, together with a smooth atlas $\mathcal{A}$. We say that M is of dimension n if its charts take values in $\mathbb{R}^n$.

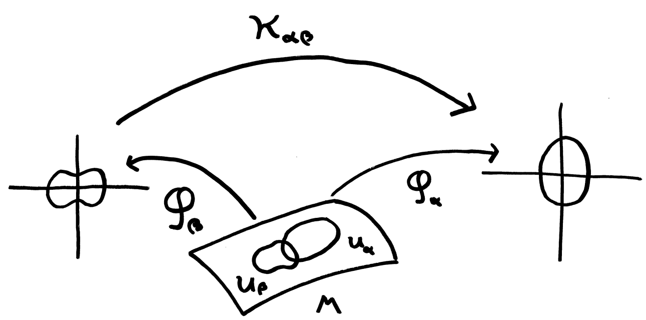


Figure 1: Transition function between two charts (U_α, ϕ_α) and (U_β, ϕ_β)

Remark 2.5 (Equivalent atlases)**.** Two smooth atlases $\mathcal{A}_1$ and $\mathcal{A}_2$ are called *equivalent* if their charts are mutually compatible. We will usually not distinguish the corresponding manifold structures on M.

Remark 2.6. Sometimes non-Hausdorff manifolds also make an appearance in the literature (e.g. in connection to foliations), but this remains somewhat exceptional. In contrast to the usual definition in the literature, we will *not* require M to be second countable.

2.2 Smooth functions

For a topological space X, there is a natural notion of *continuous* functions $f\colon X \to \mathbb{R}$ (cf. Appendix A.2). For a manifold M, there is even a natural notion of *smooth* functions $f\colon M \to \mathbb{R}$, which is not available in general topological spaces. Using a chart (U_α, ϕ_α) around $p \in M$, we can locally represent the function f on the subset U_α of M by the *coordinate representation* $f_\alpha\colon \phi_\alpha(U_\alpha) \to \mathbb{R}$, defined by

$$f_\alpha := f \circ \phi_\alpha^{-1}.$$

The coordinate representation f_α of f tells us what f looks like if you use the coordinates (ϕ_α, U_α). Indeed, if the coordinates of p are $\phi_\alpha(p) = (x^1, \ldots, x^n)$, then $f(p) = f_\alpha(x^1, \ldots, x^n)$. Since f_α is defined on the open subset $\phi_\alpha(U_\alpha)$ of $\mathbb{R}^n$, it makes perfect sense to ask for f_α to be smooth at $(x^1, \ldots, x^n)$!

Definition 2.7 (Differentiable and smooth functions)**.** A function $f\colon M \to \mathbb{R}$ is *differentiable/smooth* at $p \in M$ if for some chart U_α containing p, the coordinate representation $f_\alpha\colon \mathbb{R}^n \supset \phi_\alpha(U_\alpha) \to \mathbb{R}$ is differentiable/smooth at $\phi_\alpha(p) \in \mathbb{R}^n$.

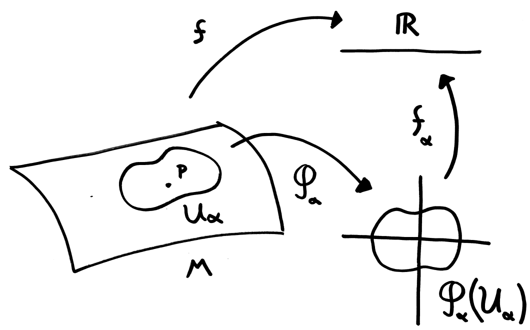


Figure 2: Coordinate representation $f_\alpha = f \circ \phi_\alpha^{-1}$

In fact, the above definition does not depend on the choice of coordinates. In order to show this, we need the smoothness of the transition functions. Suppose that $p \in U_\alpha \cap U_\beta$ lies in the intersection of two coordinate patches U_α and U_β, and suppose that the coordinate representation $f_\alpha = f \circ \phi_\alpha^{-1}$ is smooth at $\phi_\alpha(p)$. Then since the transition function $\kappa_{\beta\alpha} := \phi_\alpha \circ \phi_\beta^{-1}$ is smooth at $\phi_\beta(p)$, the chain rule applied to

$$(f \circ \phi_\alpha^{-1}) \circ \kappa_{\beta\alpha} = (f \circ \phi_\alpha^{-1}) \circ (\phi_\alpha \circ \phi_\beta^{-1}) = f \circ (\phi_\alpha^{-1} \circ \phi_\alpha) \circ \phi_\beta^{-1} = f \circ \phi_\beta^{-1} \quad (13)$$

shows that $f_\beta = f \circ \phi_\beta^{-1}$ is smooth at $\phi_\beta(p)$ as well. A similar argument with α and β interchanged shows that f_α is smooth at $\phi_\alpha(p)$ if and only if f_β is smooth at $\phi_\beta(p)$.

We say that $f\colon M \to \mathbb{R}$ is *smooth* if it is smooth at any point $p \in M$, and we denote the set of smooth functions by $C^\infty(M)$. As on $\mathbb{R}^n$, adding or multiplying smooth functions yields smooth functions again.

Proposition 2.8 ($C^\infty(M)$ is an algebra)**.** *If $f, g \in C^\infty(M)$ and $\lambda \in \mathbb{R}$, then*

(1) $\lambda f \in C^\infty(M)$,

(2) $f + g \in C^\infty(M)$,

(3) $fg \in C^\infty(M)$.

Proof. For (3), note that if f and g are smooth in $p \in M$, then their coordinate representations $f \circ \phi_\alpha^{-1}$ and $g \circ \phi_\alpha^{-1}$ with respect to the chart (U_α, ϕ_α) are smooth at $\phi_\alpha(p)$. But then the coordinate representation $(fg) \circ \phi_\alpha^{-1} = (f \circ \phi_\alpha^{-1}) \cdot (g \circ \phi_\alpha^{-1})$ is smooth as well, so fg is smooth at p. The proof of property (2) is similar, and (1) follows from (3) applied to the constant function $g = \lambda \mathbf{1}$. □

2.3 Examples of manifolds

To get a feeling for the definition of a manifold, let us consider two examples: *spheres* and *tori*.

2.3.1 The sphere

To get a feeling for the precise definition of a manifold, we consider the 2-sphere

$$\mathbb{S}^2 := \left\{(\xi, \eta, \zeta) \in \mathbb{R}^3 \,;\, \xi^2 + \eta^2 + \zeta^2 = 1\right\}.$$

Since $\mathbb{S}^2$ is a subset of $\mathbb{R}^3$, it is a topological space with the subset topology. We show that $\mathbb{S}^2$ is a smooth manifold that is covered by *two* coordinate charts.

Let $n := (0, 0, 1)$ be the North Pole in $\mathbb{S}^2$, and let $s := (0, 0, -1)$ be the South Pole. The first way of assigning coordinates $(x, y) \in \mathbb{R}^2$ to a point p in $\mathbb{S}^2$ is good for every point $p \in \mathbb{S}^2$ *except* the north pole. Draw a straight line $\overline{pn}$ from p to the north pole n, and let q be the intersection of $\overline{pn}$ with the xy-plane. If $q = (x, y, 0)$, then we say that the coordinates of p are (x, y).

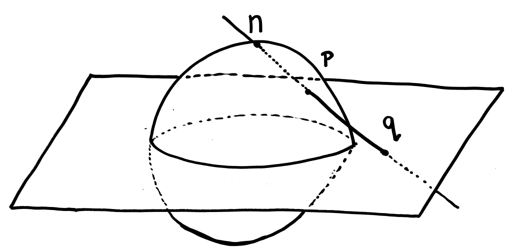


Figure 3: A chart for $\mathbb{S}^2$ centered at the south pole

Problem 2.9. The coordinates (x, y) of a point $p = (\xi, \eta, \zeta)$ in $\mathbb{S}^2$ are

$$(x, y) = \left(\frac{\xi}{1-\zeta}, \frac{\eta}{1-\zeta}\right).$$

The open set $U_n := \mathbb{S}^2 \setminus \{n\}$ where these coordinates make sense the *coordinate neighbourhood*, and the map $\phi_n \colon \mathbb{S}^2 \supseteq U_n \to \mathbb{R}^2$ that assigns to the point $p = (\xi, \eta, \zeta)$ its coordinates

$$\phi_1(p) = \left(\frac{\xi}{1-\zeta}, \frac{\eta}{1-\zeta}\right) \tag{14}$$

is called a *chart*. This type of chart is useful for a cartographer living at the South Pole, but it has terrible distortion close to the North Pole.

Problem 2.10. Show that $\phi_n \colon U_n \to \mathbb{R}^2$ is a homeomorphism. (Problem A.20 may be of use here.)

The second way of assigning coordinates to p is analogous, except that n and s are reversed. We now draw a line $\overline{ps}$ from p to the South Pole s, and determine the coordinates $(\overline{x}, \overline{y})$ from the intersection $q' = (\overline{x}, \overline{y}, 0)$ of $\overline{ps}$ with the xy-plane. These coordinates make sense in the coordinate neighbourhood $U_s = \mathbb{S}^2 \setminus \{s\}$, and the corresponding coordinate chart $\phi_s \colon U_s \to \mathbb{R}^2$ is

$$\phi_s(p) = \left(\frac{\xi}{1+\zeta}, \frac{\eta}{1+\zeta} \right) . \tag{15}$$

Needless to say, this type of chart is useful for a cartographer living at the North Pole.

We have covered $\mathbb{S}^2$ with two coordinate neighbourhoods $U_n = \mathbb{S}^2 \setminus \{n\}$ and $U_s = \mathbb{S}^2 \setminus \{s\}$. To check that it is a smooth manifold, it remains to show that the transition functions are smooth. If $p \in \mathbb{S}^2$ has coordinates $(x, y) \in \mathbb{R}^2$ with respect to the chart (U_1, ϕ_1), and coordinates $(\overline{x}, \overline{y}) \in \mathbb{R}^2$ with respect to the chart (U_2, ϕ_2), then κ_{ns} maps (x, y) to $(\overline{x}, \overline{y})$. Note that this transition function only makes sense for points $p \in U_n \cap U_s$ that have a coordinate representation in both charts, so $\kappa_{ns} := \phi_s \circ \phi_n^{-1}$ is a map from $\phi_n(U_n \cap U_s)$ to $\phi_s(U_n \cap U_s)$.

Problem 2.11. Show that $\phi_n(U_n \cap U_s) = \phi_s(U_n \cap U_s) = \mathbb{R}^2 \setminus \{(0,0)\}$, and give an explicit expression for $\kappa_{ns} \colon (\mathbb{R}^2 \setminus \{(0,0)\}) \to (\mathbb{R}^2 \setminus \{(0,0)\})$. Conclude that both transition functions κ_{ns} and κ_{sn} are smooth.

Now that we have covered $\mathbb{S}^2$ by the two coordinate neighbourhoods U_n and U_s, we can define what it means for a function $f \colon \mathbb{S}^2 \to \mathbb{R}$ to be differentiable. From the charts $\phi_n \colon U_n \to \mathbb{R}^2$ and $\phi_s \colon U_s \to \mathbb{R}^2$, we get two coordinate representations $f_n := f \circ \phi_n^{-1}$ and $f_s := f \circ \phi_s^{-1}$ of the function f. They tell us what f looks like for that choice of coordinates.

For the north pole $p = n$, we *have* to use the chart (U_s, ϕ_2) to see if f is differentiable at n. For the south pole s, we *have* to use the chart (U_n, ϕ_1). But for $p \in U_n \cap U_s = \mathbb{S}^2 \setminus \{n, s\}$, we can use *either* (U_n, ϕ_n) *or* (U_s, ϕ_s). This does not introduce any ambiguity: because the transition functions are smooth, $f \circ \phi_n^{-1}$ is smooth at $\phi_n(p) = (x, y)$ if and only if $f \circ \phi_s^{-1}$ is smooth at $\phi_s(p) = (\overline{x}, \overline{y})$. If we think of the two different coordinate charts (U_n, ϕ_n) and (U_s, ϕ_s) as two different cartographers trying to describe the globe, then we conclude that they agree on the differentiability of $f \colon \mathbb{S}^2 \to \mathbb{R}$ at p as soon as p occurs on both of their maps.

Problem 2.12. Is the function $f(\xi, \eta, \zeta) = \zeta$ differentiable on all of $\mathbb{S}^2$?

Problem 2.13 ($\mathbb{S}^n$ as a manifold)**.** Find coordinate charts for the n-sphere

$$\mathbb{S}^n := \{(x^0, \ldots, x^n) \in \mathbb{R}^{n+1} \, ; \, (x^0)^2 + \ldots + (x^n)^2 = 1\},$$

and prove that it is a smooth manifold of dimension n.

2.3.2 The torus

The 2-torus $\mathbb{T}^2 := \mathbb{R}^2/{\sim}$ is the quotient of $\mathbb{R}^2$ by the relation that identifies $x, x' \in \mathbb{R}^2$ if and only if $x - x' \in \mathbb{Z}^2$. Since it is the quotient of the topological space $\mathbb{R}^2$ by an equivalence relation, it is a topological space itself. It is not hard to see that $\mathbb{T}^2$ is Hausdorff.

Problem 2.14. Let $\mathbb{S}^1 := \{e^{i\phi}\,;\, \phi \in [0, 2\pi)\}$ be the unit circle. Show that the map $\mathbb{T}^2 \to \mathbb{S}^1 \times \mathbb{S}^1$ defined by $[(x, y)] \mapsto (e^{2\pi i x}, e^{2\pi i y})$ is well defined, and show that it is a bijection.

To show that $\mathbb{T}^2$ is a manifold, we introduce charts as follows. For $c \in \mathbb{R}^2$ and $0 < r < 1/2$, let $B_c(r) \subseteq \mathbb{R}^2$ be the open ball in $\mathbb{R}^2$ with centre c and radius r, that is, $B_c(r) = \{v \in \mathbb{R}^2\,;\, \|v - c\| < r\}$. For $\alpha := (c, r)$, we define $U_\alpha \subseteq \mathbb{T}^2$ as $U_\alpha := \{[v]\,;\, v \in B_c(r)\}$, and we define the inverse chart $\phi_\alpha^{-1} \colon B_c(r) \to U_\alpha$ by $\phi_\alpha^{-1}(v) = [v]$. It is bijective because $\phi_\alpha^{-1}(v) = \phi_\alpha^{-1}(w)$ implies that $v - w \in \mathbb{Z}^2$, and hence that $v = w$ because $\|v-w\| \leq \|v-c\|+\|c-w\| < 2r < 1$. One readily checks that ϕ_α is a homeomorphism, so it remains to check that the transition functions are smooth.

Problem 2.15 ($\mathbb{T}^2$ as a manifold)**.** Show that the transition functions are smooth, and conclude that $\mathbb{T}^2$ is a manifold.

The 2-torus $\mathbb{T}^2$ can be visualized as follows. Since every $[(x, y)] \in \mathbb{T}^2$ has precisely one representative (x, y) in $[0, 1) \times [0, 1)$, we can think of $\mathbb{T}^2$ as the square $[0, 1)\times[0, 1)$. Since $[(x, 0)] = [(x, 1)]$, the bottom of the square is identified with the top, and since $[(0, y)] = [(1, y)]$, the left and right hand sides of the square are identified.

Problem 2.16 ($\mathbb{T}^n$ as a manifold)**.** The n-torus is defined as $\mathbb{T}^n = \mathbb{R}^n/\!\sim$, where $v \sim w$ if $v - w \in \mathbb{Z}^n$. Find charts that cover $\mathbb{T}^n$, and show that the transition functions are smooth. Show that the charts are homeomorphisms onto their image, and conclude that $\mathbb{T}^n$ is a manifold.

Problem 2.17. The *Möbius strip* $M = \{[(x, y)]\,;\, (x, y) \in \mathbb{R}^2\}$ is the quotient of $\mathbb{R}^2$ by the following relation: $(x, y) \sim (x', y')$ if there exists an $n \in \mathbb{Z}$ such that $(x', y') = (x + n, (-1)^n y)$.

a) Make a sketch of the Möbius strip – or at least the part with $y \in (-1, 1)$.

b) Show that the quotient map $\psi\colon \mathbb{R}^2 \to M$ defined by $\psi(x, y) = [(x, y)]$ is not injective. Show that its restriction to any open ball $B_c(r)$ of radius $r < 1/4$ around $c = (x_0, y_0)$ is injective.

c) The inverse $\phi_{c,r} \colon M \supset \psi(B_c(r)) \to B_c(r) \subset \mathbb{R}^2$ of ψ on $B_c(r)$ is a homeomorphism, and serves as a chart for M. Show that the transition function between $(\psi(B_c(r)), \phi_{c,r})$ and $(\psi(B_{c'}(r')), \phi_{c',r'})$ is smooth.

Problem 2.18 (Open submanifolds)**.** Prove that an open subset $U \subseteq M$ of a manifold M is a manifold of the same dimension.

Problem 2.19 (Products of manifolds are manifolds)**.** Prove that a product of two manifolds M and N of dimension m and n is a manifold of dimension $m+n$. (Hint: if (U_α, ϕ_α) and $(V_{\alpha'}, \psi_{\alpha'})$ are coordinate charts for M and N, try $\phi_\alpha \times \psi_{\alpha'} \colon U_\alpha \times V_{\alpha'} \to \phi_\alpha(U_\alpha) \times \psi_{\alpha'}(V_{\alpha'}) \subseteq \mathbb{R}^n \times \mathbb{R}^m$ as a chart for $M \times N$.)

2.4 Smooth maps

If M and N are smooth manifolds, we can define not only smooth maps from M to $\mathbb{R}$, but also from M to N.

2.4.1 Smooth maps between manifolds

Recall that a function $f\colon M \to \mathbb{R}$ on a manifold M is smooth at $p \in M$ if, for some chart (U_α, ϕ_α) around p, the coordinate representation $f_\alpha := f \circ \phi_\alpha^{-1}$ of f is smooth at the coordinate $\phi_\alpha(p)$ of p.

The definition for a map $F\colon M \to N$ from a manifold M to a manifold N to be smooth at $p \in M$ is similar, except that we now need *two* charts. One chart (U_α, ϕ_α) in M around $p \in M$, and one chart (V_β, ψ_β) in N around $F(p) \in N$.

If $F(U_\alpha) \subseteq V_\beta$, then the coordinate representation of F around p is

$$F_{\alpha\beta} := \psi_\beta \circ F \circ \phi_\alpha^{-1}\,, \tag{16}$$

considered as a map from $\phi_\alpha(U_\alpha) \subseteq \mathbb{R}^n$ to $\psi_\beta(V_\beta) \subseteq \mathbb{R}^m$. If $F(p) = q$, then $F_{\alpha\beta}$ maps the coordinate $\phi_\alpha(p) = (x^1, \ldots, x^n)$ of p to the coordinate $\psi_\beta(q) = (y^1, \ldots, y^m)$ of q.

Definition 2.20 (Smooth maps)**.** A continuous map $F\colon M \to N$ is differentiable (or smooth) at $p \in M$ if there exist coordinate charts (U_α, ϕ_α) in M and (V_β, ψ_β) in N such that $p \in U_\alpha$, $F(p) \in V_\beta$, and the coordinate representation $F_{\alpha\beta}$ is differentiable (or smooth) at the coordinate $\phi_\alpha(p) \in \mathbb{R}^n$.

This definition is independent of the choice of charts. If $(U_{\alpha'}, \phi_{\alpha'})$ and $(V_{\beta'}, \psi_{\beta'})$ are other charts around $p \in M$ and $F(p) \in N$, respectively, then $F_{\beta'\alpha'}$ can be expressed in terms of $F_{\beta\alpha}$ as

$$F_{\alpha'\beta'} := \psi_{\beta'} \circ F \circ \phi_{\alpha'}^{-1} = (\psi_{\beta'} \circ \psi_\beta^{-1}) \circ (\psi_\beta \circ F \circ \phi_\alpha^{-1}) \circ (\phi_\alpha \circ \phi_{\alpha'}^{-1}) = \kappa_{\beta\beta'} \circ F_{\alpha\beta} \circ \kappa_{\alpha'\alpha}.$$

The new coordinate representation is differentiable (or smooth) at p because both transition functions are so.

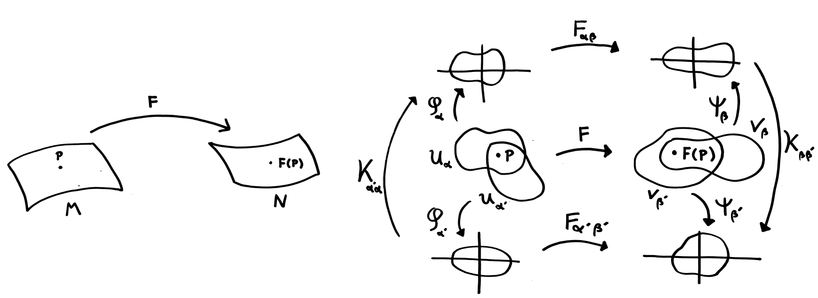


Figure 4: Smoothness of a function does not depend on the charts

If $(x^1, \ldots, x^n)$ are the coordinates of $p \in M$ and $(y^1, \ldots, y^m)$ are the coordinates of $F(p) \in N$ for the same point p, then y^ν (with $\nu = 1, \ldots, m$) is a smooth function of the x^μ (with $\mu = 1, \dots, n$). Indeed, $y^\nu(x^1, \dots, x^n)$ is simply the νth coordinate of $F_{\alpha\beta}(x^1, \dots, x^n)$.

Proposition 2.21. *If $F\colon M \to N$ is differentiable (or smooth) at $p \in M$, and $G\colon N \to L$ is differentiable (or smooth) at $F(p) \in N$, then $G \circ F\colon M \to L$ is differentiable (or smooth) at p.*

Proof. With respect to charts (U_α, ϕ_α) for M around p, (V_β, ψ_β) for N around $F(p)$ and (W_γ, χ_γ) for L around $G \circ F(p)$, we have $(G \circ F)_{\alpha\gamma} = G_{\beta\gamma} \circ F_{\alpha\beta}$. □

Now that we have a good notion of smooth maps between smooth manifolds, we can formulate what it means for two smooth manifolds to be 'similar'.

Definition 2.22. A *diffeomorphism* $\phi\colon M \to N$ is a smooth bijection for which the inverse $\phi^{-1}\colon N \to M$ is smooth as well.

We call M and N *diffeomorphic* if there exists a diffeomorphism between them. Diffeomorphic manifolds are 'the same' as far as their manifold properties are concerned – in much the same way that vector spaces are 'the same' if they are linearly isomorphic.

Problem 2.23. The curve $\gamma\colon \mathbb{R} \to \mathbb{S}^2$ with $\gamma(t) = (\cos(t), \sin(t), 0)$ is smooth. The same holds for the curve $\tilde{\gamma}(t) = (\cos(t), 0, \sin(t))$.

Problem 2.24 (Cartesian products of smooth maps are smooth)**.** If the maps $F\colon M_1 \to M_2$ and $G\colon N_1 \to N_2$ are smooth, then their cartesian product $(m, n) \mapsto (F(m), G(n))$ is a smooth map $M_1 \times N_1 \to M_2 \times N_2$. (See also Problem 2.19).

2.4.2 Index notation

It is instructive to look at equation (13) in a little more detail. If we write

$$\phi_\alpha(p) = (x^1, \dots, x^n) \quad \text{and} \quad \phi_\beta(p) = (x^{\overline{1}}, \dots, x^{\overline{n}}) \tag{17}$$

for the coordinates of $p \in M$ with respect to the charts (U_α, ϕ_α) and (U_β, ϕ_β), then the transition function $\kappa_{\alpha\beta} = \phi_\beta \circ \phi_\alpha^{-1}$ maps $(x^1, \dots, x^n)$ to $(x^{\overline{1}}, \dots, x^{\overline{n}})$. For every $\overline{\mu} = 1, \dots, n$, we can therefore consider

$$x^{\overline{\mu}}(x^1, \dots, x^n) \tag{18}$$

as a *smooth* function of $x^1, \dots, x^n$, where it is understood that x^μ and $x^{\overline{\mu}}$ designate the *same* point $p \in M$ with respect to different coordinate systems.

In the same vein, the coordinate representation $f \circ \phi_\alpha^{-1}$ of $f\colon M \to \mathbb{R}$ is often written as

$$f_\alpha(x^1, \dots, x^n) := f \circ \phi_\alpha^{-1}(x^1, \dots, x^n), \tag{19}$$

or even as $f(x^1, \dots, x^n)$ if the context is clear. Similarly, we write

$$f_\beta(x^{\overline{1}}, \dots, x^{\overline{n}}) := f \circ \phi_\beta^{-1}(x^{\overline{1}}, \dots, x^{\overline{n}}) \tag{20}$$

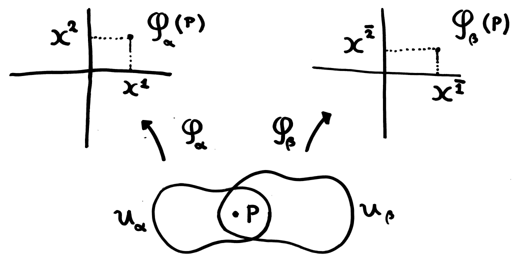


Figure 5: The points $p \in M$, $(x^1, x^2) \in \phi_\alpha(U_\alpha)$ and $(x^{\overline{1}}, x^{\overline{2}}) \in \phi_\beta(U_\beta)$

for the coordinate representation with respect to (U_β, ϕ_β). Combining (18), (19) and (20), we find

$$f_\alpha(x^1, \ldots, x^n) = f_\beta(x^{\overline{1}}(x^1, \ldots, x^n), \ldots, x^{\overline{n}}(x^1, \ldots, x^n)). \tag{21}$$

This is of course just a different guise of equation (13)! Indeed, $x^{\overline{\mu}}(x^1, \ldots, x^n)$ is the μth component of $\kappa_{\alpha\beta}(x^1, \ldots, x^n)$, so equation (21) is just $f_\alpha = f_\beta \circ \kappa_{\alpha\beta}$ evaluated at the coordinate $\phi_\alpha(p) = (x^1, \ldots, x^n)$.

Applying the chain rule, we find that for every $\mu = 1, \ldots, n$, we have

$$\frac{\partial}{\partial x^\mu} f_\alpha = \sum_{\overline{\mu}=1}^{n} \left(\frac{\partial x^{\overline{\mu}}}{\partial x^\mu} \right) \frac{\partial}{\partial x^{\overline{\mu}}} f_\beta \,. \tag{22}$$

Here, the $n \times n$-matrix J whose entry in row $\overline{\mu}$ and column μ is $J^{\overline{\mu}}_\mu = \frac{\partial x^{\overline{\mu}}}{\partial x^\mu}$ is the *Jacobian matrix* of the transition function $\kappa_{\alpha\beta}$ at the coordinates $(x^1, \ldots x^n)$. Writing (22) as

$$\frac{\partial}{\partial x^\mu} f_\alpha = \sum_{\overline{\mu}=1}^{n} J^{\overline{\mu}}_\mu \frac{\partial}{\partial x^{\overline{\mu}}} f_\beta \,, \tag{23}$$

we see that the partial derivatives $\partial_\mu f_\alpha := \frac{\partial}{\partial x^\mu} f_\alpha$ are a pointwise *linear* combinations of the partial derivatives $\partial_{\overline{\mu}} f_\beta$.

Considering different coordinate representations of the same function can even be useful for manifolds that can be covered with a single coordinate patch. For example, on $M = \mathbb{R}^2$, we can describe a function $f \colon M \to \mathbb{R}$ with respect to the cartesian coordinates or polar coordinates.

Cartesian coordinates are given on the coordinate neighbourhood $U_1 = \mathbb{R}^2$ by $\phi_1(p) = (x, y)$, where x and y are the projections of the point p on the x and y axis, respectively. Polar coordinates are given on $U_2 = \mathbb{R} \setminus \{(x, 0)\,;\, x \leq 0\}$ by $\phi_2(p) = (r, \theta)$, where r is the distance from p to the origin $\mathcal{O}$ and $\theta \in (-\pi, \pi)$ is the angle between the line $\overline{\mathcal{O}p}$ and the x-axis.

Note that the polar coordinates (r, θ) take the values $r > 0$ and $-\pi < \theta < \pi$ for $p \in U_2$, so $\phi_2(U_2) = (0, \infty) \times (-\pi, \pi)$. The transition function $\kappa_{21} \colon (0, \infty) \times (-\pi, \pi) \to \mathbb{R} \setminus \{(x, 0)\,;\, x \leq 0\}$ is of course the familiar formula for the coordinate transformation from polar to Cartesian coordinates:

$$(x, y) = (r\cos(\theta), r\sin(\theta))\,.$$

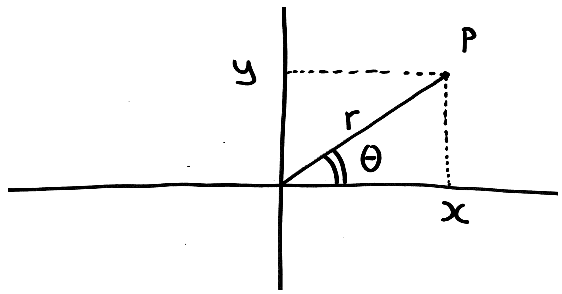


Figure 6: The manifold $M = \mathbb{R}^2$, with Cartesian and polar coordinates

Problem 2.25. Calculate the Jacobian matrix $J^{\overline{\mu}}_{\mu}$ for the transformation from polar coordinates $(x^1, x^2) = (r, \theta)$ to Cartesian coordinates $(x^{\overline{1}}, x^{\overline{2}}) = (x, y)$.

a) Why is $J^{\overline{\mu}}_{\mu}$ invertible?

b) Calculate $\frac{\partial}{\partial r} f(r\cos(\theta), r\sin(\theta))$ and $\frac{\partial}{\partial \theta} f(r\cos(\theta), r\sin(\theta))$, and compare with (23).

2.5 Complex manifolds

In the definition of a smooth manifold, we require that the transition functions are smooth. A *complex manifold* is defined in the same way as a smooth manifold, except that we now require the charts ϕ_α to be homeomorphisms from $U_\alpha \subseteq M$ to an open subset of $\mathbb{C}^n$, and we require the transition functions $\kappa_{\alpha\beta} = \phi_\beta \circ \phi_\alpha^{-1}$ between open subsets of $\mathbb{C}^n$ to be *holomorphic*.

Since every holomorphic map between open subsets of $\mathbb{C}^n$ can be considered as a smooth map between open subsets of $\mathbb{R}^{2n}$, every complex manifold of (complex) dimension n can be considered as a smooth manifold of (real) dimension $2n$. If M and N are complex manifolds, then a function $F \colon M \to N$ is called *holomorphic* if its coordinate representations $F_{\alpha\beta}$ are holomorphic. Since the transition functions are holomorphic, this is independent of the choice of charts.

2.5.1 Complex projective space

The *complex projective space* $\mathbb{CP}^n$ is the set of all *rays* in $\mathbb{C}^{n+1}$, that is, the set of all 1-dimensional complex linear subspaces of $\mathbb{C}^{n+1}$ with the origin deleted. If we denote the ray through a nonzero vector $v \in \mathbb{C}^{n+1}$ by

$$[v] := \{\lambda v\,;\, \lambda \in \mathbb{C}^\times\},$$

then we have $\mathbb{CP}^n := \{[v]\,;\, v \in \mathbb{C}^{n+1} \setminus \{0\}\}$.

Remark 2.26. In quantum mechanics, *pure states* are usually described by *vectors* in a Hilbert space. In a quantum system with $n+1$ degrees of freedom, the relevant Hilbert space is $\mathbb{C}^{n+1}$. Note that two nonzero vectors $v, v' \in \mathbb{C}^{n+1}$

yield the same physical state if $v' = \lambda v$ for some $\lambda \in \mathbb{C}^\times$. Indeed, for any observable $A = A^\dagger \in M_{n+1}(\mathbb{C})$, the expectation values $\frac{\langle v', Av' \rangle}{\langle v', v' \rangle}$ and $\frac{\langle v, Av \rangle}{\langle v, v \rangle}$ coincide, so there is no experiment that can tell the difference between v and v'. A *pure state* is therefore properly described by a ray $[v]$ through a nonzero vector v, and the pure state space is $\mathbb{C}\mathrm{P}^n$. In particular, the state space of a two-level system (qubit) is described by $\mathbb{C}\mathrm{P}^1$.

Complex projective space is a topological space, since it is the quotient of the topological space $\mathbb{C}^{n+1} \setminus \{0\}$ by the equivalence relation where $v \sim v'$ if $v' = \lambda v$ for some $\lambda \in \mathbb{C}^\times$. See example A.28 for more details. In example A.49, we show that $\mathbb{C}\mathrm{P}^n$ is Hausdorff.

Proposition 2.27. *The complex projective space $\mathbb{C}\mathrm{P}^n$ is a complex manifold of (complex) dimension n.*

Proof. A nonzero vector $v = (v_0, \ldots, v_n)$ in $\mathbb{C}^{n+1}$ has at least one nonzero component, say v_α. If we define for every $\alpha = 0, \ldots, n$ the coordinate neighbourhood

$$U_\alpha = \{[v] \in \mathbb{C}\mathrm{P}^n \,;\, v_\alpha \neq 0\},$$

then the U_α cover $\mathbb{C}\mathrm{P}^n$. We define the charts $\phi_\alpha \colon U_\alpha \to \mathbb{C}^n$ by

$$\phi_\alpha([(v_0, \ldots, v_n)]) = \left(\frac{v_0}{v_\alpha}, \ldots, \frac{v_{\alpha-1}}{v_\alpha}, \frac{v_{\alpha+1}}{v_\alpha}, \ldots, \frac{v_n}{v_\alpha} \right).$$

One can think of $\phi_\alpha([v])$ as the intersection of the ray $[v] \subseteq \mathbb{C}^{n+1}$ with the complex hyperplane $\{(z_0, \ldots, z_n) \,;\, z_\alpha = 1\} \simeq \mathbb{C}^n$. Suppose that $\alpha > \beta$. (The case $\alpha < \beta$ is similar.) Then the transition function $\kappa_{\alpha\beta}$ is given by

$$\kappa_{\alpha\beta}(z_1, \ldots, z_n) = \left(\frac{z_1}{z_\beta}, \ldots, \frac{z_{\alpha-1}}{z_\beta}, \frac{1}{z_\beta}, \frac{z_{\alpha+1}}{z_\beta}, \ldots, \frac{z_{\beta-1}}{z_\beta}, \frac{z_{\beta+1}}{z_\beta}, \ldots, \frac{z_n}{z_\beta} \right). \quad (24)$$

This transition function is holomorphic in the n complex variables $z_1, \ldots, z_n$.

Having shown that the transition functions $\kappa_{\alpha\beta}$ are holomorphic, it remains to show that the chart $\phi_\alpha \colon U_\alpha \to \mathbb{C}^n$ is a homeomorphism. Let $\pi \colon \mathbb{C}^{n+1} \setminus \{0\} \to \mathbb{C}\mathrm{P}^n$ be the quotient map $\pi(v) = [v]$. We need to show that $W \subseteq \mathbb{C}^n$ is open if and only if $\phi_\alpha^{-1}(W) \subseteq U_\alpha$ is open. By definition, the latter is the case if and only if

$$\pi^{-1}\phi_\alpha^{-1}(W) := \{(\lambda x_1, \ldots, \lambda x_{\alpha-1}, \lambda, \lambda x_\alpha, \ldots, \lambda x_n) \,;\, \lambda \in \mathbb{C}^\times, \vec{x} \in W\}$$

is open in $\mathbb{C}^{n+1} \setminus \{0\}$.

If $\pi^{-1}\phi_\alpha^{-1}(W)$ is open, then for every $v \in \pi^{-1}\phi_\alpha^{-1}(W)$ there is an open, $n+1$ dimensional ball $B_v(r)$ inside $\pi^{-1}\phi_\alpha^{-1}(W)$ that is centered at v. Then $B_{v/v_\alpha}(r/|v_\alpha|)$ is an open ball inside $\pi^{-1}\phi_\alpha^{-1}(W)$ as well. Since it is centered at a point with α coordinate 1, its intersection with the hyperplane $x_\alpha = 1$ in $\mathbb{R}^{n+1}$ is an open ball of dimension n that lies entirely within W.

Conversely, suppose that $B_u(r) \subseteq \mathbb{R}^n$ is an open ball around $u \in W$. Consider W as a subset of the hyperplane $z_\alpha = 1$ in $\mathbb{C}^{n+1}$. Then an open ball around $(u_1, \dots, u_{\alpha-1}, 1, u_\alpha, \dots u_n)$ with radius $r/\sqrt{1+(r+\|u\|)^2}$ is contained entirely within $\pi^{-1}\phi_\alpha^{-1}(W)$. (Convince yourself of this with a picture of the triangle with vertices 0, $(0, \dots, 1, \dots, 0)$, and the point in $\overline{B}_u(r)$ that is furthest from the origin.) $\square$

Problem 2.28. The above proof is somewhat lacking in detail. Check that the transition functions are indeed given by (24). What is their domain $\phi_\alpha(U_\alpha \cap U_\beta)$?

In particular, $\mathbb{C}\mathrm{P}^1$ is a complex manifold of dimension 1. We show that as a *real* manifold of dimension 2, it is diffeomorphic to the 2-sphere $\mathbb{S}^2$. With respect to the two charts (U_0, ϕ_0) and (U_1, ϕ_1), a complex line $[v_0, v_1] \in \mathbb{C}\mathrm{P}^1$ receives the coordinates $\phi_0([v_0, v_1]) = v_1/v_0 \in \mathbb{C}$ and $\phi_1([v_0, v_1]) = v_0/v_1 \in \mathbb{C}$. The transition function $\kappa_{01} \colon \mathbb{C} \setminus \{0\} \to \mathbb{C} \setminus \{0\}$ is therefore $z \mapsto 1/z$.

If we identify $z = x + iy \in \mathbb{C}$ with $(x, y) \in \mathbb{R}^2$, then the transition function $z \mapsto 1/z$ takes the form

$$(x, y) \mapsto \frac{1}{x^2 + y^2}(x, -y)\,. \tag{25}$$

This is reminiscent of the 2-dimensional sphere $\mathbb{S}^2$, which (recall Problem 2.11) is covered by two coordinate patches with transition function

$$(x, y) \mapsto \frac{1}{x^2 + y^2}(x, y)\,. \tag{26}$$

To make this correspondence precise, we introduce on $\mathbb{C}\mathrm{P}^1$ the anti-holomorphic coordinates $(U_\alpha, \phi_{\overline{\alpha}})$ with $\phi_{\overline{\alpha}}([v_0, v_1]) = \overline{\phi_\alpha([v_0, v_1])}$. Since $\mathbb{C}\mathrm{P}^1$ is covered by the coordinate neighbourhoods (U_0, ϕ_0) and $(U_1, \phi_{\overline{1}})$ with transition function $\kappa_{0,\overline{1}} z = 1/\overline{z}$ equal to (26), we infer that $\mathbb{C}\mathrm{P}^1$ is diffeomorphic to $\mathbb{S}^2$.

Remark 2.29 (Riemann sphere)**.** Using the chart ϕ_0, we can identify $z \in \mathbb{C}$ with the point $[1, z] \in \mathbb{C}\mathrm{P}^1$. This way, we are able to describe all points $[v_0, v_1] = [1, v_1/v_0] \in \mathbb{C}\mathrm{P}^1$ except the single point $[0, 1] \in \mathbb{C}\mathrm{P}^1$. It is reasonable to define $\infty := [0, 1]$, because $\lim_{z \to \infty} \phi_0^{-1}(z) = [0, 1]$ in $\mathbb{C}\mathrm{P}^1$. Since $\mathbb{C}\mathrm{P}^1$ is a complex manifold, we can now define what it means for a function $f \colon \mathbb{C} \cup \{\infty\} \to \mathbb{C}$ to be *holomorphic* at ∞. The picture of $\mathbb{C} \cup \{\infty\}$ as a complex manifold is often called the *Riemann sphere*.

On compact complex manifolds, holomorphic functions are surprisingly rare. To see this, we need some basic facts on holomorphic functions of several variables that are collected in Appendix C.

Theorem 2.30. *Let M be a compact, connected, complex manifold. Then every holomorphic function $f \colon M \to \mathbb{C}$ is constant.*

Proof. The absolute value $|f| \colon M \to \mathbb{R}$ is a continuous function on a compact topological space, so by Corollary A.58 there exists a point $p_0 \in M$ where $|f|$ achieves a maximal value $|f(p_0)|$. Since f is continuous, the set $S := \{p \in M\,;\, f(p) = f(p_0)\}$ is closed. We show that S is also open. Let $p \in S$, and let

$U_\alpha \subseteq M$ be a connected coordinate neighbourhood of p. Then the coordinate representation f_α of f has maximal modulus at the coordinates $\phi_\alpha(p)$ of p. By the maximum modulus principle (Corollary C.6), the holomorphic function f_α is constant on its domain $\phi_\alpha(U_\alpha) \subseteq \mathbb{C}^n$. It follows that f is constant on U_α, so that $U_\alpha \subseteq S$. In particular, every $p \in S$ has an open neighbourhood contained in S, so that $S \subseteq M$ is open as well as closed. Since M is connected (cf. Definition A.9) and S is nonempty, we conclude that $S = M$. □

Problem 2.31. The projection $\pi\colon \mathbb{C}^{n+1}\setminus\{0\} \to \mathbb{CP}^n$ with $\pi(v) = [v]$ is smooth.

Problem 2.32. Let $f\colon \mathbb{C} \setminus \{z_1, \ldots, z_r\} \to \mathbb{C}$ be a holomorphic function.

a) Then f extends to a continuous function

$$\hat{f}\colon \mathbb{CP}^1 \setminus \{\phi_0^{-1}(z_1), \ldots, \phi_0^{-1}(z_r)\} \to \mathbb{C}$$

if and only if the limit $f(\infty) := \lim_{z\to 0} f(1/z)$ exists.

b) The resulting function $\hat{f}\colon \mathbb{CP}^1\setminus\{\phi_0^{-1}(z_1), \ldots \phi_0^{-1}(z_r)\} \to \mathbb{C}$ is holomorphic if and only if the function $h\colon \mathbb{C} \setminus \{z_1, \ldots z_r\} \to \mathbb{C}$ with

$$h(z) := \begin{cases} f(1/z) & \text{for } |z| > 0 \\ f(\infty) & \text{for } z = 0 \end{cases}$$

is holomorphic at 0.

Problem 2.33. Let $p(z)$ and $q(z)$ be polynomials in $z \in \mathbb{C}$ which are not identically zero. Then by the fundamental theorem of algebra, $p(z)/q(z)$ is a *meromorphic* function on $\mathbb{C}$ with finitely many poles and finitely many zeros. Identify $\mathbb{CP}^1$ with the Riemann sphere $\mathbb{C} \cup \{\infty\}$ as in Remark 2.29, and show that $p(z)/q(z)$ extends to a *holomorphic* function $\mathbb{CP}^1 \to \mathbb{CP}^1$.

Problem 2.34. Let $\mathrm{SL}(2,\mathbb{C})$ be the group of complex 2×2 matrices of determinant 1, and let $A \in \mathrm{SL}(2,\mathbb{C})$.

a) The expression $[v] \mapsto [Av]$ yields a well-defined map $F_A\colon \mathbb{CP}^1 \to \mathbb{CP}^1$, which is trivial if and only if $A = \pm\mathbf{1}$.

b) The map $F_A\colon \mathbb{CP}^1 \to \mathbb{CP}^1$ is holomorphic.

c) For $A, B \in \mathrm{SL}(2,\mathbb{C})$, we have $F_{AB} = F_A \circ F_B$. In particular, every map $F_A\colon \mathbb{CP}^1 \to \mathbb{CP}^1$ is invertible.

d) So $F_A\colon \mathbb{CP}^1 \to \mathbb{CP}^1$ is a holomorphic diffeomorphism.

Remark 2.35. In fact, one can show that every holomorphic diffeomorphism of $\mathbb{CP}^1$ is of this form. The group $\mathrm{PSL}(2,\mathbb{C}) := \mathrm{SL}(2,\mathbb{C})/\{\pm\mathbf{1}\}$ is the *automorphism group* of the complex manifold $\mathbb{CP}^1$.

2.6 Embedded submanifolds in $\mathbb{R}^n$

In linear algebra, a system of $n-k$ independent *linear* equations

$$
\begin{array}{ccccccc}
a_1^1 x^1 & + & \dots & + & a_n^1 x^n & = & c^1 \\
\vdots & + & \dots & + & \vdots & = & \vdots \\
a_1^{n-k} x^1 & + & \dots & + & a_n^{n-k} x^n & = & c^{n-k}
\end{array}
$$

defines an affine subspace of dimension k. This space of solutions can equivalently be described as the preimage $A^{-1}(\{c\})$ of $c \in \mathbb{R}^{n-k}$ under the linear map $A\colon \mathbb{R}^n \to \mathbb{R}^{n-k}$ obtained from the coefficients. The independence of the equations translates to the $(n-k) \times n$ matrix A being *surjective*.

Proposition 2.36. *For a $(n-k) \times n$ matrix A, the following are equivalent.*

a) All $(n-k)$ rows $r^\mu := (a_1^\mu, \dots a_n^\mu)$ are linearly independent.

b) There are $(n-k)$ independent columns $k_\nu := (a_\nu^1, \dots a_\nu^{n-k})^T$.

c) The linear map $A\colon \mathbb{R}^n \to \mathbb{R}^{n-k}$ is surjective.

Proof. All three properties are conserved under pivot operations, because every pivot operation on A can be expressed by left multiplication with an invertible $(n-k) \times (n-k)$ matrix. Thus, we may as well assume that our matrix is in reduced row echelon form, say

$$
A = \begin{pmatrix}
1 & a_2^1 & 0 & a_4^1 & a_5^1 & \cdots \\
0 & 0 & 1 & a_4^2 & a_5^2 & \cdots \\
\vdots & & & \ddots & & \\
0 & 0 & \dots & 1 & a_{n-1}^{n-k} & a_n^{n-k}
\end{pmatrix}.
$$

For a matrix in row echelon form, all three properties are equivalent to the appearance of at least one nonzero entry on the bottom row. □

In applications, one is often interested in the space $\Sigma \subseteq \mathbb{R}^n$ of solutions for a system of *nonlinear* equations. This will in general no longer be a vector space, but we will show that under suitable nondegeneracy conditions, it will still be a *manifold* of dimension k.

More precisely, let $F^1, \dots, F^{n-k}$ be $n-k$ smooth $\mathbb{R}$-valued functions on $\mathbb{R}^n$, and let Σ be the set of solutions of

$$
\begin{array}{rcl}
F^1(x^1, \dots x^n) & = & c^1 \\
\vdots & = & \vdots \\
F^{n-k}(x^1, \dots, x^n) & = & c^{n-k}.
\end{array}
$$

Using the inverse function theorem, we will show that if the derivative $D_pF\colon \mathbb{R}^n \to \mathbb{R}^{n-k}$ is *surjective* for every solution $p = (x^1, \dots, x^n)$ of the above equations,

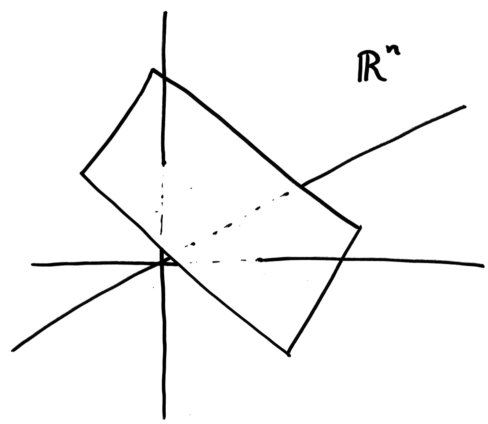


Figure 7: Affine space of solutions for $(n-k)$ linear equations

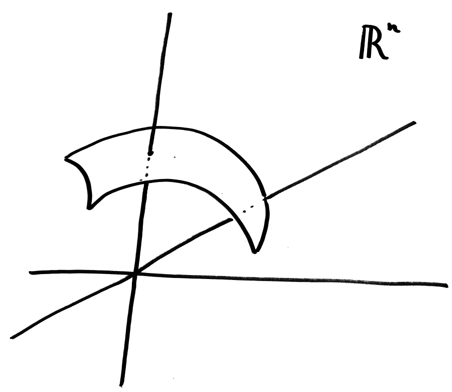


Figure 8: Manifold of solutions for $(n-k)$ nonlinear equations

then the solution space $\Sigma = F^{-1}(\{c\})$ is a k-dimensional *embedded submanifold* of $\mathbb{R}^n$. A value $c \in \mathbb{R}^{n-k}$ with this property is called a *regular value.*

This yields a way of showing that a subset $\Sigma \subseteq M$ is a manifold without explicitly constructing charts. For instance, $\mathbb{S}^n \subseteq \mathbb{R}^{n+1}$ is a manifold because it is the solution to the quadratic equation

$$(x^0)^2 + \ldots + (x^n)^2 = 1\,,$$

where $c = 1$ is a regular value because $D_pF = (2x^0, \ldots, 2x^n)$ is surjective for all points $p = (x^0, \ldots, x^n)$ satisfying $(x^0)^2 + \ldots + (x^n)^2 = 1$.

2.6.1 Embedded submanifolds

If Σ is a subset of $\mathbb{R}^n$, then we can use the fact that $\mathbb{R}^n$ is a manifold to help us prove that Σ is a manifold as well. More generally, consider a subset Σ of an

n-dimensional manifold M.

Definition 2.37 (Embedded submanifolds)**.** The subset $\Sigma \subseteq M$ is called a k-dimensional *embedded submanifold* of M if for every point $p \in \Sigma$, there exists a coordinate chart (U_α, ϕ_α) around p in M such that

$$\phi_\alpha(U_\alpha \cap \Sigma) = \phi_\alpha(U_\alpha) \cap (\mathbb{R}^k \oplus \{0\}),$$

where $\mathbb{R}^k \oplus \{0\} \subseteq \mathbb{R}^k \oplus \mathbb{R}^{n-k}$ is a k-dimensional hyperplane in $\mathbb{R}^n$. A chart with this property is called a *slice chart.*

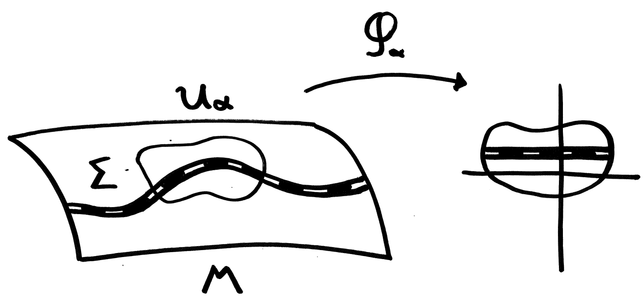


Figure 9: Slice chart

For example, an embedded 1-dimensional submanifold $\ell \subseteq \mathbb{R}^3$ is locally diffeomorphic to the straight line $\{(x, 0, 0)\,;\, x \in \mathbb{R}\}$. Around any point $p \in \mathbb{R}^3$, the part of the line ℓ that lies inside the coordinate neighbourhood $U_\alpha \subseteq \mathbb{R}^3$ around p is parameterised by $x \mapsto \phi_\alpha^{-1}(x, 0, 0)$.

Problem 2.38. The meridian $\{(\xi, \eta, \zeta) \in \mathbb{S}^2\,;\, \eta = 0\}$ is a 1-dimensional embedded submanifold of the sphere $\mathbb{S}^2$.

Problem 2.39. For any smooth function $f \colon \mathbb{R}^2 \to \mathbb{R}$, the graph

$$\Gamma = \{(x, y, z) \in \mathbb{R}^3\,;\, z = f(x, y)\}$$

is a 2-dimensional, embedded submanifold of $\mathbb{R}^3$. (Hint: show that $\phi_\alpha \colon \mathbb{R}^3 \to \mathbb{R}^3$ with $\phi_\alpha(x, y, z) = (x, y, z - f(x, y))$ is a chart of $\mathbb{R}^3$.)

Problem 2.40. Show that $\mathbb{S}^2$ is an embedded submanifold of $\mathbb{R}^3$. (Hint: show that (U_{001}, ϕ_{001}) with $U_{001} = \{(x, y, z \in \mathbb{R}^3\,;\, x^2 + y^2 < 1, z > 0)\}$ and $\phi_{001}(x, y, z) = (x, y, z - \sqrt{1 - x^2 - y^2})$ is a slice chart around $(0, 0, 1)$.)

Since the definition of an embedded submanifold says nothing about smoothness of transition functions, it is not *a priori* clear that an embedded submanifold is a manifold. The following proposition shows that the name is nonetheless justified.

Proposition 2.41. *Every embedded submanifold $\Sigma \subseteq M$ is a smooth manifold.*

Proof. The idea is to define the charts (ψ_α, V_α) for Σ as restrictions of the charts (ϕ_α, U_α) for M, so that smoothness of the transition functions on Σ follows from smoothness of the transition functions for M.

More precisely, let k and n be the dimensions of Σ and M, respectively. With $\mathbb{R}^n = \mathbb{R}^k \oplus \mathbb{R}^{n-k}$, let $\pi\colon \mathbb{R}^n \to \mathbb{R}^k$ be the linear projection $\pi(x \oplus y) = x$, and let $j\colon \mathbb{R}^k \hookrightarrow \mathbb{R}^n$ be the inclusion of $\mathbb{R}^k$ into $\mathbb{R}^n$.

Since Σ is an embedded submanifold, there exists a chart $U_\alpha \subseteq M$ around $p \in \Sigma$ such that $\phi_\alpha(U_\alpha \cap \Sigma) = \phi_\alpha(U_\alpha) \cap (\mathbb{R}^k \oplus \{0\})$. If we set $V_\alpha := U_\alpha \cap \Sigma$, then the restriction $\psi_\alpha := \pi \circ \phi_\alpha|_{V_\alpha} \colon V_\alpha \to \mathbb{R}^k$ is therefore injective. It is continuous because ϕ_α and π are continuous, and it is a homeomorphism onto its image because $\phi_\alpha^{-1} \circ j\colon \psi_\alpha(V_\alpha) \to V_\alpha$ is a continuous inverse.

The transition functions $\kappa^{\Sigma}_{\alpha\beta}$ for Σ are given in terms of the transition functions $\kappa^{M}_{\alpha\beta}$ for M by $\kappa^{\Sigma}_{\alpha\beta} = \pi \circ \kappa^{M}_{\alpha\beta} \circ j$. Since the three maps on the right hand side are smooth, $\kappa^{\Sigma}_{\alpha\beta}$ is smooth as well, and we conclude that Σ is a manifold. □

Embedded submanifolds behave well with respect to smooth maps. In particular, the canonical inclusion $\Sigma \hookrightarrow M$ is always smooth.

Proposition 2.42. *Let M and N be smooth manifolds, let $\Sigma \subseteq M$ be an embedded submanifold.*

a) If $F\colon M \to N$ is smooth, then so is the restriction $F|_\Sigma \colon \Sigma \to N$.

b) A map $F\colon N \to \Sigma$ is smooth if and only if it is smooth as a map into M.

Proof. Let Σ, M and N be of dimension k, n and l, respectively. Choose slice coordinates $x^1, \dots, x^k; y^1, \dots, y^{n-k}$ for $U_\alpha \subseteq M$, so $\Sigma \cap U_\alpha$ corresponds to the locus $y^\mu = 0$. Let $z^1, \dots, z^l$ be coordinates on $V_\beta \subseteq N$.

For (a), note that if $F\colon M \to N$ admits the coordinate representation $F_{\alpha\beta}(x^1, \dots, x^k; y^1, \dots, y^{n-k})$, then $F|_\Sigma \colon \Sigma \to N$ has coordinate representation

$$(x^1, \dots, x^k) \mapsto (F|_\Sigma)_{\alpha\beta}(x^1, \dots, x^k; 0, \dots, 0).$$

If $F_{\alpha\beta}$ is smooth, then so is $(F|_\Sigma)_{\alpha\beta}$.

For (b), note that a smooth function $F\colon N \to M$ with values in Σ takes the local form $F_{\beta\alpha}(z^1, \dots, z^l) = (F^1_{\alpha\beta}(z), \dots, F^k_{\alpha\beta}(z); 0, \dots, 0)$, the smoothness of which depends only on the first k entries. □

Problem 2.43. Revisit Problem 2.12 using Proposition 2.42 and Problem 2.40.

2.6.2 Regular values and embedded submanifolds of $\mathbb{R}^n$

In order to show that a subset $\Sigma \subseteq \mathbb{R}^n$ is a smooth manifold, it therefore suffices to show that it is an embedded submanifold of $\mathbb{R}^n$.

Let $U \subseteq \mathbb{R}^n$, and let $F\colon U \to \mathbb{R}^{(n-k)}$ be a smooth map. A point $c \in \mathbb{R}^{n-k}$ is called a *regular value* if for all $p \in F^{-1}(\{c\})$, the differential $D_pF\colon \mathbb{R}^n \to \mathbb{R}^{n-k}$ is *surjective.*

Theorem 2.44 (Regular Level Sets)**.** *If c is a regular value of $F\colon U \to \mathbb{R}^{n-k}$, then $\Sigma := F^{-1}(\{c\})$ is a closed, embedded submanifold of $\mathbb{R}^n$ of dimension k.*

This result is based on the Inverse Function Theorem for smooth functions $\phi\colon U \to \mathbb{R}^n$ from an open subset $U \subseteq \mathbb{R}^n$ to $\mathbb{R}^n$.

Theorem 2.45 (Inverse Function Theorem)**.** *If $D_p\phi$ is invertible at $p \in U$, then there exist open neighbourhoods $U_0 \subseteq \mathbb{R}^n$ of p and $V_0 \subseteq \mathbb{R}^n$ of $\phi(p)$ such that $\phi|_{U_0}\colon U_0 \to V_0$ is a diffeomorphism.*

In other words, if $D_p\phi$ is invertible, then (U_0, ϕ) is a coordinate chart for $\mathbb{R}^n$ around p. We prove the Inverse Function Theorem in Appendix B. Using the inverse function theorem, we prove Theorem 2.44.

Proof of Theorem 2.44. Since F is continuous, the preimage $F^{-1}(\{c\})$ of the closed set $\{c\}$ is closed. To show that it is an embedded subanifold, we use the fact that $D_pF\colon \mathbb{R}^n \to \mathbb{R}^{n-k}$ is surjective at $p \in \Sigma$ to construct a slice chart $\phi\colon U_0 \to \phi(U_0) \subseteq \mathbb{R}^n$ for $\mathbb{R}^n$ around the point $p \in U_0 \subseteq U$.

By changing F to $\tilde{F}(x) := F(x) - c$, we may as well assume that $c = 0$. To see that ϕ is a slice chart, it then suffices to show that

$$F \circ \phi^{-1}(x^1, \ldots, x^k; y^1, \ldots, y^{n-k}) = (y^1, \ldots, y^{n-k}). \tag{27}$$

Indeed, if the chart ϕ satisfies (27), then the preimage of $c = 0$ under the coordinate representation $F \circ \phi^{-1}$ is the hyperplane $\mathbb{R}^k \oplus 0 \subseteq \mathbb{R}^n$.

It is always possible to choose coordinates

$$x^1, \ldots, x^k\,;\, y^1, \ldots, y^{n-k}$$

on $\mathbb{R}^n$ in which the $(n-k) \times n$ matrix D_pF takes the form

$$D_pF = \left(\frac{\partial F^{\overline{\mu}}}{\partial x^\nu}, \frac{\partial F^{\overline{\mu}}}{\partial y^\mu}\right), \tag{28}$$

where $(\frac{\partial F^{\overline{\mu}}}{\partial x^\nu})_{\overline{\mu}\nu}$ is an $(n-k) \times k$ matrix and $(\frac{\partial F^{\overline{\mu}}}{\partial y^\mu})_{\overline{\mu}\mu}$ is an $(n-k) \times (n-k)$ matrix that is *invertible*. Indeed, since $D_pF\colon \mathbb{R}^n \to \mathbb{R}^{(n-k)}$ is surjective, there are always $n-k$ independent columns in (28), which we can put on the right hand side by reshuffling the coordinates if necessary.

We split $\mathbb{R}^n = \mathbb{R}^k \oplus \mathbb{R}^{n-k}$ using the above coordinates, and write $x \oplus y$ instead of $(x^1, \ldots, x^k\,;\, y^1, \ldots, y^{n-k})$. Define the smooth map $\phi\colon \mathbb{R}^n \to \mathbb{R}^n$ by

$$\phi(x, y) := x \oplus F(x, y).$$

Since

$$D_p\phi = \begin{pmatrix} \mathbf{1}_k & 0 \\ \frac{\partial F^{\overline{\mu}}}{\partial x^\nu} & \frac{\partial F^{\overline{\mu}}}{\partial y^\mu} \end{pmatrix}$$

has invertible right lower block, $D_p\phi$ is invertible. Indeed, since $\det(\frac{\partial F^{\overline{\mu}}}{\partial y^\mu}) \neq 0$, we have

$$\det(D_p\phi) = \det(\mathbf{1}_k) \det\left(\frac{\partial F^{\overline{\mu}}}{\partial y^\mu}\right) \neq 0$$

as well. Applying the Inverse Function Theorem to ϕ, we conclude that there exists an open neighbourhood U_0 of p such that $\phi\colon U_0 \to \phi(U_0) \subseteq \mathbb{R}^n$ is a diffeomorphism onto its image. In other words, (U_0, ϕ_0) is a chart for $\mathbb{R}^n$ around p. Since $\phi \circ \phi^{-1}(x \oplus y) = x \oplus F(\phi^{-1}(x \oplus y)) = x \oplus y$, we have $F \circ \phi^{-1}(x \oplus y) = y$ as required. $\square$

In the course of the proof, we found the following normal form for F, which we record for future use.

Corollary 2.46. *Let $p \in \mathbb{R}^n$ be a regular point for $F\colon \mathbb{R}^n \subseteq U \to \mathbb{R}^{n-k}$. Then there exist coordinates $(x^1, \ldots, x^k; y^1, \ldots, y^{n-k})$ in a neighbourhood of p in $\mathbb{R}^n$ for which F has the coordinate representation*

$$(x^1, \ldots, x^k; y^1, \ldots, y^{n-k}) \mapsto (y^1, \ldots, y^{n-k})\,.$$

Moreover, we obtain the following characterization of embedded submanifolds $\Sigma \subseteq \mathbb{R}^n$ as subsets which are locally given by solutions to $(n-k)$ smooth equations $F^\nu(x^1, \ldots, x^n) = c^\nu$.

Corollary 2.47. *A subset $\Sigma \subseteq \mathbb{R}^n$ is an embedded submanifold of dimension k if and only if for every $p \in \Sigma$, there exists an open neighbourhood $U \subseteq \mathbb{R}^n$ of p and a smooth function $F\colon \mathbb{R}^n \supseteq U \to \mathbb{R}^{n-k}$ such that $\Sigma \cap U = F^{-1}(\{c\})$ and c is a regular value for F.*

Proof. If a function F has these properties, then $\Sigma \cap U$ is an embedded submanifold by by Theorem 2.44. It follows that Σ is an embedded submanifold (why?). Conversely, if Σ is an embedded submanifold, then every point $p \in \Sigma$ has a slice chart $\phi_\alpha\colon \mathbb{R}^n \supseteq U_\alpha \to \mathbb{R}^k \oplus \mathbb{R}^{n-k}$. Since $\phi_\alpha(U_\alpha \cap \Sigma) = \phi_\alpha(U_\alpha) \cap (\mathbb{R}^k \oplus \{0\})$, we can take $F\colon U_\alpha \to \mathbb{R}^{n-k}$ to be the projection of ϕ_α on $\mathbb{R}^{n-k}$. The differential D_pF is surjective because $D_p\phi_\alpha$ is. $\square$

Problem 2.48. Show that the helix

$$x = \cos(z), \quad y = \sin(z)$$

is a 1-dimensional, embedded submanifold of $\mathbb{R}^3$. Hint: show that $(0,0)$ is a regular value for the smooth function $F\colon \mathbb{R}^3 \to \mathbb{R}^2$ defined by $F(x,y,z) = (x - \cos(z), y - \sin(z))$.

Problem 2.49. A torus in $\mathbb{R}^3$ with 'large radius' 5 and 'small radius' 2 is described by the equation

$$(r-5)^2 + z^2 = 4,$$

where $r(x,y) = \sqrt{x^2 + y^2}$ is smooth on $\mathbb{R}^3 \setminus \{(0,0,z)\,;\, z \in \mathbb{R}\}$.

a) Show that 4 is a regular value of $F(x,y,z) = (r-5)^2 + z^2$.

b) Conclude that the torus is an embedded submanifold of $\mathbb{R}^3$.

Problem 2.50. The *light cone* in $\mathbb{R}^4$ is given by

$$\mathbf{L} = \{(t, x, y, z) \in \mathbb{R}^4 \,;\, x^2 + y^2 + z^2 = c^2t^2\},$$

where c is the speed of light.

a) Prove that $\mathbf{L}\backslash\{(0, 0, 0, 0)\}$ is an embedded submanifold of $\mathbb{R}^4$. (For example by using the regular value theorem.)

b) Do you think $\mathbf{L}$ is an embedded submanifold? (No need for a proof, enough to draw a picture.)

Problem 2.51. Let $F\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by $F(x, y) = x^3 + xy + y^3$. What are the regular values of F?

Problem 2.52. A *conic section* $\Sigma_{a,b} := C \cap P_{a,b}$ is the intersection of the cone

$$C := \{(x, y, z) \in \mathbb{R}^3 \,;\, x^2 + y^2 = z^2\}$$

with the plane

$$P_{a,b} = \{(x, y, z) \in \mathbb{R}^3 \,;\, z = ax + b\}.$$

Show that every conic section with $b \neq 0$ is an embedded submanifold of $\mathbb{R}^3$. What is its dimension?

Problem 2.53 (Surface of revolution)**.** Let $H\colon \mathbb{R} \to \mathbb{R}$ be a smooth function, and let

$$\Sigma := \{(x, y, z) \in \mathbb{R}^3 \,;\, x^2 + y^2 = H(z)\}$$

be the surface of revolution for the curve $H(z) = x^2$.

a) Sketch the set Σ for $H = z$ and $H = z^2$. Which of these two do you think is a submanifold of $\mathbb{R}^3$?

b) Calculate D_pF for the function $F\colon \mathbb{R}^3 \to \mathbb{R}$ defined by

$$F(x, y, z) := x^2 + y^2 - H(z)\,.$$

c) Show that Σ is a 2-dimensional, embedded submanifold of $\mathbb{R}^3$ if $\frac{d}{dz}H(z) \neq 0$ for all $z \in \mathbb{R}$ with $H(z) = 0$. Return to (a) and prove your guess.

Problem 2.54. If we consider the sphere $\mathbb{S}^2 \subseteq \mathbb{R}^3$ as the level set $F^{-1}(\{1\})$ of the smooth function $F(x, y, z) = x^2 + y^2 + z^2$, then the proof of the Regular Level Set Theorem yields slice charts around every point $p \in \mathbb{S}^2$.

a) Revisit the proof of Theorem 2.44, and determine these slice charts for $p = (1, 0, 0)$, $p = (0, 1, 0)$ and $p = (0, 0, 1)$.

b) Calculate the inverse of these charts.

c) The slice charts from a) are charts for $\mathbb{R}^3$. What are their restrictions to charts for $\mathbb{S}^2$?

Problem 2.55 (Closed embedded submanifolds)**.** Let $\Sigma \subseteq \mathbb{R}^2$ be the subset given by $\Sigma = \{(t, \sin 1/t)\,;\, t > 0\}$.

a) Show that Σ is an embedded submanifold of $\mathbb{R}^2$.

b) Show that Σ is not closed.

c) Give a smooth function $F\colon \mathbb{R}^2 \supset \{(x, y) \in \mathbb{R}^2\,;\, x > 0\} \to \mathbb{R}$ that has Σ as a regular level set.

d) There do not exist smooth functions $F\colon \mathbb{R}^2 \to \mathbb{R}$ that have Σ as a regular level set.

e) Show that $\Sigma \subseteq M$ is is a *closed* embedded submanifold if and only if for every point $p \in M$, there exists a coordinate chart (U_α, ϕ_α) around p in M such that
$$\phi_\alpha(U_\alpha \cap \Sigma) = \phi_\alpha(U_\alpha) \cap (\mathbb{R}^k \oplus \{0\}).$$
(Note that for 'ordinary' embedded submanifolds, one only requires this for $p \in \Sigma$.)

Problem 2.56. Let $p\colon \mathbb{C}^{n+1} \to \mathbb{C}$ be the polynomial
$$p(v_1, \ldots, v_{n+1}) = (v_1)^3 + \ldots + (v_{n+1})^3.$$

a) Show that if $p(v) = 0$, then $p(\lambda v) = 0$ for all $\lambda \in \mathbb{C}$. It follows that $\Sigma := \{[v] \in \mathbb{C}\mathrm{P}^n\,;\, p(v) = 0\}$ is well defined.

b) Show that Σ is an embedded submanifold of $\mathbb{C}\mathrm{P}^n$.

(Hint: It suffices to show that $U_\alpha \cap \Sigma$ is an embedded submanifold of U_α for a collection of coordinate neighbourhoods that cover $\mathbb{C}\mathrm{P}^n$.)

2.7 Lie groups

In both geometry and physics, the groups of symmetries that one encounters are often *Lie groups*. A Lie group is a group G which is at the same time a smooth manifold, and for which the multiplication $(g, h) \mapsto gh$ is a smooth map $\mu\colon G \times G \to G$. The inverse $\iota\colon G \to G$ with $\iota(g) = g^{-1}$ is then automatically smooth as well (cf. Problem 5.10).

An important example of a Lie group is the group $G = \mathrm{GL}(n, \mathbb{R})$ of invertible $n \times n$ matrices. Since the determinant is a continuous function,
$$\mathrm{GL}(n, \mathbb{R}) := \{g \in M(n, \mathbb{R})\,;\, \det(g) \neq 0\}$$
is an open subset of $M_n(\mathbb{R})$, and hence a smooth manifold by Problem 2.18. The matrix multiplication $\mu\colon M(n, \mathbb{R}) \times M(n, \mathbb{R}) \to M(n, \mathbb{R})$ is smooth because it is quadratic in each entry, so its restriction to the open submanifold $\mathrm{Gl}(n, \mathbb{R}) \times \mathrm{Gl}(n, \mathbb{R})$ is smooth as well.

2.7.1 The Euclidean motion group

Both the orthogonal group $O(n)$ and the Euclidean motion group $E(n)$ are Lie groups. We show that $\mathrm{O}(n)$ of is a Lie group of dimension dimension $\frac{1}{2}n(n-1)$, and then use this to see that $E(n)$ is a Lie group of dimension $\frac{1}{2}n(n+1)$.

To show that it is a Lie group, it suffices to show that $\mathrm{O}(n)$ is a smooth, embedded submanifold of $M(n,\mathbb{R})$, as the multiplication $\mu\colon \mathrm{O}(n)\times\mathrm{O}(n)\to\mathrm{O}(n)$ will then be automatically smooth by Proposition 2.42. (It is the restriction of the smooth multiplication $\mu\colon M_n(\mathbb{R})\times M_n(\mathbb{R})\to M_n(\mathbb{R})$ to an embedded submanifold.)

To apply Theorem 2.6.2, identify the vector space $M(n,\mathbb{R})$ of real $n\times n$ matrices with $\mathbb{R}^{n^2}$, and identify the vector space $\mathrm{Sym}(n,\mathbb{R})$ of symmetric $n\times n$ matrices with $\mathbb{R}^{\frac{1}{2}n(n+1)}$. Then $\mathrm{O}(n)$ is the preimage of the identity $\mathbf{1}_n$ under the map $F\colon\mathbb{R}^{n^2}\to\mathbb{R}^{\frac{1}{2}n(n+1)}$ defined by $F(R)=R^TR$. Note that the derivative at R is given by

$$D_RF(X) := \frac{d}{d\varepsilon}\Big|_{\varepsilon=0}F(R+\varepsilon X) = X^TR + R^TX.$$

To show that $\mathbf{1}_n$ is a regular value of F, we need to show that D_RF is surjective for all $R\in\mathrm{O}(n)$. Indeed, since $R^T=R^{-1}$ for an orthogonal matrix R, we have $D_R(RX)=X^T+X$, so the image of D_R is the entire space $\mathrm{Sym}(n,\mathbb{R})$ of symmetric $n\times n$ matrices. By Theorem 2.6.2, $\mathrm{O}(n)$ is a closed, embedded submanifold of $M(n,\mathbb{R})$ of dimension $k=n^2-\frac{1}{2}n(n+1)=\frac{1}{2}n(n-1)$.

Problem 2.57. Show that

$$\mathrm{Sl}(2,\mathbb{R}) := \{g\in M(2,\mathbb{R})\,;\,\det(g)=1\}$$

is a smooth manifold of dimension 3. Hint: consider the determinant as a smooth function from $\mathbb{R}^4$ to $\mathbb{R}$, and show that 1 is a regular value of

$$\det\begin{pmatrix} x & y\\ z & w\end{pmatrix} = xw-yz.$$

Recall from Theorem 1.3 that every isometry can be decomposed in an orthogonal transformation $R\in\mathrm{O}(n)$ and a translation T_v over a vector $v\in\mathbb{R}^n$, yielding a bijection $\mathrm{O}(n)\times\mathbb{R}^n\simeq E(n)$. Since $\mathrm{O}(n)$ is a manifold of dimension $\frac{1}{2}n(n-1)$ and $\mathbb{R}^n$ is a manifold of dimension n, their product is a manifold of dimension $\frac{1}{2}n(n-1)+n=\frac{1}{2}n(n+1)$, cf. Problem 2.19.

Note that concatenation of the isometries $x\mapsto Rx+v$ and $x\mapsto R'x+v'$ is the isometry $x\mapsto R''x+v''$ with $R''=R'R$ and $v''=R'v+v'$. If we identify $E(n)$ with $\mathrm{O}(n)\times\mathbb{R}^n$, the multiplication is therefore given by

$$(R',v')\cdot(R,v) = (R'R, R'v+v').$$

Since $\mathrm{O}(n)\times\mathbb{R}^n$ is an embedded submanifold of $M_n(\mathbb{R})\times\mathbb{R}^n$ (why?), and since the multiplication is the restriction of the map $(M_n(\mathbb{R})\times\mathbb{R}^n)\times(M_n(\mathbb{R})\times\mathbb{R}^n)\to(M_n(\mathbb{R})\times\mathbb{R}^n)$ which is at most quadratic in every entry, the multiplication is smooth and $E(n)$ is a Lie group.

Problem 2.58. Show that the group of affine, area-preserving transformations of $\mathbb{R}^2$ is a Lie group of dimension 5.

2.7.2 The Poincaré group

To prove that the Poincaré group P is a Lie group of dimension 10, it suffices to prove that $\mathrm{O}(3,1)$ is a Lie group of dimension 6. Indeed, by Theorem 1.10 we have a bijection $P \simeq \mathrm{O}(3,1) \times \mathbb{R}^4$, so we can consider P as a manifold of dimension $6 + 4 = 10$. The multiplication on this product is smooth by an argument similar to the one in §(2.7.1).

The proof that $\mathrm{O}(3,1)$ is a Lie group is analogous to the proof for $\mathrm{O}(n)$. We realize $\mathrm{O}(3,1)$ as the level set $F^{-1}(\{H\})$ for the smooth function

$$F\colon M(4,\mathbb{R}) \to \mathrm{Sym}(4,\mathbb{R}), \quad F(\Lambda) = \Lambda^T H \Lambda,$$

with $M(4,\mathbb{R}) \simeq \mathbb{R}^{16}$ and $\mathrm{Sym}(4,\mathbb{R}) \simeq \mathbb{R}^{10}$. To see that H is a regular value, we need to show that for all Λ with $\Lambda^T H \Lambda = H$, the Jacobian

$$D_\Lambda F(X) = \frac{d}{dt}\Big|_{t=0} F(\Lambda + tX) = X^T H \Lambda + \Lambda^T H X$$

is surjective as a map from $M(4,\mathbb{R})$ to the symmetric matrices $\mathrm{Sym}(4,\mathbb{R})$. One readily checks that for any $Y^T = Y \in \mathrm{Sym}(4,\mathbb{R})$, the element $X = \frac{1}{2}\Lambda H Y$ does the trick, $D_\Lambda F(X) = Y$. By the Regular Level Set Theorem 2.44, we conclude that $\mathrm{O}(3,1)$ is an embedded submanifold of $M(4,\mathbb{R})$ of dimension $16 - 10 = 6$. It is a Lie group because the multiplication is the restriction of a smooth map $\mu\colon M(4,\mathbb{R}) \times M(4,\mathbb{R}) \to M(4,\mathbb{R})$.

Problem 2.59. A function $f\colon \mathbb{R}^N \to \mathbb{R}$ is *homogeneous of degree d* if

$$f(\lambda v) = \lambda^d f(v)$$

for all $v \in \mathbb{R}^N$ and for all $\lambda > 0$.

a) Show that if f is homogeneous of degree $d > 0$, then every nonzero $c \in \mathbb{R}$ is a regular value.

b) Show that the *special linear group* $\mathrm{Sl}(n,\mathbb{R}) := \{g \in M(n,\mathbb{R})\,;\, \det(g) = 1\}$ is a Lie group. What is its dimension?

c) The group of affine transformations of $\mathbb{R}^n$ that preserve both volume and orientation is generated by $\mathrm{Sl}(n,\mathbb{R})$, together with the translations $T_v(x) = x + v$. Show that this is a Lie group. What is its dimension?

3 Tangent bundles for embedded submanifolds

Using the definition of a manifold in terms of coordinate patches, we defined the notion of a smooth map $F\colon M \to N$. In order to do calculus on manifolds, we will also need to define the *derivative* of F. To handle these in a coordinate-invariant way, we will need *tangent vectors*, *tangent bundles* and *vector fields*.

Before giving the definitions of these objects for general manifolds M, we first look at the easier case of embedded submanifolds $\Sigma \subseteq \mathbb{R}^n$. Here we can make use of the linear structure of the surrounding space $\mathbb{R}^n$, which is *not* available in the general case. This is why the current section is rather shorter than the next one.

3.1 Tangent vectors

Let $I \subseteq \mathbb{R}$ be an open interval containing zero. A *smooth curve* in M through $p \in M$ is a smooth function $\gamma\colon \mathbb{R} \supseteq I \to M$ such that $\gamma(0) = p$. For *embedded* submanifolds $\Sigma \subseteq \mathbb{R}^n$, we can define tangent vectors as follows.

Definition 3.1 ($T_p\Sigma$, embedded case)**.** Let $\Sigma \subseteq \mathbb{R}^n$ be an embedded submanifold. A vector $v_p \in \mathbb{R}^n$ is called a *tangent vector* to Σ at p if there exists a smooth curve $\gamma\colon \mathbb{R} \supseteq I \to \Sigma$ through p such that

$$v_p = \dot{\gamma}(0) = \lim_{h\to 0} \frac{\gamma(h) - \gamma(0)}{h}. \tag{29}$$

The set $T_p\Sigma \subseteq \mathbb{R}^n$ of tangent vectors is called the *tangent space* of Σ at p.

Remark 3.2. Note that by subtracting $\gamma(0) \in \Sigma$ from $\gamma(h) \in \Sigma$, we are making essential use of the fact that Σ lies inside $\mathbb{R}^n$. For general manifolds (not embedded in $\mathbb{R}^n$), we will need a *different* definition.

3.2 Tangent bundles and vector fields

For *embedded* submanifolds $\Sigma \subseteq \mathbb{R}^n$, we can define the tangent bundle as follows.

Definition 3.3 ($T\Sigma$, embedded case)**.** The *tangent bundle* $T\Sigma \subseteq \mathbb{R}^n \times \mathbb{R}^n$ is defined as

$$T\Sigma = \{(p, v) \in \mathbb{R}^n \times \mathbb{R}^n\,;\, p \in \Sigma \text{ and } v \in T_p\Sigma\}.$$

The *canonical projection* is the map $\pi\colon T\Sigma \to \Sigma$ with $\pi(p, v) = p$.

If we identify $(p, v) \in T\Sigma$ with $v \in T_p\Sigma$, we can consider $T\Sigma = \bigcup_{p\in\Sigma} T_p\Sigma$ as the disjoint union of the tangent spaces $T_p\Sigma$. Note that every tangent space $T_p\Sigma$ is a fibre of the canonical projection, $T_p\Sigma = \pi^{-1}(\{p\})$.

If $\Sigma \subseteq \mathbb{R}^n$ arises as the preimage of a regular value, then the tangent bundle $T\Sigma$ admits the following useful description.

Proposition 3.4. *Let $\Sigma \subseteq \mathbb{R}^n$ be an embedded submanifold that arises as the preimage under $F\colon \mathbb{R}^n \supseteq U \to \mathbb{R}^{n-k}$ of a regular value c. Then*

$$T\Sigma = \{(p, v) \in U \times \mathbb{R}^n\,;\, F(p) = c \text{ and } D_pF(v) = 0\}. \tag{30}$$

Proof. Suppose that (p, v) is in $T\Sigma$. Then there exists a curve γ in Σ through p with tangent vector $v \in T_p\Sigma$. Then since $F(\gamma(t)) = c$, we have $D_pF(v) = \frac{d}{dt}|_0 F(\gamma(t)) = 0$. Conversely, suppose that $D_pF(v) = 0$. To show that v is a tangent vector at p, we exhibit a smooth curve γ in Σ through p such that $\dot\gamma(0) = v$. Choose a slice chart (ϕ_α, U_α) around p, and define $F_\alpha := F \circ \phi_\alpha^{-1}$ and $v_\alpha := D_p\phi_\alpha(v)$. Since $F = F_\alpha \circ \phi_\alpha$, the chain rule yields $D_{\phi_\alpha(p)}F_\alpha(v_\alpha) = 0$. Since $\phi_\alpha(U_\alpha \cap \Sigma) = \phi_\alpha(U_\alpha) \cap (\mathbb{R}^k \times \{0\})$ is an open subset of the linear subspace $\mathbb{R}^k \subseteq \mathbb{R}^n$, it contains the straight line segment $\gamma_\alpha(t) = \phi_\alpha(p) + tv_\alpha$ for t sufficiently close to 0. It follows that $\gamma(t) = \phi_\alpha^{-1} \circ \gamma_\alpha(t)$ is a smooth curve in Σ with tangent vector v. ☐

Proposition 3.5. *Let $\Sigma \subseteq \mathbb{R}^n$ be an embedded submanifold. Then also the tangent bundle $T\Sigma \subseteq \mathbb{R}^n \times \mathbb{R}^n$ is an embedded submanifold, and the canonical projection $\pi\colon T\Sigma \to \Sigma$ is a smooth map.*

Proof. By Corollary 2.47, every embedded submanifold $\Sigma \subseteq \mathbb{R}^n$ is *locally* given by the preimage under $F\colon \mathbb{R}^n \supseteq U \to \mathbb{R}^{n-k}$ of a regular value $c \in \mathbb{R}^{n-k}$. It follows that every $p \in \Sigma$ admits a neighbourhood $U \subseteq \mathbb{R}^n$ such that $T(\Sigma \cap U) = \pi^{-1}(\Sigma \cap U)$ is given by (30). In particular, it is the preimage of $(c, 0)$ under the smooth map $DF\colon U \times \mathbb{R}^n \to \mathbb{R}^{n-k} \times \mathbb{R}^{n-k}$ defined by $(p, v) \mapsto D_pF(v)$. Since $(c, 0)$ is a regular value for this map (why?), $T(\Sigma \cap U) = (DF)^{-1}(c, 0)$ is an embedded submanifold. Since $T\Sigma$ is covered by the embedded submanifolds $T(\Sigma \cap U)$, it is an embedded submanifold itself. The canonical projection is smooth because it is the restriction to $T\Sigma \subseteq \mathbb{R}^n \times \mathbb{R}^n$ of the smooth map $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ which projects on the first factor. ☐

Remark 3.6. Note that $T\mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$. So if Σ is an embedded submanifold of $\mathbb{R}^n$, the $T\Sigma$ is an embedded submanifold of $T\mathbb{R}^n$.

This description of the tangent space is often very explicit. For example, consider the 2-sphere $\mathbb{S}^2_r$ of radius r, which is an embedded submanifold of $\mathbb{R}^3$. It is the level set of the function $F\colon \mathbb{R}^3 \to \mathbb{R}$ with $F(x, y, z) = x^2 + y^2 + z^2$ at the regular value $r^2 > 0$. Since $DF_p = (2x, 2y, 2z)$, we have

$$T\mathbb{S}^2_r = \{(\vec r, \vec v) \in \mathbb{R}^3 \times \mathbb{R}^3\,;\, \vec r \in \mathbb{S}^2_r \text{ and } \vec r \cdot \vec v = 0\}.$$

Definition 3.7. A *vector field* on an embedded submanifold $\Sigma \subseteq \mathbb{R}^n$ is a smooth map $v\colon \Sigma \to T\Sigma$ with $\pi \circ v = \mathrm{id}_\Sigma$. We denote the space of vector fields by $\mathrm{Vec}(M)$.

A vector field assigns to every point $p \in \Sigma$ a tangent vector v_p in the tangent space $T_p\Sigma$ *at the point* p.

Problem 3.8. If c is a regular value of the smooth function $f\colon \mathbb{R}^n \to \mathbb{R}$, then the tangent space $T\Sigma$ of $\Sigma := f^{-1}(\{c\})$ can be identified with

$$T\Sigma \simeq \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^n\,;\, x \in \Sigma \text{ and } v \cdot \nabla_x f = 0\}.$$

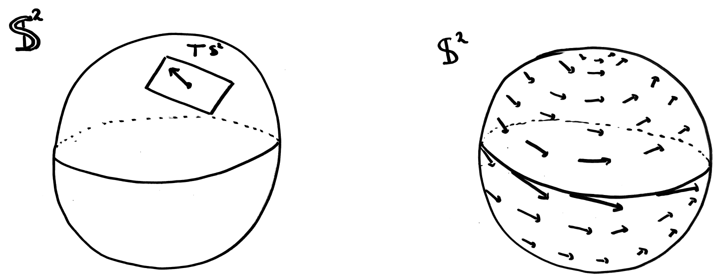


Figure 10: The tangent space (left) and a vector field (right) on the sphere.

Problem 3.9. Describe the tangent space of the ellipsoid

$$\{(x, y, z) \in \mathbb{R}^3\,;\,(x/a)^2 + (y/b)^2 + (z/c)^2 = 1\}$$

with $a, b, c > 0$ as an embedded submanifold of $\mathbb{R}^3 \times \mathbb{R}^3$.

Problem 3.10. Let c be a regular value of the smooth function $f\colon \mathbb{R}^n \to \mathbb{R}$, and let $\Sigma = f^{-1}(\{c\})$. Then the space $\mathrm{Vec}(\Sigma)$ of vector fields can be identified with

$$\mathrm{Vec}(\Sigma) \simeq \{v\colon \Sigma \to \mathbb{R}^n\,;\, v \text{ is smooth and } v(x)\cdot \nabla_x f = 0\}.$$

Problem 3.11. Describe the vector fields on the ellipsoid

$$\Sigma_{abc} := \{(x, y, z) \in \mathbb{R}^3\,;\,(x/a)^2 + (y/b)^2 + (z/c)^2 = 1\}$$

with $a, b, c > 0$ in terms of smooth maps from Σ_{abc} to $\mathbb{R}^3$.

Problem 3.12. For every $\vec{\Omega} \in \mathbb{R}^3$, the function $\widetilde{v}_\Omega(\vec{x}) := (\vec{x}, \vec{\Omega} \times \vec{x})$ yields a vector field on $\mathbb{S}^2$.

Problem 3.13. Let Σ be an embedded submanifold of $\mathbb{R}^n$.

a) Show that $\mathrm{Vec}(\Sigma)$ is a vector space.

b) If the dimension of Σ is at least 1, then $\mathrm{Vec}(\Sigma)$ is an infinite dimensional vector space.

4 Tangent bundles for smooth manifolds

For two smooth manifolds M and N, we have defined what it means for a map $F\colon M \to N$ to be *differentiable*. We have not yet defined its *derivative*. In order to do this, we will need *tangent vectors* and *tangent bundles*.

To define tangent vectors on *embedded* manifolds $\Sigma \subseteq \mathbb{R}^n$, we made essential use of the linear structure of the envelopping space $\mathbb{R}^n$, which is unfortunately not available for a general manifold M. There are (at least) three equivalent ways to define tangent vectors – all of which have their advantages and disadvantages. The definition that we will use is based on smooth curves.

4.1 Tangent vectors and the tangent space T_pM

Let M be a smooth manifold of dimension n, and let $\gamma\colon \mathbb{R} \supseteq I \to M$ be a smooth curve through $p \in M$. With respect to a chart (U_α, ϕ_α) that contains p, the coordinate representation $\gamma_\alpha = \phi_\alpha \circ \gamma$ is defined on an open neighbourhood of $0 \in I \subseteq \mathbb{R}$, and takes values in $\mathbb{R}^n$. Since $\mathbb{R}^n$ is a vector space, we can define

$$v^\mu_\alpha := \dot\gamma^\mu_\alpha(0) = \lim_{h\to 0} \frac{\gamma^\mu_\alpha(h) - \gamma^\mu_\alpha(0)}{h}. \tag{31}$$

This expression depends on the chart in a controlled way. Since $\gamma_\beta = \kappa_{\alpha\beta} \circ \gamma_\alpha$, we infer from the chain rule that $\dot\gamma_\beta(0)$ depends linearly on $\dot\gamma_\alpha(0)$,

$$\dot\gamma_\beta(0) = \Big(D_{\phi_\alpha(p)}\kappa_{\alpha\beta}\Big)\dot\gamma_\alpha(0). \tag{32}$$

Explicitly, if we define $v^\mu_\alpha := \dot\gamma^\mu_\alpha(0)$ and $v^{\overline{\mu}}_\beta := \dot\gamma^{\overline{\mu}}_\beta(0)$, and if we let $J^{\overline{\mu}}_\mu = \frac{\partial x^{\overline{\mu}}}{\partial x^\mu}$ be the Jacobian matrix of $\kappa_{\alpha\beta}$ at $\phi_\alpha(p) = (x^1, \ldots, x^n)$, then the above equation has coordinate expression

$$v^{\overline{\mu}}_\beta = \sum_{\mu=1}^n J^{\overline{\mu}}_\mu v^\mu_\alpha. \tag{33}$$

This has the following simple, but important, consequence.

Proposition 4.1. *Suppose that two smooth curves γ and l through $p \in M$ satisfy $\dot\gamma_\alpha(0) = \dot l_\alpha(0)$ with respect to a chart (U_α, ϕ_α) around p. Then they satisfy $\dot\gamma_\beta(0) = \dot l_\beta(0)$ with respect to any other chart (U_β, ϕ_β) as well.*

Proof. If $v^\mu_\alpha := \dot\gamma^\mu_\alpha(0)$ is equal to $w^\mu_\alpha := \dot l^\mu_\alpha(0)$, then

$$v^{\overline{\mu}}_\beta := \dot\gamma^{\overline{\mu}}_\beta(0) = \sum_{\mu=1}^n J^{\overline{\mu}}_\mu v^\mu_\alpha = \sum_{\mu=1}^n J^{\overline{\mu}}_\mu w^\mu_\alpha = \dot l^{\overline{\mu}}_\beta(0) =: w^{\overline{\mu}}_\beta. \qquad \square$$

We can therefore define an equivalence relation on curves through p as follows. Two curves γ and l are equivalent, $\gamma \sim_p l$, if there exists a chart (U_α, ϕ_α) around p in which the first order derivatives $v^\mu_\alpha = \dot\gamma^\mu_\alpha(0)$ and $w^\mu_\alpha = \dot l^\mu_\alpha(0)$ agree. By the above proposition, this will then automatically hold in any other chart around p as well!

Definition 4.2 (Tangent space (general case))**.** A tangent vector at $p \in M$ is an equivalence class $v_p = [\gamma]$ of curves through p with respect to the relation $\sim_p$. The set T_pM of tangent vectors is called the *tangent space* of M at p.

For a tangent vector $v_p \in T_pM$, we call $(v_\alpha^1, \dots v_\alpha^n) \in \mathbb{R}^n$ the *coordinate expression* of v_p with respect to the chart (U_α, ϕ_α). If we fix the chart, then v_p is uniquely determined by its coordinates $(v_\alpha^1, \dots v_\alpha^n)$ Conversely, every n-tuple $(v_\alpha^1, \dots v_\alpha^n) \in \mathbb{R}^n$ gives rise to a unique tangent vector v_p.

Proposition 4.3. *The map*

$$\phi_{\alpha *} \colon T_pM \to \mathbb{R}^n, \quad \phi_{\alpha *}(v_p) = (v_\alpha^1, \dots, v_\alpha^n)$$

is a bijection.

Proof. Injectivity follows directly from definition 4.2 (why?). To see that $\phi_{\alpha *}$ is surjective, note that $(v_\alpha^1, \dots, v_\alpha^n)$ is the derivative at zero of the curve $\gamma_\alpha(t) = \phi_\alpha(p) + (tv_\alpha^1, \dots, tv_\alpha^n)$ in $\mathbb{R}^n$ through $\phi_\alpha(p)$. To construct a vector $v_p = [\gamma] \in T_pM$ with $\phi_{\alpha *}(v_p) = (v_\alpha^1, \dots, v_\alpha^n)$, we simply take the curve in M through p defined by $\gamma(t) = \phi_\alpha^{-1} \circ \gamma_\alpha(t)$. Its domain is the open interval $\gamma_\alpha^{-1}(\phi_\alpha(U_\alpha))$ in $\mathbb{R}$. □

This allows us to consider T_pM as a *vector space.* To define the sum of $v_p \in T_pM$ and $w_p \in T_pM$, we simply add the corresponding coordinates in $\mathbb{R}^n$ and define $v_p + w_p \in T_pM$ to be the unique vector with coordinates $(v_\alpha^1 + w_\alpha^1, \dots, v_\alpha^n + w_\alpha^n)$. Similarly, we define $\lambda v_p \in T_pM$ to be the unique tangent vector with coordinates $(\lambda v_\alpha^1, \dots, \lambda v_\alpha^n)$.

Although the addition and scalar multiplication on T_pM are defined using the chart (U_α, ϕ_α), the resulting vector space structure is in fact independent of the chart. Indeed, equation (32) shows that the bijections $\phi_{\alpha *} \colon T_pM \to \mathbb{R}^n$ and $\phi_{\beta *} \colon T_pM \to \mathbb{R}^n$ are related by the *linear* map $D\kappa_{\alpha\beta}(\phi_\alpha(p)) \colon \mathbb{R}^n \to \mathbb{R}^n$. Since the two bijections differ by a linear map, they define the same vector space structure on T_pM.

Problem 4.4. If the above description is not sufficiently rigorous to you (I am not calling anyone a nitpicker), one can proceed as follows. Define

$$v +_\alpha w := \phi_{\alpha *}^{-1}\big(\phi_{\alpha *}(v) + \phi_{\alpha *}(w)\big) \quad \text{and} \quad \lambda \cdot_\alpha v := \phi_{\alpha *}^{-1}\big(\lambda \cdot \phi_{\alpha *}(v)\big)\,.$$

Prove that these operations make T_pM into a vector space. Prove that if we analogously define $v +_\beta w := \phi_{\beta *}^{-1}(\phi_{\beta *}(v) + \phi_{\beta *}(w))$ and $\lambda \cdot_\beta v := \phi_{\beta *}^{-1}(\lambda \cdot \phi_{\beta *}(v))$, then $v +_\alpha w = v +_\beta w$ and $\lambda \cdot_\alpha v = \lambda \cdot_\beta v$.

Problem 4.5. Prove that if $v_p = [\gamma]$, then $\lambda v_p = [\gamma(\lambda\,\cdot\,)]$, where $\gamma(\lambda\,\cdot\,)$ denotes the curve $t \mapsto \gamma(\lambda t)$.

Problem 4.6. We saw that $\mathbb{S}^2$ can be covered by two coordinate charts (U_α, ϕ_α) and (U_β, ϕ_β) with $U_\alpha = \mathbb{S}^2 \setminus \{n\}$, $U_\beta = \mathbb{S}^2 \setminus \{s\}$, and transition function $\kappa_{\alpha\beta}(x, y) = \frac{1}{x^2+y^2}(x, y)$.

(a) Give the relation between (v^1_α, v^2_α) and (v^1_β, v^2_β).

(b) If (v^1_α, v^2_α) is constant in p, what happens to (v^1_β, v^2_β) as p approaches $n = (0, 0, 1)$? And as p approaches $s = (0, 0, -1)$?

Remark 4.7. For $M = \mathbb{R}^n$, there exists a *canonical* global chart $\phi \colon \mathbb{R}^n \to \mathbb{R}^n$ which is simply the identity. We will use this chart to identify $T_p\mathbb{R}^n$ with $\mathbb{R}^n$ without further comment.

Definition 4.8. For a smooth curve $\gamma \colon \mathbb{R} \supseteq I \to M$, we define the *tangent vector* $\dot\gamma(t_0)$ at the point $p = \gamma(t_0)$ as the equivalence class modulo $\sim_p$ of the curve $t \mapsto \gamma(t + t_0)$.

We thus have two different ways of looking at tangent vectors. We can describe them coordinate-invariantly as equivalence classes $v_p = [\gamma]$ of curves, and upon choosing coordinates (U_α, ϕ_α), we can describe them by an n-tuple of real numbers $(v^1_\alpha, \ldots, v^n_\alpha)$. If we use *different* coordinates (U_β, ϕ_β) to describe the *same* vector v_p, then the coordinates change according to formula (33). This is called *contravariant* transformation behaviour.

4.2 Derivations

Here is a third, more algebraic way of viewing tangent vectors. Recall from Prop. 2.8 that the space $C^\infty(M)$ of smooth functions $f \colon M \to \mathbb{R}$ is an algebra.

Definition 4.9 (Derivations at a point)**.** A *derivation* of $C^\infty(M)$ at the point $p \in M$ is a linear map $D_p \colon C^\infty(M) \to \mathbb{R}$ that satisfies the *Leibniz rule*

$$D_p(fg) \quad = \quad D_p(f)g(p) + f(p)D_p(g). \tag{34}$$

We denote by $\mathrm{Der}_p(C^\infty(M))$ the set of derivations of $C^\infty(M)$ at p.

The derivations $\mathrm{Der}_p(C^\infty(M))$ constitute a *vector space*. That is, for any two derivations D_p and D'_p and any $\alpha, \beta \in \mathbb{R}$, the linear combination $\alpha D_p + \beta D'_p$ is again a derivation.

Problem 4.10. Prove that $\mathrm{Der}_p(C^\infty(M))$ is a vector space.

From a smooth curve $\gamma \colon \mathbb{R} \supseteq I \to M$ through $p \in M$, we obtain a derivation $D^\gamma_p \colon C^\infty(M) \to \mathbb{R}$ by differentiating along γ,

$$D^\gamma_p(f) := \frac{d}{dt}\Big|_0 f(\gamma(t)). \tag{35}$$

Proposition 4.11. *The map $D^\gamma_p \colon C^\infty(M) \to \mathbb{R}$ is a derivation.*

Proof. The map D^γ_p is linear because differentiation is linear:

$$\begin{aligned} D^\gamma_p(\alpha f + \beta g) &= \frac{d}{dt}\Big|_0 \Big((\alpha f + \beta g)(\gamma(t))\Big) \\ &= \alpha \frac{d}{dt}\Big|_0 f(\gamma(t)) + \beta \frac{d}{dt}\Big|_0 g(\gamma(t)) = \alpha D^\gamma_p(f) + \beta D^\gamma_p(g). \end{aligned}$$

The Leibniz rule for D_p^γ comes from the product rule for differentiation,

$$\begin{aligned}
D_p^\gamma(fg) &= \frac{d}{dt}\Big|_0 f(\gamma(t))\cdot g(\gamma(t)) \\
&= \left(\frac{d}{dt}\Big|_0 f(\gamma(t))\right) g(p) + f(p)\left(\frac{d}{dt}\Big|_0 (g(\gamma(t))\right) \\
&= D_p^\gamma(f)g(p) + f(p)D_p^\gamma(g).
\end{aligned}$$

□

To express D_p^γ in local coordinates, note that

$$f\circ\gamma = (f\circ\phi_\alpha^{-1})\circ(\phi_\alpha\circ\gamma) = f_\alpha\circ\gamma_\alpha$$

for t sufficiently close to zero. Since $\dot\gamma_\alpha^\mu = v_\alpha^\mu$, we have

$$D_p^\gamma(f) = \frac{d}{dt}\Big|_0 f_\alpha\big(\gamma_\alpha^1(t),\dots,\gamma_\alpha^n(t)\big) = \sum_{\mu=1}^n v_\alpha^\mu \frac{\partial}{\partial x^\mu} f_\alpha(x^1,\dots,x^n), \tag{36}$$

where $(x^1,\dots,x^n)$ are the coordinates of p with respect to (U_α,ϕ_α). In particular, the derivation D_p^γ depends on the curve γ only through its tangent vector $v_p = [\gamma]$! We conclude that a tangent vector $v_p = [\gamma]\in T_pM$ gives rise to the derivation

$$D_p^v(f) = \sum_{\mu=1}^n v_\alpha^\mu \frac{\partial}{\partial x^\mu} f_\alpha(x^1,\dots,x^n). \tag{37}$$

The following theorem allows us to identify tangent vectors at p with derivations.

Theorem 4.12 (Tangent vectors as derivations)**.** *The map $T_pM\to \mathrm{Der}_p(C^\infty(M))$ defined by $v_p\mapsto D_p^v$ is a linear isomorphism of vector spaces.*

Proof. The proof uses the *Hausdorff property* of M in an essential way. We will come back to the proof of this theorem in Section 7. □

Problem 4.13 (Hadamard lemma)**.** Let $f\colon\mathbb{R}^n\to\mathbb{R}$ be a smooth function. Show that f can be written as

$$f(x) = f(p) + \sum_{\mu=1}^n (x^\mu - p^\mu)\rho_\mu(x)$$

for smooth functions $\rho_\mu\colon\mathbb{R}^n\to\mathbb{R}$ with $\rho_\mu(p) = \frac{\partial}{\partial x^\mu}f(p)$.
(Hint: evaluate $h(1) = h(0) + \int_0^1 h'(t)dt$ for the function $h(t) := f(p+t(x-p))$.)

Problem 4.14. Prove that the vector space $\mathrm{Der}_p(C^\infty(\mathbb{R}^n))$ of derivations at $p\in\mathbb{R}^n$ is isomorphic to $T_p\mathbb{R}^n$.

a) Show that every derivation satisfies $D_p(1) = 0$. Conclude that $D_p(f) = 0$ for any constant function f.

b) Using Problem 4.13 or otherwise, show that every derivation D_p of $C^\infty(\mathbb{R}^n)$ at p is given by $D_p(f) = \sum_{\mu=1}^n v^\mu \frac{\partial f}{\partial x^\mu}(p)$ for certain constants $v^\mu \in \mathbb{R}$.

c) Conclude that $T_p\mathbb{R}^n \simeq \mathrm{Der}_p(C^\infty(\mathbb{R}^n))$.

Problem 4.15. Let A be a commutative algebra with unit over a field $\mathbb{K}$, and let $I \subseteq A$ be an ideal.

a) The $\mathbb{K}$-vector space A/I is an A-module with action $A \times (A/I) \to (A/I)$ given by $a \cdot [b] = [ab]$.

b) A *derivation at* I is a linear map $D\colon A \to A/I$ that satisfies the Leibniz rule $D(ab) = aD(b) + bD(a)$. Show that the set $\mathrm{Der}_I(A)$ of derivations at I is a vector space over $\mathbb{K}$.

c) Let $\mathbb{K} = \mathbb{R}$, let $A = C^\infty(M)$, and let $I_p \subseteq C^\infty(M)$ be the set of smooth functions that vanish at $p \in M$. Show that I_p is an ideal. Give an isomorphism between A/I_p and the A-module $\mathbb{R}$ with action $A \times \mathbb{R} \to \mathbb{R}$ given by $f \cdot x = f(p)x$. Identify derivations of $C^\infty(M)$ at I_p with derivations of $C^\infty(M)$ at $p \in M$ in the sense of definition 4.9.

d) Let $A = \mathbb{C}[z_1, \ldots, z_n]$ be the algebra over $\mathbb{K} = \mathbb{C}$ of complex polynomials in n variables. Show that the set I_0 of polynomials that vanish at $(0, \ldots, 0) \in \mathbb{C}^n$ is an ideal. Give an isomorphism between A/I_0 and the A-module $\mathbb{C}$ with action $A \times \mathbb{C} \to \mathbb{C}$ defined by $p \cdot z = p(0)z$. Prove that $\mathrm{Der}_{I_0}(\mathbb{C}[z_1, \ldots, z_n])$ is an n-dimensional complex vector space with basis $D_i(p) := [\frac{d}{dz_i}p]$.

4.3 Einstein summation convention

Since we will have to write down quite a lot of expressions like (37), we introduce some notational conventions to make them a bit shorter. First of all, we will often denote the derivation D_p^v simply by v_p,

$$v_p(f) = \sum_{\mu=1}^n v_\alpha^\mu \frac{\partial}{\partial x^\mu} f_\alpha(x^1, \ldots, x^n). \tag{38}$$

Secondly, denote by $\partial_\mu \in T_pM$ the tangent vector whose coordinates with respect to the chart (U_α, ϕ_α) are all zero, except for the μ-coordinate, which is one. Since the vectors ∂_μ with $\mu = 1, \ldots, n$ correspond to the standard basis in $\mathbb{R}^n$, they constitute a basis of T_pM called the *coordinate basis*. The reason for denoting these vectors by ∂_μ is that the corresponding derivation is simply the partial derivative in the x^μ-direction,

$$\partial_\mu(f) = \frac{\partial}{\partial x^\mu} f_\alpha(x^1, \ldots, x^n).$$

With this notation, (38) reduces to

$$v_p(f) = \sum_{\mu=1}^{n} v_\alpha^\mu \partial_\mu(f). \tag{39}$$

Finally, we use the *Einstein summation convention* that repeated indices are summed if one is a subscript and the other a superscript. This turns equation (39) for the derivation into the (rather shorter) expression

$$v_p(f) = v_\alpha^\mu \partial_\mu(f)\,.$$

This corresponds to

$$v_p = v_\alpha^\mu \partial_\mu \tag{40}$$

at the level of tangent vectors.

4.4 Three different ways to view a tangent vector

Summarizing, we now have *three* different ways to describe a tangent vector $v_p \in T_pM$ for a manifold M.

(1) A tangent vector $v_p \in T_pM$ is, by definition, an equivalence class $v_p = [\gamma]$ of smooth curves through p. We can think of v_p as the derivative of $\gamma(t)$ at $t = 0$.

(2) With respect to a chart (U_α, ϕ_α) around p, every tangent vector $v_p \in T_pM$ is of the form

$$v_p = v_\alpha^\mu \partial_\mu$$

for an n-vector $(v_\alpha^1, \ldots, v_\alpha^n)$. If $v_p = v_\beta^{\overline{\mu}} \partial_{\overline{\mu}}$ with respect to a different chart (U_β, ϕ_β), then the coefficients $v_\beta^{\overline{\mu}}$ are related to v_α^μ by

$$v_\beta^{\overline{\mu}} = \frac{\partial x^{\overline{\mu}}}{\partial x^\mu} v_\alpha^\mu,$$

where $x^{\overline{\mu}}(x^1, \ldots x^n)$ is the transition function $\kappa_{\alpha\beta}$ between the charts.

(3) A tangent vector $v_p \in T_pM$ can be viewed as a derivation of $C^\infty(M)$ at $p \in M$, that is, a linear map $v_p \colon C^\infty(M) \to \mathbb{R}$ that satisfies the Leibniz rule $v_p(fg) = v_p(f)g(p) + f(p)v_p(g)$.

Each of these descriptions has its advantages and disadvantages. We have chosen description (1) as the definition of a tangent vector and derived the other two, but in the literature one also encounters definitions based on (2) and (3).

4.5 The derivative of a smooth function

Now that we have defined tangent vectors, we can finally define the derivative of a smooth function $F \colon M \to N$. In the context of differential geometry, the derivative is often called the *pushforward*. If F maps $p \in M$ to $F(p) \in N$, then the pushforward is a linear map $F_* \colon T_pM \to T_{F(p)}N$.

Definition 4.16 (Pushforward)**.** The pushforward of a smooth map $F\colon M \to N$ at $p \in M$ is the linear map

$$F_*\colon T_pM \to T_{F(p)}N$$

that takes $v_p = [\gamma]$ to $F_*(v_p) = [F \circ \gamma]$.

To see that F_* is linear and well defined, express $v_p = [\gamma]$ in coordinates (U_α, ϕ_α) around $p \in M$ as $v_p = v^\mu_\alpha \partial_\mu$. Similarly, express $w_{F(p)}$ in coordinates (V_β, ψ_β) around $F(p) \in N$ as $w_{F(p)} = w^\nu_\beta \partial_\nu$. Since $(F \circ \gamma)_\beta = F_{\alpha\beta} \circ \gamma_\alpha$, the chain rule yields

$$w^\nu_\beta = \tfrac{d}{dt}\big|_0 \big(F^\nu_{\alpha\beta}(\gamma^1_\alpha(t), \ldots, \gamma^n_\alpha(t))\big) = \partial_\mu F^\nu_{\alpha\beta}\big(x^1(p), \ldots, x^n(p)\big) v^\mu_\alpha, \tag{41}$$

where $(\gamma^1_\alpha(0), \ldots, \gamma^n_\alpha(0)) = (x^1(p), \ldots, x^n(p))$ are the coordinates of the point p. In particular, w^ν_β depends on γ only through the coordinates v^μ_α, and it does so in a linear fashion. Since the coordinate expression (41) for the derivative is quite useful, we record it in a proposition.

Proposition 4.17 (Pushforward and Jacobian)**.** *With respect to the coordinate basis ∂_μ of T_pM and ∂_ν of $T_{F(p)}N$, the linear map $F_*\colon T_pM \to T_{F(p)}N$ is represented by the Jacobian matrix*

$$J^\nu_\mu = \partial_\mu F^\nu_{\alpha\beta}(x^1(p), \ldots, x^n(p)).$$

Remark 4.18 (Pushforward in $\mathbb{R}^n$)**.** For a smooth map $F\colon \mathbb{R}^n \to \mathbb{R}^m$, we can canonically identify $T_p\mathbb{R}^n$ with $\mathbb{R}^n$ and $T_{F(p)}\mathbb{R}^m$ with $\mathbb{R}^m$. In this case (and in this case only!), we will identify the pushforward $F_*\colon T_p\mathbb{R}^n \to T_{F(p)}\mathbb{R}^m$ with the total derivative $D_pF\colon \mathbb{R}^n \to \mathbb{R}^m$.

The pushforward enjoys the following familiar properties.

Proposition 4.19 (Chain rule)**.** *Let M and N be smooth manifolds, and let $F\colon M \to N$ be a smooth map.*

a) The chain rule holds. If $F\colon M \to N$ and $G\colon L \to M$ are smooth maps, then

$$(F \circ G)_* = F_* \circ G_*.$$

More precisely, the linear map $(F \circ G)_*\colon T_pL \to T_{F(G(p))}N$ is the concatenation of $G_*\colon T_pL \to T_{G(p)}M$ and $F_*\colon T_{G(p)}M \to T_{F(G(p))}N$.

b) The pushforward of the identity is the identity, $\mathrm{Id}_{M*} = \mathrm{Id}_{T_pM}$.

c) The pushforward of a diffeomorphism $\phi\colon M \to N$ is an isomorphism of vector spaces $\phi_*\colon T_pM \to T_{\phi(p)}N$.

Proof. The chain rule (a) immediately follows from $(F \circ G)_*([\gamma]) = [(F \circ G) \circ \gamma]$ and $F_*(G_*[\gamma]) = [F \circ (G \circ \gamma)]$ for $[\gamma] \in T_pM$. For part (b), simply note that $\mathrm{Id}_*([\gamma]) = [\mathrm{Id} \circ \gamma] = [\gamma]$. From (a) and (b), we have $(\phi^{-1})_* \circ \phi_* = \mathrm{Id}_{T_pM}$ and $\phi_* \circ (\phi^{-1})_* = \mathrm{Id}_{T_{\phi(p)}M}$, so $\phi_*\colon T_pM \to T_{\phi(p)}M$ is a linear isomorphism with inverse $(\phi_*)^{-1} = (\phi^{-1})_*$. □

Definition 4.20 (Pullback)**.** The *pullback* along $F\colon M \to N$ of a function $f \in C^\infty(N)$ is the function $F^*f \in C^\infty(M)$ defined by $F^*f := f \circ F$.

This is compatible with the pushforward of vector fields in the sense that

$$v(F^*f) = F_*v(f) \tag{42}$$

for all $v \in T_pM$, cf. Problem 4.24.

Problem 4.21. Let $\gamma\colon \mathbb{R} \supseteq I \to M$ be a smooth curve. Then the pushforward $\gamma_*(\partial_t)$ of $\partial_t \in T_{t_0}\mathbb{R}$ is the tangent vector $\dot\gamma(t_0) \in T_{\gamma(t_0)}M$ of Definition 4.8.

Problem 4.22. Let $\gamma\colon \mathbb{R} \to \mathbb{S}^1$ be the smooth curve $\gamma(t) = (\cos(t), \sin(t))$, and let $R_\theta\colon \mathbb{S}^1 \to \mathbb{S}^1$ be the rotation over angle θ,

$$R_\theta \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Show that $R_{\theta *}\dot\gamma(t_0) = \dot\gamma(t_0 - \theta)$.

Problem 4.23. Let $\chi \in \mathbb{R}$, and let $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the smooth function $F(x,y) = (\cosh(\chi)x + \sinh(\chi)y, \sinh(\chi)x + \cosh(\chi)y)$.

a) Calculate the pullback F^*f of the functions $f(x,y) = x$ and for $f(x,y) = y$.

b) Calculate the pushforward F_*v of the vectors $v = \partial_x$ and $v = \partial_y$.

Problem 4.24 (Push forward, pull back)**.** Let $F\colon M \to N$ be a smooth map.

a) The pullback $F^*\colon C^\infty(N) \to C^\infty(M)$ is an algebra homomorphism.

b) For every $D_p \in \mathrm{Der}_p(C^\infty(M))$, the pushforward $F_*D_p := D_p \circ F^*$ is an element of $\mathrm{Der}_{F(p)}(C^\infty(N))$.

c) The pushforward $F_*\colon \mathrm{Der}_p(C^\infty(M)) \to \mathrm{Der}_{F(p)}(C^\infty(N))$ of derivations is compatible with the pushforward $F_*\colon T_pM \to T_{F(p)}N$ of tangent vectors in the sense of (42).

4.6 The tangent bundle

For a smooth manifold M, the tangent space T_pM is the set of tangent vectors v_p at the point $p \in M$. The *tangent bundle* TM is the set of *all* tangent vectors v, regardless where they are based. In other words, TM is the disjoint union of the sets T_pM, where p runs over the entire manifold M,

$$TM := \bigsqcup_{p \in M} T_pM. \tag{43}$$

For $M = \mathbb{R}^n$, we have a bijective correspondence $T\mathbb{R}^n \simeq \mathbb{R}^n \times \mathbb{R}^n$. Indeed, $\mathbb{R}^n$ has a canonical global chart, so we can identify every tangent vector $v_x \in T\mathbb{R}^n$ with a unique point $(x, v) \in \mathbb{R}^n \times \mathbb{R}^n$ such that $v_x = v^\mu \partial_\mu$. Similarly, for an open subset $U \subseteq \mathbb{R}^n$, we have a bijective correspondence $TU \simeq U \times \mathbb{R}^n$. This allows us to consider TU as a *smooth manifold*, where the bijective correspondence $TU \simeq U \times \mathbb{R}^n$ serves as a (global) chart for TU. Using the following lemma, we will define a smooth manifold structure on TM for arbitrary smooth manifolds M.

Lemma 4.25. *If $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ are open subsets, and $F\colon U \to V$ is a smooth map, then the pushforward $F_*\colon TU \to TV$ is a smooth map as well.*

Proof. If we identify $TU \simeq U \times \mathbb{R}^n$ and $TV \simeq V \times \mathbb{R}^m$, then the coordinate representation of $F_*\colon TU \to TV$ is the map

$$(x^1, \ldots, x^n; v^1, \ldots, v^n) \mapsto (F^1(x), \ldots, F^m(x); \partial_\mu F^1(x) v^\mu, \ldots \partial_\mu F^m(x) v^\mu),$$

which is smooth in x^μ because F is smooth, and in v^μ because the expression is linear in v. $\square$

Theorem 4.26. *If M is a smooth manifold of dimension n, then TM is a smooth manifold of dimension $2n$. The canonical projection is a smooth map $\pi\colon TM \to M$, and for any smooth map $F\colon M \to N$, the pushforward $F_*\colon TM \to TN$ is a smooth map as well.*

Proof. If M is covered by coordinate neighbourhoods U_α, then TM is covered by the sets TU_α. Since $\phi_\alpha\colon U_\alpha \to \phi_\alpha(U_\alpha)$ is a diffeomorphism, the pushforward

$$\phi_{\alpha*}\colon TU_\alpha \to T\phi_\alpha(U_\alpha) \simeq \phi_\alpha(U_\alpha) \times \mathbb{R}^n$$

is a bijection by part (3) of Proposition 4.19. Since $\phi_\alpha(U_\alpha)$ is an open subset of $\mathbb{R}^n$, $T\phi_\alpha(U_\alpha)$ is a manifold with a global chart $T\phi_\alpha(U_\alpha) \simeq \phi_\alpha(U_\alpha) \times \mathbb{R}^n$. The strategy is to use the sets $TU_\alpha \subseteq TM$ as coordinate neighbourhoods for TM, to use the maps $\phi_{\alpha*}\colon TU_\alpha \to \phi_\alpha(U_\alpha) \times \mathbb{R}^n \subseteq \mathbb{R}^{2n}$ as coordinates. For this, we need to endow TM with a topology that makes the pushforward $\phi_{\alpha*}$ into a homeomorphism, and show that the transition functions between $\phi_{\alpha*}$ and $\phi_{\beta*}$ are smooth.

Step 1 is to define a Hausdorff topology on TM, and to show that the intended charts $\phi_{\alpha*}$ are diffeomorphisms. For this, declare $W \subseteq TM$ to be *open* if its coordinate image $\phi_{\alpha*}(W \cap TU_\alpha) \subseteq \mathbb{R}^n \times \mathbb{R}^n$ is open for every chart $(TU_\alpha, \phi_{\alpha*})$. In Problem 4.27 you are asked to check that this is indeed a topology.

To see that the intended chart $\phi_{\alpha*}\colon TU_\alpha \to \phi_\alpha(U_\alpha) \times \mathbb{R}^n$ is a homeomorphism, we need to check that $W \subseteq TU_\alpha$ is open if and only if $\phi_{\alpha*}(W) \subseteq \mathbb{R}^n \times \mathbb{R}^n$ is open. If $\phi_{\alpha*}(W) \subseteq \mathbb{R}^n \times \mathbb{R}^n$ is open, then for any other chart $(TU_\beta, \phi_{\beta_*})$, the set

$$\phi_{\beta_*}(W \cap TU_\beta) = \kappa_{\alpha\beta_*}\Big(\phi_{\alpha*}(W) \cap \big(\phi_\alpha(U_\alpha \cap U_\beta) \times \mathbb{R}^n\big)\Big)$$

is open because $\kappa_{\alpha\beta_*}$ is a homeomorphism and $\phi_\alpha(U_\alpha \cap U_\beta) \times \mathbb{R}^n$ is open. It follows that W is open in TM. The converse implication follows straight from the definition.

To show that the topology on TM is Hausdorff, note that since M is Hausdorff, two different points $p, q \in M$ can be separated by open neighbourhoods

$p \in U_p \subseteq M$ and $q \in U_q \subseteq M$ with $U_p \cap U_q = \emptyset$. Two vectors $v_p, v_q \in TM$ with different base points $p, q \in M$ can therefore be separated by the open sets $TU_p \subseteq TM$ and $TU_q \subseteq TM$. If $p = q$ but $v_p \neq v_q$, then v_p and v_q reside in a single coordinate chart $(TU_\alpha, \phi_{\alpha *})$, and their coordinates $(x, v) \in \phi_\alpha(U_\alpha) \times \mathbb{R}^n$ and $(x, w) \in \phi_\alpha(U_\alpha) \times \mathbb{R}^n$ can be separated by open neighbourhoods because $\mathbb{R}^n$ is Hausdorff.

Step 2 is to check that the transition functions between $\phi_{\alpha *}$ and $\phi_{\beta *}$ are smooth. This is the case because the transition function between the push-forward charts $(TU_\alpha, \phi_{\alpha *})$ and $(TU_\beta, \phi_{\beta_*})$ is just the pushforward $\kappa_{\alpha\beta_*}$ of the transition function $\kappa_{\alpha\beta}$ between the charts (U_α, ϕ_α) and (U_β, ϕ_β). Indeed, on the intersection $TU_\alpha \cap TU_\beta = T(U_\alpha \cap U_\beta)$, the transition function is given by $\phi_{\beta *} \circ (\phi_{\alpha *})^{-1} = (\phi_\beta \circ \phi_\alpha^{-1})_* = \kappa_{\alpha\beta_*}$, considered as a map from $\phi_\alpha(U_\alpha \times U_\beta) \times \mathbb{R}^n$ to $\phi_\beta(U_\alpha \times U_\beta) \times \mathbb{R}^n$,

$$\begin{tikzcd} & T(U_\alpha \cap U_\beta) \arrow[dl, "\phi_{\alpha_*}"'] \arrow[dr, "\phi_{\beta_*}"] & \\ \phi_\alpha(U_\alpha \cap U_\beta) \times \mathbb{R}^n \arrow[rr, "\kappa_{\alpha\beta_*}"] & & \phi_\beta(U_\alpha \cap U_\beta) \times \mathbb{R}^n. \end{tikzcd}$$

Since $\kappa_{\alpha\beta}$ is a smooth map $\phi_\alpha(U_\alpha \cap U_\beta) \to \phi_\beta(U_\alpha \cap U_\beta)$, its pushforward $\kappa_{\alpha\beta_*}$ is smooth by Lemma 4.25.

We conclude that TM is a smooth manifold with charts $\phi_{\alpha *}$. To see that $\pi\colon TM \to M$ is smooth, note that in local coordinates (U_α, ϕ_α) around $p \in M$ and $(TU_\alpha, \phi_{\alpha *})$ around $v_p \in TM$, the projection takes the form

$$\pi_{\alpha_*\alpha}\colon \phi_\alpha(U_\alpha) \times \mathbb{R}^n \to \phi_\alpha(U_\alpha), \quad (x, v) \mapsto x\,,$$

which is clearly smooth. It remains to check that that for any smooth map $F\colon M \to N$, the pushforward $F_*\colon TM \to TN$ is smooth as well. In local coordinates $(TU_\alpha, \phi_{\alpha *})$ around $v_p \in TM$ and $(TV_\alpha, \psi_{\beta_*})$ around $F_* v_p \in TN$, the pushforward F_* admits the coordinate representation

$$F_{\alpha_*\beta_*}\colon \phi_\alpha(U_\alpha) \times \mathbb{R}^n \to \psi_\beta(U_\beta) \times \mathbb{R}^n, \quad (x, v) \mapsto (F_{\alpha\beta}(x), \partial_\mu F^\nu_{\alpha\beta}(x) v^\mu)\,,$$

which is smooth by Lemma 4.25. □

Problem 4.27. Check that the open sets in TM indeed constitute a topology.

Problem 4.28. Show that $T\mathbb{S}^1$ is diffeomorphic to $\mathbb{S}^1 \times \mathbb{R}$.

Problem 4.29. If $\phi\colon M \to N$ is a diffeomorphism, then $\phi_*\colon TM \to TN$ is a diffeomorphism as well.

5 Embedded submanifolds revisited

In Section 2.6, we mainly studied embedded submanifolds Σ of $\mathbb{R}^n$. Now that we defined derivatives of smooth functions between general manifolds, we can extend these results to embedded submanifolds $\Sigma \subseteq M$. We use this to study the *Hopf fibration.*

5.1 The Regular Level Set Theorem

For a smooth function $F\colon M \to N$, the pushforward $F_*\colon T_pM \to T_{F(p)}N$ takes the role of the Jacobian matrix. The notion of a regular value therefore makes perfect sense for arbitrary smooth maps.

Definition 5.1 (Regular value)**.** A *regular value* for a smooth map $F\colon M \to N$ is a point $c \in N$ such that for all $p \in M$ with $F(p) = c$, the pushforward $F_*\colon T_pM \to T_cN$ is surjective.

Note that F_* is surjective if and only if its Jacobian matrix $\partial_\mu F^\nu_{\alpha\beta}$ with respect to charts on M and N is surjective as a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$. (Or, equivalently, if $\partial_\mu F^\nu_{\alpha\beta}$ satisfies any of the 3 equivalent properties in Proposition 2.36.) Just like for smooth functions on $\mathbb{R}^n$, the preimage $\Sigma := F^{-1}(\{c\})$ of a regular value is an embedded submanifold.

Theorem 5.2 (Regular level sets)**.** *If c is a regular value of a smooth function $F\colon M \to N$, then the level set $\Sigma := F^{-1}(\{c\})$ is a closed, embedded submanifold of M. If M is of dimension n and N is of dimension $(n-k)$, then Σ is of dimension k.*

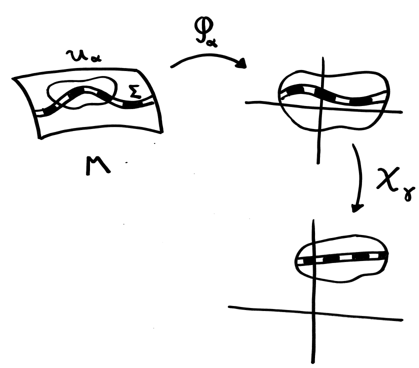


Figure 11: The composition of a chart with a slice chart is a slice chart.

Proof. This is a straightforward consequence of Theorem 2.44. If $c \in N$ is a regular value and $p \in F^{-1}(\{c\})$, choose coordinate charts (U_α, ϕ_α) in M around p, and (V_β, ψ_β) in N around c. Since $\psi_\beta(c)$ is a regular value for

$F_{\alpha\beta}\colon \phi_\alpha(U_\alpha) \to \mathbb{R}^m$ (why?), Theorem 2.44 yields a slice chart (W_γ, χ_γ) for the subset $F^{-1}_{\alpha\beta}(\psi_\beta(c)) \subseteq \phi_\alpha(U_\alpha)$. The composition $\chi_\gamma \circ \phi_\alpha \colon \phi_\alpha^{-1}(W_\gamma) \to \chi_\gamma(W_\gamma)$ is then a slice chart for Σ around $p \in M$. $\square$

In the same way, many interesting results for smooth functions on $\mathbb{R}^n$ can be 'ported' to the setting of manifolds by formulating them in terms of pushforwards (rather than Jacobi matrices) and choosing charts. A case in point is the inverse function theorem for manifolds.

Theorem 5.3 (Inverse Function Theorem)**.** *Let $F\colon M \to N$ be a smooth map, and suppose that for some $p \in M$, the pushforward $F_*\colon T_pM \to T_{F(p)}N$ is a linear isomorphism. Then there exists an open neighbourhood $U \subseteq M$ of p such that $F|_U\colon U \to F(U)$ is a diffeomorphism.*

Proof. Choose coordinates (U_α, ϕ_α) around $p \in M$ and (V_β, ψ_β) around $F(p) \in N$. Since $F_*\colon T_pM \to T_{F(p)}N$ is an isomorphism at p, the Jacobian matrix $\partial_\mu F^\nu_{\alpha\beta}$ of the coordinate representation of F is invertible at $\phi_\alpha(p)$. By the 'ordinary' inverse function theorem 2.45 applied to $F_{\alpha\beta}\colon \phi_\alpha(U_\alpha) \to \mathbb{R}^m$, we conclude that $F_{\alpha\beta}$ is a diffeomorphism when restricted to an open neighbourhood U of $\phi_\alpha(p)$. If follows that F is a diffeomorphism when restricted to $\phi_\alpha^{-1}(U) \subseteq M$. $\square$

From this, we obtain a handy tool to check whether a smooth bijection is a diffeomorphism.

Corollary 5.4. *If $F\colon M \to N$ is bijective, smooth, and if $F_*\colon T_pM \to T_{F(p)}N$ is a linear isomorphism for each $p \in M$, then F is a diffeomorphism.*

Problem 5.5. The function $F\colon \mathbb{R} \to \mathbb{R}$ with $F(x) = x^3$ is a smooth bijection, but not a diffeomorphism.

Problem 5.6. Let $\mathbb{S}^n \subseteq \mathbb{R}^{n+1}$ be the sphere of radius 1. Let $F\colon \mathbb{S}^n \to \mathbb{R}$ be the projection $F(x^0, \ldots, x^n) = x^0$.

a) Every $c \neq \pm 1$ is a regular value of F.

b) The $n-1$ dimensional sphere $\{(x^0, \ldots, x^n) \in \mathbb{S}^n\,;\, x^0 = c\}$ of radius $\sqrt{1-c^2} > 0$ is a closed, embedded submanifold of dimension $n-1$.

c) Is the preimage $F^{-1}(\{c\})$ an embedded submanifold if $c = \pm 1$? If so, what is its dimension?

In Section 3, we defined the tangent bundle of an embedded submanifold Σ of $\mathbb{R}^n$ to be the set of all n-component vectors $(v^1, \ldots, v^n)$ based at $p \in \Sigma \subseteq \mathbb{R}^n$, and pointing in a direction tangent to Σ (Definition 3.1). The following result shows that this is compatible with the definition of the tangent bundle for general manifolds that we gave in Section 4.6.

Proposition 5.7. *Let $\Sigma \subseteq M$ be an embedded submanifold, and let $\iota\colon \Sigma \hookrightarrow M$ be the canonical inclusion. Then $\iota_*(T\Sigma) \subseteq TM$ is an embedded submanifold, and $\iota_* : T\Sigma \hookrightarrow TM$ is a diffeomorphism onto its image.*

Proof. With respect to a slice chart (U_α, ϕ_α) for M and its restriction (V_β, ψ_β) to Σ, the inclusion $\iota\colon \Sigma \to M$ has the simple coordinate representation

$$\iota_{\beta\alpha}(x^1, \ldots, x^k) = (x^1, \ldots, x^k; 0, \ldots, 0).$$

With respect to the charts $(TU_\alpha, \phi_{\alpha*})$ for TM and $(TV_\beta, \psi_{\beta_*})$ for $T\Sigma$, the pushforward ι_* reads

$$(\iota_*)_{\beta_*\alpha_*}(x^1, \ldots, x^k; v^1, \ldots, v^k) = (x^1, \ldots, x^k, 0, \ldots, 0; v^1, \ldots, v^k, 0, \ldots, 0)$$

Up to reordering of indices, this is again a slice chart, so $\iota_*(T\Sigma)$ is an embedded submanifold of TM. Since ι is injective and $\iota_*\colon T_p\Sigma \to T_pM$ is injective for every $p \in M$ by the above coordinate representation, we conclude that $\iota_*\colon T\Sigma \to TM$ is injective. The coordinate representation shows that $\iota_*\colon TV_\beta \to TU_\alpha$ is a diffeomorphism onto its image. $\square$

Remark 5.8 (Whitney Embedding Theorem)**.** In fact, the *Whitney embedding theorem* ([L03, Theorems 6.15 and 6.19]) states that *any* manifold of dimension n admits an embedding into $\mathbb{R}^{2n}$.

Problem 5.9 (Elliptic curves)**.** An *elliptic curve* is a (complex) manifold of the form $\Sigma = \{(x, y) \in \mathbb{C}^2\,;\, y^2 = x^3 + ax + b\}$.

a) Show that if $4a^3 + 27b^2 \neq 0$, then Σ is an embedded submanifold of $\mathbb{C}^2$.

b) Let $\widetilde{\Sigma} = \{[x, y, z] \in \mathbb{C}\mathrm{P}^2\,;\, y^2z = x^3 + axz^2 + bz^3\}$. Show that the inclusion $\iota\colon \Sigma \to \widetilde{\Sigma}$ defined by $\iota(x, y) = [x, y, 1]$ is injective, and that the complement $\widetilde{\Sigma} \setminus \iota(\Sigma)$ of its image consists of a single point 'at infinity'.

c) Show that if $4a^3 + 27b^2 \neq 0$, then $\widetilde{\Sigma}$ is an embedded submanifold of $\mathbb{C}\mathrm{P}^2$.
Hint: every point $[x, y, z] \in \widetilde{\Sigma}$ with $z \neq 0$ admits a slice chart by part (a), so it remains to verify this for the single point 'at infinity'.

d) Conclude that $\widetilde{\Sigma}$ is a smooth one-point compactification of Σ. (In fact, one can use a complex version of the regular value theorem to show that $\widetilde{\Sigma}$ is a complex manifold.)

Problem 5.10 (Automatic smoothness of inversion)**.** For a Lie group G, the multiplication $\mu\colon G \times G \to G$ is smooth by definition. Prove that the inversion $\iota\colon G \to G$ with $\iota(g) = g^{-1}$ is automatically smooth as well.

a) Let $\lambda_g\colon G \to G$ be left multiplication by g, $\lambda_g(h) = gh$. Similarly, let $\rho_h(g) = gh$ be right multiplication by h. Show that $\lambda_{g*}\colon T_hG \to T_{gh}G$ and $\rho_{h*}\colon T_gG \to T_{gh}G$ are linear isomorphisms.

b) Show that $\mu_*\colon T_{(g,h)}(G \times G) \simeq T_gG \times T_hG \to T_{gh}G$ is given by $\mu_*(v, w) = \rho_{h*}(v) + \lambda_{g*}(w)$.

c) Show that the map $\phi\colon G \times G \to G \times G$ defined by $\phi(g, h) = (g, gh)$ is a diffeomorphism.
Hint: use the inverse function theorem.

d) Prove that the inverse $\iota(g) = g^{-1}$ is a smooth map from G to G.

Problem 5.11 (Transversal submanifolds)**.** Two embedded submanifolds Σ_1 and Σ_2 of M are called *transversal* if $T_p\Sigma_1 + T_p\Sigma_2 = T_pM$ for all $p \in \Sigma_1 \cap \Sigma_2$. Show that the intersection $\Sigma_1 \cap \Sigma_2$ of transversal embedded submanifolds is an embedded submanifold. *Hint: you can assume without loss of generality that $M = \mathbb{R}^n$, and that Σ_1 and Σ_2 are regular level sets for the smooth functions $F_1 \colon \mathbb{R}^n \to \mathbb{R}^{n-k_1}$ and $F_2 \colon \mathbb{R}^n \to \mathbb{R}^{n-k_2}$.*

5.2 Submersions and the Hopf fibration

A *submersion* is a smooth map $F \colon M \to N$ such that $F_* \colon T_pM \to T_{F(p)}N$ is surjective for all $p \in M$. If F itself is surjective, this is equivalent with surjectivity of the pushforward $F_* \colon TM \to TN$. Since every value $c \in N$ is a regular value, every fibre $\Sigma_c := F^{-1}(\{c\}) \subseteq M$ is an embedded submanifold.

For example, let $F \colon \mathbb{R}^3 \to \mathbb{R}$ be defined by $F(x, y, z) = z$. This map is surjective, and since $D_pF = (0, 0, 1)$ is nonzero for all points $p = (x, y, z)$, it is a submersion. The fibres are precisely the 2-dimensional planes $F^{-1}(\{c\}) = \{(x, y, z) \in \mathbb{R}^3\,;\, z = c\}$.

Problem 5.12. Let $F \colon \mathbb{S}^1 \times \mathbb{S}^1 \to \mathbb{S}^1$ be defined by $F(x, y) = x$. Show that F is a surjective submersion, and describe the fibres. (A drawing is helpful.)

More generally, if M_1 and M_2 are smooth manifolds, then the projection $F \colon M_1 \times M_2 \to M_1$ with $F(p_1, p_2) = p_1$ is a surjective submersion. By the regular value theorem, every surjective submersion is *locally* of this form. However, surjective submersions can behave quite differently on a *global* level.

Proposition 5.13. *The canonical projection $\pi \colon \mathbb{C}^{n+1} \setminus \{0\} \to \mathbb{C}\mathrm{P}^n$, defined by $\pi(v) = [v]$, is a surjective submersion.*

Proof. With respect to the coordinates (U_α, ϕ_α) on $\mathbb{C}\mathrm{P}^n$, the projection is represented by $\pi_\alpha(v_0, \ldots, v_n) = \phi_\alpha([v])$, so that

$$\pi_\alpha(v_0, \ldots, v_n) = \left(\frac{v_0}{v_\alpha}, \ldots, \frac{v_{\alpha-1}}{v_\alpha}, \frac{v_{\alpha+1}}{v_\alpha}, \ldots, \frac{v_n}{v_\alpha}\right)$$

on its domain $\pi^{-1}(U_\alpha) = \{(v_0, \ldots, v_n) \in \mathbb{C}^{n+1}\,;\, v_\alpha \neq 0\}$. This map is holomorphic, and hence smooth. To show that $\pi_* \colon T_v\mathbb{C}^{n+1} \to T_{[v]}\mathbb{C}\mathrm{P}^n$ is surjective, assume without loss of generality that $\alpha = n$, so the last coordinate v_n is nonzero. Then

$$D_v\pi_\alpha = \frac{1}{v_n}\begin{pmatrix} 1 & 0 & \ldots & 0 & -v_0/v_n \\ 0 & \ddots & & & -v_1/v_n \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \ldots & 0 & 1 & -v_{n-1}/v_n \end{pmatrix}$$

is surjective as a linear map $D_v\pi_\alpha \colon \mathbb{C}^{n+1} \to \mathbb{C}^n$. If we identify $\mathbb{C}^{n+1}$ with $\mathbb{R}^{2(n+1)}$ and $\mathbb{C}^n$ with $\mathbb{R}^{2n}$, then the real $2n \times 2(n+1)$ matrix corresponding to $D_v\pi_\alpha$ is still surjective. $\square$

Problem 5.14. If $F\colon M \to N$ is a surjective submersion from an n-dimensional manifold M to an $(n-k)$-dimensional manifold N, then M is the disjoint union of k-dimensional submanifolds $F^{-1}(\{c\})$ indexed by $c \in N$.

5.2.1 The Hopf fibration

An interesting example of a map which is *almost*, but not quite, a submersion is the map $F\colon \mathbb{S}^2 \to \mathbb{R}$ with $F(x, y, z) = z$. In Problem 5.6 we saw that every $c \neq \pm 1$ is a regular value, but it turns out that the values $c = \pm 1$ are not regular.

Problem 5.15. Show that for the north pole $n := (0, 0, 1)$ in $\mathbb{S}^2$, the pushforward $F_*\colon T_n\mathbb{S}^2 \to T_1\mathbb{R}$ is identically zero. Conclude that 1 is not a regular value of F.

If we remove the north pole $n := (0, 0, 1)$ and the south pole $s := (0, 0, -1)$ from $\mathbb{S}^2$, then the remainder $\mathbb{S}^2 \setminus \{n, s\}$ is a disjoint union of embedded circles $\{(x, y, z) \in \mathbb{S}^2\,;\, z = c\}$, with $|c| < 1$. One may wonder whether it is possible to cover the *entire* 2-sphere $\mathbb{S}^2$ by circles which can be parameterized in a smooth way, without 'leftover points'.

It turns out that the answer is no; one can prove that every smooth map $F\colon \mathbb{S}^2 \to N$ to a 1-dimensional manifold has a singular point $p \in \mathbb{S}^2$ where the pushforward $F_*\colon T_p\mathbb{S}^2 \to T_{F(p)}N$ vanishes. Perhaps surprisingly, the situation for the 3-dimensional sphere $\mathbb{S}^3$ is quite different:

Theorem 5.16 (Hopf fibration)**.** *There exists a surjective submersion*

$$F\colon \mathbb{S}^3 \to \mathbb{S}^2$$

such that every fibre $F^{-1}(\{c\}) \subseteq \mathbb{S}^3$ is diffeomorphic to $\mathbb{S}^1$.

In other words, the 3-sphere $\mathbb{S}^3$ *can* be covered by embedded circles, and, moreover, these circles can be smoothly parameterized by points on the 2-sphere $\mathbb{S}^2$. The surjective submersion $F\colon \mathbb{S}^3 \to \mathbb{S}^2$ is called the *Hopf fibration.*

5.2.2 The Hopf fibration in terms of $\mathbb{C}\mathrm{P}^1$.

To construct the Hopf fibration $F\colon \mathbb{S}^3 \to \mathbb{S}^2$, the basic ingredient we need is the canonical projection from $\mathbb{C}^2 \setminus \{0\}$ to $\mathbb{C}\mathrm{P}^1 \simeq \mathbb{S}^2$.

Note that $\mathbb{S}^3$ is an embedded submanifold of $\mathbb{C}^2 \simeq \mathbb{R}^4$. Indeed, we can consider $\mathbb{S}^3$ as the set of all $(v_0, v_1) \in \mathbb{C}^2$ for which

$$|v_0|^2 + |v_1|^2 = 1.$$

Writing $v_0 = x_0 + iy_0$ and $v_1 = x_1 + iy_1$, this is equivalent to the defining relation

$$x_0^2 + x_1^2 + y_0^2 + y_1^2 = 1$$

for the 3-sphere.

The Hopf fibration $F\colon \mathbb{S}^3 \to \mathbb{S}^2$ is just the restriction of the canonical projection $\pi\colon \mathbb{C}^2 \setminus \{0\} \to \mathbb{C}\mathrm{P}^1$ to the 3-sphere $\mathbb{S}^3 \subseteq \mathbb{C}^2 \setminus \{0\}$. The fibres $F^{-1}([v]) \subseteq \mathbb{S}^3$ of this map are circles of the form

$$F^{-1}([v_0, v_1]) = \{e^{i\phi}(v_0, v_1)/\sqrt{|v_0|^2 + |v_1|^2} \in \mathbb{C}^2\,;\, \phi \in \mathbb{R}\}. \tag{44}$$

To show that these circles are embedded submanifolds, we prove that F is a surjective submersion.

Since $\mathbb{S}^3 \subseteq \mathbb{C}^4$ is embedded, we can realize $T_p\mathbb{S}^3$ inside $T_p\mathbb{C}^2 \simeq \mathbb{C}^2 \simeq \mathbb{R}^4$ as

$$T\mathbb{S}^3 \simeq \{v \in \mathbb{R}^4\,;\, v \perp p\}$$

(cf. §3). We can thus write the 4-dimensional vector space $T_p\mathbb{C}^2$ as a direct sum

$$T_p\mathbb{C}^2 = T_p\mathbb{S}^3 \oplus N_p\mathbb{S}^3$$

of the 3 dimensional tangent space $T_p\mathbb{S}^3$ and the 1-dimensional vector space

$$N_p\mathbb{S}^3 := \{v \in \mathbb{R}^3\,;\, v = \lambda p \text{ for some } \lambda \in \mathbb{R}\},$$

called the *normal* bundle at $p \in \mathbb{S}^3$.

Recall that for all $p \in \mathbb{S}^3$, the pushforward $\pi_*\colon T_p\mathbb{C}^2 \to T_{[p]}\mathbb{C}\mathrm{P}^1$ is surjective. Writing $v \in N_p\mathbb{S}^3$ as $v = \dot\gamma(0)$ for the curve $\gamma(t) = p + \lambda tp$ in $\mathbb{C}^2 \setminus \{0\}$, we have $\pi_*(v) = \frac{d}{dt}|_0\pi((1+\lambda t)p) = \frac{d}{dt}[p] = 0$. We conclude that $\pi_*|_{N_p\mathbb{S}^3} = \{0\}$, and hence that $\pi_*|_{T_p\mathbb{S}^3}$ is surjective. Since this is precisely $F_* = \pi_*|_{T_p\mathbb{S}^3}$, this shows that F is a surjective submersion. This concludes the proof of Theorem 5.16

Problem 5.17. Show that the fibres of the Hopf fibration are given by (44). Why is this independent of the representative $(v_0, v_1) \in \mathbb{C}^2$ of $[v_0, v_1] \in \mathbb{C}\mathrm{P}^1$?

5.2.3 Decompositions of the 3-sphere

The 2-sphere $\mathbb{S}^2$ is the disjoint union of the northern and southern hemisphere

$$\mathbb{S}^2_N := \{(x,y,z) \in \mathbb{S}^2\,;\, z > 0\} \quad \text{and} \quad \mathbb{S}^2_S := \{(x,y,z) \in \mathbb{S}^2\,;\, z < 0\},$$

and the equator $\overline{\mathbb{S}}^2_N \cap \overline{\mathbb{S}}^2_S = \{(x,y,z) \in \mathbb{S}^2\,;\, z = 0\}$. In view of the following exercise, we can think of $\mathbb{S}^2$ as two copies of the disk

$$D^2 = \{(x,y) \in \mathbb{R}^2\,;\, x^2 + y^2 < 1\},$$

glued along their boundary $\mathbb{S}^1 = \{(x,y) \in \mathbb{R}^2\,;\, x^2 + y^2 = 1\}$.

Problem 5.18. Show that both hemispheres $\mathbb{S}^2_N$ and $\mathbb{S}^2_S$ are diffeomorphic to a disk D^2, and that the equator is diffeomorphic to $\mathbb{S}^1$.

In the same vein, the 3-sphere $\mathbb{S}^3$ consists of two copies of the solid ball

$$B^3 = \{(x,y,z) \in \mathbb{R}^3\,;\, x^2 + y^2 + z^2 < 1\},$$

glued along their boundary $\mathbb{S}^2 = \{(x,y,z) \in \mathbb{R}^3\,;\, x^2 + y^2 + z^2 = 1\}$.

Problem 5.19. Show that both hemispheres

$$\mathbb{S}^3_N := \{(x,y,z,t) \in \mathbb{S}^3\,;\, t > 0\} \quad \text{and} \quad \mathbb{S}^3_S := \{(x,y,z,t) \in \mathbb{S}^3\,;\, t < 0\}$$

are diffeomorphic to the solid ball B_3, and that their common boundary $\overline{\mathbb{S}}^3_N \cap \overline{\mathbb{S}}^3_S = \{(x,y,z,t) \in \mathbb{S}^3\,;\, t = 0\}$ is diffeomorphic to $\mathbb{S}^2$.

Problem 5.20. Formulate and prove a version of Problem 5.19 for the n-sphere $\mathbb{S}^n$.

In the following problem, we show that $\mathbb{S}^3$ is also the union of two solid *tori* which are glued together along their boundary. The solid tori are the preimages under the Hopf fibration $F \colon \mathbb{S}^3 \to \mathbb{S}^2$ of the northern and southern hemisphere $\overline{\mathbb{S}}^2_N$ and $\overline{\mathbb{S}}^2_S$ of the 2-sphere $\mathbb{S}^2$.

Problem 5.21. Show that $\mathbb{S}^3$ is the union of two solid tori, glued together along their boundary.

a) Show that the Hopf fibration $F \colon \mathbb{S}^3 \to \mathbb{C}\mathrm{P}^1$ maps the open subset

$$U := \{(v_0, v_1) \in \mathbb{C}^2\,;\, |v_0|^2 + |v_1|^2 = 1, v_0 \neq 0\} \subseteq \mathbb{S}^3$$

into the domain $U_0 \subseteq \mathbb{C}\mathrm{P}^1$ of the coordinate map $\phi_0([v_0, v_1]) = v_1/v_0$. Show that $\phi_0 \circ F \colon U \to \mathbb{C}$ is given by $(v_0, v_1) \mapsto v_1/v_0$.

b) The preimage under $\phi_0 \circ F$ of the disc $D^2 := \{z \in \mathbb{C}\,;\, |z| < 1\}$ is the solid torus

$$T_+ := \{(\sqrt{1-\rho^2}e^{i\theta}, \rho e^{i\phi})\,;\, \phi, \theta \in \mathbb{R}, \rho \in [0, \tfrac{1}{2}\sqrt{2}\,)\}.$$

c) Exchange the roles of v_0 and v_1 in the above, and show that the preimage of the disk D^2 under $\phi_1 \circ F$ is the solid torus

$$T_- := \{(\rho e^{i\phi}, \sqrt{1-\rho^2}e^{i\theta})\,;\, \phi, \theta \in \mathbb{R}, \rho \in [0, \tfrac{1}{2}\sqrt{2})\}.$$

d) Conclude that $\mathbb{S}^3$ is the union of two solid tori T_+ and T_-, whose common boundary

$$\mathbb{T}^2 = \{(\tfrac{1}{2}\sqrt{2}e^{i\phi}, \tfrac{1}{2}\sqrt{2}e^{i\theta})\,;\, \phi, \theta \in \mathbb{R}\}.$$

is the 2-torus arising as the preimage of the equator $\mathbb{S}^1 \simeq \overline{\mathbb{S}}^2_N \cap \overline{\mathbb{S}}^2_S \subseteq \mathbb{S}^2$ under the Hopf fibration.

Apparently, gluing two solid tori along their boundary $\mathbb{T}^2$ yields the same result as gluing two solid balls along their boundary $\mathbb{S}^2$! Either way, you obtain the 3-sphere $\mathbb{S}^3$.

6 Vector fields, Lie brackets and flows

In §3.2, we defined vector fields for embedded submanifolds Σ of $\mathbb{R}^n$. Now that we defined the tangent bundle TM for a general smooth manifold M (cf. § 4.6), we can define vector fields om M in much the same way as before.

Definition 6.1 (Vector fields)**.** A *vector field* on M is a smooth map

$$v \colon M \to TM$$

such that $\pi \circ v = \mathrm{Id}_M$. We denote the space of vector fields on M by $\mathrm{Vec}(M)$.

A vector field v assigns to every point p in M an element v_p of the tangent bundle TM. The requirement that $\pi \circ v = \mathrm{Id}_M$ means that v_p is in fact an element of the tangent space T_pM *at the point* p.

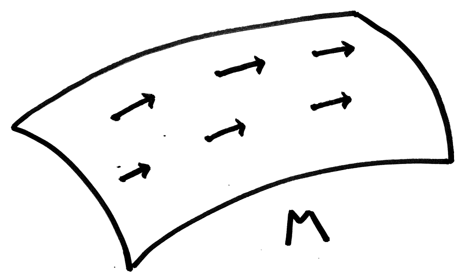


Figure 12: A vector field assigns to each $p \in M$ a tangent vector at the point p

Remark 6.2 (Vector fields are sections)**.** A *section* of a surjective map $\pi \colon X \to Y$ is a map $\sigma \colon Y \to X$ such that $\pi \circ \sigma = \mathrm{Id}_Y$. A vector field is thus a smooth section of the canonical projection $\pi \colon TM \to M$ associated to the tangent bundle.

6.1 Three ways of looking at vector fields

Just like for tangent vectors, there are essentially three equivalent ways to describe vector fields. We can view vector fields as smooth sections of the canonical projection $\pi \colon TM \to M$, as we did above. Further, we have a *local* description using coordinates, and an *algebraic* description using derivations.

6.1.1 Vector fields on $\mathbb{R}^n$

Vector fields are particularly easy to describe if M is an open subset U of $\mathbb{R}^n$. In that case the tangent bundle is $TU = U \times \mathbb{R}^n$, and the projection $\pi \colon TU \to U$ is simply the projection onto the first factor,

$$\pi(x^1, \ldots, x^n; v^1, \ldots, v^n) = (x^1, \ldots, x^n).$$

In this case, the requirement $\pi \circ v = \mathrm{Id}_M$ simply means that

$$v_x = (x^1, \ldots, x^n; v^1(x), \ldots, v^n(x))$$

for some n-tuple of functions $v^\mu \colon U \to \mathbb{R}$. Since $v \colon U \to TU$ is smooth, the functions v^μ are smooth as well.

Note that the real numbers $v^\mu(x)$ are the coefficients of the tangent vector $v_x \in T_xU$ with respect to the standard basis ∂_μ of T_xU, obtained from the global chart, $v(x) = v^\mu(x)\partial_\mu$. We therefore call the functions v^μ the *coefficients* of the vector field v with respect to the canonical global chart.

6.1.2 Vector fields in local coordinates

Now let M be *any* smooth manifold. With respect to the local coordinates (U_α, ϕ_α) on M and $(TU_\alpha, \phi_{\alpha*})$ on TM, the vector field $v \colon M \to TM$ is represented by

$$v_\alpha(x) = (x^1, \ldots, x^n; v^1_\alpha(x), \ldots, v^n_\alpha(x)).$$

The n-tuple of smooth functions $v^\mu \colon \phi_\alpha(U_\alpha) \to \mathbb{R}$ form the *coordinate representation* of v with respect to the chart (U_α, ϕ_α).

If (U_β, ϕ_β) is a different set of coordinates, then on the overlap $U_\alpha \cap U_\beta$, the coordinate representation $v^{\overline{\mu}}$ of v with respect to (U_β, ϕ_β) is related to v^μ by

$$v^{\overline{\mu}}(x^{\overline{1}}, \ldots, x^{\overline{n}}) = \left(\frac{\partial x^{\overline{\mu}}}{\partial x^\mu}\right) v^\mu(x^1, \ldots, x^n). \tag{45}$$

Here the Einstein summation convention implies a sum over the repeated index μ on the right hand side. The coordinates x^σ with respect to (U_α, ϕ_α) should be considered as smooth functions of the coordinates $x^{\overline{\sigma}}$ with respect to (U_β, ϕ_β).

Since the *vector* $v_p \in T_pM$ can be expressed in the coordinate basis ∂_μ of T_pM as $v_p = v^\mu_\alpha(x)\partial_\mu$, we will often denote the *vector field* by

$$v = v^\mu_\alpha \partial_\mu.$$

Here, by a slight abuse of notation, ∂_μ denotes the vector field on U_α that assigns to every point $p \in U_\alpha$ the coordinate vector $\partial_\mu \in T_pM$.

Problem 6.3. Let $F \colon \mathbb{R}^2 \to \mathbb{R}^2$ be the smooth function

$$F(x, y) = \Big((x^2 + y^2)x, (x^2 + y^2)y\Big),$$

and let $v \in \mathrm{Vec}(\mathbb{R}^2)$ be the smooth vector field $v = x\partial_x + y\partial_y$. Show that $F_*(v_p) = Cv_{F(p)}$ for some constant C, and calculate this constant.

If $v \colon M \to TM$ is a vector field and $f \colon M \to \mathbb{R}$ is a smooth function, then we define the product $fv \colon M \to TM$ by $(fv)_p := f(p)v_p$. Since this is again a vector field, the space $\mathrm{Vec}(M)$ of vector fields is a *module* over the smooth functions. This means that the product $(f, v) \mapsto fv$ is linear in f as well as v, that $f(gv) = (fg)v$ for all f and g in $C^\infty(M)$, and that $1v = v$.

Problem 6.4 ($\mathrm{Vec}(M)$ as a module over $C^\infty(M)$)**.** Show that the product $fv \colon M \to TM$ of f and v is smooth, and that $\mathrm{Vec}(M)$ is a module over $C^\infty(M)$.

6.1.3 Derivations

So far, we have seen two ways of looking at vector fields. Using the definition, we can view a vector field as a smooth function $v\colon M \to TM$ that assigns to every point p in M a tangent vector v_p in the tangent space T_pM at p. But in coordinates, we can also describe a vector field *locally* by an n-tuple of smooth functions v^μ, where different coordinate representations are related by the Jacobian matrix of the coordinate transformation (45).

Just like for tangent vectors, there is a third, more algebraic way of describing vector fields. It uses derivations $D\colon C^\infty(M) \to C^\infty(M)$, which are defined analogously to the derivations $D\colon C^\infty(M) \to \mathbb{R}$ of Definition 4.9.

Definition 6.5 (Derivations of $C^\infty(M)$)**.** A *derivation* of $C^\infty(M)$ is a linear operator $D\colon C^\infty(M) \to C^\infty(M)$ that satisfies the *Leibniz rule*

$$D(fg) \quad = \quad D(f)g + fD(g) \tag{46}$$

for all $f, g \in C^\infty(M)$. The space of derivations of the algebra $C^\infty(M)$ is denoted by $\mathrm{Der}(C^\infty(M))$.

Every vector field $v \in \mathrm{Vec}(M)$ gives rise to the *Lie derivative*

$$\mathcal{L}_v\colon C^\infty(M) \to C^\infty(M),$$

defined by

$$\mathcal{L}_v(f)(p) := v_p(f). \tag{47}$$

Note that if $f\colon M \to \mathbb{R}$ is smooth, then the function $\mathcal{L}_v(f)$, which maps p to $v_p(f)$, is smooth as well. Indeed, in local coordinates it maps x to $v^\mu_\alpha(x)\partial_\mu f_\alpha(x)$, which depends smoothly on x because the coordinate representation $v^\mu(x)$ of v does so.

To see that $\mathcal{L}_v$ is a derivation from $C^\infty(M)$ to $C^\infty(M)$, we need to check that $\mathcal{L}_v(fg) = f\mathcal{L}_v(g) + g\mathcal{L}_v(f)$. For this, evaluate the above expression in $p \in M$ and use that every value v_p of the vector field v is a derivation from $C^\infty(M)$ to $\mathbb{R}$,

$$\begin{aligned}
\mathcal{L}_v(fg)(p) &= v_p(fg) \\
&= f(p)v_p(g) + g(p)v_p(f) \\
&= (f\mathcal{L}_v(g))(p) + (g\mathcal{L}_v(f))(p).
\end{aligned}$$

Problem 6.6 (Derivations as a vector space)**.** Let $D, E\colon C^\infty(M) \to C^\infty(M)$ be derivations, and let $\lambda \in \mathbb{R}$. Then

a) λD is a derivation

b) $D + E$ is a derivation.

Problem 6.7 (Derivations as a module over $C^\infty(M)$)**.** Let D be a derivation from $C^\infty(M)$ to $C^\infty(M)$, and let $f \in C^\infty(M)$ be a smooth function. Then the operator $fD\colon C^\infty(M) \to C^\infty(M)$ defined by $(fD)(g) := fD(g)$ is again a derivation.

In fact, it turns out that *every* derivation $D: C^\infty(M) \to C^\infty(M)$ is of the form $\mathcal{L}_v$ for some vector field $v \in \mathrm{Vec}(M)$.

Theorem 6.8 (Algebraic characterization of vector fields)**.** *Every derivation $D: C^\infty(M) \to C^\infty(M)$ is of the form $D(f) = \mathcal{L}_v(f)$ for some vector field $v \in \mathrm{Vec}(M)$.*

Proof. We prove this in §7. □

In view of the above result, one often *identifies* a vector field on M with a derivation of $C^\infty(M)$. Rather than $\mathcal{L}_v(f)$, one then simply writes $v(f)$.

6.1.4 The three descriptions of vector fields

Summarizing, we now have three equivalent ways of describing vector fields.

(1) As a smooth map $v: M \to TM$ such that $v(p) \in T_pM$ for all $p \in M$.

(2) The restriction of v to a coordinate neighbourhood $U_\alpha \subseteq M$ can be described as $v = v_\alpha^\mu \partial_\mu$, using the n-tuple of smooth functions v_α^μ. The coordinate representation $v_\beta{}^{\overline{\mu}}$ with respect to a different coordinate system (U_β, ϕ_β) is related to v_α^μ by $v_\beta^{\overline{\mu}} = \frac{\partial x^{\overline{\mu}}}{\partial x^\mu} v_\alpha^\mu$ on the overlap $U_\alpha \cap U_\beta$.

(3) A vector field v is uniquely determined by the derivation $\mathcal{L}_v$, and every derivation is of this form.

Each of these descriptions has its advantages and disadvantages, and it depends on the context which one is more convenient to use.

6.2 The Lie bracket

For two derivations $D, E: C^\infty(M) \to C^\infty(M)$, we define their *Lie bracket* by

$$[D, E] := D \circ E - E \circ D. \tag{48}$$

We say that D and E *commute* if $[D, E] = 0$.

Proposition 6.9. *The Lie bracket $[D, E]$ of two derivations is a derivation.*

Proof. Since $D: C^\infty(M) \to C^\infty(M)$ and $E: C^\infty(M) \to C^\infty(M)$ are linear, their concatenations $D \circ E$ and $E \circ D$ are linear. It follows that $D \circ E - E \circ D$ is linear as well. The main thing to show is that $[D, E] := D \circ E - E \circ D$ satisfies the Leibniz rule

$$[D, E](fg) = f([D, E](g)) + ([D, E](f))g. \tag{49}$$

Using the Leibniz rule for D and E to expand (49), we find

$$\begin{aligned}
[D, E](fg) &= D(E(fg)) - E(D(fg)) \\
&= D(E(f)g + fE(g)) - E(D(f)g + fD(g)) \\
&= D(Ef)g + fD(Eg) - E(D(f))g - fE(D(g)) \\
&\quad + E(f)D(g) + D(f)E(g) - D(f)E(g) - E(f)D(g) \\
&= [D, E](f)g + f[D, E](g),
\end{aligned}$$

as required. Note that $D \circ E$ and $E \circ D$ separately do *not* satisfy the Leibniz rule! The fact that $[D, E]$ is a derivation hinges on the cancellation in the fourth line, which is due to the minus sign in the definition of the Lie bracket. $\square$

6.2.1 Lie bracket of vector fields

If D and E are given by the Lie derivative along the vector fields v and w, then the derivation $[D, E]$ is again the Lie derivative along a vector field. If we denote this vector field by $[v, w]$, we have

$$[\mathcal{L}_v, \mathcal{L}_w] = \mathcal{L}_{[v,w]}.$$

In local coordinates, the Lie bracket $[v, w]$ of two vector fields v and w can be calculated as follows.

Proposition 6.10. *If $v = v^\mu \partial_\mu$ and $w = w^\nu \partial_\nu$ with respect to the coordinates (U_α, ϕ_α), then*

$$[v^\mu \partial_\mu, w^\nu \partial_\nu] = (v^\mu \partial_\mu w^\nu - w^\mu \partial_\mu v^\nu)\partial_\nu.$$

Proof. Since

$$v^\mu \partial_\mu (w^\nu \partial_\nu (f)) = (v^\mu \partial_\mu w^\nu)\partial_\nu f + v^\mu w^\nu \partial_\mu \partial_\nu f$$

and

$$w^\nu \partial_\nu (v^\mu \partial_\mu (f)) = (w^\nu \partial_\nu v^\mu)\partial_\mu f + v^\mu w^\nu \partial_\nu \partial_\mu f,$$

and since $\partial_\mu \partial_\nu f = \partial_\nu \partial_\mu f$ for the smooth function f, we find

$$[v^\mu \partial_\mu, w^\nu \partial_\nu](f) = (v^\mu \partial_\mu w^\nu)\partial_\nu f - (w^\nu \partial_\nu v^\mu)\partial_\mu f. \tag{50}$$

Since we sum over μ and ν in both terms, we may exchange the labels μ and ν, so that $(w^\nu \partial_\nu v^\mu)\partial_\mu f = (w^\mu \partial_\mu v^\nu)\partial_\nu f$. To see why this is true, it may help to relabel μ and ν by i and j first, and then relabel i and j by ν and μ in the converse order;

$$\sum_{\nu=1}^{n}\sum_{\mu=1}^{n}(w^\nu \partial_\nu v^\mu)\partial_\mu f = \sum_{j=1}^{n}\sum_{i=1}^{n}(w^j \partial_j v^i)\partial_i f = \sum_{\mu=1}^{n}\sum_{\nu=1}^{n}(w^\mu \partial_\mu v^\nu)\partial_\nu f.$$

From (50), we then find $[v^\mu \partial_\mu, w^\nu \partial_\nu](f) = (v^\mu \partial_\mu w^\nu)\partial_\nu f - (w^\mu \partial_\mu v^\nu)\partial_\nu f$, as required. $\square$

6.2.2 Algebraic properties of the Lie bracket

The Lie bracket satisfies the following algebraic properties.

Proposition 6.11. *The Lie bracket $[\,\cdot\,,\,\cdot\,]\colon \mathrm{Vec}(M) \times \mathrm{Vec}(M) \to \mathrm{Vec}(M)$ is bilinear, skew-symmetric, and it satisfies the Jacobi identity*

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0. \tag{51}$$

Proof. The skew-symmetry $[v, w] = -[w, v]$ is clear from the definition, as is the bilinearity

$$\begin{aligned} [u, av + bw] &= a[u, v] + b[u, w] \\ [av + bw, u] &= a[v, u] + b[w, u]. \end{aligned}$$

For the Jacobi identity, we identify the vector fields u, v and w with the corresponding derivations and note that

$$\begin{aligned} [u, [v, w]] + [v, [w, u]] + [w, [u, v]] &= \quad uvw - uwv - vwu + wvu \\ &\quad +vwu - wvu - wuv + vuw \\ &\quad +wuv - vuw - uvw + uwv = 0 \end{aligned}$$

because the 12 terms cancel pairwise. $\square$

A vector space $\mathfrak{g}$ with a skew-symmetric, bilinear form $[\,\cdot\,;\,\cdot\,]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying the Jacobi identity is called a *Lie algebra*. By Proposition 6.11, $\mathrm{Vec}(M)$ is a Lie algebra.

Problem 6.12. For $M = \mathbb{R}^3$, calculate:

a) $[\partial_x + x\partial_y, x^2\partial_y + x\partial_z]$.

b) $[\partial_x, x^{12}\partial_x + e^x\partial_y]$

Problem 6.13. For $v = x^2y\partial_x + y\partial_y$ and $w = x\partial_x + y^2\partial_y$, calculate:

a) The term $v^\mu\partial_\mu w^\nu$ for $\nu = 1$ and $\nu = 2$

b) The term $w^\mu\partial_\mu v^\nu$ for $\nu = 1$ and $\nu = 2$

c) The Lie bracket $(v^\mu\partial_\mu w^\nu - w^\mu\partial_\mu v^\nu)\partial_\nu$

Problem 6.14. For $M = \mathbb{R}^3$, let $E := x\partial_x + y\partial_y + z\partial_z$ be the *Euler vector field*. Calculate:

a) $[E, x^{37}y^{59}z^4\partial_x]$

b) $[E, x^{11}y^{38}z^{51}\partial_y]$

c) $[E, x^{37}y^{59}z^4\partial_x + x^{11}y^{38}z^{51}\partial_y + x^{63}y^{14}z^{23}\partial_z]$

d) A *homogeneous polynomial* of degree N is a polynomial of the form

$$p(x, y, z) = \sum_{n_1+n_2+n_3=N} a_{n_1,n_2,n_3} x^{n_1} y^{n_2} z^{n_3}.$$

Show that if $p^1(x, y, z)$, $p^2(x, y, z)$ and $p^3(x, y, z)$ are homogeneous polynomials of degree N, then $[E, p^\mu(x, y, z)\partial_\mu] = (N - 1)(p^\mu\partial_\mu)$.

6.3 Flows and the Lie bracket

A vector field v gives rise to an ordinary differential equation (ODE) on the manifold M. If we think of v as the velocity field of a moving fluid or gas, and if we think of $\gamma(t) \in M$ as the the position of a particle drifting along the flow of v, then the curve $\gamma \colon \mathbb{R} \to M$ satisfies the ODE

$$\dot{\gamma}(t) = v_{\gamma(t)}. \tag{52}$$

Together with the boundary condition $\gamma(0) = p$, this specifies the *flow* along the vector field v, starting at $p \in M$. Indeed, the tangent vector $\dot{\gamma}(t)$ to the curve at time t is equal to the vector field v at the position $\gamma(t)$ of the particle.

Remark 6.15. In local coordinates, equation (52) reads

$$\dot{\gamma}^{\mu}_{\alpha}(t) = v^{\mu}_{\alpha}(\gamma^{1}_{\alpha}(t), \ldots, \gamma^{n}_{\alpha}(t)).$$

Since the functions v^{μ}_{α} are smooth, they are certainly Lipschitz on every closed subset of $\phi_{\alpha}(U_{\alpha})$. The local existence and uniqueness results for first order ODE's therefore apply, and yield local existence and uniqueness for (52).

If the flow starts at a different point p', then, of course, we obtain a different curve γ'. We denote by $\phi_t(p)$ the flow along the vector field v starting at $p \in M$. In other words, we define $t \mapsto \phi_t(p)$ to be the solution of the ODE

$$\frac{d}{dt}\phi_t(p) = v_{\phi_t(p)} \tag{53}$$

$$\phi_0(p) = p. \tag{54}$$

One can show that for each $p \in M$, there exists an interval (a, b), possibly depending on p, on which the solution $t \mapsto \phi_t(p)$ is well defined and smooth. A vector field is called *complete* or *integrable* if its integral curves exist for all time $t \in \mathbb{R}$, and for all starting points $p \in M$.

Problem 6.16. Let $M = \mathbb{R}^2$.

a) Show that the vector field $v = x\partial_y - y\partial_x$ is integrable, and determine the flow $\phi_t(x, y)$.

b) Show that the vector field $v = x^2\partial_x - y\partial_y$ is not integrable. For which values of t is the flow $\phi_t(x, y)$ well defined?

For complete vector fields v, one can prove that $(t, p) \mapsto \phi_t(p)$ is a smooth map $\mathbb{R} \times M \to M$, and that for each $t \in \mathbb{R}$, the map $\phi_t \colon M \to M$ is a diffeomorphism with inverse ϕ_{-t}. (See [L03, Chapter 9].)

6.3.1 Transporting a function along the flow

The flow equation (52) yields a first order PDE on $f \in C^{\infty}(M)$ by 'dragging a function along the flow'. Simply apply the left and right hand side of equation (53) to a smooth function $f \in C^{\infty}(M)$ to find that $\frac{d}{dt}f(\phi_t(p)) = v_{\phi_t(p)}(f)$.

Evaluating this at zero, we find

$$\frac{d}{dt}\Big|_{t=0} f(\phi_t(p)) = \mathcal{L}_v(f)(p). \tag{55}$$

It follows that the Lie derivative $\mathcal{L}_v(f)$ of f along the vector field v is just the derivative of f along the flow generated by v.

Remark 6.17. Suppose that $f_t(p)$ represents, say, the temperature of a fluid in $p \in M$ at a time t. If the heat transfer is dominated by convection rather than diffusion, the temperature at p at time 0 will be the same as the temperature at $\phi_t(p)$ at time t. We therefore have $f_t(\phi_t(p)) = f_0(p)$ for all p, and hence $f_t(p) = f_0(\phi_t^{-1}(p))$. Since $\phi_t^{-1}(p) = \phi_{-t}(p)$, the convection equation reads

$$\tfrac{\partial}{\partial t} f_t = -\mathcal{L}_v(f),$$

or $\partial_t f_t(x) = -v^\mu(x)\partial_\mu f(x)$ in local coordinates.

6.3.2 The Lie bracket in terms of flows

If ϕ_t^v and ϕ_t^w are the flows generated by complete vector fields v and w on M, respectively, then

$$\begin{aligned} v_p(f) &= \frac{d}{dt}\Big|_0 f(\phi_t^v(p)), \quad \text{and} \\ w_p(f) &= \frac{d}{ds}\Big|_0 f(\phi_s^w(p)). \end{aligned}$$

It follows that

$$\begin{aligned} [v,w]_p(f) &= v\left(\frac{d}{ds}\Big|_0 f(\phi_s^w(p))\right) - w\left(\frac{d}{dt}\Big|_0 f(\phi_t^v(p))\right) \\ &= \frac{d}{dt}\Big|_0 \frac{d}{ds}\Big|_0 \Big(f\big(\phi_s^w \circ \phi_t^v(p)\big) - f\big(\phi_t^v \circ \phi_s^w(p)\big)\Big). \end{aligned}$$

In other words, the Lie bracket is a measure of the difference between flowing first along v and then along w, or flowing first along w and then along v. In particular, if the flows of v and w commute, $\phi_t^{\partial_x} \circ \phi_s^{\partial_y} = \phi_s^{\partial_y} \circ \phi_t^{\partial_x}$, then $[v,w] = 0$.

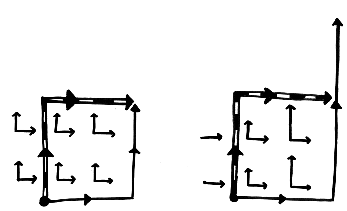


Figure 13: Left: Flow of the commuting vector fields $v = \partial_x$ and $w = \partial_y$. Right: Flow of the noncommuting vector fields $v = \partial_x$ and $w = x\partial_y$.

Problem 6.18. Let $u = x\partial_x + y\partial_y$, $v = x\partial_y - y\partial_x$, and $w = \partial_x$ be vector fields on $\mathbb{R}^2$. The three vector fields u, v, w give rise to the three flows $\phi_t^A(x, y) = (x+t, y)$, $\phi_t^B(x, y) = (xe^t, ye^t)$, and $\phi^C(x, y) = (x\cos(t) - y\sin(t), x\sin(t) + y\cos(t))$. Which vector field belongs to which flow?

Problem 6.19 (Pushforward of vector fields)**.** Let $\phi\colon M \to N$ be a diffeomorphism from M to N.

a) Let $v\colon M \to TM$ be a vector field on M. Show that $v^\phi := \phi_* \circ v \circ \phi^{-1}$ is a vector field on N.

b) Show that $(fv)^\phi = (f \circ \phi^{-1})v^\phi$ for any $f \in C^\infty(M)$.

c) Let $D\colon C^\infty(M) \to C^\infty(M)$ be a derivation on M. Show that the map $D^\phi(f) := D(f \circ \phi) \circ \phi^{-1}$ is a derivation on N.

d) Let $\mathcal{L}_v\colon C^\infty(M) \to C^\infty(M)$ be the Lie derivative along the vector field v. Show that $(\mathcal{L}_v)^\phi = \mathcal{L}_{v^\phi}$.

e) Show that $[D^\phi, E^\phi] = [D, E]^\phi$ for all derivations D, E on M.

f) Show that $[v^\phi, w^\phi] = [v, w]^\phi$ for all vector fields $v, w \in \mathrm{Vec}(M)$.

In the following problem, you can use that if $f\colon M_1 \to M_2$ and $g\colon N_1 \to N_2$ are smooth maps, then $(m, n) \mapsto (f(m), g(n))$ is a smooth map $M_1 \times N_1 \to M_2 \times N_2$ (cf. Problem 2.24).

Problem 6.20. Let G be a Lie group of dimension n.

a) Prove that for each $g \in G$, the left multiplication $\lambda_g(h) := gh$ is a diffeomorphism $\lambda_g\colon G \to G$. What is its inverse?

b) For a vector field $v \in \mathrm{Vec}(G)$, we define $v^{\lambda_g} := (\lambda_g)_* \circ v \circ (\lambda_g)^{-1}$ as in Problem 6.19. Define $\mathfrak{g}$ to be the vector space of all left-invariant vector fields,
$$\mathfrak{g} := \{v \in \mathrm{Vec}(G)\,;\, v^{\lambda_g} = v \text{ for all } g \in G\}.$$
Let $\mathbf{1} \in G$ be the identity in G, and let $\mathrm{ev}_\mathbf{1}\colon \mathfrak{g} \to T_\mathbf{1}G$ be the evaluation at the identity, $\mathrm{ev}_\mathbf{1}(v) := v_\mathbf{1}$. Show that $\mathrm{ev}_\mathbf{1}$ is injective.

c) Show that $\mathrm{ev}_\mathbf{1}$ is surjective.

Hint: Let $\gamma\colon \mathbb{R} \to G$ be a smooth curve in G with $\gamma(0) = \mathbf{1}$. Show that the map $G \times \mathbb{R} \to G$ defined by $(g, t) \mapsto g\gamma(t)$ is then smooth as well. It follows that $v_g := \frac{d}{dt}|_0 g\gamma(t)$ is a smooth map $G \to TG$. Show that v is a vector field, and that $v \in \mathfrak{g}$.

d) What is the dimension of $\mathfrak{g}$? Show that for $v, w \in \mathfrak{g}$, the Lie bracket $[v, w]$ is again an element of $\mathfrak{g}$. The vector space $\mathfrak{g}$ with the bracket $[\,\cdot\,,\,\cdot\,]\colon \mathfrak{g}\times\mathfrak{g} \to \mathfrak{g}$ is called the *Lie algebra* of the Lie group G.

7 Derivations

Every tangent vector $v_p \in T_pM$ gives rise to a derivation $D_p^v \colon C^\infty(M) \to \mathbb{R}$ at $p \in M$ given by $D_p^v(f) := v_\alpha^\mu \partial_\mu(f)$. Similarly, every vector field $v \in \mathrm{Vec}(M)$ gives rise to a derivation $\mathcal{L}_v \colon C^\infty(M) \to C^\infty(M)$ with $\mathcal{L}_v(f)(p) := v_\alpha^\mu(x)\partial_\mu f_\alpha(x)$. In this section, we show that this yields a bijective correspondence: all derivations are of this type.

7.1 Evaluation of derivations

If $D \colon C^\infty(M) \to C^\infty(M)$ is a derivation and p is a point in M, then the operator $D_p \colon C^\infty(M) \to \mathbb{R}$ defined by $D_p(f) := (D(f))(p)$ is a derivation at p. We call D_p the *evaluation* of D at p.

Proposition 7.1. *A derivation $D \colon C^\infty(M) \to C^\infty(M)$ is zero if and only if $D_p = 0$ for all $p \in M$.*

Proof. The 'only if' direction is clear. For the converse direction, suppose that $D_p = 0$ for all $p \in M$, and let $f \in C^\infty(M)$ be a smooth function. Then $D_p(f) := (D(f))(p) = 0$ for all $p \in M$, so $D(f) = 0$. But f was arbitrary, so D is zero. □

Since two derivations D and D' are equal if and only if $D - D' = 0$, it follows that $D = D'$ if and only if $D_p = D_p$ for all $p \in M$. In other words, a derivation $D \colon C^\infty(M) \to C^\infty(M)$ is completely determined by its evaluations $D_p \colon C^\infty(M) \to \mathbb{R}$.

7.2 Vectors and vector fields on $\mathbb{R}^n$

We show that for $M = \mathbb{R}^n$, all derivations at $a \in \mathbb{R}^n$ are derived from tangent vectors, and all derivations of $C^\infty(M)$ from vector fields.

Proposition 7.2. *Every derivation $D_a C^\infty(\mathbb{R}^n) \to \mathbb{R}$ at $a \in \mathbb{R}^n$ is of the form $D_a(f) = v_a^\mu \partial_\mu f(a)$, where v_a^μ is obtained by applying D_a to the coordinate function $x^\mu \in C^\infty(\mathbb{R}^n)$, $v_a^\mu = D_a(x^\mu)$.*

Proof. By the Hadamard Lemma (cf. Problem 4.13), we can write

$$f(x) = f(a) + (x^\mu - a^\mu)\rho_\mu(x) \tag{56}$$

for smooth functions $\rho_\mu \colon \mathbb{R}^n \to \mathbb{R}$ with $\rho_\mu(a) = \frac{\partial}{\partial x^\mu} f(a)$. We apply D_a to f, and consider the two terms in (56) separately. The first term vanishes. Indeed, we have $D_a(1) = 0$ since $D_a(1) = D_a(1 \cdot 1) = D_a(1) \cdot 1 + 1 \cdot D_a(1)$. By linearity, we conclude that D_a vanishes on all constant functions, so in particular $D_a(f(a)) = 0$. Since $x^\mu - a^\mu$ evaluates to zero at a, the second term yields

$$D_a(f) = D_a\big(\rho_\mu(x)(x^\mu - a^\mu)\big) = \rho_\mu(a) D_a(x^\mu - a^\mu) = \partial_\mu f(a) D_a(x^\mu - a^\mu)\,.$$

Since D_a vanishes on the constants a^μ, this yields $D_a(f) = D_a(x^\mu)\partial_\mu f(a)$ as required. □

Since $\mathrm{Der}_a(C^\infty(\mathbb{R}^n))$ is isomorphic to $T_a\mathbb{R}^n$, it is a vector space of dimension n, with basis $\partial_\mu|_a$ defined by $\partial_\mu|_a(f) := \frac{\partial}{\partial x^\mu} f(a)$.

Proposition 7.3. *Every derivation $D\colon C^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ is of the form $D(f) = v^\mu \partial_\mu f$, with $v^\mu \in C^\infty(\mathbb{R}^n)$ given by $v^\mu = D(x^\mu)$.*

Proof. By Proposition 7.2, the derivation $D_a(f) := D(f)(a)$ at the point $a \in \mathbb{R}^n$ is of the form $D_a(f) = v^\mu_a \partial_\mu f(a)$ with $v^\mu_a = D_a(x^\mu) = D(x^\mu)(a)$. Note that $v^\mu \colon a \mapsto v^\mu_a$ is a smooth function, since it is the image under D of the smooth function x^μ. Since $D(f)(a) = v^\mu_a \partial_\mu f(a)$ for every $a \in \mathbb{R}^n$, we have $D(f) = v^\mu \partial_\mu f$. $\square$

7.3 Smooth bump functions

In order to transport the results from Section 7.2 from $\mathbb{R}^n$ to M, we have to prove the existence of *bump functions* on M.

Lemma 7.4 (Bump functions)**.** *Let $U \subset M$ be an open neighbourhood of the point $p_0 \in M$. Then there exists a smooth function $\psi\colon M \to \mathbb{R}$ and closed neighbourhoods $V_1 \subseteq V_2 \subseteq U$ of p_0 such that:*

(1) $\psi(p) = 1$ for p in V_1,

(2) $\psi(p) = 0$ for $p \in M - V_2$,

and $\psi(p) \in [0,1]$ for all $p \in M$.

The reason that ψ is called a *bump function* should be clear from Figure 14.

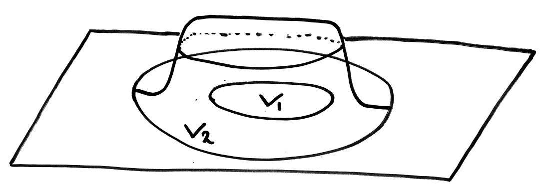


Figure 14: Bump function on $U \subseteq M$.

We start by constructing a bump function on $\mathbb{R}^n$. The following is the special case of Lemma 7.4 where $M = \mathbb{R}^n$, $U = \mathbb{R}^n$, $p_0 = 0$, $V_2 = \overline{B}_\varepsilon(0)$ is the (closed) ball of radius ε around 0, and $V_1 = \overline{B}_{\varepsilon/2}(0)$ is the (closed) ball of radius $\varepsilon/2$.

Lemma 7.5. *For every $\varepsilon > 0$, there exists a smooth function $\psi_n\colon \mathbb{R}^n \to \mathbb{R}$ with $\psi_n(x) = 1$ for $\|x\| \leq \varepsilon/2$, $\psi_n(x) = 0$ for $\|x\| \geq \varepsilon$, and $0 < \psi_n(x) < 1$ for $\varepsilon/2 < \|x\| < \varepsilon$.*

Proof. First, define $f\colon \mathbb{R} \to \mathbb{R}$ by

$$f(x) := \begin{cases} e^{-1/x} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0\,. \end{cases} \tag{57}$$

Since this is a smooth function by Problem 7.6, the function

$$h(x) := \frac{f(2-x)}{f(2-x)+f(x-1)} \tag{58}$$

is smooth as well. It satisfies $h(x) = 1$ for $x \leq 1$, $0 < h(x) < 1$ for $1 < x < 2$, and $h(x) = 0$ for $x \geq 2$. It is well defined because the denominator is positive. Indeed, for every x either the expression $x-1$ or the expression $2-x$ is positive, yielding a positive value of f.

Using this function h, we obtain a bump function on $\mathbb{R}^n$ by setting

$$\psi_n(x) := h(\tfrac{2}{\varepsilon}\|x\|)\,.$$

One checks that $\psi_n(x)$ is zero for $\|x\| \geq \varepsilon$, one for $\|x\| \leq \varepsilon/2$, and that $0 < \psi_n(x) < 1$ for $\varepsilon/2 < \|x\| < \varepsilon$. It is smooth in 0 because it is identically 1 on a ball of radius $\varepsilon/2$ around 0. □

Problem 7.6. Prove that f in equation (57) is a smooth function.

a) Prove that $f'(0) = 0$, and that f' is continuous on $\mathbb{R}$.

b) Show by induction that

$$f^{(k)}(x) := \begin{cases} \frac{p_k(x)}{x^{2k}} e^{-1/x} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0\,, \end{cases}$$

where $p_k(x)$ is a polynomial of degree $k-1$ given by $p_1 = 1$ and by $p_{k+1} = x^2 p_k'(x) + (1-2kx)p_k(x)$.

c) Conclude that f is smooth.

We now have an obvious candidate for the bump function ψ on M. Since M locally looks like $\mathbb{R}^n$, we simply transport the bump function ψ_n on $\mathbb{R}^n$ to a bump function ψ on M, using a coordinate chart $\phi_\alpha \colon M \supset U_\alpha \to \mathbb{R}^n$ centered around $p_0 \in M$.

Proof of Proposition 7.4. Choose a chart (U_α, ϕ_α) that contains p_0, and choose an $\varepsilon > 0$ such that $\overline{B}_\varepsilon(\phi_\alpha(p_0)) \subseteq \phi_\alpha(U \cap U_\alpha) \subseteq \mathbb{R}^n$. Define

$$\psi(p) := \begin{cases} \psi_n(\phi_\alpha(p) - \phi_\alpha(p_0)) & \text{if } p \in U_\alpha, \\ 0 & \text{if } p \notin U_\alpha\,, \end{cases}$$

and define the closed neighbourhoods V_1, V_2 by $V_1 := \phi_\alpha^{-1}(\overline{B}_{\varepsilon/2}(\phi_\alpha(p_0)))$ and $V_2 := \phi_\alpha^{-1}(\overline{B}_\varepsilon(\phi_\alpha(p_0)))$. One checks that $\psi(p) = 1$ for $p \in V_1$, and $\psi(p) = 0$ for $p \in M - V_2$. Since ψ_n is smooth, ψ is smooth in any point $p \in U_\alpha$. It remains to show that ψ is smooth for $p \in M - U_\alpha$. For this, it suffices to show that $V_2 \subseteq U$ is closed. Indeed, since ψ is then zero on the *open* set $M - V_2$ containing p, ψ is smooth at $p \in M - U_\alpha$.

The proof that V_2 is closed uses the Hausdorff property of M. First of all, the closed ball $\overline{B}_\varepsilon(\phi_\alpha(p_0)) \subseteq \mathbb{R}^n$ is compact because it is a closed, bounded subset of $\mathbb{R}^n$ (Cf. Corollary A.55). Since the restriction of ϕ_α^{-1} to $\overline{B}_\varepsilon(\phi_\alpha(p_0)) \subseteq \phi_\alpha(U_\alpha)$ is a continuous map from the compact ball to the Hausdorff space M, its image is closed by Corollary A.56(b). □

7.4 Vectors and vector fields on M

In order to transport the results from Section 7.2 from $\mathbb{R}^n$ to M, we need to investigate how derivations transform under smooth maps.

For a smooth map $F\colon M \to N$, the *pullback* $F^*\colon C^\infty(N) \to C^\infty(M)$ defined by $F^*f := f \circ F$ is an algebra homomorphism. That is,

$$\begin{aligned} F^*(\alpha f + \beta g) &= \alpha F^* f + \beta F^* g \quad \text{and} \\ F^*(fg) &= (F^*f)(F^*g) \end{aligned}$$

for all $f, g \in C^\infty(N)$ and $\alpha, \beta \in \mathbb{R}$. We use this to define the *pushforward* on the level of derivations by $F_*D(f) := D(F^*f)$.

Proposition 7.7. *If D_p is a derivation of $C^\infty(M)$ at $p \in M$, then the pushforward F_*D_p along $F\colon M \to N$ is a derivation of $C^\infty(N)$ at $F(p) \in N$. The resulting map $F_*\colon \mathrm{Der}_p(C^\infty(M)) \to \mathrm{Der}_{F(p)}(C^\infty(N))$ is linear.*

Proof. To prove the Leibniz rule (34) for F_*D_p, use the Leibniz rule for D_p and the fact that $F^*(gh) = F^*(g)F^*(h)$,

$$\begin{aligned} F_*D_p(gh) &= D_p\big(F^*(g \cdot h)\big) = D_p\big((F^*g) \cdot (F^*h)\big) \\ &= D_p(F^*g) \cdot (F^*h)(p) + (F^*g)(p) \cdot D_p(F^*h) \\ &= F_*D_p(g) \cdot h(F(p)) + g(F(p)) \cdot F_*D_p(h)\,. \end{aligned}$$

That is F_*D_p satisfies the Leibniz rule for $C^\infty(N)$ at the point $F(p) \in N$. Since $F^*\colon C^\infty(N) \to C^\infty(M)$ and $D_p\colon C^\infty(M) \to \mathbb{R}$ are linear maps, their composition $F_*D_p(f) = D_p \circ F^*(f)$ is linear in f as well as D_p. $\square$

The chain rule for derivations is just as easy to prove as the chain rule for tangent vectors.

Proposition 7.8 (Chain rule)**.** *Let M and N be smooth manifolds.*

(a) *If $F\colon M \to N$ and $G\colon N \to L$ are smooth maps, then $G_* \circ F_*(D_p) = (G \circ F)_*(D_p)$.*

(b) $\mathrm{Id}_*\colon \mathrm{Der}_p(C^\infty(M)) \to \mathrm{Der}_{\phi(p)}(C^\infty(M))$ *is the identity.*

(c) *If $\phi\colon M \to N$ is a diffeomorphism, then the pushforward on derivations $\phi_*\colon \mathrm{Der}_p(C^\infty(M)) \to \mathrm{Der}_{\phi(p)}(C^\infty(N))$ is a linear isomorphism.*

Proof. For (a), note that both $(G \circ F)_*(D_p)(f)$ and $F_*(G_*D_p)(f)$ are equal to $D_p(f \circ G \circ F)$. Part (*b*) is trivial. Part (c) follows by applying (a) and (b) to the equations $\phi^{-1} \circ \phi = \mathrm{Id}$ and $\phi \circ \phi^{-1} = \mathrm{Id}$. $\square$

Although the following result looks rather innocuous, its proof actually uses the existence of bump functions, and hence the Hausdorff property of M.

Theorem 7.9. *Let $U \subseteq M$ be an open subset, let $p \in U$, and let $\iota\colon U \hookrightarrow M$ be the canonical inclusion. Then $\iota_*\colon \mathrm{Der}_p(C^\infty(U)) \xrightarrow{\sim} \mathrm{Der}_p(C^\infty(M))$ is a linear isomorphism.*

First we prove that for a derivation D_p of $C^\infty(M)$ at $p \in M$, the value of $D_p(f)$ depends only on the values of f on an arbitrarily small neighbourhood $U \subseteq M$ of p.

Lemma 7.10. *Let $D_p \in \mathrm{Der}_p(C^\infty(M))$, let $f, g \in C^\infty(M)$, and suppose that $f|_U = g|_U$ on an open neighbourhood $U \subseteq M$ of p. Then $D_p(f) = D_p(g)$.*

Proof. Let $\psi\colon M \to \mathbb{R}$ be a bump function around $p \in U$ in the sense of Lemma 7.4. Then $1 - \psi$ is equal to 1 on $M - U$. Since $f - g$ is equal to 0 on U, we have

$$D_p(f - g) = D_p\big((1 - \psi)(f - g)\big).$$

By the Leibniz rule, we then find

$$D_p(f - g) = D_p(1 - \psi)(f(p) - g(p)) + (1 - \psi(p))D_p(f - g),$$

which is 0 since $\psi(p) = 1$ and $f(p) = g(p)$. It follows that $D_p(f) = D_p(g)$. $\square$

Having seen that $D_p(f)$ is determined by the restriction $f|_U$ of $f \in C^\infty(M)$ to $U \subseteq M$, it is not hard to prove that $\iota_*\colon \mathrm{Der}_p(C^\infty(U)) \to \mathrm{Der}_p(C^\infty(M))$ is a linear isomorphism.

Proof of Theorem 7.9. First we show that $\iota_*\colon \mathrm{Der}_p(C^\infty(U)) \to \mathrm{Der}_p(C^\infty(M))$ is injective. Let $D_p \in \mathrm{Der}_p(C^\infty(U))$ with $\iota_* D_p = 0$. Then $\iota_* D_p(f) = D_p(\iota^* f) = 0$ for all $f \in C^\infty(M)$. Since $\iota^* f = f|_U$, it follows that $D_p(g) = 0$ for any $g \in C^\infty(U)$ which is the restriction $g = f|_U$ of a smooth function on M. Unfortunately, not every smooth function $g\colon U \to \mathbb{R}$ extends to M. (See figure 15) However, for any $g \in C^\infty(U)$, we can construct a smooth function f on M by

$$f(q) := \begin{cases} \psi(q)g(q) & \text{for } q \in U \\ 0 & \text{for } q \in M - U. \end{cases}$$

This function is smooth because it is the product of two smooth functions on the open set $U \subseteq M$, and it is identically zero on the open set $M - V \subseteq M$. Since ψ is 1 in a neighbourhood V of p, the functions f and g agree on $V \subseteq U$. It follows that $D_p(g) = D_p(f|_U) = D_p(\iota^* f) = 0$, so D_p vanishes on all $g \in C^\infty(U)$. Thus $D_p = 0$, and ι_* is injective.

Next we show that ι_* is surjective. For a given derivation $D^M_p\colon C^\infty(M) \to \mathbb{R}$, we construct a derivation $D^U_p\colon C^\infty(U) \to \mathbb{R}$ such that $\iota_* D^U_p = D^M_p$. On $g \in C^\infty(U)$, we define $D^U_p(g)$ to be $D^M_p(f)$ for *any* $f \in C^\infty(M)$ that agrees with g on some (arbitrarily small) neighbourhood V of p. The existence of such functions follows from the above construction, and the value of $D^M_p(f)$ is independent of the choice of f by Lemma 7.10. Since D^M_p is linear and satisfies the Leibniz identity, the same holds for D^U_p. By construction, we now have $\iota_* D^U_p(f) = D^U_p(\iota^* f) = D^U_p(f|_U) = D^M_p(f)$, and it follows that ι_* is surjective. $\square$

Recall from Section 4.2 that every tangent vector $v_p \in T_pM$ gives rise to a derivation D^v_p of $C^\infty(M)$ at $p \in M$ that satisfies $D_p(f) = \frac{d}{dt}|_0 f(\gamma(t))$ for any smooth curve γ through p with tangent vector v_p.

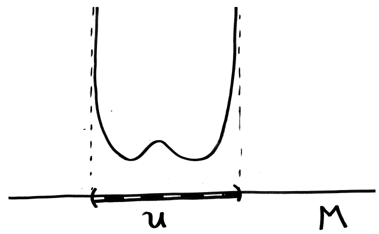


Figure 15: Not every smooth function on $U \subseteq M$ extends to M

Theorem 7.11. *Every derivation $D_p \colon C^\infty(M) \to \mathbb{R}$ at $p \in M$ is of the form D_p^v for a tangent vector $v \in T_pM$. The resulting map $T_pM \to \mathrm{Der}_p(C^\infty(M))$ is a linear isomorphism.*

Proof. Let $\phi_\alpha \colon U_\alpha \to M \supseteq \phi_\alpha(U_\alpha) \subseteq \mathbb{R}^n$ be a chart around $p \in M$. By Theorem 7.9, $\mathrm{Der}_p(C^\infty(M))$ is canonically isomorphic to $\mathrm{Der}_p(U_\alpha)$. Since ϕ_α is a diffeomorphism from U_α to $\phi_\alpha(U_\alpha)$, it induces a linear isomorphism from $\mathrm{Der}_p(U_\alpha)$ to $\mathrm{Der}_{\phi_\alpha(p)}(\phi_\alpha(U_\alpha))$. By another application of Theorem 7.9, $\mathrm{Der}_{\phi_\alpha(p)}(\phi_\alpha(U_\alpha))$ is canonically isomorphic to $\mathrm{Der}_{\phi_\alpha(p)}(\mathbb{R}^n)$. It follows that every derivation D_p is of the form ${\phi_\alpha}_* D'_{\phi_\alpha(p)}$ for some $D'_{\phi_\alpha(p)} \in \mathrm{Der}_{\phi_\alpha(p)}(\mathbb{R}^n)$. Since $D'_{\phi_\alpha(p)}(f_\alpha) = v^\mu_\alpha \partial_\mu f_\alpha$ by Proposition 7.2, the result follows. □

Theorem 7.12. *Every derivation $D \colon C^\infty(M) \to C^\infty(M)$ is of the form $D(f) = \mathcal{L}_v(f)$ for some vector field $v \in \mathrm{Vec}(M)$. The resulting map $v \mapsto \mathcal{L}_v$ is is a $C^\infty(M)$-linear isomorphism from $\mathrm{Vec}(M)$ to $\mathrm{Der}(C^\infty(M))$.*

Proof. To see that the map is linear over $C^\infty(M)$, simply check that $\mathcal{L}_{fv}(g) = (f\mathcal{L}_v)(g)$. We prove that the map is an isomorphism. By Proposition 7.1, the derivation D is uniquely determined by its evaluation D_p at all points $p \in M$. For every p, Theorem 7.11 yields a vector $v_p \in T_pM$ such that $D(f)(p) = v_p(f)$. To show that the resulting section $v \colon M \to TM, p \mapsto v_p$ is smooth, let $v^\mu_\alpha(x(p))$ be the coordinates of v_p with respect to a chart (U_α, ϕ_α) around p. If we choose a bump function ψ on U_α that is 1 in a neighbourhood V of $p \in M$, then $v^\mu_\alpha = D(\psi x^\mu)$ on V. Since the latter is a smooth function on M, the coordinates $v^\mu_\alpha(x)$ are smooth in a neighbourhood of p. Since p was arbitrary, the result follows. □

8 Tensors, tensor fields and metrics

Recall that the mathematical structures underlying Riemannian and Minkowski geometry are essentially symmetric bilinear forms. In the case of Riemannian geometry on $\mathbb{R}^n$, this is an inner product $(\,\cdot\,,\,\cdot\,)\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. For Minkowski geometry, the relevant bilinear form $\eta\colon \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$ has signature $(-+++)$.

In order to generalise Riemannian and Minkowski geometry to the setting of smooth manifolds, we will equip every tangent space T_pM with a bilinear form $g_p\colon T_pM \times T_pM \to \mathbb{R}$ in such a way that g_p depends smoothly on p. More generally, we will need *tensors* on M.

8.1 Covariant tensors

Tensors are most easily described using multilinear algebra. If V and W are vector spaces, then a *multilinear map*

$$F\colon \underbrace{V \times \ldots \times V}_{k \text{ times}} \to W$$

is a map which is linear in each of its entries. In other words, we have

$$F(v_1, \ldots, \alpha v_i + \beta v_i', \ldots, v_k) = \alpha F(v_1, \ldots, v_i, \ldots, v_k) + \beta F(v_1, \ldots, \beta v_i', \ldots, v_k)$$

for all $v_i, v_i' \in V$, and for all $\alpha, \beta \in \mathbb{R}$.

8.1.1 Definition of covariant tensors

Let M be a manifold, and let p be a point in M. A *covariant tensor* of rank k at $p \in M$ is a multilinear map

$$\tau_p\colon \underbrace{T_pM \times \ldots \times T_pM}_{k \text{ times}} \to \mathbb{R}.$$

An important special case are the tensors of rank $k = 2$, the *bilinear forms* $T_pM \times T_pM \to \mathbb{R}$. A covariant tensor of rank 1 is simply a linear map $T_pM \to \mathbb{R}$, often called a *linear form*.

We denote the set of covariant tensors of rank k at $p \in M$ by $T_k(M)_p$. Equipped with the usual addition and scalar multiplication, this is a vector space of dimension n^k. To see this, we calculate $\tau(v_1, \ldots, v_k)$ with respect to a coordinate basis ∂_μ of T_pM. For $i = 1, \ldots, k$ we express $v_i = v_i^\mu \partial_\mu$ as a linear combination of basis vectors. Using the fact that τ is linear in each of its k entries, we find that

$$\begin{aligned} \tau(v_1, \ldots, v_k) &= \tau(v_1^{\mu_1}\partial_{\mu_1}, \ldots, v_k^{\mu_k}\partial_{\mu_k}) \\ &= v_1^{\mu_1} \cdots v_k^{\mu_k} \tau(\partial_{\mu_1}, \ldots, \partial_{\mu_k}). \end{aligned} \tag{59}$$

Note that the Einstein summation convention implies a sum $\Sigma_{\mu_1=1}^n \ldots \Sigma_{\mu_k=1}^n$! With respect to the basis ∂_μ, the tensor τ is therefore completely determined by the n^k real numbers

$$\tau_{\mu_1\ldots\mu_k} := \tau(\partial_{\mu_1}, \ldots, \partial_{\mu_k}), \tag{60}$$

which depend linearly on τ. They are the *coefficients* of the tensor τ_p with respect to the coordinate basis.

8.1.2 The covariant tensor bundle

The *tensor bundle* $T_k(M)$ is the set of *all* covariant tensors on M. In other words, it is the disjoint union

$$T_k(M) := \bigsqcup_{p \in M} T_k(M)_p \tag{61}$$

of the vector spaces $T_k(M)_p$. Note that the bundle $T_k(M)$ is *not* a vector space, since it does not make sense to add tensors that are defined on different tangent spaces. The bundle $T_k(M)$ comes with the canonical projection $\pi\colon T_k(M) \to M$, defined by $\pi(\tau_p) = p$.

To construct a smooth manifold structure on $T_k(M)$, we can proceed along the same lines as in §4.6, where we constructed a manifold structure on the tangent bundle TM. We will skip over some of the details, but if you reread §4.6, you should be able to fill the gaps yourself.

Recall that local coordinates x^μ on $U_\alpha \subseteq M$ yield a coordinate basis ∂_μ for every tangent space T_pM over $p \in U_\alpha$. This allows us to describe a tensor τ_p at a point $p \in U_\alpha$ by the $n + n^k$ coordinates x^μ and $\tau_{\mu_1\ldots\mu_k}$. Specifically, x^μ are the coordinates of the point p, and if the vectors $v_1, \ldots, v_k \in T_pM$ are given by $v_i = v_i^\mu \partial_\mu$, then

$$\tau_p(v_1, \ldots, v_k) = \tau_{\mu_1\ldots\mu_k} v_1^{\mu_1} \ldots v^{\mu_k}.$$

We thus obtain a chart on the set $\pi^{-1}(U_\alpha) \subseteq T_kM$ of all tensors τ_p whose base point p lies in U_α.

To show that the coordinate transitions are smooth, suppose that $x^{\overline{\mu}}$ are local coordinates on a different coordinate neighbourhood $U_\beta \subseteq M$. Then on the intersection $U_\alpha \cap U_\beta$, we have

$$\tau_{\overline{\mu}_1\ldots\overline{\mu}_k} v_1^{\overline{\mu}_1} \cdots v_k^{\overline{\mu}_k} = \tau_p(v_1, \ldots, v_k) = \tau_{\mu_1\ldots\mu_k} v_1^{\mu_1} \cdots v_k^{\mu_k}. \tag{62}$$

Recall from (33) that $v^\mu = J^\mu_{\overline{\mu}} v^{\overline{\mu}}$, where

$$J^\mu_{\overline{\mu}} := \left(\tfrac{\partial x^\mu}{\partial x^{\overline{\mu}}}\right) \tag{63}$$

is the Jacobian matrix of the coordinate transformation $x^\mu(x^{\overline{1}}, \ldots, x^{\overline{n}})$. Equation (62) then yields

$$\tau_{\overline{\mu}_1\ldots\overline{\mu}_k} v_1^{\overline{\mu}_1} \cdots v_k^{\overline{\mu}_k} = \tau_{\mu_1\ldots\mu_k} J^{\mu_1}_{\overline{\mu}_1} \cdots J^{\mu_k}_{\overline{\mu}_k} v_1^{\overline{\mu}_1} \cdots v_k^{\overline{\mu}_k}. \tag{64}$$

Since this holds for all vectors v_1 through v_k, the coordinates of τ_p with respect to x^μ and $x^{\overline{\mu}}$ are related by

$$\tau_{\overline{\mu}_1\ldots\overline{\mu}_k} = \tau_{\mu_1\ldots\mu_k} J^{\mu_1}_{\overline{\mu}_1} \cdots J^{\mu_k}_{\overline{\mu}_k}. \tag{65}$$

Since the Jacobian matrix $J^{\mu}_{\overline{\mu}}$ depends smoothly on the point x^μ, the coordinates $x^{\overline{\mu}}$ and $\tau_{\overline{\mu}_1\ldots\overline{\mu}_k}$ depend smoothly on x^μ and $\tau_{\mu_1\ldots\mu_k}$.

This shows that T_kM can be covered with coordinates in such a way that the transition functions are smooth. In the same way as in §4.6, we can define a Hausdorff topology on $T_k(M)$ for which the coordinate charts are homeomorphisms onto their image. In these coordinates, it is easy to see that the canonical projection is a surjective submersion: in local coordinates, it simply projects on the variables x^1 through x^n.

Proposition 8.1 ($T_k(M)$ as a smooth manifold)**.** *The covariant tensor bundle $T_k(M)$ is a smooth manifold, and the canonical projection $\pi\colon T_k(M) \to M$ is a surjective submersion.*

8.2 Covariant tensor fields

Recall that a vector field is a smooth section of the canonical projection from TM to M. Similarly, we define a *covariant tensor field* as a smooth section of the canonical projection $\pi\colon T_k(M) \to M$.

Definition 8.2. A *covariant tensor field* om M is a smooth map $\tau\colon M \to T_k(M)$ such that $\pi \circ \tau = \mathrm{Id}_M$.

In other words, the tensor field τ assigns to each point $p \in M$ a multilinear map

$$\tau_p\colon T_pM \times \ldots \times T_pM \to \mathbb{R}.$$

In local coordinates x^μ, a covariant tensor field is described by n^k smooth functions $\tau_{\mu_1\ldots\mu_k}(x^1, \ldots, x^n)$. If the same tensor field is described using different coordinates $x^{\overline{\mu}}$, then the components are related by

$$\tau_{\overline{\mu}_1\ldots\overline{\mu}_k} = J^{\mu_1}_{\overline{\mu}_1} \cdots J^{\mu_k}_{\overline{\mu}_k} \tau_{\mu_1\ldots\mu_k}, \tag{66}$$

where on the right hand side, $\tau_{\mu_1\ldots\mu_k}$ is considered as a function of the variables $x^{\overline{\mu}}$ via the coordinate transformation $x^\mu(x^{\overline{1}}, \ldots, x^{\overline{n}})$.

8.2.1 Three different ways of describing covariant tensor fields

A covariant tensor field can thus be described either as a smooth section of the canonical projection $\pi\colon T_k(M) \to M$, or by its components $\tau_{\mu_1\ldots\mu_k}$ in local coordinates x^μ. A third, more algebraic description of tensor fields is obtained as follows.

A covariant tensor field $\tau\colon M \to T_k(M)$ yields a multilinear map

$$\widetilde{\tau}\colon \underbrace{\mathrm{Vec}(M) \times \cdots \times \mathrm{Vec}(M)}_{k \text{ times}} \to C^\infty(M) \tag{67}$$

by $\widetilde{\tau}(v_1, \ldots, v_k)(p) := \tau_p(v_1(p), \ldots, v_k(p))$. If the vector fields v_i are given by $v_i^\mu \partial_\mu$ in local coordinates, then the function $\widetilde{\tau}(v_1, \ldots, v_k)$ is described in local coordinates by the contraction

$$v_1^{\mu_1} \cdots v_k^{\mu_k} \tau_{\mu_1 \ldots \mu_k}.$$

Recall that $\mathrm{Vec}(M)$ is a module over the algebra $C^\infty(M)$ of smooth functions, meaning that the product fv of a smooth function f with a vector field v is again a vector field. The map $\widetilde{\tau}$ is multilinear *over the smooth functions*, meaning that

$$\widetilde{\tau}(v_1, \ldots, fv_i, \ldots v_k) = f\widetilde{\tau}(v_1, \ldots, v_i, \ldots v_k)$$

for all $f \in C^\infty(M)$. In fact, one can show that every map $\widetilde{\tau}$ that is multilinear over $C^\infty(M)$ comes from a tensor field of rank k.

8.2.2 Transformation of covariant tensor fields

Let $F \colon M \to N$ be a smooth map. Then there is a *pullback* map F^* that takes covariant tensor fields on N to covariant tensor fields on M. The pullback of $\tau \colon N \to T_k(N)$ is the tensor field $F^*\tau \colon M \to T_k(M)$ defined by

$$(F^*\tau)_p(v_1, \ldots, v_k) := \tau_{F(p)}(F_*v_1, \ldots, F_*v_k). \tag{68}$$

Suppose that τ has coefficients $\tau_{\nu_1 \ldots \nu_k}(y^1, \ldots, y^m)$ with respect to local coordinates y^ν on N. If F has coordinate representation $y^\nu(x^1, \ldots, x^n)$ with respect to local coordinates x^μ on M, then

$$(F^*\tau)_{\mu_1 \ldots \mu_k} = \left(\frac{\partial y^{\nu_1}}{\partial x^{\mu_1}}\right) \cdots \left(\frac{\partial y^{\nu_k}}{\partial x^{\mu_k}}\right) \tau_{\nu_1 \ldots \nu_k}. \tag{69}$$

In particular, this shows that $F^*\tau$ is smooth.

Problem 8.3. Show that (69) holds.

8.3 Symmetric and alternating covariant tensors

A covariant tensor τ_p is called *symmetric* if

$$\tau_p(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = \tau_p(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k), \tag{70}$$

for all $1 \leq i < j \leq k$, and *alternating* if

$$\tau_p(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -\tau_p(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k). \tag{71}$$

Accordingly, the coefficients of a symmetric tensor are invariant under permutation,

$$\tau_{\mu_1 \ldots \mu_i \ldots \mu_j \ldots \mu_k} = \tau_{\mu_1 \ldots \mu_j \ldots \mu_i \ldots \mu_k},$$

whereas the coefficients of an alternating tensor satisfy

$$\tau_{\mu_1 \ldots \mu_i \ldots \mu_j \ldots \mu_k} = -\tau_{\mu_1 \ldots \mu_j \ldots \mu_i \ldots \mu_k}.$$

We denote the vector spaces of symmetric and alternating tensors of rank k at $p \in M$ by $S^k(T_p^*M)$ and $\bigwedge^k(T_p^*M)$, respectively. The bundles

$$S^k(T^*M) := \bigsqcup_{p\in M} S^k(T_p^*M) \quad \text{and} \quad \bigwedge^k(T^*M) := \bigsqcup_{p\in M} \bigwedge^k(T_p^*M)$$

are smooth manifolds, and the canonical projections

$$\pi\colon \bigwedge^k(T^*M) \to M \quad \text{and} \quad \pi\colon \bigwedge^k(T_p^*M) \to M$$

are smooth. An alternating tensor field $\tau\colon M \to \bigwedge^k(T^*M)$ is often called a k-form.

An important example it the case $k = 2$. A covariant tensor of rank 2 is a bilinear map $\tau_p\colon T_pM \times T_pM \to \mathbb{R}$, represented by the matrix $\tau_{\mu\nu}$ with respect to a basis ∂_μ of T_pM. Symmetric tensors of rank 2 are symmetric bilinear forms $\tau_p(v,w) = \tau_p(w,v)$, represented by a symmetric matrix $\tau_{\mu\nu} = \tau_{\nu\mu}$. Alternating tensors of rank 2 are skew-symmetric bilinear forms $\tau_p(v,w) = -\tau_p(w,v)$, which are represented by a skew-symmetric matrix $\tau_{\mu\nu} = -\tau_{\nu\mu}$.

Problem 8.4. An alternating tensor $\tau \in \bigwedge^k(T_p^*M)$ is determined by the coefficients $\tau_{\mu_1\ldots\mu_k}$ with $1 \le \mu_1 < \ldots < \mu_k \le n$. It follows that $\bigwedge^k(T_p^*M)$ is of dimension $\binom{n}{k}$.

Problem 8.5. A symmetric tensor $\tau \in S^k(T_p^*M)$ is determined by the coefficients $\tau_{\mu_1\ldots\mu_k}$ with $1 \le \mu_1 \le \ldots \le \mu_k \le n$. It follows that $S^k(T_p^*M)$ is of dimension $\binom{n+k-1}{k}$.

Problem 8.6 (Determinants and volume forms)**.** An alternating tensor of rank $k = n$ is called a *volume form*.

a) The space $\bigwedge^n T_p^*M$ of volume forms has dimension 1. Every volume form is a multiple of

$$\tau_p(v_1, \ldots, v_n) = \sum_{\sigma\in S_n} \operatorname{sg}(\sigma) v_1^{\sigma(1)} \cdots v_n^{\sigma(n)},$$

where the sign $\operatorname{sg}(\sigma)$ of the permutation σ is $+1$ if σ is a product of an even number of 'swaps', and -1 if σ is the product of an odd number of swaps.

b) Let $F\colon M \to M$ be a smooth map with $F(p) = p$, and let τ_p be a nonzero volume form at p. Then $(F^*\tau)_p = \lambda\tau_p$ for some $\lambda \in \mathbb{R}$. This number is independent of the choice of τ_p. In fact, it is the determinant of the linear map $F_*\colon T_pM \to T_pM$,

$$\det(F_*) = \sum_{\sigma\in S_n} \operatorname{sg}(\sigma) \frac{\partial F^1}{\partial x^{\sigma(1)}} \cdots \frac{\partial F^n}{\partial x^{\sigma(n)}}.$$

8.4 Contravariant and mixed tensors

A linear functional $\alpha_p \colon T_pM \to \mathbb{R}$ is called a *covector* at $p \in M$. The set T_p^*M of covectors at p is a vector space of dimension n, called the *cotangent space* of M at p. The basis ∂_μ of T_pM gives rise to the dual basis dx^μ of T_p^*M defined by

$$dx^\mu(\partial_\nu) = \delta^\mu_\nu.$$

Every covector α_p is determined by its values $\alpha_\mu := \alpha_p(\partial_\mu)$ on the coordinate basis, $\alpha_p = \alpha_\mu dx^\mu$.

Since the basis $\partial_{\overline{\mu}}$ is related to the basis ∂_μ by $\partial_{\overline{\mu}} = \frac{\partial x^\mu}{\partial x^{\overline{\mu}}}\partial_\mu$, the bases $dx^{\overline{\mu}}$ and dx^μ are related by

$$dx^{\overline{\mu}} = \left(\frac{\partial x^{\overline{\mu}}}{\partial x^\mu}\right) dx^\mu. \tag{72}$$

This motivates the suggestive notation dx^μ for the dual coordinate basis. Note that since covectors are precisely covariant tensors of rank 1, the cotangent bundle $T^*M := \bigsqcup_{p\in M} T_p^*M$ is a smooth manifold.

Problem 8.7. Prove equation 72.

8.4.1 Contravariant tensors

A *contravariant tensor* of rank l at p is a multilinear map

$$\tau \colon \underbrace{T_p^*M \times \ldots \times T_p^*M}_{l\,\text{times}} \to \mathbb{R}.$$

We denote the vector space of contravariant tensors of rank l at p by by $T^l(M)_p$. By a line of reasoning similar to (59), every contravariant tensor of rank l is determined by the n^l real numbers

$$\tau^{\nu_1\ldots\nu_l} = \tau(dx^{\nu_1}, \ldots, dx^{\nu_l}), \tag{73}$$

the *coefficients* of τ with respect to the basis dx^ν of T_p^*M. Since the coefficients of τ with respect to different coordinates are related by

$$\tau^{\overline{\nu}_1\ldots\overline{\nu}_l} = \left(\frac{\partial x^{\overline{\nu}_1}}{\partial x^{\nu_1}}\right)\cdots\left(\frac{\partial x^{\overline{\nu}_l}}{\partial x^{\nu_l}}\right)\tau^{\nu_1\ldots\nu_l},$$

the bundle $T^l(M) := \bigsqcup_{p\in M} T^l(M)_p$ is a smooth manifold of dimension $n + n^l$. A smooth section $\tau \colon M \to T^l(M)$ of the canonical projection $\pi \colon T^l(M) \to M$ is called a *contravariant tensor field* of rank l.

Note that for $l = 1$, we have a canonical linear isomorphism $T_pM \to T^1(M)_p$. Every tangent vector $v_p \in T_pM$ defines a linear functional $T_p^*M \to \mathbb{R}$ by $\alpha_p \mapsto \alpha_p(v_p)$. We therefore identify $T^1(M)$ with TM. In particular, a contravariant tensor field of rank 1 is just a vector field on M.

8.4.2 Mixed tensors

A *mixed tensor* of rank (k, l) at $p \in M$ is a multilinear map

$$\tau_p \colon \underbrace{T_pM \times \ldots \times T_pM}_{k\text{ times}} \times \underbrace{T_p^*M \times \ldots \times T_p^*M}_{l\text{ times}} \to \mathbb{R}.$$

We denote the space of mixed tensors at p by $T_k^l(M)_p$. With respect to coordinates x^μ, a mixed tensor τ is given by the n^{k+l} coefficients

$$\tau_{\mu_1\ldots\mu_k}^{\nu_1\ldots\nu_l} := \tau(\partial_{\mu_1}, \ldots, \partial_{\mu_k}, dx^{\nu_1}, \ldots, dx^{\nu_l}).$$

We identify a mixed tensor of rank $(1, 1)$ with the linear map $T_pM \to T_pM$ defined by $v^\mu \partial_\mu \mapsto v^\mu \tau_\mu^\nu \partial_\nu$. More generally, a mixed tensor of rank $(k, 1)$ can be identified with the multilinear map

$$\underbrace{T_pM \times \ldots \times T_pM}_{k\text{ times}} \to T_pM \quad \text{defined by} \quad (v_1, \ldots, v_k) \mapsto v_1^{\mu_1} \cdots v_k^{\mu_k} \tau_{\mu_1\ldots\mu_k}^\nu \partial_\nu.$$

Problem 8.8. The image of $(v_1, \ldots, v_k)$ is the unique vector v that satisfies $\alpha(v) = \tau(v_1, \ldots, v_k, \alpha)$ for all $\alpha \in T_p^*M$. In particular, the above expression is independent of the choice of basis.

The components of a mixed tensor transform as

$$\tau_{\overline{\mu}_1\ldots\overline{\mu}_k}^{\overline{\nu}_1\ldots\overline{\nu}_l} = \left(\frac{\partial x^{\mu_1}}{\partial x^{\overline{\mu}_1}}\right) \cdots \left(\frac{\partial x^{\mu_k}}{\partial x^{\overline{\mu}_k}}\right) \cdot \left(\frac{\partial x^{\overline{\nu}_1}}{\partial x^{\nu_1}}\right) \cdots \left(\frac{\partial x^{\overline{\nu}_l}}{\partial x^{\nu_l}}\right) \cdot \tau_{\mu_1\ldots\mu_k}^{\nu_1\ldots\nu_l}. \tag{74}$$

Using this, one shows that the bundle $T_k^l(M) := \bigsqcup_{p\in M} T_k^l(M)_p$ of mixed tensors is a smooth manifold, with smooth canonical projection $\pi \colon T_k^l(M) \to M$. A smooth section $\tau \colon M \to T_k^l(M)$ of the canonical projection is called a *mixed tensor field.*

8.4.3 Pullback along diffeomorphisms

Unlike covariant tensor fields, contravariant and mixed tensor fields can only be pulled back along *diffeomorphisms.* Let $\Phi \colon M \to N$ be a diffeomorphism, and let $\tau \colon N \to T_k^l(N)$ be a mixed tensor field of rank (k, l) on N. Then the pullback $\Phi^*\tau \colon M \to T_k^l(M)$ is defined by

$$(\Phi^*\tau)_p(v_1, \ldots, v_k; \alpha_1, \ldots, \alpha_l) := \tau_{\Phi(p)}(\Phi_* v_1, \ldots, \Phi_* v_k; \Phi^{-1\,*}\alpha_1, \ldots, \Phi^{-1\,*}\alpha_l).$$

The reason that this definition only works for diffeomorphisms is that although the tangent vectors $v_1, \ldots, v_k \in T_pM$ can be pushed forward along the map $\Phi \colon M \to N$, the cotangent vectors $\alpha_1, \ldots, \alpha_l \in T_p^*M$ can only be *pulled back.* For this, we need that the inverse map $\Phi^{-1} \colon N \to M$ exists and is smooth.

9 Riemannian geometry

We turn to the study of *Riemannian geometry*. Whereas the notion of a smooth manifold allows us to deal with differentiability, the notion of a *Riemannian metric* allows us to handle problems involving distances and angles in a nonlinear setting.

9.1 Riemannian metrics

Let M be a manifold. An *inner product* on T_pM is a symmetric, bilinear form $g_p\colon T_pM \times T_pM \to \mathbb{R}$ such that $g_p(v,v) > 0$ for all nonzero $v \in T_pM$.

Definition 9.1 (Riemannian metric)**.** A *Riemannian metric* g on M is a covariant tensor field of rank 2 such that for every point $p \in M$, the bilinear form $g_p\colon T_pM \times T_pM \to \mathbb{R}$ is an inner product.

A *Riemannian manifold* (M, g) is a smooth manifold equipped with a Riemannian metric. On a Riemannian manifold, then, every tangent space T_pM comes equipped with an inner product g_p that varies smoothly with p.

Example 9.2 (Euclidean metric on $\mathbb{R}^n$)**.** Let $M = \mathbb{R}^n$, and let $x^1, \ldots, x^n$ be the Cartesian coordinates on $\mathbb{R}^n$. This choice of local coordinates allows us to write $v, w \in T_p\mathbb{R}^n$ as $v = v^\mu \partial_\mu$ and $w = w^\mu \partial_\mu$. The *Euclidean metric* on $M = \mathbb{R}^n$ is given by

$$g_p^E(v^\mu \partial_\mu, w^\nu \partial_\nu) = v^1 w^1 + \ldots + v^n w^n.$$

9.1.1 Coordinate expressions

It will often be convenient to describe Riemannian metrics in local coordinates. If $v = v^\mu \partial_\mu$ and $w = w^\nu \partial_\nu$ are the coordinate expressions of $v, w \in T_pM$ with respect to local coordinates x^μ on $U_\alpha \subseteq M$, then

$$g_p(v, w) = g_{\mu\nu}(x) v^\mu w^\nu \tag{75}$$

for an $n \times n$ matrix $g_{\mu\nu}(x)$ that depends smoothly on $x \in \phi_\alpha(U_\alpha)$. Since g_p is an inner product, the matrices $g_{\mu\nu}(x)$ are symmetric and positive definite,

$$g_{\mu\nu}(x) = g_{\nu\mu}(x) \quad \text{and} \quad g_{\mu\nu}(x) v^\mu v^\nu > 0$$

for all nonzero vectors $(v^1, \ldots, v^n) \in \mathbb{R}^n$.

For different coordinates $x^{\overline{\mu}}$ on $U_\beta \subseteq M$, the functions $g_{\overline{\mu}\overline{\nu}}$ are related to $g_{\mu\nu}$ on the overlap $U_\alpha \cap U_\beta$ by the covariant transformation rule

$$g_{\overline{\mu}\overline{\nu}} = \left(\frac{\partial x^\mu}{\partial x^{\overline{\mu}}}\right)\left(\frac{\partial x^\nu}{\partial x^{\overline{\nu}}}\right) g_{\mu\nu}. \tag{76}$$

Example 9.3. Consider the Euclidean metric g^E on $M = \mathbb{R}^2$ (cf. example 9.2). With respect to the Cartesian coordinates x, y on $\mathbb{R}^2$, it is given by

$$\begin{pmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In polar coordinates r, ϕ with $x = r\cos(\phi)$ and $y = r\sin(\phi)$, the 2×2 Jacobian matrix $\frac{\partial x^\mu}{\partial x^{\overline{\mu}}}$ has entries

$$\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \phi} & \frac{\partial y}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -r\sin(\phi) & r\sin(\phi) \end{pmatrix}.$$

The coordinate transformation (76) therefore yields

$$\begin{aligned} g_{rr} &= \frac{\partial x}{\partial r}\frac{\partial x}{\partial r} g_{xx} + 2\frac{\partial x}{\partial r}\frac{\partial y}{\partial r} g_{xy} + \frac{\partial y}{\partial r}\frac{\partial y}{\partial r} g_{yy} = 1, \\ g_{\phi\phi} &= \frac{\partial x}{\partial \phi}\frac{\partial x}{\partial \phi} g_{xx} + 2\frac{\partial x}{\partial \phi}\frac{\partial y}{\partial \phi} g_{xy} + \frac{\partial y}{\partial \phi}\frac{\partial y}{\partial \phi} g_{yy} = r^2, \\ g_{r\phi} = g_{\phi r} &= \frac{\partial x}{\partial r}\frac{\partial x}{\partial \phi} g_{xx} + 2\frac{\partial x}{\partial r}\frac{\partial y}{\partial \phi} g_{xy} + \frac{\partial y}{\partial r}\frac{\partial y}{\partial \phi} g_{yy} = 0. \end{aligned}$$

A different (and shorter!) way to calculate the same result is by expressing the coordinate vector fields $\partial_r, \partial_\phi$ in terms of ∂_x and ∂_y as

$$\begin{aligned} \partial_r &= \cos(\phi)\partial_x + \sin(\phi)\partial_y \\ \partial_\phi &= -r\sin(\phi)\partial_x + r\cos(\phi)\partial_y \end{aligned}$$

and calculate $g_{rr} = g(\partial_r, \partial_r)$, $g_{\phi\phi} = g(\partial_\phi, \partial_\phi)$, and $g_{r\phi} = g_{\phi r} = g(\partial_r, \partial_\phi)$.

Problem 9.4. Let $M = \mathbb{R}^3$, let (U_α, ϕ_α) be the cartesian coordinates on $U_\alpha = \mathbb{R}^3$, and let (U_β, ϕ_β) be spherical coordinates on

$$U_\beta = \mathbb{R}^3 \setminus \{(x, y, z) \in \mathbb{R}^3\,;\, x = 0, y \leq 0\}.$$

The transition functions on $U_\alpha \cap U_\beta$ are then explicitly given by

$$x = \rho\sin(\theta)\sin(\phi), \quad y = \rho\sin(\theta)\cos(\phi), \quad \text{and} \quad z = \rho\cos(\theta).$$

With respect to (U_α, ϕ_α), the Euclidean metric has coefficients $g_{xx} = g_{yy} = g_{zz} = 1$ and $g_{xy} = g_{yz} = g_{xz} = 0$. Calculate the corresponding coefficients $g_{\rho\rho}, g_{\theta\theta}, g_{\phi\phi}$ and $g_{\rho\phi}, g_{\rho\theta}, g_{\phi\theta}$ in spherical coordinates.

9.1.2 Infinitesimal length and angles

If (M, g) is a Riemannian manifold, then the inner product g_p on T_pM allows us to define the lenght and angles at the level of tangent vectors. The *lenght* of $v \in T_pM$ is defined as

$$\|v\| := \sqrt{g_p(v, v)}, \tag{77}$$

and the *angle* between $v \in T_pM$ and $w \in T_pM$ is defined by

$$\theta = \arccos\left(\frac{g(v,w)}{\sqrt{g(v,v)g(w,w)}}\right). \tag{78}$$

Example 9.5. The *Poincaré disk* is given by the open set

$$D^2 := \{(x,y) \in \mathbb{R}^2\,;\, x^2 + y^2 < 1\},$$

with the metric $g_{xx} = g_{yy} = 4/(1 - x^2 - y^2)^2$, and $g_{xy} = g_{yx} = 0$. At the point $p = (x,y)$, the coordinate vector $\partial_x(p)$ has length

$$\|\partial_x\| = \sqrt{g(\partial_x, \partial_x)} = \sqrt{g_{xx}} = 2/(1 - x^2 - y^2).$$

Note that $\|\partial_x\|$ approaches ∞ as $x^2 + y^2 \uparrow 1$.

9.1.3 Pullback and isometries

Suppose that $F\colon M \to N$ is a smooth map, and that τ is a covariant tensor of rank 2 on N. Recall from §8.2.2 that the pullback $(F^*\tau)_p(v_p, w_p) := \tau_{F(p)}(F_*v_p, F_*w_p)$ is again a covariant tensor of rank 2 on M.

Proposition 9.6. *Let g be a Riemannian metric on N, and let $F\colon M \to N$ be a smooth map such that $F_*\colon T_pM \to T_{F(p)}N$ is injective for all $p \in M$. Then the pullback F^*g is a metric on M.*

Proof. We need to show that $(F^*g)_p$ is an inner product, i.e., that $(F^*g)_p(v,v) := g_{F(p)}(F_*v, F_*v) > 0$ for all nonzero $v \in T_pM$. Since F_*v is nonzero if v is nonzero, this follows from the fact that $g_{F(p)}(w,w) > 0$ for all nonzero $w \in T_{F(p)}N$. $\square$

Definition 9.7 (Isometries)**.** Let (M,g) and (N,h) be Riemannian manifolds. An *isometry* is a diffeomorphism $\phi\colon M \to N$ such that

$$g = \phi^*h.$$

We call (M,g) and (N,h) *isometric* if they admit an isometry. Since the infinitesimal length and angles with respect to g_p on T_pM then correspond to the infinitesimal length and angles for $h_{\phi(p)}$ on $T_{\phi(p)}N$, we consider (M,g) and (N,h) as 'the same' as far as their metric properties are concerned.

9.1.4 Musical isomorphisms

Recall that the *cotangent space* T_p^*M is the linear dual of T_pM, and a *covector* $\alpha_p \in T_p^*M$ is a linear map $\alpha_p\colon T_pM \to \mathbb{R}$. Using the inner product g_p on T_pM, we can assign to every vector $v \in T_pM$ the covector $v^\flat \in T_p^*M$ with

$$v^\flat(w) := g_p(v,w), \tag{79}$$

yielding a linear isomorphism

$$T_pM \to T_p^*M\colon \quad v \mapsto v^\flat. \tag{80}$$

If the tangent vector v is given by $v = v^\mu \partial_\mu$, then the coefficients of the covector $v^\flat$ are given by $v^\mu g_{\mu\nu}$, since $v^\flat_\nu = v^\flat(\partial_\nu) = g(v, \partial_\nu)$. Since coefficients of vectors carry a superindex and coefficients of covectors carry a subindex, one often simply writes

$$v_\nu = g_{\mu\nu} v^\mu$$

for the coefficients of $v^\flat$. We say that the metric $g_{\mu\nu}$ is used to *lower the index.*

Conversely, let $\alpha = \alpha_\nu dx^\nu$ be a covector. Denote by $g^{\mu\nu}$ the inverse of the matrix $g_{\mu\nu}$, which has the property that $g_{\mu\sigma} g^{\sigma\nu} = \delta^\nu_\mu$. We can use $g^{\mu\nu}$ to *raise the index* of α_ν by setting $\alpha^\mu := g^{\mu\nu}\alpha_\nu$. The vector $\alpha^\sharp := \alpha^\mu \partial_\mu$ is the unique vector in T_pM that satisfies

$$\alpha(w) = g_p(\alpha^\sharp, w) \quad \text{for all} \quad w \in T_pM.$$

To see this, note that

$$g_p(\alpha^\sharp, w) = g_{\mu\nu}\alpha^\mu w^\nu = g_{\mu\nu} g^{\mu\sigma} \alpha_\sigma w^\nu = \delta^\sigma_\mu \alpha_\sigma w^\nu = \alpha_\nu w^\nu.$$

The corresponding linear isomorphism

$$T_p^*M \to T_pM\colon \quad \alpha \mapsto \alpha^\sharp, \tag{81}$$

which maps $\alpha = \alpha_\nu dx^\nu$ to $\alpha^\sharp = \alpha^\mu \partial_\mu$, is of course inverse to (80).

Problem 9.8. Show that

$$g_{\mu\nu} v^\mu w^\nu = v_\nu w^\nu = v^\mu w_\mu = g^{\nu\mu} v_\nu w_\mu\,.$$

Problem 9.9. The linear maps

$$T_pM \to T_p^*M\colon v \mapsto v^\flat \quad \text{and} \quad T_p^*M \to T_pM\colon \alpha \mapsto \alpha^\sharp$$

are inverse to each other. Can you explain the notation for these *musical isomorphisms*?

9.1.5 Metrics on embedded submanifolds

If (M, g) is a Riemannian manifold and $\Sigma \subseteq M$ is an embedded submanifold, then Σ is itself a Riemannian manifold in a natural way. The metric on Σ is simply the pullback $\iota^* g$ of g along the canonical inclusion $\iota\colon \Sigma \hookrightarrow M$. Since $\iota_*\colon T_\sigma\Sigma \to T_\sigma M$ is injective, this is indeed a Riemannian metric by Proposition 9.6.

This gives us a rich source of Riemannian manifolds. Indeed, since $\mathbb{R}^n$ comes equipped with the Euclidean metric g^E, every k-dimensional embedded submanifold $\Sigma \subseteq \mathbb{R}^n$ comes with a natural metric $\iota^* g^E$. Since $T_\sigma\Sigma$ can be identified with a k-dimensional linear subspace of $T_\sigma\mathbb{R}^n \simeq \mathbb{R}^n$, every tangent vector $v \in T_\sigma\Sigma$ can be written as $v = (v^1, \ldots, v^n)$ with respect to the cartesian coordinates on $\mathbb{R}^n$. The inner product $\iota^* g^E_\sigma(v, w)$ is then simply $v^1 w^1 + \ldots + v^n w^n$.

Example 9.10 (Round metric on the sphere)**.** The n-sphere $\mathbb{S}^n$ is an embedded submanifold of $\mathbb{R}^{n+1}$, with n-dimensional tangent space

$$T_p\mathbb{S}^n \simeq \{v \in \mathbb{R}^{n+1}\,;\, v \perp p\}.$$

The *round metric* is the restriction $\iota^* g^E_\sigma$ of the Euclidean metric to the sphere,

$$\iota^* g^E_\sigma(v, w) = v^1 w^1 + \ldots + v^{n+1} w^{n+1}.$$

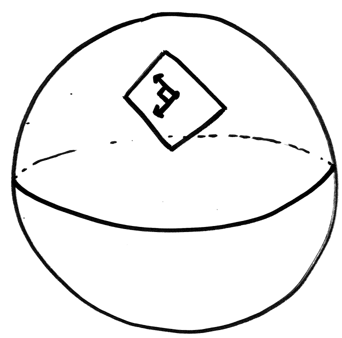


Figure 16: The round metric on the 2-sphere.

The following problem illustrates that although diffeomorphisms only care about the 'shape' of a manifold, isometries also care about 'size'.

Problem 9.11. Let $\mathbb{S}^2_r$ be the sphere in $\mathbb{R}^3$ of radius $r > 0$. Show that the diffeomorphism $\phi\colon \mathbb{S}^2_r \to \mathbb{S}^2_R$ with $\phi(\vec{x}) = (R/r)(\vec{x})$ is an isometry between $\mathbb{S}^2_r$ and $\mathbb{S}^2_R$ if and only if $r = R$.

Problem 9.12. Let $\mathbb{S}^2$ be the unit sphere in $\mathbb{R}^3$ of radius 1. Recall from §2.3.1 that $\mathbb{S}^2$ is covered by two charts (U_1, ϕ_1) and (U_2, ϕ_2).

a) Determine the coefficients $g_{\mu\nu}$ of the round metric with respect to the coordinates $(x, y) = \phi_1(\xi, \eta, \zeta) = (\frac{\xi}{1-\zeta}, \frac{\eta}{1-\zeta})$ on $\mathbb{S}^2 \setminus \{(0, 0, 1)\}$.

b) Determine the coefficients $g_{\overline{\mu\nu}}$ of the same metric with respect to the coordinates $(\overline{x}, \overline{y}) = \phi_2(\xi, \eta, \zeta) = (\frac{\xi}{1+\zeta}, \frac{\eta}{1+\zeta})$ on $\mathbb{S}^2 \setminus \{(0, 0, -1)\}$.

c) Recall from Problem 2.11 that the transition function between the two charts is $(x, y) = \frac{1}{\overline{x}^2+\overline{y}^2}(\overline{x}, \overline{y})$. Check that $g_{\overline{\mu\nu}} = \left(\frac{\partial x^\mu}{\partial x^{\overline{\mu}}}\right)\left(\frac{\partial x^\nu}{\partial x^{\overline{\nu}}}\right) g_{\mu\nu}$.

Problem 9.13 (Hyperbolic geometry)**.** Let $H = \{(u, v, w) \in \mathbb{R}^3\,;\, u^2 - v^2 - w^2 = 1\}$ be the hyperboloid in $\mathbb{R}^3$, and let $H^+ = \{(u, v, w) \in H\,;\, u \geq 1\}$ be a the upper sheet of H. Let $D = \{(x, y) \in \mathbb{R}^2\,;\, x^2 + y^2 < 1\}$ be the open unit disc.

a) Show that H^+ is an embedded submanifold of $\mathbb{R}^3$.

b) The line ℓ through $(u,v,w) \in H^+$ and $(-1,0,0) \in \mathbb{R}^3$ intersects the plane $u = 0$ in $(0,x,y)$. Determine x and y as a function of (u,v,w), show that $(x,y) \in D$, and show that the resulting stereographic projection $\phi\colon H^+ \to D$ is a smooth map.

c) For $(x,y) \in D$, the line ℓ through $(-1,0,0) \in \mathbb{R}^3$ and $(0,x,y) \in \mathbb{R}^3$ intersects H^+ in (u,v,w). Determine (u,v,w) as a function of (x,y), and show that $\psi(x,y) = (u,v,w)$ is a smooth map $D \to H^+$. Conclude that $\psi\colon D \to H^+$ is a diffeomorphism.

d) Let η be the covariant tensor field of rank 2 defined in cartesian coordinates by
$$\begin{pmatrix} \eta_{uu} & \eta_{uv} & \eta_{uw} \\ \eta_{vu} & \eta_{vv} & \eta_{vw} \\ \eta_{wu} & \eta_{wv} & \eta_{ww} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$
It is sometimes called the *Minkowski metric* on $\mathbb{R}^3$. Calculate the components
$$\begin{pmatrix} \psi^*\eta_{xx} & \psi^*\eta_{xy} \\ \psi^*\eta_{yx} & \psi^*\eta_{yy} \end{pmatrix}$$
of the tensor $\psi^*\eta$ with respect to the cartesian coordinates (x,y) on D.

e) Let $\iota\colon H^+ \hookrightarrow \mathbb{R}^3$ be the inclusion of H^+ into $\mathbb{R}^3$. Show that $(H^+, \iota^*\eta)$ is a Riemannian manifold which is isometric to the Poincaré disc (D,g). Explain the name "hyperbolic geometry" for the geometry of the Poincaré disc.

9.2 Length

Let $[a,b]$ be a closed interval. Then a curve $\gamma\colon [a,b] \to M$ is called *regular* if it is the restriction to $[a,b]$ of a smooth curve with $\dot\gamma(t) \neq 0$ for $t \in [a,b]$. The curve γ is called *piecewise regular* if there exists a finite subdivision $a = a_0 < a_1 < \ldots < a_N = b$ such that each $\gamma|_{[a_i,a_{i+1}]}$ is regular.

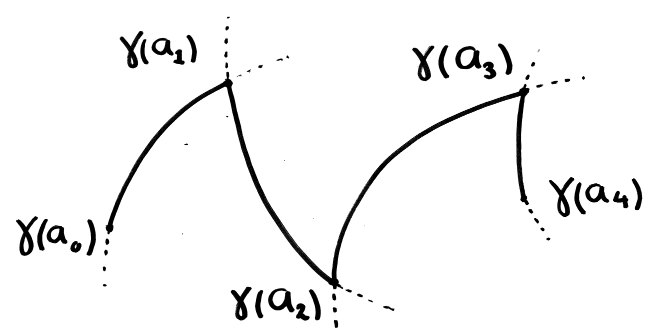


Figure 17: A piecewise smooth curve in $\mathbb{R}^2$.

Just like for curves in $\mathbb{R}^n$, we define the *length* of γ by

$$\mathcal{L}(\gamma) := \int_a^b \|\dot{\gamma}(t)\| dt\,, \tag{82}$$

where $\|\dot{\gamma}(t)\| := \sqrt{g_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))}$.

The length of a curve does not depend on the parameterization. Indeed, suppose that $\widetilde{\gamma} = \gamma \circ \phi$ for a piecewise smooth reparameterization $\phi\colon [c, d] \to [a, b]$ which is either increasing or decreasing. Then the chain rule yields

$$\dot{\widetilde{\gamma}}(t) = \dot{\gamma}(\phi(t))\phi'(t),$$

so that

$$\mathcal{L}(\widetilde{\gamma}) = \int_c^d \|\dot{\widetilde{\gamma}}(t)\| dt = \int_c^d \|\dot{\gamma}(\phi(t))\| |\phi'(t)| dt = \int_a^b \|\dot{\gamma}(s)\| ds = \mathcal{L}(\gamma).$$

For embedded submanifolds $\Sigma \subseteq \mathbb{R}^n$ with the Euclidean metric, this coincides with the usual definition of length.

Example 9.14. Let $\mathbb{S}^2_r$ be the 2-sphere of radius r with the round metric, and let $\gamma\colon [0, \Theta] \to \mathbb{S}^2$ be the curve

$$\gamma(t) := (r\cos(t), r\sin(t), 0).$$

Then $\mathcal{L}(\gamma) = \int_0^\Theta \|\dot{\gamma}(t)\| dt = \int_0^\Theta \|(-r\sin(t), r\cos(t), 0)\| dt = r\Theta$.

Example 9.15. Let D^2 be the Poincaré disk from Example 9.5, and let $\gamma\colon [0, r] \to D^2$ be the curve $\gamma(t) = (t, 0)$. Note that with respect to the metric on D^2, this is *not* a path of constant speed. Indeed, at $(x, y) = (t, 0)$, the tangent vector $\dot{\gamma}(t) = \partial_x$ has length $2/(1 - t^2)$. It follows that

$$\mathcal{L}(\gamma) = \int_0^r \|\partial_x\| dt = \int_0^r 2/(1 - t^2) dt = \log\left(\frac{1 + r}{1 - r}\right).$$

To make $\gamma\colon [a, b] \to M$ into a path of constant speed, we can always reparametrize by the arc length

$$s(t) := \int_a^t \|\dot{\gamma}(t)\| dt.$$

Indeed, if γ is regular, then $ds/dt = \|\dot{\gamma}(t)\| \neq 0$, so the strictly increasing function $s(t)$ has a smooth, strictly increasing inverse $t(s)$. Then $\tilde{\gamma}(s) := \gamma(t(s))$ is a unit speed parameterization, since $\frac{d}{ds}\tilde{\gamma}(s) = \|\dot{\gamma}(t(s))\|(dt/ds) = 1$.

Problem 9.16. Calculate the unit speed parameterization of the curve in Example 9.14.

Problem 9.17. Show that the unit speed reparameterization of the path in Example 9.15 is $(x(s), y(s)) = (\frac{e^s - 1}{e^s + 1}, 0)$.

a) Use $s(t) = \log((1+t)/(1-t))$ to calculate $t(s)$.

b) Calculate $\tilde{\gamma}(s) := \gamma(t(s))$, and check that its speed $\|\dot{\tilde{\gamma}}(s)\|$ of $\tilde{\gamma}(s) := \gamma(t(s))$ is indeed 1.

c) Conclude that γ approaches the boundary of the disk in finite time, but with infinite speed. On the other hand $\widetilde{\gamma}$ has unit speed, but it needs an infinite amount of time to reach the boundary.

Definition 9.18. The *distance* $d(p,q)$ between two points p and q on M is defined as

$$d(p,q) := \inf\{\mathcal{L}(\gamma)\,;\, \gamma(a) = p \text{ and } \gamma(b) = q\}, \tag{83}$$

where the infimum is taken over all piecewise regular curves $\gamma\colon [a,b] \to M$ that start at p and end at q.

Proposition 9.19. *For a Riemannian manifold, the distance $d\colon M \times M \to \mathbb{R}^+$ is a metric.*

Proof. If γ is a path from p to q, then then $\gamma(-t)$ is a path from q to p with the same length. It follows that $d(p,q) = d(q,p)$.

If $\gamma_1\colon [a_1,b_1] \to M$ is a path from p to x and $\gamma_2\colon [a_2,b_2] \to M$ is a path from x to q, then the concatenated path from p to q via x, defined by

$$\gamma_2 * \gamma_1 := \begin{cases} \gamma_1(t) & \text{for } t \in [a_1,b_1] \\ \gamma_2(t - b_1 + a_2) & \text{for } t \in [b_1, b_1 + b_2 - a_2], \end{cases}$$

has length $\mathcal{L}(\gamma_1) + \mathcal{L}(\gamma_2)$. It follows that $d(p,q) \le d(p,x) + d(x,q)$.

Finally, we show that $d(p,q) = 0$ implies $p = q$. Suppose that $p \ne q$. Let $U_\alpha \subseteq M$ be a coordinate neighbourhood of p, and let $\overline{B}_\varepsilon$ be a closed coordinate ball centered at $\phi_\alpha(p)$ which does not contain $\phi_\alpha(q)$. We prove that there exist constants $C_1, C_2 > 0$ such that for all $v \in TM$ with $\phi_\alpha(\pi(v)) \in \overline{B}_\varepsilon$,

$$C_1\sqrt{(v^1)^2 + \ldots + (v^n)^2} \le \|v\|_g \le C_2\sqrt{(v^1)^2 + \ldots + (v^n)^2}. \tag{84}$$

Once we have proven (84), we can conclude that $d(p,q) > C_1\varepsilon$. Indeed, the length with respect to g of a path from p to q is at least C_1 times the *Euclidean* length of any segment of the path – or at least of the segment of the path that stays within $\phi_\alpha(\overline{B}_\varepsilon)$. Since the path eventually leaves the coordinate ball, its Euclidean length must be at least ε, so

$$d(p,g) = \inf\{\mathcal{L}(\gamma)\,;\, \gamma(a) = p, \gamma(b) = q\} \ge C_1\varepsilon\,.$$

Since $\|\lambda v\|_g = |\lambda|\|v\|_g$ for all $\lambda \in \mathbb{R}$, it suffices to prove (84) for vectors of Euclidean length $(v^1)^2 + \ldots + (v^n)^2 = 1$. Let

$$\overline{B}_\varepsilon \times \mathbb{S}^n \subseteq \mathbb{R}^n \times \mathbb{R}^n$$

be the set of coordinates of such vectors, with the additional condition that their base point lies in the coordinate ball. Since the length $\|v\|_g = \sqrt{g_{\mu\nu}v^\mu v^\nu}$

with respect to the *other* metric g is a smooth, positive map on the compact set $\overline{B}_\varepsilon \times \mathbb{S}^n$, it has a minimum value $C_1 > 0$ and a maximum value $C_2 > 0$. Since $C_1 \leq \|v\|_g \leq C_2$ for vectors of Euclidean length 1, we have (84) for vectors of arbitrary length. $\square$

If $\phi\colon M \to N$ is an isometry between (M, g) and (N, h), then the length with respect to g of $\gamma\colon [a, b] \to M$ is the same as the length with respect to h of $\phi^*\gamma\colon [a, b] \to N$. Indeed, since $\|\phi_* v\|_h = \|v\|_g$ for all $v \in T_M$, we have

$$\mathcal{L}(\phi^*\gamma) = \int_a^b \|\tfrac{d}{dt}\phi(\gamma(t))\|_h dt = \int_a^b \|\phi_*\dot{\gamma}(t)\|_h dt = \int_a^b \|\dot{\gamma}(t)\|_g dt = \mathcal{L}(\gamma).$$

In particular, $d(\phi(p), \phi(q)) = d(p, q)$. An isometry of Riemannian manifolds is therefore also an isometry of the corresponding metric spaces.

Remark 9.20 (Metrizability)**.** If g is a Riemannian metric on a manifold M, then the corresponding metric d induces a *metric topology* on M, cf. Def. A.15. Using the estimate (84), one can show that this metric topology is homeomorphic to the original topology on M (see [L97, Thm. 13.29] for details). In particular, M is *metrizable* as a topologial space, meaning that there exists a metric whose open balls generate the topology. Since every smooth manifold admits a Riemannian metric (see [L97, Prop. 13.3]), it follows that every smooth manifold is metrizable as a topological space.

9.3 The geodesic equation

For a Riemannian manifold (M, g), we are interested in the shortest path γ between a pair of points $p, q \in M$. Since the length of a path is invariant under reparametrization, we may as well assume that our paths are parametrized by arc length. A *geodesic* is a curve γ with speed 1 which is *locally* the shortest path between all the points that it connects.

Since γ moves at unit speed, the length of the curve segment from t to t' is

$$\mathcal{L}(\gamma|_{[t,t']}) = \int_t^{t'} \|\dot{\gamma}\| dt = |t' - t|.$$

It follows that $\gamma|_{[t,t']}$ is the shortest path from $\gamma(t)$ to $\gamma(t')$ if and only if $d(\gamma(t), \gamma(t')) = |t' - t|$.

Definition 9.21 (Geodesics)**.** A *geodesic* on M is a piecewise regular curve $\gamma\colon [a, b] \to M$ with unit speed $\|\dot{\gamma}(t)\| = 1$, such that for all $t_0 \in [a, b]$, there exists an open interval $I \subseteq [a, b]$ around t_0 with

$$d(\gamma(t), \gamma(t')) = |t - t'| \tag{85}$$

for all $t, t' \in I$.

Example 9.22. The curve $\gamma\colon \mathbb{R} \to \mathbb{S}^2$ with $\gamma(t) = (\cos(t), \sin(t), 0)$ is a geodesic on $\mathbb{S}^2$ with respect to the round metric. As long as $|t' - t| < \pi$, the curve $\gamma|_{[t,t']}$ is *the* shortest path from $\gamma(t)$ to $\gamma(t')$. For $|t - t'| = \pi$, this curve is *a* shortest path. And for $|t' - t| > \pi$, it is no longer a shortest path.

Theorem 9.23 (Geodesic equation)**.** *Every geodesic is regular. In local coordinates, it satisfies the second order ODE*

$$\ddot{\gamma}^\mu + \Gamma^\mu_{\sigma\tau}\dot{\gamma}^\sigma\dot{\gamma}^\tau = 0, \tag{86}$$

with the so-called Christoffel symbols $\Gamma^\mu_{\sigma\tau}$ *given by*

$$\Gamma^\mu_{\sigma\tau} = \tfrac{1}{2}\, g^{\mu\alpha}(\partial_\sigma g_{\tau\alpha} + \partial_\tau g_{\alpha\sigma} - \partial_\alpha g_{\sigma\tau}). \tag{87}$$

This is a central result in geometry because it gives us insight into the nature of geodesics. For example, since they are solutions to a second order ODE, they are locally determined by their initial position and velocity alone! But perhaps more importantly, the variational techniques in the proof are ubiquitous not just in geometry, but also in other areas of mathematics and mathematical physics.

The idea of this variational principle is the following. If γ is the shortest path from p to q, we can deform the path a little bit while keeping the endpoints fixed. This yields a 1-parameter family of paths $\gamma_\varepsilon\colon [a, b] \to M$, where $\gamma_0 = \gamma$ is the original path γ and the paths γ_ε approach γ as ε tends to zero. Since γ_0 is the shortest of the paths γ_ε from p to q, the function $\varepsilon \mapsto \mathcal{L}(\gamma_\varepsilon)$ has a minimum at $\varepsilon = 0$. We thus have $\frac{d}{d\varepsilon}\mathcal{L}(\gamma_\varepsilon) = 0$. This yields an integral equation for $\gamma(t)$ involving the infinitesimal deformation $\delta\gamma(t) := \frac{d}{d\varepsilon}\gamma_\varepsilon(t)$ in $T_{\gamma(t)}M$. From the fact that this integral equation must hold for *all* variations $\delta\gamma$ that vanish at the endpoints, we then obtain the geodesic equation.

In general, a (partial) differential equation obtained in this manner is called *Euler–Lagrange equation.* Interestingly, dynamical equations in physics are virtually always Euler–Lagange equations.

Proof. Choose a subdivision $a = a_0 < a_1 < \ldots < a_n = b$ such that the restriction of γ to every interval $[a_i, a_{i+1}]$ is regular, $\gamma([a_i, a_{i+1}])$ lies entirely within a coordinate neighbourhood U_{α_i}, and, most importantly, γ is is the shortest path from $\gamma(a_i)$ to $\gamma(a_{i+1})$.

In local coordinates, the length of $\gamma|_{[a_i,a_{i+1}]}$ is given by

$$\mathcal{L}(\gamma|_{[a_i,a_{i+1}]}) = \int_{a_i}^{a_{i+1}} \sqrt{g_{\mu\nu}(\gamma(t))\dot{\gamma}^\mu\dot{\gamma}^\nu}dt.$$

Let $\gamma_\varepsilon(t)\colon [a, b] \to M$ be a regular path that starts and ends at the same point as γ, and whose restriction to $[a_i, a_{i+1}]$ has coordinate expression

$$\gamma^\mu_\varepsilon(t) = \gamma^\mu(t) + \varepsilon\delta\gamma^\mu(t)$$

for some smooth map $\delta\gamma^\mu\colon [a_i, a_{i+1}] \to \mathbb{R}^n$. We view this as a smooth deformation of $\gamma_0(t) = \gamma(t)$ in the direction of $\delta\gamma(t)$.

If $\delta\gamma^\mu(a_i) = \delta\gamma^\mu(a_{i+1}) = 0$, then the starting point $\gamma_\varepsilon(a_i)$ and the endpoint $\gamma_\varepsilon(a_{i+1})$ are fixed. Since γ is the shortest path from $\gamma(a_i)$ to $\gamma(a_{i+1})$, the smooth map $\varepsilon \mapsto \mathcal{L}(\gamma_\varepsilon|_{[a_i,a_{i+1}]})$ has a minimum at $\varepsilon = 0$. It follows that

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\mathcal{L}\left(\gamma_\varepsilon|_{[a_i,a_{i+1}]}\right) = \int_{a_i}^{a_{i+1}} \frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\sqrt{g_{\mu\nu}(\gamma_\varepsilon(t))\dot{\gamma}_\varepsilon^\mu\dot{\gamma}_\varepsilon^\nu}\,dt = 0. \tag{88}$$

For brevity, we will write $g_{\mu\nu}$ instead of $g_{\mu\nu}(\gamma(t))$ in the following. Since $\|\dot{\gamma}_0\| = \sqrt{g_{\mu\nu}\dot{\gamma}^\mu\dot{\gamma}^\nu} = 1$, we have

$$\frac{d}{d\varepsilon}\Big|_{\varepsilon=0}\sqrt{g_{\mu\nu}(\gamma_\varepsilon(t))\dot{\gamma}_\varepsilon^\mu\dot{\gamma}_\varepsilon^\nu} = \tfrac{1}{2}\Big((\partial_\alpha g_{\mu\nu})\dot{\gamma}^\mu\dot{\gamma}^\nu\delta\gamma^\alpha + g_{\mu\nu}\delta\dot{\gamma}^\mu\dot{\gamma}^\nu + g_{\mu\nu}\dot{\gamma}^\mu\delta\dot{\gamma}^\nu\Big). \tag{89}$$

By partial integration, we find

$$\int_{a_i}^{a_{i+1}}\Big(g_{\mu\nu}\dot{\gamma}^\mu\Big)\tfrac{d}{dt}\delta\gamma^\nu dt = -\int_{a_i}^{a_{i+1}}\tfrac{d}{dt}\Big(g_{\mu\nu}\dot{\gamma}^\mu\Big)\delta\gamma^\nu dt + \big[g_{\mu\nu}\dot{\gamma}^\mu(t)\delta\gamma^\nu(t)\big]_{a_i}^{a_{i+1}}. \tag{90}$$

If $\delta\gamma(a_i) = \delta\gamma(a_{i+1}) = 0$, the last term in (90) vanishes. Note that in (89), the last two terms are equal, $g_{\mu\nu}\delta\dot{\gamma}^\mu\dot{\gamma}^\nu = g_{\mu\nu}\dot{\gamma}^\mu\delta\dot{\gamma}^\nu$. Integrating (89) and substituting (90) for the last two terms, we find

$$0 = \int_{a_i}^{a_{i+1}}\Big(\tfrac{1}{2}(\partial_\alpha g_{\mu\nu})\dot{\gamma}^\mu\dot{\gamma}^\nu - \tfrac{d}{dt}\big(g_{\mu\alpha}\dot{\gamma}^\mu\big)\Big)\delta\gamma^\alpha dt. \tag{91}$$

Since this is true for *every* choice of $\delta\gamma^\alpha$ vanishing at a_i and a_{i+1}, we obtain the *Euler–Lagrange* equation

$$0 = \tfrac{1}{2}(\partial_\alpha g_{\mu\nu})\dot{\gamma}^\mu(t)\dot{\gamma}^\nu(t) - \tfrac{d}{dt}\big(g_{\mu\alpha}\dot{\gamma}^\mu(t)\big). \tag{92}$$

Indeed, if we write $A_\alpha(t)$ for the right hand side of (92) and choose $\delta\gamma^\alpha := A_\alpha(t)$, then (91) yields $\int_{a_i}^{a_{i+1}}\sum_{\alpha=1}^n |A_\alpha(t)|^2 = 0$, so the continuous function $A_\alpha(t)$ must be zero for all $t \in [a_i, a_{i+1}]$.

The Euler–Lagrange equation (92) is equivalent to the geodesic equation (86). To derive the latter, use $\frac{d}{dt}g_{\mu\alpha}(\gamma(t)) = \partial_\nu g_{\mu\alpha}\dot{\gamma}^\nu$ to find

$$\frac{d}{dt}(g_{\mu\alpha}\dot{\gamma}^\mu) = g_{\mu\alpha}\ddot{\gamma}^\mu + \partial_\nu g_{\mu\alpha}\dot{\gamma}^\mu\dot{\gamma}^\nu = g_{\mu\alpha}\ddot{\gamma}^\mu + \tfrac{1}{2}\Big(\partial_\nu g_{\mu\alpha}\dot{\gamma}^\mu\dot{\gamma}^\nu + \partial_\mu g_{\alpha\nu}\dot{\gamma}^\mu\dot{\gamma}^\nu\Big),$$

where the last step is the relabelling $\mu \leftrightarrow \nu$ of indices. The Euler–Lagrange equation (92) is therefore equivalent to

$$g_{\alpha\mu}\ddot{\gamma}^\mu + \tfrac{1}{2}\left(\partial_\nu g_{\mu\alpha} + \partial_\mu g_{\alpha\nu} - \partial_\alpha g_{\mu\nu}\right)\dot{\gamma}^\nu\dot{\gamma}^\mu = 0. \tag{93}$$

Contracting this with the inverse matrix $g^{\sigma\alpha}$ yields

$$\ddot{\gamma}^\sigma + \tfrac{1}{2}g^{\sigma\alpha}\left(\partial_\nu g_{\mu\alpha} + \partial_\mu g_{\alpha\nu} - \partial_\alpha g_{\mu\nu}\right)\dot{\gamma}^\nu\dot{\gamma}^\mu = 0,$$

which is precisely the geodesic equation (86) if we substitute $\sigma \leftrightarrow \mu$ and $\tau \leftrightarrow \nu$.

To see what happens at the points a_i, we repeat the argument for the whole interval $[a, b]$, with $\delta\gamma(a) = 0$ and $\delta\gamma(b) = 0$, but with $\delta\gamma(a_i) \neq 0$ for $i = 1, \ldots, n-1$. Taking into account the boundary terms in (90) and summing from $i = 0$ to $i = n$, we find

$$0 = \sum_{i=1}^{n-1} g_{a_i}(\dot{\gamma}_-(a_i) - \dot{\gamma}_+(a_i), \delta\gamma(a_i)).$$

Since this is zero for all choices of $\delta\gamma(a_i)$, we conclude that the left and right limits $\dot{\gamma}_-(a_i)$ and $\dot{\gamma}_+(a_i)$ of $\dot{\gamma}$ at a_i are equal, so the coordinate coefficients γ^μ are continuously differentiable around a_i. It follows that on the entire interval $[a, b]$, the curve γ is a C^1 solution to a second order ODE with smooth coefficients, so it must be smooth at a_i as well. $\square$

Example 9.24. Let D^2 be the Poincaré disk from Example 9.5, with the metric

$$\begin{pmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{pmatrix} = \begin{pmatrix} \frac{4}{(1-x^2-y^2)^2} & 0 \\ 0 & \frac{4}{(1-x^2-y^2)^2} \end{pmatrix}.$$

Since this is a diagonal matrix, the inverse matrix is easily computed:

$$\begin{pmatrix} g^{xx} & g^{xy} \\ g^{yx} & g^{yy} \end{pmatrix} = \begin{pmatrix} \frac{(1-x^2-y^2)^2}{4} & 0 \\ 0 & \frac{(1-x^2-y^2)^2}{4} \end{pmatrix}.$$

In order to compute the Christoffel symbols $\Gamma^\mu_{\sigma\tau}$, we need the partial derivatives of the metric. With $r^2 = x^2 + y^2$, we have $\partial_x g_{xy} = \partial_y g_{xy} = 0$, and

$$\begin{aligned} \partial_x g_{xx} &= \frac{16x}{(1-r^2)^3} & \qquad \partial_x g_{yy} &= \frac{16x}{(1-r^2)^3} \\ \partial_y g_{xx} &= \frac{16y}{(1-r^2)^3} & \qquad \partial_y g_{yy} &= \frac{16y}{(1-r^2)^3}. \end{aligned}$$

To calculate the Christoffel symbol Γ^x_{xy}, we have

$$\begin{aligned} \Gamma^x_{xy} &= \tfrac{1}{2} g^{xx}(\partial_x g_{xy} + \partial_y g_{xx} - \partial_x g_{xy}) + \tfrac{1}{2} g^{xy}(\partial_x g_{yy} + \partial_y g_{xy} - \partial_y g_{xy}) \\ &= \tfrac{1}{2} g^{xx} \partial_y g_{xx} = \frac{2y}{1-r^2}. \end{aligned}$$

The other Christoffel symbols are calculated in a similar fashion:

$$\begin{pmatrix} \Gamma^x_{xx} & \Gamma^x_{xy} \\ \Gamma^x_{yx} & \Gamma^x_{yy} \end{pmatrix} = \begin{pmatrix} \frac{2x}{1-r^2} & \frac{2y}{1-r^2} \\ \frac{2y}{1-r^2} & \frac{-2x}{1-r^2} \end{pmatrix}, \qquad \begin{pmatrix} \Gamma^y_{xx} & \Gamma^y_{xy} \\ \Gamma^y_{yx} & \Gamma^y_{yy} \end{pmatrix} = \begin{pmatrix} \frac{-2y}{1-r^2} & \frac{2x}{1-r^2} \\ \frac{2x}{1-r^2} & \frac{2y}{1-r^2} \end{pmatrix}.$$

The geodesic equation for $\gamma(t) = (x(t), y(t))$ then reads

$$\begin{aligned} \ddot{x} + \Gamma^x_{xx}\dot{x}^2 + 2\Gamma^x_{xy}\dot{x}\dot{y} + \Gamma^x_{yy}\dot{y}^2 &= 0 \\ \ddot{y} + \Gamma^y_{xx}\dot{x}^2 + 2\Gamma^y_{xy}\dot{x}\dot{y} + \Gamma^y_{yy}\dot{y}^2 &= 0, \end{aligned}$$

so we obtain

$$\begin{aligned} \ddot{x} + \frac{2}{1-r^2}\left(x\dot{x}^2 + 2y\dot{x}\dot{y} - x\dot{y}^2\right) &= 0 \\ \ddot{y} + \frac{2}{1-r^2}\left(-y\dot{x}^2 + 2x\dot{x}\dot{y} + y\dot{y}^2\right) &= 0. \end{aligned}$$

In Problem 9.17, we saw that $(x(s), y(s)) = (\frac{e^s-1}{e^s+1}, 0)$ is a regular path of unit speed. It is even a geodesic. Since $y = \dot{y} = 0$, the geodesic equation reduces to $\ddot{x} + \frac{2x\dot{x}^2}{1-x^2} = 0$, which is indeed satisfied by $x(s) = \frac{e^s-1}{e^s+1} = \tanh(s/2)$.

Problem 9.25 (Geodesics of $\mathbb{S}^2$)**.** Prove that on $\mathbb{S}^2$ with the round metric g, the geodesics are precisely the great circles.

a) A part of the sphere can be described by spherical coordinates (ϕ, θ). Calculate the coefficients
$$\begin{pmatrix} g_{\phi\phi} & g_{\phi\theta} \\ g_{\theta\phi} & g_{\theta\theta} \end{pmatrix}$$
for the round metric g in these coordinates, and calculate its inverse matrix
$$\begin{pmatrix} g^{\phi\phi} & g^{\phi\theta} \\ g^{\theta\phi} & g^{\theta\theta} \end{pmatrix}.$$

b) Show that $\partial_\theta g_{\phi\phi}$ is the only nonzero partial derivative among the $\partial_\alpha g_{\beta\gamma}$, and calculate the Christoffel symbols $\Gamma^\mu_{\sigma\tau}$.

c) Let $\gamma(t) = (\cos(t), \sin(t), 0)$ be the unit speed parameterization of the meridian. In spherical coordinates this curve is described by $(\phi(t), \theta(t)) = (t, \pi/2)$. Derive the geodesic equation for the sphere, and show that it is satisfied by this curve.

d) Now let $\gamma'(t)$ be *any* great circle on $\mathbb{S}^2$, parameterized by arc length. Find coordinates (ϕ', θ') in which $\gamma'(t)$ is given by $(\phi'(t), \theta'(t)) = (t, \pi/2)$, and show that $\gamma'(t)$ is a geodesic. (This should not take more than two lines of text and a sketch.)

e) Let $\gamma(t)$ be a geodesic. Let $\gamma'(t)$ be the unique great circle with $\gamma(0) = \gamma'(0)$ and $\frac{d}{dt}\gamma(0) = \frac{d}{dt}\gamma'(0)$. Use the uniqueness of solutions to second order ODE's to show that $\gamma = \gamma'$.

f) Conclude that the geodesics on $\mathbb{S}^2$ are precisely the great circles parameterized by arc length.

The following problem was posed by Johann Bernouilli in 1696.

Problem 9.26 (The Brachistochrone)**.** Suppose that a particle of mass m moves in $\mathbb{R}^2$ under the influence of a gravitational potential $V(x, y) = -gy$. If it traverses a curve $\gamma(s) = (x(s), y(s))$ with $\gamma(0) = \dot{\gamma}(0) = (0, 0)$, then its kinetic energy at $\gamma(s) = (x(s), y(s))$ is $\frac{1}{2}m\|\dot{\gamma}\|^2 = mgy$, so its speed is $\|\dot{\gamma}(s)\| = \sqrt{2gy}$. The total time it takes to move along γ from $\gamma(s_i)$ to $\gamma(s_f)$ is therefore

$$T = \int_{s_i}^{s_f} \frac{1}{\sqrt{2gy}}\sqrt{\dot{x}^2 + \dot{y}^2}ds.$$

a) The *brachistochrone* is the path γ that minimizes the travel time T. Show that the brachistochrone satisfies

$$\frac{d}{ds}\left(\frac{\dot{x}}{\sqrt{y(\dot{x}^2+\dot{y}^2)}}\right) = 0 \tag{94}$$

$$\frac{d}{ds}\left(\frac{\dot{y}}{\sqrt{y(\dot{x}^2+\dot{y}^2)}}\right) = -\frac{1}{2}\frac{\sqrt{\dot{x}^2+\dot{y}^2}}{y^{3/2}}. \tag{95}$$

b) The *cycloid* is the path that a point on a rigid circle traverses if the circle rolls over the x-axis without slipping. Show that for a circle with radius r, the cycloid is given by $\gamma(s) = (x(s), y(s))$ with

$$\begin{aligned} x(s) &= r(s - \sin(s)) \\ y(s) &= r(1 - \cos(s)). \end{aligned}$$

c) Show that the cycloid satisfies (94) and (95).

Problem 9.27 (Isometries preserve geodesics)**.** Let $\phi\colon M \to N$ be an isometry from (M, g^M) to (N, g^N).

a) Show that $\mathcal{L}(\phi \circ \gamma) = \mathcal{L}(\gamma)$ for every piecewise regular path $\gamma\colon [a,b] \to M$.

b) The path γ is a geodesic in M if and only if $\phi \circ \gamma$ is a geodesic in N.

Problem 9.28 (Geodesics of the Poincaré disc)**.** Recall from Problem 9.13 that the Poincaré disc (D, g) is isometric to the upper sheet $H^+ = \{(u,v,w) \in \mathbb{R}^3\,;\, u^2 - v^2 - w^2 = 1, u \geq 1\}$ of the hyperboloid in $\mathbb{R}^3$, equipped with the pullback $g = \iota^*\eta$ of the Minkowski metric.

a) The orthochronous Lorentz group $\mathrm{O}^{\uparrow}(1,2)$ from Problem 1.14 acts by isometries on H^+.

b) This induces an action of $\mathrm{O}^{\uparrow}(1,2)$ on TH^+ by pushforward, $g \cdot v = g_* v$. This action is transitive on the unit vectors $\{v \in TH^+\,;\, g(v,v) = 1\}$.

c) Let $g \in \mathrm{O}^{\uparrow}(1,2)$. Using Problem 9.27 or otherwise, show that if $t \mapsto \gamma(t)$ is a geodesic in H^+, then so is $t \mapsto g\gamma(t)$.

d) Conclude that $\mathrm{O}^{\uparrow}(1,2)$ acts transitively on the geodesics.

e) The intersection of H^+ with the plane P_0 defined by $w = 0$ is (the image of) a geodesic.
Hint: consider the image under the reflection $R(u,v,w) = (u,v,-w)$ *of the unique geodesic through* $(1,0,0) \in H^+$ *in the direction* $(0,1,0) \in T_{1,0,0}H^+$.

f) Every plane P through the origin in $\mathbb{R}^3$ that intersects H^+ is of the form $P = g(P_0)$ for an element $g \in \mathrm{O}^{\uparrow}(1,2)$.

g) Every geodesic in H^+ has as image the intersection of H^+ with a plane through the origin.

h) By rotating around the u-axis, every plane through the origin that intersects H^+ is of the form $u = \lambda v$ for some $\lambda > 1$. The corresponding geodesic is then given by the equations $u^2 - v^2 - w^2 = 1$, $u = \lambda v$. Its image under the isometry $\phi\colon H^+ \to D$ from Problem 9.13b is given by the segment of the circle $(x-\lambda)^2 + y^2 = \lambda^2 - 1$ that lies within D. It intersects $x^2 + y^2 = 1$ at a right angle.

9.4 The Levi–Civita connection

The requirement that γ be a geodesic is, of course, entirely independent of the coordinates that one chooses to describe it. In our derivation of the geodesic equation, however, we made rather extensive use of local coordinates. In particular the Christoffel symbols $\Gamma^{\mu}_{\sigma\tau}$ we discovered depend very much on the coordinates one chooses. We are led to wonder whether all Christoffel symbols together – each one defined in its own coordinate neighbourhood – might represent a coordinate-independent object defined on all of M. This is indeed the case, and this object is called the *Levi–Civita connection*.

Just like a vector field is a rule for differentiating functions, a *connection* is a rule for differentiating vector fields. A connection assigns to every pair of vector fields $v, w \in \mathrm{Vec}(M)$ the vector field $\nabla_v w \in \mathrm{Vec}(M)$, called the *covariant derivative* of w along v.

Definition 9.29 (Connections)**.** A *connection* is a bilinear map

$$\nabla\colon\ \mathrm{Vec}(M) \times \mathrm{Vec}(M) \to \mathrm{Vec}(M),$$

denoted $(v, w) \mapsto \nabla_v w$, such that

$$\nabla_{fv} w = f \nabla_v w, \quad \text{and} \tag{96}$$

$$\nabla_v (fw) = f \nabla_v w + v(f) w \tag{97}$$

for all $f \in C^{\infty}(M)$ and for all $v, w \in \mathrm{Vec}(M)$.

The second condition (97) is a Leibniz rule, which is what one expects since w is being differentiated. Note that the first condition (96) is not a Leibniz rule; it expresses that the covariant derivative $(\nabla_v w)_p$ at the point $p \in M$ only depends on the value $v_p \in T_p M$ of the vector field v at p, and not on any of its derivatives.

If $\nabla_v w = 0$, then we say that w is *covariantly constant* in the direction of v. Note that without a connection, the statement that a vector field w is constant does not make coordinate-invariant sense. For instance, the vector field $w = \partial_{\phi}$ appears to be constant when viewed in polar coordinates r, ϕ. But the same vector field $w = x\partial_y - y\partial_x$ has nonconstant coefficients when expressed in Cartesian coordinates x, y! Essentially, the reason that we cannot decide whether w is constant or not, is that we cannot compare vectors $w_p \in T_pM$ and $w_{p'} \in T_{p'}M$ at different points $p, p' \in M$ to see if they are 'the same'. By telling us what it means for a vector field w to be *constant* in the direction of v, a connection 'connects' the tangent spaces T_pM at different points p in M.

Example 9.30. Let ∇ be a connection on an open subset $U \subseteq \mathbb{R}^n$. Then we can express $\nabla_{\partial_\sigma}\partial_\tau$ (the covariant derivative of ∂_τ along ∂_σ) in terms of coordinate vector fields as

$$\nabla_{\partial_\sigma}\partial_\tau = A^\mu_{\sigma\tau}\partial_\mu. \tag{98}$$

The smooth functions $A^\mu_{\sigma\tau}$ on U completely determine the connection ∇. Indeed, we can calculate $\nabla_v w$ in terms of $A^\mu_{\sigma\tau}$ as follows:

$$\begin{aligned}
\nabla_v w &= \nabla_{v^\sigma\partial_\sigma}(w^\tau\partial_\tau) \\
&= (v^\sigma\partial_\sigma w^\tau)\partial_\tau + w^\tau\nabla_{v^\sigma\partial_\sigma}\partial_\tau \quad \text{(by the Leibniz rule (97))} \\
&= (v^\sigma\partial_\sigma w^\tau)\partial_\tau + v^\sigma w^\tau\nabla_\sigma\partial_\tau \quad \text{(by equation (96))} \\
&= (v^\sigma\partial_\sigma w^\mu)\partial_\mu + v^\sigma w^\tau A^\mu_{\sigma\tau}\partial_\mu \\
&= v^\sigma(\partial_\sigma w^\mu + A^\mu_{\sigma\tau}w^\tau)\partial_\mu.
\end{aligned}$$

In particular, this formula shows that at $p \in U$, the value $(\nabla_v w)_p \in T_pM$ of the covariant derivative of w along v depends only on:

- the *value* of the coefficients v^σ_p of the vector field v at p
- the *value* of the coefficients w^τ_p of the vector field w at p
- the *derivative* $v^\sigma_p\partial_\sigma w^\tau_p$ of the coefficients of w in the direction of v at p.

Note the lack of symmetry between v and w.

Proposition 9.31. *A connection ∇ on M restricts to a connection $\nabla|_U$ on any open subset $U \subseteq M$.*

Proof. Let $v, w \in \mathrm{Vec}(U)$, and let $p \in U$. To define $\nabla_v w$ at the point $p \in U$, choose a bump function $\phi \in C^\infty(M)$ with $\phi(p') = 1$ for p' in an open neighbourhood $V \subseteq U$ of p, and $\phi(p') = 0$ for $p' \notin U$ (see Lemma 7.4). This allows us to define $((\nabla|_U)_v w)_p := (\nabla_{\phi v}(\phi w))_p$. One checks that the definition of $((\nabla|_U)_v w)_p$ does not depend on the choice of ϕ, and that $\nabla|_U$ is indeed a connection on U. $\square$

Problem 9.32. Check that the definition of $((\nabla|_U)_v w)_p$ does not depend on the choice of ϕ, and that $\nabla|_U$ is indeed a connection on U.

In view of Proposition 9.31 and Example 9.30, a connection ∇ is completely described by its coordinate functions $A^\mu_{\sigma\tau}$ with respect a collection of charts (U_α, ϕ_α) that cover M. The transformation behaviour of the coordinate functions $A^\mu_{\sigma\tau}$ is a bit awkward, and involves *derivatives* of the Jacobi matrix.

Problem 9.33 (Transformation behaviour of $A^{\mu}_{\sigma\tau}$). Show that

$$A^{\overline{\mu}}_{\overline{\sigma}\overline{\tau}} = \frac{\partial}{\partial x^{\overline{\sigma}}}\Big(\frac{\partial x^{\mu}}{\partial x^{\overline{\tau}}}\Big)\Big(\frac{\partial x^{\overline{\mu}}}{\partial x^{\mu}}\Big) + \Big(\frac{\partial x^{\overline{\mu}}}{\partial x^{\mu}}\Big)\Big(\frac{\partial x^{\sigma}}{\partial x^{\overline{\sigma}}}\Big)\Big(\frac{\partial x^{\tau}}{\partial x^{\overline{\tau}}}\Big)A^{\mu}_{\sigma\tau}.$$

One way to do this is to express $\nabla_v w$ in the two different coordinate systems as $\nabla_v w = (v^{\sigma}\partial_{\sigma}w^{\mu} + A^{\mu}_{\sigma\tau}v^{\sigma}w^{\tau})\partial_{\mu}$ and $\nabla_v w = (v^{\overline{\sigma}}\partial_{\overline{\sigma}}w^{\overline{\mu}} + A^{\overline{\mu}}_{\overline{\sigma}\overline{\tau}}v^{\overline{\sigma}}w^{\overline{\tau}})\partial_{\overline{\mu}}$. Then use the transformation formulæ for v^{σ}, w^{μ}, and ∂_{μ}.

9.4.1 Torsion-free connections

Recall from §6.3 that we already have a way to determine the derivative of a vector field w along the flow of a vector field v, namely the Lie bracket

$$[v, w](f) := v(w(f)) - w(v(f)). \tag{99}$$

The most important difference between the Lie bracket $[v, w]$ and the covariant derivative $\nabla_v w$ is that $\nabla_{v_p} w$ is well defined if we only know the *value* of the vector field v at p, whereas the Lie bracket $[v, w]$ at p uses both the value and the derivative of v.

The connection ∇ is called *torsion-free* if

$$[v, w] = \nabla_v w - \nabla_w v. \tag{100}$$

More generally, $T(v, w) := \nabla_v w - \nabla_w v - [v, w]$ is called the *torsion* of the connection, but we will only consider torsion-free connections in these notes.

Remark 9.34. Equations (99) and (100) are similar in the sense that the Lie bracket is expressed as a difference of operators involving v and w. An important difference is that in (100), each of the two terms on the right hand side is a vector field, whereas in (99), the difference is a vector field but the two individual terms are *not* vector fields.

9.4.2 Metric-preserving connections

If (M, g) is a Riemannian manifold, it is natural to impose the following compatibility condition.

Definition 9.35. A connection ∇ on a Riemannian manifold (M, g) is called *metric preserving* if for all $u, v, w \in \mathrm{Vec}(M)$,

$$u(g(v, w)) = g(\nabla_u v, w) + g(v, \nabla_u w). \tag{101}$$

If one blindly applies the chain rule on the left hand side of (101), one may expect to find an extra term $(\nabla_u g)(v, w)$ involving the derivative of the metric g along u. This term being absent, we can consider g to be constant as far as ∇ is concerned. Hence the name 'metric preserving connection'.

Although connections are plentiful, the requirement that ∇ be both torsion-free and metric-preserving is quite restrictive. In fact, for every Riemannian manifold there exists a unique connection with these properties. It is called the *Levi-Civita connection.*

Proposition 9.36 (Levi-Civita connection)**.** *For every Riemannian manifold (M, g), there is a unique connection that is both torsion-free and metric-preserving.*

Proof. Since ∇ is torsion-free, we have $\nabla_{\partial_\sigma}\partial_\tau - \nabla_{\partial_\tau}\partial_\sigma = [\partial_\sigma, \partial_\tau] = 0$. If we write D_σ for the covariant derivative ∇_{∂_σ} along ∂_σ, then this becomes

$$D_\sigma \partial_\tau = D_\tau \partial_\sigma. \tag{102}$$

Since ∇ is metric-preserving, we have the following expressions for $\partial_\sigma g_{\alpha\tau}$, $\partial_\tau g_{\sigma\alpha}$ and $\partial_\alpha g_{\sigma\tau}$.

$$\begin{aligned}
\partial_\sigma g(\partial_\alpha, \partial_\tau) &= \underline{g(D_\sigma\partial_\alpha, \partial_\tau)} + \underline{g(\partial_\alpha, D_\sigma\partial_\tau)} \\
\partial_\tau g(\partial_\sigma, \partial_\alpha) &= \underline{g(D_\tau\partial_\sigma, \partial_\alpha)} + \underline{g(\partial_\sigma, D_\tau\partial_\alpha)} \\
\partial_\alpha g(\partial_\sigma, \partial_\tau) &= \underline{g(D_\alpha\partial_\sigma, \partial_\tau)} + \underline{g(\partial_\sigma, D_\alpha\partial_\tau)}
\end{aligned}$$

By (102) and the symmetry of g, the marked expressions in the above formulæ are equal. With $D_\sigma\partial_\tau = A^\nu_{\sigma\tau}\partial_\nu$, it follows that

$$\partial_\sigma g_{\alpha\tau} + \partial_\tau g_{\sigma\alpha} - \partial_\alpha g_{\sigma\tau} = 2g(\partial_\alpha, A^\nu_{\sigma\tau}\partial_\nu) = 2g_{\alpha\nu}A^\nu_{\sigma\tau},$$

and hence, since $g^{\mu\alpha}g_{\alpha\nu}A^\nu_{\sigma\tau} = A^\mu_{\sigma\tau}$, that

$$A^\mu_{\sigma\tau} = \frac{1}{2}g^{\mu\alpha}(\partial_\sigma g_{\alpha\tau} + \partial_\tau g_{\sigma\alpha} - \partial_\alpha g_{\sigma\tau}).$$

In other words, the coefficients of the Levi-Civita connection are precisely the Christoffel symbols! Since the coefficients $A^\mu_{\sigma\tau} = \Gamma^\mu_{\sigma\tau}$ determine the connection completely, the connection is unique. □

By Problem 9.33, the Christoffel symbols transform as

$$\Gamma^{\overline{\mu}}_{\overline{\sigma\tau}} = \frac{\partial}{\partial x^{\overline{\sigma}}}\Big(\frac{\partial x^\mu}{\partial x^{\overline{\tau}}}\Big)\Big(\frac{\partial x^{\overline{\mu}}}{\partial x^\mu}\Big) + \Big(\frac{\partial x^{\overline{\mu}}}{\partial x^\mu}\Big)\Big(\frac{\partial x^\sigma}{\partial x^{\overline{\sigma}}}\Big)\Big(\frac{\partial x^\tau}{\partial x^{\overline{\tau}}}\Big)\Gamma^\mu_{\sigma\tau}.$$

Note that this involves *second derivatives* of the transition functions.

Problem 9.37 (The gradient)**.** Let (M, g) be a Riemannian manifold, and let $h\colon M \to \mathbb{R}$ be a smooth function. The *gradient* of h in p is defined as the unique vector $\mathrm{grad}_p(h) \in T_pM$ with $g_p(v_p, \mathrm{grad}_p(h)) = v_p(h)$ for all $v_p \in T_pM$. Determine the coefficients of $\mathrm{grad}_p(h)$ in local coordinates, and show that $\mathrm{grad}(h)$ is a smooth vector field on M.

Problem 9.38 (The Brachistochrone in Riemannian geometry)**.** Let (M, g) be a Riemannian manifold, and let $h\colon M \to \mathbb{R}^{>0}$ be a smooth function.

a) Let ∇ be a connection on M. Show that $\widetilde{\nabla}\colon \mathrm{Vec}(M) \times \mathrm{Vec}(M) \to \mathrm{Vec}(M)$ defined by

$$\widetilde{\nabla}_u v = \nabla_u v + \frac{1}{2h}\Big(g(u, \mathrm{grad}(h))v + g(v, \mathrm{grad}(h))u - g(u, v)\mathrm{grad}(h)\Big)$$

is again a connection.

b) The metric $\widetilde{g} = hg$ on M is defined by pointwise multiplication, $\widetilde{g}_p(v_p, w_p) := h(p)g_p(v_p, w_p)$. Show that if ∇ is the Levi-Civita connection for g, then $\widetilde{\nabla}$ is the Levi-Civita connection for $\widetilde{g}$.

c) If a particle moving on M is constrained to have speed $c(p) > 0$ at $p \in M$, then the time it takes to traverse a (regular) curve $\gamma\colon [s_i, s_f] \to M$ is

$$T = \int_{s_i}^{s_f} \frac{1}{c(\gamma(s))} \sqrt{g_{\gamma(s)}(\dot{\gamma}(s), \dot{\gamma}(s))} ds.$$

A curve which minimizes the travel time T is called a *brachistochrone*. Show that a brachistochrone γ satisfies

$$\nabla_{\dot{\gamma}}\dot{\gamma} = \frac{1}{c}\Big(2\left(\frac{d}{ds}c(\gamma(s))\right)\dot{\gamma}(s) - g(\dot{\gamma}, \dot{\gamma})\mathrm{grad}(c)\Big).$$

d) Explain how this pertains to Problem 9.26.

Problem 9.39 (Covariant derivatives)**.** Let τ be a covariant tensor field of rank k on M, and let ∇ be a connection. For $w, v_1, \ldots, v_k \in \mathrm{Vec}(M)$, we define the *covariant derivative*

$$\nabla_w \tau(v_1, \ldots, v_k) := \mathcal{L}_w \tau(v_1, \ldots, v_k) - \sum_{i=1}^{k} \tau(v_1, \ldots, \nabla_w v_i, \ldots, v_k).$$

a) Show that this expression is $C^\infty(M)$-linear in all $k+1$ arguments:

$$\nabla_{gw}\tau(f_1 v_1, \ldots, f_k v_k) = g f_1 \cdots f_k \nabla_w \tau(v_1, \ldots, v_k)$$

for all $g, f_1, \ldots, f_k \in C^\infty(M)$.

b) Express $\tau_{\mu_1 \ldots \mu_k;\nu} := \nabla_{\partial_\nu}\tau(\partial_{\mu_1}, \ldots, \partial_{\mu_k})$ in terms of the partial derivatives $\partial_\nu \tau_{\sigma_1, \ldots, \sigma_k}$ of the coefficient functions of the tensor τ, and the coordinate functions $A^\sigma_{\mu\nu}$ of the connection ∇.

c) Express $\nabla_w \tau(v_1, \ldots, v_k)$ in terms of the functions $\tau_{\mu_1 \ldots \mu_k;\nu}$, and the coordinate coefficients w^ν, $v_1^{\mu_1}, \ldots, v_k^{\mu_k}$ of the vector fields. Conclude that $\nabla\tau$ is a covariant tensor field of rank $k+1$.

d) Let g be a Riemannian metric, and ∇ the corresponding Levi-Civita connection. Then $\nabla g = 0$.

9.5 Parallel transport

Let (M, g) be a Riemannian manifold, let $\gamma\colon [a, b] \to M$ be a smooth curve. A *vector field over* γ is a smooth map $v\colon [a, b] \to TM$ such that $v(t) \in T_{\gamma(t)}M$ for all $t \in [a, b]$.

We say that v is *covariantly constant* if the covariant derivative of $v(t)$ in the direction of the velocity $\dot{\gamma}(t)$ is zero,

$$\nabla_{\dot{\gamma}(t)} v(t) = 0. \tag{103}$$

In that case, $v(b)$ is called the *parallel transport* of $v(a)$ along the curve γ.

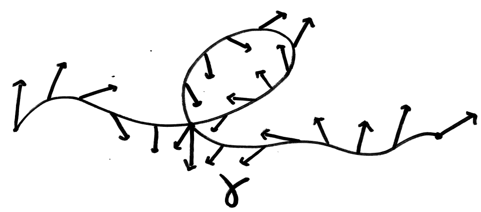


Figure 18: A vector field v over γ.

Remark 9.40. There is a slight subtlety when evaluating (103). We defined ∇ as a map from $\mathrm{Vec}(M) \times \mathrm{Vec}(M)$ to $\mathrm{Vec}(M)$, but neither v nor $\dot\gamma$ is defined on all of M. However, suppose for a second that v and $\dot\gamma$ extend to smooth vector fields $\tilde v$ and $\tilde{\dot\gamma}$ on an open subset of M, with $\tilde v|_{\gamma(t)} = v(t)$ and $\tilde{\dot\gamma}|_{\gamma(t)} = \dot\gamma(t)$ on the curve γ. Then the covariant derivative is

$$\begin{aligned} \nabla_{\tilde{\dot\gamma}} \tilde v|_{\gamma(t)} &= (\tilde{\dot\gamma}^\sigma \partial_\sigma \tilde v^\mu + \Gamma^\mu_{\sigma\tau} \tilde{\dot\gamma}^\sigma \tilde v^\tau)|_{\gamma(t)} \partial_\mu \\ &= \left(\tfrac{d}{dt} v^\mu(t) + \Gamma^\mu_{\sigma\tau}(\gamma(t)) \dot\gamma^\sigma(t) v^\tau(t)\right) \partial_\mu, \end{aligned}$$

where the last step uses the chain rule applied to $v^\mu(t) = \tilde v^\mu(\gamma(t))$. Note that $\nabla_{\tilde{\dot\gamma}} \tilde v|_{\gamma(t)}$ is entirely independent of the way in which we extend v and $\dot\gamma$ to smooth vector fields on M!

We would like to define $\nabla_{\dot\gamma(t)} v(t)$, as $\nabla_{\tilde{\dot\gamma}} \tilde v|_{\gamma(t)}$, but unfortunately, not every vector field on γ extends to a smooth vector field on M. Therefore, we simply use local coordinates to *define* the covariant derivative along γ by

$$\nabla_{\dot\gamma(t)} v(t) := \Big(\tfrac{d}{dt} v^\mu(t) + v^\sigma(t) \dot\gamma^\tau(t) \Gamma^\mu_{\sigma\tau}(\gamma(t))\Big) \partial_\mu. \tag{104}$$

It turns out that this expression is independent of the choice of coordinates.

Problem 9.41. Show that (104) is independent of the choice of coordinates, for example by using the solution to Problem 9.33. (For a more elegant proof of this fact, see [L97, Lemma 4.9].)

We can view $\nabla_{\dot\gamma(t)} \dot\gamma(t)$ as the *acceleration* of the curve γ. Since

$$\nabla_{\dot\gamma(t)} \dot\gamma(t) = \Big(\ddot\gamma^\mu + \dot\gamma^\sigma \dot\gamma^\tau \Gamma^\mu_{\sigma\tau}\Big) \partial_\mu,$$

we see that the vanishing of the acceleration, $\nabla_{\dot\gamma(t)} \dot\gamma(t) = 0$, is equivalent to the geodesic equation

$$\ddot\gamma^\mu + \dot\gamma^\sigma \dot\gamma^\tau \Gamma^\mu_{\sigma\tau} = 0.$$

Apparently, the curves with covariantly constant velocity $\dot\gamma(t)$ are precisely the geodesics! This gives us two equivalent ways to describe a geodesic:

(1) A geodesic is, locally, the shortest path between the points that it connects.

(2) A geodesic is a path with covariantly constant velocity.

9.6 Curvature and the Riemann tensor

In general, the result of parallel transport of a vector $v \in T_pM$ to a vector $v' \in T_{p'}M$ depends on the path γ from p to p'.

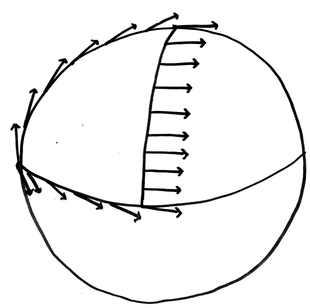


Figure 19: Parallel transport of the same vector along different paths.

So although the coordinate vector fields commute when acting on *functions*, $\partial_\mu \partial_\nu = \partial_\nu \partial_\nu$, they do not commute when acting on *vectors*, $D_\mu D_\nu \neq D_\nu D_\mu$. The failure of covariant derivatives to commute is measured by the *Riemann curvature tensor.*

Definition 9.42 (Riemann tensor)**.** The *Riemann curvature tensor* is the multilinear map

$$R\colon \operatorname{Vec}(M) \times \operatorname{Vec}(M) \times \operatorname{Vec}(M) \to \operatorname{Vec}(M)$$

defined by

$$R(u,v)w := (\nabla_u \nabla_v - \nabla_v \nabla_u - \nabla_{[u,v]})w. \tag{105}$$

Since $[\partial_\mu, \partial_\nu] = 0$ and $D_\mu := \nabla_{\partial_\mu}$, we have $R(\partial_\mu, \partial_\nu) = D_\mu D_\nu - D_\nu D_\mu$, so indeed the Riemann tensor captures the noncommutativity of covariant derivatives.

9.6.1 Coordinate representation of the Riemann tensor

Since the Riemann tensor is linear in each of its three arguments, you can pull a scalar $\lambda \in \mathbb{R}$ out of the Riemann tensor in three different ways, $R(\lambda u, v)w = R(u, \lambda v)w = R(u,v)(\lambda w) = \lambda R(u,v)w$. It turns out that you can even pull out a *function* in the same way. This is rather remarkable, because the Riemann tensor is constructed out of the Levi-Civita connection, which does *not* have this property – remember the Leibniz rule (97). The third term in (105) is added precisely to achieve this effect.

Proposition 9.43. *For all $u, v, w \in \operatorname{Vec}(M)$ and $f \in C^\infty(M)$, we have*

$$R(fu,v)w = fR(u,v)w \tag{106}$$

$$R(u,fv)w = fR(u,v)w \tag{107}$$

$$R(u,v)(fw) = fR(u,v)w. \tag{108}$$

Proof. We use (96) and (97) repeatedly to take the function f out of the brackets. Since

$$[u, fv] = u(f)v + f[u, v], \tag{109}$$

we find:

$$\begin{aligned}
R(u, fv)w &= \nabla_u \nabla_{fv} w - \nabla_{fv} \nabla_u w - \nabla_{[u,fv]} w \\
&= \nabla_u f \nabla_v w - f \nabla_v \nabla_u w - \nabla_{u(f)v+f[u,v]} w \\
&= \nabla_u f \nabla_v w - f \nabla_v \nabla_u w - \nabla_{u(f)v+f[u,v]} w \\
&= (f \nabla_u \nabla_v w + u(f) \nabla_v w) - f \nabla_v \nabla_u w - (u(f) \nabla_v w + f \nabla_{[u,v]} w) \\
&= f(\nabla_u \nabla_v w - \nabla_v \nabla_u w - \nabla_{[u,v]} w) = f R(u, v) w.
\end{aligned}$$

This proves (107). Equation (106) follows from this. Indeed, from the definition of the Riemann tensor, one readily sees that $R(u, v) = -R(v, u)$, so that $R(fu, v) = -R(v, fu) = -fR(v, u) = fR(u, v)$.

Finally, for (108), consider the three terms in

$$R(u, v)(fw) = \nabla_u \nabla_v (fw) - \nabla_v \nabla_u (fw) - \nabla_{[u,v]} (fw)$$

separately. We find

$$\begin{aligned}
\nabla_u \nabla_v (fw) &= \underline{u(v(f))}w + \underline{v(f)\nabla_u}w + \underline{u(f)\nabla_v w} + f \nabla_u \nabla_v w \\
-\nabla_v \nabla_u (fw) &= -\underline{v(u(f))}w - \underline{u(f)\nabla_v w} - \underline{v(f)\nabla_u}w - f \nabla_v \nabla_u w \\
-\nabla_{[u,v]} (fw) &= -\underline{[u, v](f)}w - f \nabla_{[u,v]} w,
\end{aligned}$$

and the marked symbols cancel to leave

$$\nabla_u \nabla_v (fw) - \nabla_v \nabla_u (fw) - \nabla_{[u,v]} (fw) = f\Big(\nabla_u \nabla_v (w) - \nabla_v \nabla_u (w) - \nabla_{[u,v]} (w)\Big),$$

or $R(u, v)(fw) = fR(u, v)w$, as required. □

In local coordinates, the Riemann tensor is completely determined by its values on the coordinate vector fields ∂_μ. On every coordinate patch $U_\alpha \subseteq M$, we define the functions $R^\alpha_{\beta\gamma\delta}$ as the coordinate coefficients of the vector field $R(\partial_\beta, \partial_\gamma)(\partial_\delta)$,

$$R(\partial_\beta, \partial_\gamma)(\partial_\delta) = R^\alpha{}_{\beta\gamma\delta} \partial_\alpha. \tag{110}$$

By Proposition 9.43, the Riemann tensor on arbitrary vector fields u, v, w is given in terms of $R^\alpha{}_{\beta\gamma\delta}$ by

$$\begin{aligned}
R(u, v)w &= R(u^\beta \partial_\beta, v^\gamma \partial_\gamma)(w^\delta \partial_\delta) \\
&= u^\beta v^\gamma w^\delta R(\partial_\beta, \partial_\gamma) \partial_\delta \\
&= u^\beta v^\gamma w^\delta R^\alpha{}_{\beta\gamma\delta} \partial_\alpha.
\end{aligned}$$

In particular, the *value* of the vector field $R(u,v)w$ at the point $p \in M$ depends only on the *values* $u_p, v_p, w_p \in T_pM$ at the same point p. The Riemann tensor therefore determines a multilinear map

$$R_p \colon T_pM \times T_pM \times T_pM \to T_pM \tag{111}$$

for every $p \in M$. Giving the Riemann tensor $R\colon \operatorname{Vec}(M) \times \operatorname{Vec}(M) \times \operatorname{Vec}(M) \to \operatorname{Vec}(M)$ is therefore equivalent to specifying the multilinear map $R_p \colon T_pM \times T_pM \times T_pM \to T_pM$ for every point of $p \in M$.

Since the Riemann tensor R is defined in terms of the Levi-Civita connection ∇, the coefficients $R^{\alpha}_{\beta\gamma\delta}$ of the Riemann tensor with respect to local coordinates can be expressed in terms of the Christoffel symbols $\Gamma^{\mu}_{\sigma\tau}$, which form the coordinate description of the Levi-Civita connection.

Proposition 9.44. *The coordinate expression for the Riemann tensor is*

$$R^{\alpha}{}_{\beta\gamma\delta} = \partial_\beta \Gamma^{\alpha}_{\gamma\delta} - \partial_\gamma \Gamma^{\alpha}_{\beta\delta} + \Gamma^{\alpha}_{\beta\sigma}\Gamma^{\sigma}_{\gamma\delta} - \Gamma^{\alpha}_{\gamma\sigma}\Gamma^{\sigma}_{\beta\delta}. \tag{112}$$

Proof. To calculate $R^{\alpha}{}_{\beta\gamma\delta}$ in terms of the Christoffel symbols, we evaluate

$$\begin{aligned}
R(\partial_\beta, \partial_\gamma)(\partial_\delta) &= \nabla_{\partial_\beta}\nabla_{\partial_\gamma}\partial_\delta - \nabla_{\partial_\gamma}\nabla_{\partial_\beta}\partial_\delta - \nabla_{[\partial_\beta,\partial_\gamma]}\partial_\delta \\
&= \nabla_{\partial_\beta}(\Gamma^{\sigma}_{\gamma\delta}\partial_\sigma) - \nabla_{\partial_\gamma}(\Gamma^{\sigma}_{\beta\delta}\partial_\sigma) \\
&= \partial_\beta\Gamma^{\sigma}_{\gamma\delta}\partial_\sigma + \Gamma^{\tau}_{\beta\sigma}\Gamma^{\sigma}_{\gamma\delta}\partial_\tau - \partial_\gamma\Gamma^{\sigma}_{\beta\delta}\partial_\sigma - \Gamma^{\tau}_{\gamma\sigma}\Gamma^{\sigma}_{\beta\delta}\partial_\tau \\
&= \Big(\partial_\beta\Gamma^{\alpha}_{\gamma\delta} - \partial_\gamma\Gamma^{\alpha}_{\beta\delta} + \Gamma^{\alpha}_{\beta\sigma}\Gamma^{\sigma}_{\gamma\delta} - \Gamma^{\alpha}_{\gamma\sigma}\Gamma^{\sigma}_{\beta\delta}\Big)\partial_\alpha.
\end{aligned}$$

In the first equality, we use that $[\partial_\beta, \partial_\gamma] = 0$. In the second equality, we use that $\nabla_{\partial_\gamma}\partial_\delta = \Gamma^{\sigma}_{\gamma\delta}\partial_\sigma$ by definition of the Christoffel symbols. In the third equality, we use the Leibniz rule (97), and the fourth equality is a reshuffling of the indices $\alpha \leftrightarrow \sigma$ and $\alpha \leftrightarrow \tau$. $\square$

Since the coordinate representation $R^{\alpha}_{\beta\gamma\delta}(x)$ of the multilinear map

$$R_p \colon T_pM \times T_pM \times T_pM \to T_pM$$

depends smoothly on the coordinate $x = \phi_\alpha(p)$, the Riemann curvature tensor is a mixed tensor field of rank $(3,1)$ in the sense of §8.4.2.

9.6.2 Parallel transport along an infinitesimal square

The coefficients $R^{\alpha}_{\beta\gamma\delta}$ can be interpreted as the effect of parallel transporting the tangent vector $\partial_\delta \in T_pM$ along an infinitesimal square spanned by ∂_β and ∂_γ.

Choose local coordinates in which $p \in M$ is represented by the origin $(0,\ldots,0)$, and identify a neighbourhood of $p \in M$ with a neighbourhood of the origin in $\mathbb{R}^n$. Since we only need the coordinates x^β and x^γ, we write $f(s,t)$ for a function evaluated at $x^\beta = s$ and $x^\gamma = t$, with all other coordinates equal to zero. Let

$$\gamma_{s,t} := (0,\ldots,0,s,0,\ldots,0,t,0,\ldots,0)$$

be the parameterized surface with $x^\beta = s$ and $x^\gamma = t$.

We consider the effect of parallel transport first along the line segment $\gamma_{0,t}$ for $t \in [0, \varepsilon]$, and then along the line segment $\gamma_{s,\varepsilon}$ for $s \in [0, \varepsilon]$, as opposed to first translating along $\gamma_{s,0}$ for $s \in [0, \varepsilon]$, and then along $\gamma_{\varepsilon,t}$ for $t \in [0, \varepsilon]$. The four paths are the sides of a square with area proportional to ε^2, see Fig. 20. Since the effect is the same to first order in ε, we need a second order Taylor expansion in ε to see the difference.

Proposition 9.45. *If we parallel transport the basis vector ∂_δ first along a coordinate vector field ∂_β and then along ∂_γ (for the same time ε), then the vector we get differs from the parallel transport of ∂_δ first along ∂_γ and then along ∂_β (for the same time ε) by*

$$\varepsilon^2 R^\alpha_{\beta\gamma\delta}\partial_\alpha + \mathcal{O}(\varepsilon^3).$$

Proof. By (104), the parallel transport equation $\nabla_{\dot\gamma(t)}v(t) = 0$ in local coordinates reads

$$\frac{d}{dt}v^\mu(t) = -\Gamma^\mu_{\sigma\tau}v^\sigma(t)\dot\gamma^\tau(t). \tag{113}$$

For the Taylor expansion, we will also need the second order derivative

$$\frac{d^2}{dt^2}v^\mu(t) = \Gamma^\mu_{\sigma\tau}\Gamma^\sigma_{\alpha\beta}v^\alpha(t)\dot\gamma^\beta(t)\dot\gamma^\tau(t) - \partial_\alpha\Gamma^\mu_{\sigma\tau}\dot\gamma^\alpha(t)v^\sigma(t)\dot\gamma^\tau(t) - \Gamma^\mu_{\sigma\tau}v^\sigma(t)\ddot\gamma^\tau(t), \tag{114}$$

which is obtained in a straightforward way by differentiating (113). (Remember that $\Gamma^\mu_{\sigma\tau}(t) = \Gamma^\mu_{\sigma\tau}(\gamma(t))$, so the chain rule yields $\frac{d}{dt}\Gamma^\mu_{\sigma\tau}(t) = \partial_\alpha\Gamma^\mu_{\sigma\tau}\dot\gamma^\alpha$.)

For the path $\gamma_{0,t}$ with initial condition $v(0) = \partial_\delta$, we find the Taylor expansion

$$\begin{aligned} v^\mu(0,\varepsilon) &= \delta^\mu_\delta + \varepsilon\frac{d}{dt}v^\mu|_{t=0} + \tfrac{1}{2}\varepsilon^2\frac{d^2}{dt^2}v^\mu|_{t=0} + \mathcal{O}(\varepsilon^3) \qquad (115)\\ &= \delta^\mu_\delta - \varepsilon\Gamma^\mu_{\delta\gamma}(0,0) + \tfrac{1}{2}\varepsilon^2\Big(\Gamma^\mu_{\sigma\gamma}(0,0)\Gamma^\sigma_{\delta\gamma}(0,0) - \partial_\gamma\Gamma^\mu_{\delta\gamma}(0,0)\Big) + \mathcal{O}(\varepsilon^3). \end{aligned}$$

We simply plugged in (113) and (114) in the Taylor formula, with $v^\mu(0) = \delta^\mu_\delta$, $\dot\gamma^\mu = \delta^\mu_\gamma$ and $\ddot\gamma^\mu = 0$. (There is a repeated sum over σ, but *not* over γ since both indices are downstairs.)

We now take $v^\mu(0,\varepsilon)$ as initial condition for the parallel transport equation along $\gamma_{s,\varepsilon}$, yielding the Taylor expansion

$$v^\mu(\varepsilon,\varepsilon) = v^\mu(0,\varepsilon) + \varepsilon\frac{d}{ds}\Big|_{s=0}v^\mu(s,\varepsilon) + \tfrac{1}{2}\varepsilon^2\frac{d^2}{ds^2}\Big|_{s=0}v^\mu(s,\varepsilon) + \mathcal{O}(\varepsilon^3). \tag{116}$$

In the first term in (116), we simply plug in (115). For the second term in (116), (113) yields $\varepsilon\frac{d}{ds}v^\mu(s,\varepsilon) = -\varepsilon\Gamma^\mu_{\sigma\beta}(0,\varepsilon)v^\sigma(0,\varepsilon)$. Again, we plug in (115) for $v^\sigma(0,\varepsilon)$, but now we also need the first order taylor expansion $\Gamma^\mu_{\sigma\beta}(0,\varepsilon) = \Gamma^\mu_{\sigma\beta}(0,0) + \varepsilon\partial_\gamma\Gamma^\mu_{\sigma\beta} + \mathcal{O}(\varepsilon^2)$ for the Christoffel symbols. This yields

$$\varepsilon\frac{d}{ds}v^\mu(0,\varepsilon) = -\varepsilon\Gamma^\mu_{\delta\beta}(0,0) + \varepsilon^2\Big(\Gamma^\mu_{\sigma\beta}(0,0)\Gamma^\sigma_{\delta\gamma}(0,0) - \partial_\gamma\Gamma^\mu_{\delta\beta}(0,0)\Big) + \mathcal{O}(\varepsilon^3).$$

Finally, using (114) for the third term in (116), we find

$$\tfrac{1}{2}\varepsilon^2 \frac{d^2}{ds^2} v^\mu(s,\varepsilon)|_{s=0} = \tfrac{1}{2}\varepsilon^2 \Big(\Gamma^\mu_{\sigma\beta}\Gamma^\sigma_{\delta\beta} - \partial_\beta \Gamma^\mu_{\delta\beta}\Big) + \mathcal{O}(\varepsilon^3). \tag{117}$$

Putting this together, we find the (mildly terrifying) expression

$$\begin{aligned} v^\mu(\varepsilon,\varepsilon) \quad = \quad & \delta^\mu_\delta - \varepsilon\Big(\Gamma^\mu_{\delta\gamma} + \Gamma^\mu_{\delta\beta}\Big) + \tfrac{1}{2}\varepsilon^2\Big(\Gamma^\mu_{\sigma\gamma}\Gamma^\sigma_{\delta\gamma} - \partial_\gamma\Gamma^\mu_{\delta\gamma}\Big) \\ & + \tfrac{1}{2}\varepsilon^2\Big(\Gamma^\mu_{\sigma\beta}\Gamma^\sigma_{\delta\beta} - \partial_\beta\Gamma^\mu_{\delta\beta}\Big) + \varepsilon^2\Big(\Gamma^\mu_{\sigma\beta}\Gamma^\sigma_{\delta\gamma} - \partial_\gamma\Gamma^\mu_{\delta\beta}\Big). \end{aligned} \tag{118}$$

If we translate first along the path $\gamma_{s,0}$ in the ∂_β-direction, and then along the path $\gamma_{\varepsilon,t}$ in the ∂_γ-direction, then of course we find the same expression with β and γ interchanged. Note that in (118), this only changes the last term on the right hand side. The difference between these two ways of transporting ∂_δ from $(0,0)$ to $(\varepsilon,\varepsilon)$ is therefore

$$\varepsilon^2\Big(\partial_\beta\Gamma^\mu_{\delta\gamma} - \partial_\gamma\Gamma^\mu_{\delta\beta} + \Gamma^\mu_{\sigma\beta}\Gamma^\sigma_{\delta\gamma} - \Gamma^\mu_{\sigma\gamma}\Gamma^\sigma_{\delta\beta}\Big)\partial_\mu = \varepsilon^2 R^\mu_{\beta\gamma\delta}\partial_\mu,$$

as required. □

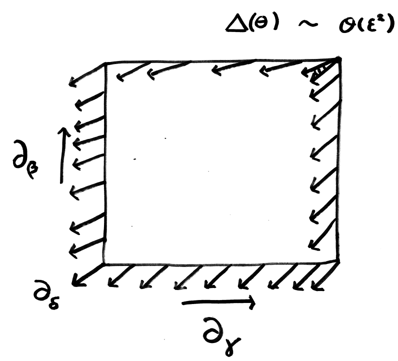


Figure 20: Parallel transport of ∂_δ along different sides of a coordinate square.

9.6.3 Symmetries of the Riemann tensor

The n^4 coefficients $R^\alpha_{\beta\gamma\delta}$ are not independent. In fact, the Riemann tensor has the following three symmetries.

Proposition 9.46 (Algebraic symmetries of the Riemann tensor)**.** *For all $u, v, w, z \in \mathrm{Vec}(M)$, we have*

$$\begin{aligned} 0 &= R(u,v)w + R(v,u)w & (119)\\ 0 &= g(R(u,v)w,z) + g(w,R(u,v)z) & (120)\\ 0 &= R(u,v)w + R(v,w)u + R(w,u)v. & (121) \end{aligned}$$

If we denote $R_{\alpha\beta\gamma\delta} := g_{\alpha\lambda}R^{\lambda}{}_{\beta\gamma\delta}$, then these three symmetries take the coordinate form

$$
\begin{aligned}
0 &= R^{\alpha}{}_{\beta\gamma\delta} + R^{\alpha}{}_{\gamma\beta\delta} && (122)\\
0 &= R_{\alpha\beta\gamma\delta} + R_{\delta\beta\gamma\alpha} && (123)\\
0 &= R^{\alpha}{}_{\beta\gamma\delta} + R^{\alpha}{}_{\gamma\delta\beta} + R^{\alpha}{}_{\delta\beta\gamma}. && (124)
\end{aligned}
$$

This cuts down the number of independent components from n^4 to $n^2(n^2-1)/12$.

Problem 9.47. Derive (122), (123) and (124) from (119), (120) and (121).

Equation (121), or its equivalent (124) in coordinates, is called the *first Bianchi identity.*

Proof. Equation (119) follows straight from the definition of the Riemann tensor. Comparing the two expressions

$$
\begin{aligned}
R(u,v)w &= \nabla_u\nabla_v w - \nabla_v\nabla_u w - \nabla_{[u,v]}w,\\
R(v,u)w &= \nabla_v\nabla_u w - \nabla_u\nabla_v w - \nabla_{[v,u]}w
\end{aligned}
$$

it follows from $[v,u]+[u,v]=0$ that $R(u,v)+R(v,u)=0$.

Equation (120) follows from the compatibility of the Levi-Civita connection ∇ with the metric. Using (101) repeatedly, we find

$$
\begin{aligned}
uvg(w,z) &= g(\nabla_u\nabla_v w,z) + g(\nabla_u w,\nabla_v z) + g(\nabla_v w,\nabla_u z) + g(w,\nabla_u\nabla_v z)\\
vug(w,z) &= g(\nabla_v\nabla_u w,z) + g(\nabla_v w,\nabla_u z) + g(\nabla_u w,\nabla_v z) + g(w,\nabla_v\nabla_u z)\\
[u,v]g(w,z) &= g(\nabla_{[u,v]w,z}) + g(w,\nabla_{[u,v]}z).
\end{aligned}
$$

Since $uvg(w,z) - vug(w,z) - [u,v]g(w,z) = 0$, we have

$$
0 = g\big((\nabla_u\nabla_v - \nabla_v\nabla_u - \nabla_{[u,v]})w,z\big) + g\big(w,(\nabla_u\nabla_v - \nabla_v\nabla_u - \nabla_{[u,v]})z\big)
$$

as required.

Just like (120) follows from the compatibility of ∇ with the metric g, the first Bianchi identity (121) follows from the compatibility of ∇ with the Lie bracket. Using the torsion-freeness relation (100) twice, we find

$$
\begin{aligned}
R(u,v)w + R(v,w)u + R(w,u)v &= \nabla_u\nabla_v w + \nabla_v\nabla_w u + \nabla_w\nabla_u v\\
&\quad -\nabla_v\nabla_u w - \nabla_w\nabla_v u - \nabla_u\nabla_w v\\
&\quad -\nabla_{[u,v]}w - \nabla_{[v,w]}u - \nabla_{[w,u]}v\\
&= \nabla_u(\nabla_v w - \nabla_w v) - \nabla_{[v,w]}u + \text{cyclic}\\
&= \nabla_u([v,w]) - \nabla_{[v,w]}u + \text{cyclic}\\
&= [u,[v,w]] + \text{cyclic} = 0.
\end{aligned}
$$

(By "+cyclic", we mean that similar terms with u,v,w replaced by their cyclic permutations v,w,u and w,u,v are added.) The last term is zero by the Jacobi identity, cf. Proposition 6.11. $\square$

The *second Bianchi identity* involves the *covariant derivative* of the Riemann tensor. For a mixed tensor of rank $(3, 1)$, the covariant derivative is defined on $u, v, w, z \in \mathrm{Vec}(M)$ by

$$(\nabla_u R)(v, w)(z) := \nabla_u(R(v, w)z) - R(\nabla_u v, w)z - R(v, \nabla_u w)z - R(v, w)\nabla_u z. \tag{125}$$

Problem 9.48. This expression is $C^\infty(M)$-linear in each of its four entries, and therefore defines a mixed tensor of rank $(4, 1)$ on M.

In local coordinates, it is therefore determined by the coefficients $R^\alpha_{\beta\gamma\delta;\eta}$ that arise from the coordinate vector fields,

$$(\nabla_{\partial_\eta} R)(\partial_\beta, \partial_\gamma, \partial_\delta) =: R^\alpha_{\beta\gamma\delta;\eta}\partial_\alpha.$$

They involve first order derivatives of the coefficients of the Riemann tensor.

Problem 9.49. Show that

$$R^\alpha{}_{\beta\gamma\delta;\eta} = \partial_\eta R^\alpha{}_{\beta\gamma\delta} + \Gamma^\alpha_{\eta\sigma} R^\sigma{}_{\beta\gamma\delta} - \Gamma^\sigma_{\eta\beta} R^\alpha{}_{\sigma\gamma\delta} - \Gamma^\sigma_{\eta\delta} R^\alpha{}_{\beta\gamma\sigma}.$$

Proposition 9.50 (Second Bianchi identity)**.**

$$(\nabla_u R)(v, w) + (\nabla_v R)(w, u) + (\nabla_w R)(u, v) = 0. \tag{126}$$

Proof. We consider ∇_u and $R(v, w)$ as linear operators from $\mathrm{Vec}(M)$ to $\mathrm{Vec}(M)$. Using (119), equation (125) can be rewritten as

$$(\nabla_u R)(v, w) = [\nabla_u, R(v, w)] + R(w, \nabla_u v) - R(v, \nabla_u w).$$

Since

$$\begin{aligned}
[\nabla_u, R(v, w)] + \text{cyclic} &= \quad [\nabla_u, [\nabla_v, \nabla_w]] - [\nabla_u, \nabla_{[v,w]}] + \text{cyclic} \\
&= -[\nabla_u, \nabla_{[v,w]}] + \text{cyclic} \\
&= -R(u, [v, w]) + \nabla_{[u,[v,w]]} + \text{cyclic} \\
&= -R(u, [v, w]) + \text{cyclic},
\end{aligned}$$

and since $[v, w] = \nabla_v w - \nabla_w, v$ by the torsion-freeness property, we find

$$\begin{aligned}
(\nabla_u R)(v, w) + \text{cyclic} &= -R(u, [v, w]) + R(w, \nabla_u v) - R(v, \nabla_u w) + \text{cyclic} \\
&= -R(u, [v, w]) + R(u, \nabla_v w - \nabla_w v) + \text{cyclic} \\
&= \quad 0.
\end{aligned}$$

□

In coordinates, the second Bianchi identity amounts to

$$R^\alpha_{\beta\gamma\delta;\eta} + R^\alpha_{\gamma\eta\delta;\beta} + R^\alpha_{\eta\beta\delta;\gamma} = 0,$$

a first order PDE in the coefficients of the Riemann tensor.

9.7 Scalar curvature and the 'Theorema Egregium'.

From the Riemann tensor, one can construct the *Ricci tensor* with coefficients

$$R_{\alpha\beta} := R^{\sigma}{}_{\sigma\alpha\beta}. \tag{127}$$

From $R^{\alpha}{}_{\beta} := g^{\alpha\sigma} R_{\sigma\beta}$, one then defines the *scalar curvature* or *Ricci scalar* as

$$R := R^{\sigma}{}_{\sigma}. \tag{128}$$

The Ricci tensor is a covariant tensor field of rank two. The scalar curvature is a covariant tensor field of rank zero, that is, a smooth function on M.

Problem 9.51. Using (124) or otherwise, show that the Ricci tensor is symmetric, $R_{\alpha\beta} = R_{\beta\alpha}$.

The following theorem, called *Theorema Egregium* (remarkable theorem) by its discoverer C.F. Gauss in 1827, lies at the very root of differential geometry.

Theorem 9.52 (Theorema Egregium)**.** *Let (M, g^M) and (N, g^N) be Riemannian manifolds with scalar curvature R^M and R^N, resectively. If M and N are isometric with isometry $\phi\colon M \to N$, then $R^M = \phi^* R^N$.*

Proof. Let (U_α, ϕ_α) be a chart for N around $q \in N$. Since $\phi\colon M \to N$ is a diffeomorphism, the composition $(\phi^{-1}(U_\alpha), \phi_\alpha \circ \phi)$ is a chart for M around $p := \phi^{-1}(q)$. The basis vectors ∂^M_μ of T_pM with respect to the chart $\phi_\alpha \circ \phi$ are related to the basis vectors ∂^N_μ of T_qN with respect to the chart ϕ_α by $\partial^N_\mu = \phi_* \partial^M_\mu$. Since $g^M = \phi^* g^N$, we have

$$g^M_p(\partial^M_\mu, \partial^M_\nu) = g^N_{\phi(p)}(\phi_*\partial^M_\mu, \phi_*\partial^M_\nu) = g^N_q(\partial^N_\mu, \partial^N_\nu).$$

It follows that the coefficients $g_{\mu\nu}(x^1, \ldots, x^n)$ of g^M with respect to the chart $(\phi^{-1}(U_\alpha), \phi_\alpha \circ \phi)$ around $p \in M$ are precisely the coefficients of g^N with respect to the chart (U_α, ϕ_α) around $q \in N$. Now the Christoffel symbols are expressed in terms of $g_{\mu\nu}(x^1, \ldots, x^n)$ by (87), the coefficients of the Riemann tensor are expressed in terms of the Christoffel symbols by (112), the coefficients of the Ricci tensor are expressed in terms of those for the Riemann tensor by (127), and, finally, the Ricci scalar is determined in terms of the coefficients of the Ricci tensor by (128). We conclude that $R^M_p = R^N_q = R^N_{\phi(p)}$. Since this holds for all $p \in M$, we find that $R^M = \phi^* R^N$ as a smooth function on M. □

This result is of major importance in cartography, where the aim is to map a portion of the earth's surface onto a flat piece of paper in such a way that – up to a fixed scale factor – the distance between two points on the surface of the earth corresponds to the distance between the images of those two points on the piece of paper. In other words, a cartographer wishes to construct an isometry between an open subset of $\mathbb{S}^2$ and an open subset of $\mathbb{R}^2$, where $\mathbb{S}^2$ carries the round metric and $\mathbb{R}^2$ carries the Euclidean metric.

The following corollary to the *Theorema Egregium* makes the life of cartographers rather difficult: it asserts that it is impossible to map a piece of the earth's surface onto a flat chart without distortion.

Corollary 9.53. *The unit sphere* $\mathbb{S}^2$ *has scalar curvature* $R = 2$. *Euclidean space has scalar curvature* $R = 0$. *Consequently, there are no isometries between open subsets of* $\mathbb{S}^2$ *and an open subsets of* $\mathbb{R}^2$.

Proof. In order to compute the Riemann tensor in local coordinates, one first computes the Christoffel symbols $\Gamma^{\mu}_{\sigma\tau}$ from the metric $g_{\mu\nu}$ using (87), and then one computes $R^{\alpha}{}_{\beta\gamma\delta}$ from the Christoffel symbols $\Gamma^{\mu}_{\sigma\tau}$ using (112).

In Euclidean space $\mathbb{R}^n$, the metric is given by the *constant* functions $g_{\mu\nu} = \delta^{\mu}_{\nu}$. Since their derivatives are zero, we find $\Gamma^{\mu}_{\sigma\tau} = 0$ by (87), and we conclude from (112) that the Riemann tensor vanishes as well. By (127) and (128), we then find $R = 0$.

To determine the Riemann tensor for the round metric on $\mathbb{S}^2$, we describe an open subset of the sphere by spherical coordinates (ϕ, θ). In Problem 9.25, we calculated the metric tensor

$$\begin{pmatrix} g_{\phi\phi} & g_{\phi\theta} \\ g_{\theta\phi} & g_{\theta\theta} \end{pmatrix} = \begin{pmatrix} \sin^2(\theta) & 0 \\ 0 & 1 \end{pmatrix}.$$

From (87) we find that the only nonzero Christoffel symbols are

$$\begin{aligned} \Gamma^{\theta}_{\phi\phi} &= \frac{1}{2} g^{\theta\theta}(\partial_{\phi} g_{\theta\phi} + \partial_{\phi} g_{\phi\theta} - \partial_{\theta} g_{\phi\phi}) = -\sin(\theta)\cos(\theta) \\ \Gamma^{\phi}_{\theta\phi} = \Gamma^{\phi}_{\phi\theta} &= \frac{1}{2} g^{\phi\phi}(\partial_{\theta} g_{\phi\phi} + \partial_{\phi} g_{\theta\phi} - \partial_{\phi} g_{\theta\phi}) = \frac{\cos(\theta)}{\sin(\theta)}. \end{aligned}$$

In order to determine the Riemann curvature tensor, it is convenient to take the symmetries (122), (123) and (124) into account. Already from the first two equations, we see that the only independent component of the Riemann tensor is $R^{\theta}_{\theta\phi\phi}$. From (112), we find

$$\begin{aligned} R^{\theta}{}_{\theta\phi\phi} &= \partial_{\theta}\Gamma^{\theta}_{\phi\phi} - \partial_{\phi}\Gamma^{\theta}_{\theta\phi} + \Gamma^{\sigma}_{\phi\phi}\Gamma^{\theta}_{\theta\sigma} - \Gamma^{\sigma}_{\theta\phi}\Gamma^{\theta}_{\phi\sigma} \\ &= \partial_{\theta}(-\sin(\theta)\cos(\theta)) - \frac{\cos(\theta)}{\sin(\theta)}(-\sin(\theta)\cos(\theta)) \\ &= \sin^2(\theta). \end{aligned}$$

From the symmetries of the Riemann tensor, we find that $R^{\theta}{}_{\phi\theta\phi} = -R^{\theta}{}_{\theta\phi\phi}$ is equal to $-\sin^2(\theta)$, and that $R^{\phi}{}_{\phi\theta\theta} = -R^{\phi}{}_{\theta\phi\theta}$ is equal to

$$R^{\phi}{}_{\phi\theta\theta} = -g^{\phi\phi} R_{\phi\theta\phi\theta} = g^{\phi\phi} R_{\theta\theta\phi\phi} = g^{\phi\phi} g_{\theta\theta} R^{\theta}{}_{\theta\phi\phi} = \frac{1}{\sin^2(\theta)} \cdot 1 \cdot \sin^2(\theta) = 1.$$

Summarizing, the four nonzero components of the Riemann tensor are

$$\begin{aligned} R^{\theta}{}_{\theta\phi\phi} &= \sin^2(\theta) \\ R^{\theta}{}_{\phi\theta\phi} &= -\sin^2(\theta) \\ R^{\phi}{}_{\phi\theta\theta} &= 1 \\ R^{\phi}{}_{\theta\phi\theta} &= -1. \end{aligned}$$

From this we find that the Ricci tensor is diagonal with

$$R_{\phi\phi} = \sin^2(\theta), \quad R_{\theta\theta} = 1,$$

so the Ricci scalar is given by

$$\begin{aligned} R &= g^{\phi\phi}R_{\phi\phi} + g^{\theta\theta}R_{\theta\theta} \\ &= \frac{\sin^2(\theta)}{\sin^2(\theta)} + 1 = 2. \end{aligned}$$

Since this holds on the open dense subset of $\mathbb{S}^2$ where $\phi \in (0, 2\pi)$ and $\theta \in (0, \pi)$, and since R is a continuous function on $\mathbb{S}^2$, we have $R = 2$ on all of $\mathbb{S}^2$. $\square$

Problem 9.54. Let (M, g^M) and (N, g^N) be Riemannian manifolds.

a) Let $\phi \colon M \to N$ be a diffeomorphism such that $\phi^* g^N = \lambda g^M$ for a scaling factor $\lambda \in \mathbb{R}^+$. Show that $\phi^* R_N = \lambda^p R_M$ for some power $p \in \mathbb{Z}$, and determine p.

b) The unit n-sphere $\mathbb{S}^n = \{\vec{x} \in \mathbb{R}^{n+1}\,;\, \|\vec{x}\| = 1\}$ has constant Ricci curvature $R = n(n-1)$. Calculate the Ricci curvature for the n-sphere $\mathbb{S}^n_r = \{\vec{x} \in \mathbb{R}^{n+1}\,;\, \|\vec{x}\| = r\}$ of radius $r > 0$.

Problem 9.55. Let $\Sigma := \{(x, y, z) \in \mathbb{R}^3\,;\, z = e^{-\frac{1}{2}(x^2+y^2)}\}$.

a) Show that Σ is an embedded submanifold of $\mathbb{R}^3$ of dimension 2.

b) Make a sketch of the surface Σ in $\mathbb{R}^3$.

c) For every $p \neq (0, 0, 1)$ in Σ, there exists a chart (ϕ_α, U_α) around p with $\phi_\alpha^{-1}(r, \phi) = (r\cos(\phi), r\sin(\phi), e^{-\frac{1}{2}r^2})$. Calculate the corresponding coordinate vectors $\partial_r, \partial_\phi \in T_p\Sigma \subseteq T_p\mathbb{R}^3$.

d) The surface Σ inherits the Euclidean metric from $\mathbb{R}^3$. Calculate the matrix

$$\begin{pmatrix} g_{rr} & g_{r\phi} \\ g_{\phi r} & g_{\phi\phi} \end{pmatrix}$$

of this metric with respect to the above coordinates, and calculate its inverse matrix

$$\begin{pmatrix} g^{rr} & g^{r\phi} \\ g^{\phi r} & g^{\phi\phi} \end{pmatrix}.$$

e) Calculate the 8 Christoffel symbols $\Gamma^\mu_{\sigma\tau}$. How many of them are zero?

f) Derive the geodesic equation for $\gamma(t) = (r(t), \phi(t))$.

g) Calculate the Riemann tensor $R^\alpha_{\beta\gamma\delta}$. (Hint: out of the $2^4 = 16$ components of the Riemann tensor, there is only one linearly independent component.)

h) Prove that there are no open neighbourhoods $U \subseteq \Sigma$ that are isometric to an open neighbourhood of the Euclidean space $\mathbb{R}^2$.

Problem 9.56 (Fubini-Study metric)**.** Let $\mathbb{C}\mathrm{P}^n$ be complex projective space. We equip $\mathbb{C}^{n+1}$ with the Hermitian inner product

$$\langle v, w \rangle := \overline{v}_1 w_1 + \ldots + \overline{v}_{n+1} w_{n+1}$$

that is linear on the *right* hand side. Define

$$N_{[v]} := \{ w \in \mathbb{C}^{n+1} \, ; \, \langle v, w \rangle = 0 \}$$

to be the orthogonal complement of the ray $[v] \in \mathbb{C}\mathrm{P}^n$ inside $\mathbb{C}^{n+1}$.

a) Let $v \in \mathbb{C}^{n+1}$ be a representative of $[v] \in \mathbb{C}\mathrm{P}^n$ with $\langle v, v \rangle = 1$. Define the linear map

$$L_v \colon N_{[v]} \to T_{[v]}\mathbb{C}\mathrm{P}^n \quad \text{by} \quad L_v(w) := \frac{d}{dt}\Big|_{t=0} [v + tw].$$

Show that L_v is a $\mathbb{C}$-linear isomorphism. Show that if $v' = zv$ for a complex number z of modulus 1, then $L_{v'}(w) = z^{-1} L_v(w)$.

b) Let (ϕ_{n+1}, U_{n+1}) be the coordinate chart $U_{n+1} = \{[v] \in \mathbb{C}\mathrm{P}^n \, ; \, v_{n+1} \neq 0\}$ and

$$\phi_{n+1} \colon U_{n+1} \to \mathbb{C}^n \simeq \mathbb{R}^{2n}, \quad \phi_{n+1}([v]) = \frac{1}{v_{n+1}}(v_1, \ldots, v_n).$$

Calculate $L_v^{-1}(\partial_\mu)$, where $\partial_\mu \in T_{[v]}\mathbb{C}\mathrm{P}^n$ with $\mu = 1, \ldots, n$ is the complex basis corresponding to the above chart.

(Hint: calculate $L_v(b_\mu)$, where $b_1, \ldots, b_n$ is the complex basis $b_\mu = e_\mu - \langle v, e_\mu \rangle v$ of $N_{[v]}$ obtained by orthogonal projection of the first n canonical basis vectors $e_\mu = (0, \ldots, 1, \ldots 0)$ in $\mathbb{C}^{n+1}$.)

c) For each $[v] \in \mathbb{C}\mathrm{P}^n$, define the inner product $g_{[v]} \colon T_{[v]}\mathbb{C}\mathrm{P}^n \times T_{[v]}\mathbb{C}\mathrm{P}^n \to \mathbb{R}$ by

$$g_{[v]}(u, w) := \mathrm{Re}\langle L_v^{-1}(u), L_v^{-1}(w) \rangle.$$

Show that if $v' = zv$, then L_v and $L_{v'}$ result in the same inner product.

d) Calculate the components $g_{\mu\nu}$ of the tensor $g_{[v]}$ with respect to the coordinates (U_α, ϕ_α) described in the notes. Show that $g_{[v]}$ depends smoothly on $[v] \in \mathbb{C}\mathrm{P}^n$, and conclude that $[v] \mapsto g_{[v]}$ is a Riemannian metric on $\mathbb{C}\mathrm{P}^n$. It is called the *Fubini-Study* metric.

(Hint: use part (b). It suffices to do the calculation for the chart (ϕ_{n+1}, U_{n+1}), the rest is similar.)

e) Every unitary transformation $U \colon \mathbb{C}^{n+1} \to \mathbb{C}^{n+1}$ induces a diffeomorphism $\phi_U \colon \mathbb{C}\mathrm{P}^n \to \mathbb{C}\mathrm{P}^n$ by $\phi_U([v]) := [Uv]$. Show that this is an *isometry* of the metric g.

(Hint: show that U maps $N_{[v]}$ to $N_{[uv]}$, and that $\phi_{U*} \circ L_v = L_{Uv} \circ U$.)

f) Prove that the scalar curvature $R := g^{\tau\nu} R^{\mu}{}_{\mu\nu\tau}$ of g is constant. (Hint: this can be done in 3 lines without any computation.)

g) Calculate the scalar curvature for $\mathbb{C}\mathrm{P}^1$.

10 General relativity

In the previous chapter, we made the transition from Euclidean geometry to Riemannian geometry. In Euclidean geometry, the underlying space is $\mathbb{R}^n$, and the notion of length is provided by the standard inner product

$$(v, w) = v^1 w^1 + \ldots + v^n w^n \tag{129}$$

on $\mathbb{R}^n$. In Riemannian geometry, the underlying space can be any smooth manifold M. Our notion of length is entirely *infinitesimal*, and it is given by an inner product

$$g_p \colon T_pM \times T_pM \to \mathbb{R}$$

on every tangent space. For a Riemannian metric g on M, every tangent space T_pM admits an orthonormal basis $e_1(p), \ldots, e_n(p)$ such that $g_p(e_i, e_j) = \delta_{ij}$. Such a basis is called an *orthonormal frame* at p.

The transition from the special theory of relativity to the general theory of relativity is very similar. In §1.2, we saw that special relativity is governed by the *Minkowski space* $\mathbb{M}^4$, with the Minkowski metric

$$\eta(v, w) = -v^0 w^0 + v^1 w^1 + v^2 w^2 + v^3 w^3.$$

This Minkowski 'metric' is not a metric at all, since the inner product of a vector with itself can be negative, or even zero. Nonetheless, we can define an 'infinitesimal version' of Minkowski space in exactly the same way as a Riemannian manifold is an 'infinitesimal version' of Euclidean space.

A *Lorentzian manifold* (M, g) is a 4-dimensional manifold M, which is equipped with a *Lorentzian metric*. A lorentzian metric is similar to a Riemannian metric, except that the symmetric bilinear form $g_p \colon T_pM \times T_pM \to \mathbb{R}$ need not be positive definite. Instead, we require that at every point p, there exists a basis $e^0(p), e^1(p), e^2(p), e^3(p)$ of T_pM such that $g_p(e_0, e_0) = -1$, $g_p(e_i, e_i) = +1$ for $i = 1, 2, 3$, and $g_p(e_i, e_j) = 0$ for $i \neq j$.

In the same way that every tangent space to a Riemannian manifold looks like a Euclidean space, every tangent space to a Lorentzian manifold now looks like a Minkowski space.

10.1 The Geodesic Principle

The dynamics of general relativity are rather aptly summarised by the following quote, usually attributed to J. A. Wheeler:

Space tells matter how to move.
Matter tells space how to curve.

Mathematically, space–time is modelled by a Lorentzian manifold (M, g). Just like for Riemannian manifolds, every Lorentzian manifold has a unique Levi-Civita connection ∇ that is both torsion-free and preserves the Lorentzian metric. And just like for Riemannian manifolds, the requirement that $\dot{\gamma}$ be constant with respect to ∇ gives rise to the geodesic equation

$$\ddot{\gamma}^{\mu} + \Gamma^{\mu}_{\sigma\tau}\dot{\gamma}^{\sigma}\dot{\gamma}^{\tau} = 0. \tag{130}$$

The first part of Wheeler's quote pertains to *Geodesic Principle*:

"The motion of a free point particle is described by a geodesic with $g(\dot{\gamma}, \dot{\gamma}) \leq 0$"

In Minkowski space, the coefficients $g_{\mu\nu} = \eta_{\mu\nu}$ of the metric tensor are constant, so the Christoffel symbols are zero. The geodesic equation

$$\ddot{\gamma}^{\mu} = 0$$

is rather easy to solve; the solutions are

$$\gamma(s) = \begin{pmatrix} ct(s) \\ x(s) \\ y(s) \\ z(s) \end{pmatrix} = \begin{pmatrix} ct(0) + p_0 s \\ x(0) + p_1 s \\ y(s) + p_2 s \\ z(s) + p_3 s \end{pmatrix}.$$

Note that $v_x = dx/dt = cp_1/p_0$, and similarly $v_y = cp_2/p_0$ and $v_z = cp_3/p_0$. The requirement $g(\dot{\gamma}, \dot{\gamma}) \leq 0$ is equivalent to $p_1^2 + p_2^2 + p_3^2 \leq p_0^2$, which, in turn, is equivalent to $\|\vec{v}\| \leq c$. In other words, the geodesic principle implies that no particle can move faster than light!

Remark 10.1 (Proper time)**.** In Riemannian geometry, a geodesic locally minimizes the distance between the points it connects. Similarly, in Lorentzian geometry, a geodesic is an extremal point of the quantity

$$S(\gamma) = \int_{s_i}^{s_f} \sqrt{-g_{\mu\nu}\dot{\gamma}^{\mu}\dot{\gamma}^{\nu}}ds.$$

This means that geodesics are still solutions to $\frac{d}{d\varepsilon}|_{\varepsilon=0}S(\gamma_{\varepsilon})$, but they can be minima, maxima or saddle points. If $g(\dot{\gamma}, \dot{\gamma}) < 0$, then $S(\gamma)$ is interpreted as the *proper time* along the curve, i.e., the time difference between the initial and final space-time points $\gamma(s_i)$ and $\gamma(s_f)$ as experienced by an observer *that moves along with the curve.*

In a curved space–time (M, g), geodesic motion is the closest thing we have to a straight line. In that sense, the Geodesic Principle rather resembles Newton's first law of motion, which states that a free particle moves at a constant speed as long as it does not experience any force.

However, if the Christoffel symbols of the metrig g are nonzero, then the effect of geodesic motion can be spectacularly different from a straight line. For example, in Einstein's theory of general relativity, the motion of the earth around the sun is a geodesic, but certainly not a straight line. According to Einstein, the reason that the earth moves around the sun is not that it experiences any gravitational force, but that the sun causes space–time to 'curve'. Due to the large mass of the sun, the space–time through which the earth moves is described by $M = \mathbb{R}^4$ not with the Minkowski metric $g_{\mu\nu} = \eta_{\mu\nu}$, but with a *different* Lorentzian metric $g_{\mu\nu}$, whose nontrivial Christoffel symbols give rise to 'curved' geodesics. It is this change in the metric – rather than any non-existent 'gravitational force' – which causes the earth to stay in orbit around the sun.

10.2 Einstein's equation

To make this effect quantitative, Einstein proposed an equation in which, in the words of Wheeler, matter "causes space–time to curve". In this formula, the relevant properties of matter are captured by the *Stress-Energy-Momentum tensor*

$$(T_{\mu\nu}) = \begin{pmatrix} T_{00} & T_{01} & T_{02} & T_{03} \\ T_{10} & T_{11} & T_{12} & T_{13} \\ T_{20} & T_{21} & T_{22} & T_{23} \\ T_{30} & T_{31} & T_{32} & T_{33} \end{pmatrix} \tag{131}$$

Since this is a symmetric tensor, it contains only 10 independent quantities. With respect to coordinates[1] in which $g_{\mu\nu}(p) = \eta_{\mu\nu}$, we can interpret $T^{00}(p)$ as the energy density, T^{0j} as the density of j-momentum (for $j = x, y, z$), T^{i0} as the flow of energy in the ∂_i direction, and T^{ij} as the flow of j-momentum in the i-direction.

To find a partial differential equation for $g_{\mu\nu}$ in terms of $T_{\mu\nu}$, it is natural to consider the Riemann tensor $R^{\alpha}{}_{\beta\gamma\delta}$. Recall that the Ricci tensor is defined by

$$R_{\alpha\beta} := R^{\sigma}{}_{\sigma\alpha\beta}, \tag{132}$$

and the Ricci scalar is defined in terms of $R^{\alpha}{}_{\beta} = g^{\alpha\sigma}R_{\sigma\beta}$ by

$$R := R^{\sigma}{}_{\sigma} = g^{\alpha\sigma}R_{\sigma\alpha}. \tag{133}$$

The *Einstein equation*, describing the effect of the Stress-Energy-Momentum tensor $T_{\mu\nu}$ on space–time, is

$$R_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}, \tag{134}$$

where G is Newton's gravitational constant and c is the speed of light.

The Einstein equation is a nonlinear second order PDE for $g_{\mu\nu}$, with $T_{\mu\nu}$ as a source term. Indeed, since the formula 87 for the Christoffel symbols $\Gamma^{\mu}_{\sigma\tau}$ involves first order derivatives of $g_{\mu\nu}$, and since the formula 112 for the Riemann tensor $R^{\alpha}{}_{\beta\gamma\delta}$ involves first order derivatives of the Christoffel symbols, the left hand side of (134) involves second order derivatives of $g_{\mu\nu}$.

Problem 10.2 (Vaccum field equations)**.** Show that on 4-dimensional space–time, the trace of the Einstein tensor $G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$ is $G^{\mu}_{\mu} = -R$. Conclude that if $T_{\mu\nu} = 0$, then (134) is equivalent to

$$R_{\mu\nu} = 0. \tag{135}$$

These are the *vacuum field equations* for the gravitational field.

[1]For each point $p \in M$, one can find coordinates x^{μ} such that $g_{\mu\nu}(p) = \eta_{\mu\nu}$ at *that* particular point p, but not necessarily for other points p' in the same coordinate patch.

10.3 Schwarzschild solution and gravitational waves

Although Einstein achieved remarkable success with approximate solutions using a Taylor series expansion, he originally did not expect to find exact solutions to the equation which bears his name. However, within a few months after Einstein published his 1915 paper, Karl Schwarzschild came up with the following useful solution.

Consider the sun as a point particle, traversing the geodesic $\gamma(t) = (t, 0, 0, 0)$. Then the Stress-Energy-Momentum tensor is a δ-function localized at $(x, y, z) = (0, 0, 0)$. We therefore try to solve the Einstein equation (134) on the manifold $\mathbb{R}^4 \setminus \{(t, 0, 0, 0)\,;\, t \in \mathbb{R}\}$, with $T_{\mu\nu} = 0$ on the right hand side.

Schwarzschild used coordinates t, r, ϕ, θ, and he used an Ansatz of the form

$$\begin{pmatrix} g_{tt} & g_{tr} & g_{t\theta} & g_{t\phi} \\ g_{rt} & g_{rr} & g_{r\theta} & g_{r\phi} \\ g_{\theta t} & g_{\theta r} & g_{\theta\theta} & g_{\theta\phi} \\ g_{\phi t} & g_{\phi r} & g_{\phi\theta} & g_{\phi\phi} \end{pmatrix} = \begin{pmatrix} -f^2(r) & 0 & 0 & 0 \\ 0 & \frac{1}{f^2(r)} & 0 & 0 \\ 0 & 0 & r^2 \sin^2(\phi) & 0 \\ 0 & 0 & 0 & r^2 \end{pmatrix}, \tag{136}$$

where the function $f(r)$ is still to be determined. Note that the lower right 2×2 block is just the round metric on a sphere with radius r, and the left upper 2×2 block is proportional to the Minkowski metric, with an r-dependent scaling factor $1/f^2(r)$ that is still to be determined.

One can calculate the Christoffel symbols for this metric, use them to calculate the Riemann tensor, and find the Ricci tensor and scalar curvature. Plugging this into the Einstein equation (134) with $T_{\mu\nu} = 0$, one finds that Schwarzschild's Ansatz is a solution if and only if $\frac{d}{dr}(rf^2(r)) = 1$. This has the solution $f^2(r) = 1 - \frac{\rho}{r}$, where the constant ρ is called the *Schwarzschild radius.*

Problem 10.3. Perform the necessary calculations.

Comparing the geodesics for the Schwarzschild metric with the solutions of Newton's equations of motion for the gravitational potential, one can show that $\rho = 2MG/c^2$, where G is the gravitational constant and M is the mass of the sun.

The Schwarzschild solution gives rise to a number of surprising phenomena. Using the Christoffel symbols derived from the metric (136), one finds the geodesic equation that describes the rotation of the planets around the sun due to the gravitational distortion of space–time. These geodesics wind around the t-axis in $\mathbb{R}^4$, and projecting them onto the x, y, z-hyperplane, one expects to find an ellipse with the sun in one of the focal points. As it turns out, this is *almost* what happens. In fact, the orbits do not exactly 'close up', but every year the ellipse traversed by the planet shifts a little. This *perihelium precession* was known before the advent of general relativity, and the fact that general relativity correctly explains this effect was historically the first experimental confirmation of the theory. (In his original paper, Einstein used Taylor expansions to find an approximate solution in order to calculate this effect, as Schwarzschild had not yet found his solution.)

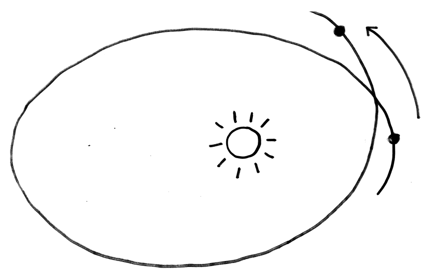


Figure 21: Precession of planetary orbits around the sun.

Another interesting effect of general relativity is that even light moves along a geodesic – albeit one with $g(\dot\gamma, \dot\gamma) = 0$. If one calculates the deflection of light caused by the gravitational effect of the sun in the Schwarzschild solution, one finds an effect that is twice as strong as the corresponding effect in the Newtonian theory of gravity. This effect was measured during a 1919 eclipse (when the sunlight is sufficiently blocked to observe the deflection of starlight by the sun), and the effects were found to be in agreement with the predictions of general relativity.

Problem 10.4. To describe a point particle falling straight into a black hole, it suffices to consider only the t and r component of the Schwarzschild metric,

$$\begin{pmatrix} g_{tt} & g_{tr} \\ g_{rt} & g_{rr} \end{pmatrix} = \begin{pmatrix} -(1-\frac{\rho}{r}) & 0 \\ 0 & \frac{1}{1-\frac{\rho}{r}} \end{pmatrix},$$

where $\rho := \frac{2MG}{c^2}$ is the Schwarzschild radius. It is convenient to parameterize the curve γ by the distance r from the origin, so $(\gamma^t(r), \gamma^r(r)) = (t(r), r)$. This way, the function $t(r)$ is the only relevant degree of freedom.

a) For Lorentzian manifolds, the geodesic equation describes extremal points of the action
$$S(\gamma) := \int_{r_i}^{r_f} \sqrt{-g_{\mu\nu}\,\tfrac{d}{dr}\gamma^\mu\,\tfrac{d}{dr}\gamma^\nu}\; dr.$$
Prove that for a small variation of the path $t_\varepsilon(r) = t(r) + \varepsilon\delta t(r)$ that vanishes at the endpoints, $\delta t(r_i) = \delta t(r_f) = 0$, the Euler-Lagrange equation for $\frac{d}{d\varepsilon}S(t_\varepsilon)|_{\varepsilon=0} = 0$ reads
$$\frac{d}{dr}\left(\frac{(1-\frac{\rho}{r})(\frac{dt}{dr})}{\sqrt{(1-\frac{\rho}{r})(\frac{dt}{dr})^2 - \frac{1}{1-\frac{\rho}{r}}}}\right) = 0.$$
(Hint: revisit the proof of the geodesic equation, but note that with our parameterization $g_{\mu\nu}\frac{d}{dr}\gamma^\mu\,\frac{d}{dr}\gamma^\nu \neq 1$!)

b) Conclude that there exists a constant K such that

$$t(r_f) = t(r_i) + \int_{r_i}^{r_f} \frac{K}{1-\frac{\rho}{r}} \sqrt{\frac{1}{\frac{\rho}{r} + K^2 - 1}}\, dr.$$

Determine K in terms of r_i and the velocity $v_i = (\frac{dt}{dr})^{-1}$ at the initial radius r_i.

c) Show that $t(r_f) \to \infty$ for $r_f \downarrow \rho$. So from the point of view of an outside observer, a point-particle falling into a black hole will never reach the Schwarzschild radius.

Finally, let us mention another phenomenon that is predicted by general relativity: gravitational waves. On $M = \mathbb{R}^4$ with coordinates x, y, $u = t - z$ and $v = t + z$, the Ansatz

$$\begin{pmatrix} g_{xx} & g_{xy} & g_{xu} & g_{xv} \\ g_{yx} & g_{yy} & g_{yu} & g_{yv} \\ g_{ux} & g_{uy} & g_{uu} & g_{uv} \\ g_{vx} & g_{vy} & g_{vu} & g_{vv} \end{pmatrix} = \begin{pmatrix} L^2(u)e^{2\beta(u)} & 0 & 0 & 0 \\ 0 & L^2(u)e^{-2\beta(u)} & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \tag{137}$$

is a solution to the vacuum ($T_{\mu\nu} = 0$) Einstein solution if and only if $L(u)$ and $\beta(u)$ satisfy the ODE

$$\frac{d^2}{dt^2}L(u) + \left(\frac{d}{du}\beta(u)\right)^2 L(u) = 0. \tag{138}$$

Problem 10.5. Verify that (137) is a solution to (134) if and only if $L(u)$ and $\beta(u)$ satisfy (138).

If $L = 1$ and $\beta = 0$, this is simply the Minkowski metric in the coordinates x, y, u, v. The solutions to the linearized equation represent *gravitational waves*, where space–time periodically expands and contracts. These gravitational waves arise in astrophysical events involving extremely massive objects undergoing rather violent dynamics (such as the merging of black holes), and they propagate through empty space. When the gravitational wave reaches the earth, the contractions of space–time, although minute, can in principle be detected by laser interferometry. This is the way in which the LIGO-observatories managed to experimentally confirm the existence of gravitational waves on September 14, 2015.

A Topological spaces

Roughly speaking, a *topology* is the minimum structure that one needs in order to talk about continuity of functions. In this appendix we define topological spaces, and prove some of their basic properties.

A.1 Continuity in $\mathbb{R}^n$

First, let us recall what it means for a function $\phi\colon \mathbb{R}^n \to \mathbb{R}^m$ to be continuous.

Definition A.1 (Continuity for $\mathbb{R}^n$, ε/δ-version)**.** A map $\phi\colon \mathbb{R}^n \to \mathbb{R}^m$ is continuous if for every $x \in \mathbb{R}^n$ and for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $\|\phi(y) - \phi(x)\| < \varepsilon$ whenever $\|x - y\| < \delta$.

This is the definition in terms of ε's and δ's that we all know and some of us love. But there is another, equivalent definition of continuity in terms of *open sets*. A subset $U \subseteq \mathbb{R}^n$ is called *open* if for every $x \in U$, there exists an $\varepsilon > 0$ such that the ball

$$B_\varepsilon(x) := \{y \in \mathbb{R}^n\,;\, \|y - x\| < \varepsilon\}$$

of radius ε around x is contained in U.

We can give the following definition of continuity in terms of open subsets.

Definition A.2 (Continuity for $\mathbb{R}^n$, open version)**.** A map $\phi\colon \mathbb{R}^n \to \mathbb{R}^m$ is continuous if the preimage $\phi^{-1}(U) \subseteq \mathbb{R}^n$ of every open set $U \subseteq \mathbb{R}^m$ is open.

The two definitions A.1 and A.2 of continuity are equivalent, but the advantage of Definition A.2 is that all ε's and δ's are hidden in the definition of an open set. This gives us an important clue on how to define continuity for functions on spaces that are more general than $\mathbb{R}^n$.

We do not need a norm or metric, we only need to know what the open sets are.

Problem A.3. Prove that Definition A.1 and Definition A.2 are equivalent.

A.2 Topological spaces and continuity

A *topological space* is a set M, together with a collection $\mathcal{T}$ of subsets $U \subseteq M$ that we choose to call *open*. These open sets are required to behave well under union and intersection.

Definition A.4 (Topology)**.** A *topological space* is a set M, together with a collection $\mathcal{T}$ of subsets $U \subseteq M$. We call a subset $U \subseteq M$ *open* if it is in $\mathcal{T}$. The open sets are required to satisfy the following rules.

1 The empty set and M itself are open.

2 Unions of open sets are open.

3 Finite intersections of open sets are open.

A subset $U \subseteq M$ is called *closed* if its complement $M \setminus U$ is open. If the topology is clear, we simply refer to $(M, \mathcal{T})$ as 'the topological space M'.

Problem A.5 (Standard topology on $\mathbb{R}^n$)**.** Prove that the collection $\mathcal{T}$ of subsets $U \subseteq \mathbb{R}^n$ that are open in the sense of §A.1 is indeed a topology on $\mathbb{R}^n$. This is called the *standard topology* of $\mathbb{R}^n$.

Topological spaces allow us to give a very clean formulation of notions that involve 'nearby points' in terms of *neighbourhoods*.

- An *open neighbourhood* of x is an open set $U \subseteq M$ that contains x.
- A *neighbourhood* $N \subseteq M$ of x is a set that contains an open neighbourhood of x. In other words, there exists an open subset $U \subseteq N$ with $x \in U \subseteq N$.

We think of a neighbourhood of x as a set that contain all points that are 'close to x'.

We would like to call a map $\phi\colon M \to N$ *continuous* if it sends nearby points to nearby points. If M and N are topological spaces, we can make this precise as follows.

Definition A.6 (Continuity)**.** Let M and N be topological spaces. The map $\phi\colon M \to N$ is continuous if the preimage $\phi^{-1}(U) \subseteq M$ of any open set $U \subseteq N$ is open in M.

Note that this is exactly the same definition as Definition A.2! The motivating example of a topological space is of course the set $M = \mathbb{R}^n$, with the topology $\mathcal{T}$ of subsets $U \subseteq \mathbb{R}^n$ that are open in the sense of §A.1.

Problem A.7. What is a neighbourhood of a point $x \in \mathbb{R}^n$? Would you think it is fair to say that a neighbourhood of x contains all points that are 'close to x'?

Problem A.8 (Concatenation)**.** Suppose that L, M and N are topological spaces. If $\phi\colon M \to N$ and $\psi\colon L \to M$ are continuous, then their concatenation $\phi \circ \psi\colon L \to N$ is continuous as well.

A topological space is called *connected* if it cannot be written as the disjoint union of two nonempty open subsets.

Definition A.9 (Connectedness)**.** A topological space M is connected if for any pair $A, B \subseteq M$ of open sets with $M = A \cup B$ and $A \cap B = \emptyset$, we have either $A = \emptyset$ or $B = \emptyset$.

Problem A.10. Equivalently, every $A \subseteq M$ that is both open and closed satisfies either $A = \emptyset$ or $A = M$.

Problem A.11 (The continuous image of a connected space is connected)**.** Suppose that M is connected, and that $\phi\colon M \to N$ is both continuous and surjective. Then N is connected.

Although they will not be very prominent in this course, it is good to realise that there exist examples of topological spaces which are not at all $\mathbb{R}^n$-like.

Problem A.12 (Discrete topology)**.** Let M be an arbitrary set, and let $\mathcal{T}$ be the *discrete topology*, consisting of *all* subsets of M.

a) What does it mean for a function $\phi\colon M \to \mathbb{R}$ to be continuous?

b) What does it mean for a function $\phi\colon \mathbb{R} \to M$ to be continuous?

Problem A.13 (Indiscrete topology)**.** Let M be an arbitrary set, and let $\mathcal{T} := \{\emptyset, M\}$ be the *indiscrete topology*.

a) What does it mean for a function $\phi\colon M \to \mathbb{R}$ to be continuous?

b) What does it mean for a function $\phi\colon \mathbb{R} \to M$ to be continuous?

Problem A.14 (A subset of a topological space is not a door)**.** A subset U of a topological space M can be open, closed, neither open nor closed, or both open and closed. Consider these options for the subsets $U \subseteq \mathbb{R}$ given by $\emptyset$, $(0,1)$, $[0,1)$, $(0,1]$, $[0,1]$ and the whole space $\mathbb{R}$.

A.3 Constructing topological spaces

We briefly explore a number of ways in which topological spaces arise in practise. First of all, every *metric space* is a topological space. Secondly, starting from a topological space, one can form *subspaces* and *quotients* to obtain new topological spaces. Finally, *products* of topological spaces are topological spaces again. The topological spaces that we encounter in this course are often quotients or subspaces of the metric space $\mathbb{R}^n$.

A.3.1 Metric spaces are topological spaces

One important source of topological spaces are *metric spaces* (M, d). Recall that a *metric* $d\colon M \times M \to \mathbb{R}$ satisfies

(1) $d(x, y) = d(y, x)$

(2) $d(x, z) \leq d(x, y) + d(y, z)$

(3) $d(x, y) = 0$ if and only if $x = y$.

If one thinks of $d(x, y)$ as the distance from x and y, then (1) says that this is the same as the distance from y to x, (2) says that going from x to z is at most as far as first going from x to y and then going from y to z, and (3) says that the only point at distance zero from x is x itself.

Definition A.15 (Topology for a metric space)**.** For a metric space (M, d), we define the topology $\mathcal{T}_d$ by declaring $U \subseteq M$ to be open if for every $x \in U$, there exists an $\varepsilon > 0$ such that the ball

$$B_\varepsilon(x) := \{y \in \mathbb{R}^n\,;\, d(x, y) < \varepsilon\}$$

of radius ε around x is contained in U.

Proposition A.16. *Every metric space is a topological space.*

Proof. We check that $\mathcal{T}_d$ has the three properties mentioned in Definition A.4.

(1) Clearly $\emptyset$ and M are open.

(2) Suppose that U_i is open for all i in an index set I. Then $\bigcup_{i\in I} U_i$ is open as well. Indeed, if $x \in \bigcup_{i\in I} U_i$, then certainly $x \in U_i$ for some $i \in I$. Since U_i is open, there exists an $\varepsilon > 0$ such that $B_\varepsilon(x)$ is contained in U_i. But then $B_\varepsilon(x)$ is contained in $\bigcup_{i\in I} U_i$ as well.

(3) Suppose that U_i is open for $i = 1, \ldots, n$. Then $\bigcap_{i=1}^n U_i$ is open as well. Indeed, if $x \in \bigcap_{i=1}^n U_i$, then $x \in U_i$ for all $i = 1, \ldots, n$. Since U_i is open, there exists an $\varepsilon_i > 0$ such that $B_{\varepsilon_i}(x)$ is contained in U_i. If we take ε to be the smallest of the $\varepsilon_1, \ldots, \varepsilon_n$, then $B_\varepsilon(x)$ is contained in all sets U_i, and hence in their intersection $\bigcap_{i=1}^n U_i$. $\square$

Note that although *finite* intersections of open sets are open, *infinite* intersections need not be open. For example, the intervals $(-1/n, 1/n)$ are open subsets of $\mathbb{R}$, but their intersection $\bigcap_{n=1}^\infty(-1/n, 1/n) = \{0\}$ is not.

A.3.2 Subspaces are topological spaces

Let M be a topological space. Then any subset $\Sigma \subseteq M$ is a topological space if we endow it with the *subspace topology.*

Definition A.17 (Subspace topology)**.** The *subspace topology* of Σ is the collection of sets of the form $U \cap \Sigma$, where $U \subseteq M$ is open in M.

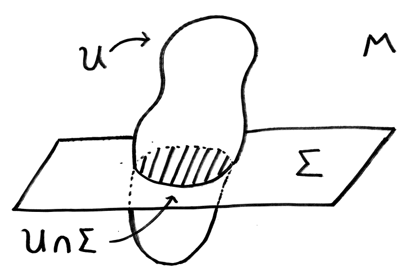


Figure 22: Open set in the subspace topology

Problem A.18. Show that the subspace topology is indeed a topology.

Problem A.19 (Topology of spheres)**.** Describe the open sets for the topology of

$$\mathbb{S}^n := \{(x^0, \ldots, x^n) \in \mathbb{R}^{n+1}\,;\, (x^0)^2 + \ldots + (x^n)^2 = 1\}\,,$$

considered as a subspace of $\mathbb{R}^{n+1}$. What does it mean for a map $\phi\colon \mathbb{S}^n \to \mathbb{R}$ to be continuous?

The following exercise shows that continuity of functions with respect to the subspace topology can often be expressed in terms of continuity in the surrounding space.

Problem A.20. Suppose that M and N are topological spaces, and that $\Sigma \subseteq M$ has the subspace topology.

a) Show that the inclusion $\iota\colon \Sigma \hookrightarrow M$ is continuous.

b) Conclude that if $\phi\colon M \to N$ is continuous, then $\phi \circ \iota\colon \Sigma \to N$ is continuous as well, cf. Problem A.8. (This is just the restriction of ϕ to Σ.)

c) The map $\phi\colon N \to \Sigma$ is continuous if and only if $\iota \circ \phi\colon N \to M$ is continuous. (This is just the map to Σ considered as a map into M.)

Problem A.21. The altitude map $\phi\colon \mathbb{S}^2 \to \mathbb{R}$ with $\phi(x, y, z) = z$ is continuous, as is the curve $c\colon \mathbb{R} \to \mathbb{S}^2$ defined by $c(t) = (\cos(t), \sin(t), 0)$. (Hint: use Problem A.20.)

A.3.3 Quotients are topological spaces

Starting from a topological space M, we can also form new topological spaces by taking *quotients*.

Recall that if $\sim$ is an equivalence relation on a set M, then the equivalence class of $x \in M$ is the set of all points x' that are related to x,

$$[x] := \{x' \in M\,;\, x' \sim x\}.$$

The *quotient* of M by the relation $\sim$ is the set of all equivalence classes,

$$M/\!\sim\; := \{[x]\,;\, x \in M\}.$$

Since $[x'] = [x]$ if and only if $x' \sim x$, one can think of $M/\!\sim$ as the space obtained from M by identifying all the points which are in relation to each other.

Example A.22. One can think of the *circle* as the space obtained from the real line by identifying all points that differ by an integer. More precisely, define the relation $\sim$ on $\mathbb{R}$ by $x' \sim x$ if and only if $x' - x \in \mathbb{Z}$. Every point $[x] \in \mathbb{R}/\!\sim$ can be represented by a unique $x \in [0, 1)$. Since $[0] = [1]$ in $\mathbb{R}/\!\sim$, we can view $\mathbb{R}/\!\sim$ as the interval $[0, 1]$ where 0 and 1 are identified.

Problem A.23 (The circle as a quotient)**.** Find a bijection between $\mathbb{R}/\!\sim$ and the circle $\{(x, y) \in \mathbb{R}^2\,;\, x^2 + y^2 = 1\}$.

If M is a topological space and $\sim$ is a relation on M, then $M/\!\sim$ is a topological space in a natural manner. We define the quotient map $\pi\colon M \to M/\!\sim$ by $\pi(x) = [x]$. By definition, we say that a subset $U \subseteq M/\!\sim$ is open in $M/\!\sim$ if and only if

$$\pi^{-1}(U) := \{x \in M\,;\, [x] \in U\}$$

is open in M. This yields the *quotient topology* on $M/\!\sim$.

Proposition A.24 (Quotient topology)**.** *The open sets $U \subseteq M/{\sim}$ constitute a topology on $M/{\sim}$.*

Proof. There are three properties to check.

(1) The empty set $U = \emptyset$ and the whole space $U = M/{\sim}$ are open in $M/{\sim}$ because $\pi^{-1}(\emptyset) = \emptyset$ and $\pi^{-1}(M/{\sim}) = M$ are open in M.

(2) Suppose that $U_i \subseteq M/{\sim}$ is open in $M/{\sim}$ for every $i \in I$. We wish to show that their union $\bigcup_{i\in I} U_i$ is open as well. Since U_i is open in $M/{\sim}$, the set $\pi^{-1}(U_i) = \{x \in M\,;\,[x] \in U_i\}$ is open in M by definition. But since M is a topological space, the union $\bigcup_{i\in I} \pi^{-1}(U_i)$ is again open in M. As $\pi^{-1}(\bigcup_{i\in I} U_i) = \bigcup_{i\in I} \pi^{-1}(U_i)$ (why?), the union $\bigcup_{i\in I} U_i$ is open in $M/{\sim}$.

(3) The proof that finite intersections of open sets are open is similar. □

Proposition A.25. *The quotient map $\pi\colon M \to M/{\sim}$ is continuous.*

Proof. If $U \subseteq M/{\sim}$ is open, then $\pi^{-1}(U)$ is open by definition. □

The definition of quotients spaces is somewhat abstract, so it is good to look at some examples that will be important in the course.

Example A.26 (Tori)**.** The the n-torus $\mathbb{T}^n := \mathbb{R}^n/{\sim}$ is the quotient of $\mathbb{R}^n$ by the relation that identifies $x, x' \in \mathbb{R}^n$ if and only if $x - x' \in \mathbb{Z}^n$. The 2-torus $\mathbb{T}^2$ can be visualized as follows. Since every $[x] \in \mathbb{T}^2$ has precisely one representative $x = (x_1, x_2)$ in $[0,1) \times [0,1)$, we can think of $\mathbb{T}^2$ as the square $[0,1) \times [0,1)$. Since $[(x_1, 0)] = [(x_1, 1)]$, the bottom of the square is identified with the top, and since $[(0, x_2)] = [(1, x_2)]$, the left and right sides are identified. Similarly, the 3-torus $\mathbb{T}^3$ can be visualized as the unit cube $[0,1) \times [0,1) \times [0,1)$ where opposite faces are identified.

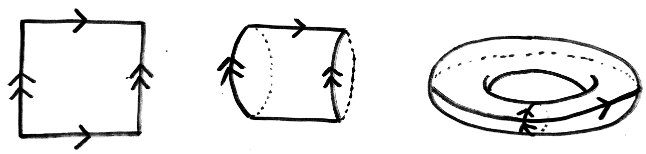


Figure 23: The 2-torus as a square where opposite sides are identified.

Problem A.27 (Open subsets of $\mathbb{T}^2$)**.** Consider the subset $U \subseteq \mathbb{T}^2$ defined by

$$U := \{[(x_1, x_2)] \in \mathbb{T}^2\,;\, x_1^2 + x_2^2 < 1/10\}\,.$$

a) Draw a picture of $U \subseteq \mathbb{T}^2$ and $\pi^{-1}(U) \subseteq \mathbb{R}^2$.

b) Show that U is open.

Example A.28 (Projective space)**.** The *complex projective space* $\mathbb{C}\mathrm{P}^n$ is the set of all *rays* in $\mathbb{C}^{n+1}$, that is, the set of all 1-dimensional complex linear subspaces of $\mathbb{C}^{n+1}$ with the origin deleted. To see that this is a topological space, we need both the subspace topology and the quotient topology. First, the set $M := \{v \in \mathbb{C}^{n+1}\,;\, v \neq 0\}$ of all nonzero vectors in $\mathbb{C}^{n+1}$ is a topological space, as it inherits the subspace topology from $\mathbb{C}^{n+1}$. We then define $\mathbb{C}\mathrm{P}^n$ to be the quotient of M by the equivalence relation that identifies two nonzero vectors $v, v' \in M$ if and only if they are parallel, that is, $v \sim v'$ if and only if $v' = \lambda v$ for some $\lambda \in \mathbb{C}^{\times}$. Considered as an equivalence class, a *point* $[v] \in \mathbb{C}\mathrm{P}^n$ is then the complex *line*

$$[v] = \{\lambda v\,;\, \lambda \in \mathbb{C}^{\times}\}$$

of all nonzero vectors that are parallel to v.

A.3.4 Products are topological spaces

Another way of making new topological spaces from old ones is by taking products.

Definition A.29 (Product topology)**.** If M and N are topological spaces, we define a subset of $M \times N$ to be open if it is a union of sets of the form $U \times V$, where U is open in M and V is open in N.

Problem A.30. Show that the product topology is indeed a topology.

Problem A.31. Show that the standard topology of $\mathbb{R}^2$ (Problem A.5) coincides with the product topology on $\mathbb{R} \times \mathbb{R}$. (Hint: squares fit inside circles and vice versa. Pictures are helpful.)

Problem A.32 (Product topology on $\mathbb{R}^n$)**.** For $\mathbb{R}^n = \mathbb{R} \times \ldots \times \mathbb{R}$, we have a metric topology (from Proposition A.16) and a product topology (from Definition A.29). Show that these two topologies coincide.

a) Show that the metric topology of $\mathbb{R}^{n+m}$ coincides with the product topology on $\mathbb{R}^n \times \mathbb{R}^m$.

b) Conclude that the n-fold product $\mathbb{R} \times \ldots \times \mathbb{R}$ has the same topology as $\mathbb{R}^n$.

c) Show that a map $\phi\colon \mathbb{R} \to \mathbb{R} \times \ldots \times \mathbb{R}$ is continuous for the *product* topology if and only if each of its components $\phi_\mu\colon \mathbb{R} \to \mathbb{R}$ is continuous.

d) Conclude that a map $\phi\colon \mathbb{R} \to \mathbb{R}^n$ is continuous for the *metric* topology if and only if each of its n components $\phi_\mu\colon \mathbb{R} \to \mathbb{R}$ is continuous.

Problem A.33. A map $\phi\colon L \to M \times N$ is continuous if and only if its components $\phi_M\colon L \to M$ and $\phi_N\colon L \to N$ are continuous.

A.4 Homeomorphisms

Suppose that a map $\phi\colon M \to N$ is bijective, and that both ϕ and ϕ^{-1} are continuous. Then ϕ does not only identify points in M with points in N, but it also identifies open sets U in M with open sets $\phi(U)$ in N.

Definition A.34 (Homeomorphisms)**.** A *homeomorphism* $\phi\colon M \to N$ is a bijection such that both ϕ and ϕ^{-1} are continuous.

Proposition A.35. *Let $\phi\colon M \to N$ be a homeomorphism. Then open subsets $U \subseteq M$ correspond bijectively with open subsets $V \subseteq N$ via $V = \phi(U)$ and $U = \phi^{-1}(V)$.*

Proof. If $\phi(U)$ is open, then $U = \phi^{-1}(\phi(U))$ is open as well, because ϕ is continuous. Conversely, if U is open, then $\phi(U) = (\phi^{-1})^{-1}(U)$ is open because ϕ^{-1} is continuous. □

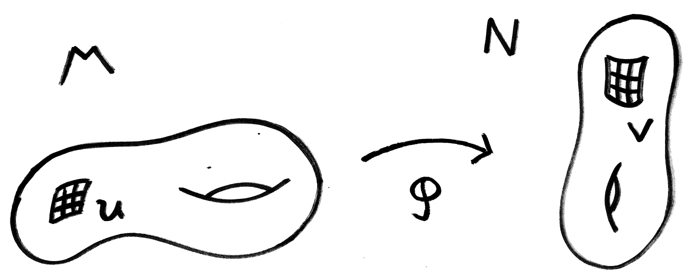


Figure 24: Open sets $U \subseteq M$ correspond to open sets $\phi(U) \subseteq N$

If there is a homeomorphism between M and N, we therefore think of M and N as 'the same space' as far as their topologies are concerned. Such spaces are called *homeomorphic*.

Problem A.36 (n-spheres with different radii are homeomorphic)**.** The n-sphere with radius $r > 0$ is the set $\mathbb{S}^n_r := \{x \in \mathbb{R}^{n+1}\,;\, (x^0)^2 + \ldots + (x^n)^2 = r^2\}$, with the subspace topology inherited from $\mathbb{R}^{n+1}$. Prove that n-spheres with different radii are homeomorphic.

Problem A.37 (Two representations of the circle)**.** Let

$$\mathbb{S}^1 := \{(x, y) \in \mathbb{R}^2\,;\, x^2 + y^2 = 1\}$$

be the unit circle in $\mathbb{R}^2$, equipped with the subspace topology. Let $\mathbb{T}^1 := \mathbb{R}/\mathbb{Z}$ be the 1-torus, equipped with the quotient topology. Prove that $\mathbb{T}^1$ and $\mathbb{S}^1$ are homeomorphic.

Problem A.38. Intuitively, would you think $\mathbb{S}^1$ and $[0, 2\pi)$ ought to be identified as topological spaces?

a) Check that the map $\phi\colon [0, 2\pi) \to \mathbb{S}^1$ with $\theta \mapsto (\cos(\theta), \sin(\theta))$ is bijective and continuous.

b) Is it a homeomorphism?

A.5 Hausdorff spaces

On a topological space M, we are able to define continuity in terms of open sets. Another notion that involves 'nearby points' is the notion of convergence. A sequence x_i *converges* to x if x_i is 'close to' x if i is sufficiently large. Again, we can make sense of this in the context of topological spaces.

Definition A.39 (Convergence)**.** A sequence x_i of points in M *converges* to $x \in M$ if for any open neighbourhood U of x, there exists an $N > 0$ such that $x_i \in U$ for all $i \geq N$.

Again, this is a definition where the ε's apear to have magically vanished. The following exercise shows that they are really hidden in the definition of a topological space.

Problem A.40 (Convergence in $\mathbb{R}^n$)**.** Show that this definition of convergence coincides with the usual definition of convergence in $\mathbb{R}^n$. Namely, a sequence of points x_i in $\mathbb{R}^n$ converges to $x \in \mathbb{R}^n$ if for any $\varepsilon > 0$, there exists an $N > 0$ such that $\|x_i - x\| < \varepsilon$ whenever $i \geq N$.

A.5.1 Hausdorff spaces

Topological spaces allow us to talk about continuity and limits. It turns out, however, that our new notion of limits has a major drawback: it is possible to construct topological spaces where a single sequence x_i converges to two different points x and x' at the same time.

Problem A.41 (A silly counterexample)**.** Let $M = \{0, 1\}$, and let $\mathcal{T}$ be the topology consisting of $\emptyset$ and $\{0, 1\}$. Check that $\mathcal{T}$ is indeed a topology, and that *every* sequence x_i converges to 0 as well as 1.

To exclude such artificial counterexamples, and to ensure that every sequence has at most *one* limit, we introduce the notion of a *Hausdorff space*.

Definition A.42 (Hausdorff spaces)**.** A topological space M is called *Hausdorff* if for any two distinct points $x, y \in M$, there exist open neighbourhoods U_x of x and U_y of y which do not intersect, $U_x \cap U_y = \emptyset$.

Proposition A.43. *In a Hausdorff space M, any sequence has at most one limit.*

Proof. Suppose that a sequence x_i converges to two distinct points x and y in M. Since M is Hausdorff, we can choose open neighbourhoods U_x of x and U_y of y such that $U_x \cap U_y = \emptyset$. Since x_i converges to x, there exists an $N > 0$ such that $x_i \in U_x$ for $i > N$. Similarly, there exists an $M > 0$ such that $x_i \in U_y$ for $i > M$. Apparently, we have $x_i \in U_x \cap U_y$ for $i \geq \max\{N, M\}$, which is a contradiction because $U_x \cap U_y = \emptyset$. $\square$

Problem A.44. The silly example of Problem A.41 is not Hausdorff.

Problem A.45. Prove that the n-torus $\mathbb{T}^n$ is Hausdorff.

Although non-Hausdorff topological spaces are used in some areas of mathematics (algebraic geometry, logic, foliations), we will not need them in this course. We will allways require a smooth manifold to be Hausdorff.

A.5.2 Construction of Hausdorff topological spaces

Now that we got acquainted with topological spaces and the Hausdorff property, let us describe a few ways to construct such spaces. One important source of Hausdorff topological spaces are *metric spaces* (M, d).

Proposition A.46. *Every topological space with a continuous metric is Hausdorff.*

Proof. Let $x, y \in M$ be distinct points in M. If we define $\varepsilon := \frac{1}{3}d(x, y)$, then the neighbourhoods $U_x := B_\varepsilon(x)$ and $U_y := B_\varepsilon(y)$ are open because the metric is continuous, and they are disjoint by the triangle inequality. Indeed, suppose that $z \in U_x \cap U_y$. Then $d(x, z) < \varepsilon$ and $d(y, z) < \varepsilon$, so that $d(x, y) < 2\varepsilon$ by the triangle inequality. But since $2\varepsilon = \frac{2}{3}d(x, y)$, this implies $d(x, y) = 0$, contrary to our assumption that x and y are distinct. □

Once we have established that a space is Hausdorff, all its subspaces will be Hausdorff as well.

Proposition A.47 (Subspaces of Hausdorff spaces are Hausdorff)**.** *If M is Hausdorff, then any subset $\Sigma \subseteq M$ equipped with the subspace topology will be Hausdorff as well.*

Proof. Since M is Hausdorff, two distinct points σ, σ' in Σ admit disjoint open neighbourhoods $U_\sigma \subseteq M$ and $U_{\sigma'} \subseteq M$ inside M. But then $U_\sigma \cap \Sigma$ and $U_{\sigma'} \cap \Sigma$ are disjoint open neighbourhoods in Σ. □

For example, since $\mathbb{R}^{n+1}$ is a Hausdorff space, the sphere

$$\mathbb{S}^n := \{(x^0, \ldots, x^n) \in \mathbb{R}^{n+1}\,;\,(x^0)^2 + \ldots + (x^n)^2 = 1\}$$

is a Hausdorff space as well.

Another way of constructing new topological spaces out of old ones is by taking products. It turns out that the product of two Hausdorff spaces will again be a Hausdorff space.

Proposition A.48. *The product of Hausdorff spaces is Hausdorff.*

Proof. If $(x, y) \in M \times N$ is distinct from $(x', y') \in M \times N$, then either $x \neq x'$ or $y \neq y'$. If $x \neq x'$, then we can choose disjoint neighbourhoods $U_x \subseteq M$ and $U_{x'} \subseteq M$ of x and x' inside M. This yields disjoint neighbourhoods $U_x \times N$ and $U_{x'} \times N$ of (x, y) and (x', y') inside $M \times N$. If $x = x'$, then $y \neq y'$ and we have a similar argument where the roles of M and N are flipped. □

Summarizing, if a topological space is a metric space, a subspace of a Hausdorff space, or a product of Hausdorff spaces, then it will be Hausdorff itself. Unfortunately, a quotient of a Hausdorff space is *not* always Hausdorff.

Example A.49 ($\mathbb{C}\mathrm{P}^n$ as a Hausdorff space)**.** Since we constructed $\mathbb{C}\mathrm{P}^n$ as a quotient space, we have to do some work to prove that it is Hausdorff. By Proposition A.46 it suffices to construct a metric on this space. The angle θ between two nonzero vectors $v, w \in \mathbb{C}^{n+1}$ satisfies

$$\cos(\theta) = \frac{|\langle v, w\rangle|}{\|v\|\|w\|}.$$

Since this angle does not depend on the representative $v \in [v]$ and $w \in [w]$, we can use it to define the *Fubini-Study metric* on $\mathbb{C}\mathrm{P}^n$, denoted d_{FS}. The distance between two rays $[v], [w] \in \mathbb{C}\mathrm{P}^n$ is simply defined to be the angle between them,

$$d_{\mathrm{FS}}([v],[w]) := \arccos\left(\frac{|\langle v, w\rangle|}{\|v\|\|w\|}\right).$$

Since this expression is continuous in v, w on $(\mathbb{C}^{n+1} \setminus \{0\}) \times (\mathbb{C}^{n+1} \setminus \{0\})$, the metric is continuous on $\mathbb{C}\mathrm{P}^n \times \mathbb{C}\mathrm{P}^n$.

A.6 Compact spaces

If a subset K of $\mathbb{R}^n$ is both closed and bounded, then it has the remarkable property that every continuous function $\phi\colon K \to \mathbb{R}$ assumes a maximal value at some point $x \in K$. In order to formulate and prove an analogous statement for general topological spaces M, we need to make sense of the phrase "closed and bounded" also if M is not equipped with a metric. In this context, the correct generalization turns out to be *compactness.*

Definition A.50 (Covers)**.** Let K be a subset of M, and let $\{U_\alpha\,;\,\alpha \in A\}$ be a (not necessarily finite) collection of subsets of M. We say that the sets U_α *cover* K if $K \subseteq \bigcup_{\alpha\in A} U_\alpha$.

Definition A.51 (Compact sets)**.** A subset $K \subseteq M$ is called *compact* if for every open cover U_α of K, there exist finitely many $U_{\alpha_1}, \ldots, U_{\alpha_n}$ that cover K.

A.6.1 Compact subsets of $\mathbb{R}^n$

We prove that the compact subsets of $\mathbb{R}^n$ are precisely the sets which are both closed and bounded. Important examples of compact sets are the closed blocks and balls in $\mathbb{R}^n$. We start by showing that the blocks are compact.

Lemma A.52. *For any $L > 0$, the closed block*

$$C_L(x) := x + [-L/2, L/2]^n \subseteq \mathbb{R}^n$$

centered at $x \in \mathbb{R}^n$ is compact.

Proof. Let $U_\alpha \subseteq \mathbb{R}^n$ be an open cover of $C_L(x)$, and suppose that $C_L(x)$ cannot be covered by any finite collection $U_{\alpha_1}, \ldots, U_{\alpha_n}$. We derive a contradiction by looking at increasingly smaller subblocks.

Since $C_L(x)$ can be covered by 2^n blocks with lenght $\frac{1}{2}L$, at least one of these blocks cannot be covered by finitely many U_α's. (Why?) Let $C_{\frac{1}{2}L}(x_1)$ be one such block. Repeating the above argument for $C_{\frac{1}{2}L}(x_1)$, we find a sub-block $C_{(\frac{1}{2})^2 L}(x_2)$ of length $(\frac{1}{2})^2 L$ that cannot be covered by finitely many U_α's. Continuing in this way, we find a nested sequence of blocks

$$C_L(x) \supset C_{\frac{1}{2}L}(x_1) \supset C_{(\frac{1}{2})^2 L}(x_2) \supset C_{(\frac{1}{2})^3 L}(x_3) \supset \ldots,$$

none of which can be covered by finitely many U_α's.

We show that x_i is a Cauchy sequence, so that it converges to a point $x_\infty \in C_L(x)$. For this, note that if $i, j \geq N$, then x_i and x_j are in the block $C_{(\frac{1}{2})^N L}(x_N)$. It follows that their distance $d(x_i, x_j)$ is at most equal to the diameter of this block, $d(x_i, x_j) \leq (\frac{1}{2})^N L\sqrt{n}$ (recall that n is the dimension). Since this tends to zero for $N \to \infty$, the sequence x_i is Cauchy, and converges to a point $x_\infty \in C_L(x)$.

Since the U_α's cover $C_L(x)$, we have $x_\infty \in U_\beta$ for at least one U_β. Since U_β is open, there exists an $\varepsilon > 0$ such that $y \in U_\beta$ if $d(x_\infty, y) < \varepsilon$. For sufficiently large i, we have that $d(x_i, x_\infty) < \varepsilon/2$, and also $d(x_i, y) < \varepsilon/2$ for any $y \in C_{(\frac{1}{2})^i L}(x_i)$, so the entire block $C_{(\frac{1}{2})^i L}(x_i)$ is contained in U_β.

But then $C_{(\frac{1}{2})^i L}(x_i)$ is covered by finitely many U_α's (it is even covered by a single one), which is a contradiction. $\square$

Other important examples of compact sets are the closed balls $\overline{B}_r(x) \subseteq \mathbb{R}^n$ of radius r. Since $\overline{B}_r(x) \subseteq C_{2r}(x)$, this follows from the following proposition.

Proposition A.53. *A closed subset C of a compact set K is compact.*

Proof. Suppose that the open sets U_α cover C. If we add the single open set $U_0 := M - C$, then we obtain a cover of M, and hence of K. Since K is compact, it is covered by U_0, together with finitely many sets $U_{\alpha_1}, \ldots, U_{\alpha_k}$ from the original cover. Since C is contained in K, these sets cover C as well. But since C does not intersect $U_0 := M - C$, we can omit U_0 from the cover, and conclude that C is already covered by the sets $U_{\alpha_1}, \ldots, U_{\alpha_k}$. $\square$

In particular, any closed, bounded subset of $\mathbb{R}^n$ is compact. To prove the converse, we need the following, surprisingly useful result.

Proposition A.54. *A compact subset K of a Hausdorff space M is closed.*

Proof. We need to prove that $M - K$ is open. Let $x \in M - K$ and $y \in K$. Since M is Hausdorff, there exist open neighbourhoods $U_{x,y}$ of x and $V_{x,y}$ of y which do not intersect, $U_{x,y} \cap V_{x,y} = \emptyset$. Since K is compact and covered by the $V_{x,y}$ with $y \in K$, it can be covered by finitely many sets $V_{x,y_1}, \ldots, V_{x,y_k}$ of this type.

Consider the finite intersection $U := \bigcap_{i=1}^k U_{x,y_i}$, and the finite union $V := \bigcup_{i=1}^k V_{x,y_i}$. Since $x \in U$, $K \subseteq V$, and $U \cap V = \emptyset$ (why?), it follows that

U is an open neighbourhood of x in $M - K$. Since we can find such open neighbourhoods for any $x \in M - K$, it follows that $M - K$ is open, and hence that K is closed. $\square$

We have already seen that every closed, bounded subset of $\mathbb{R}^n$ is compact. Conversely, we can now show that every compact subset $K \subseteq \mathbb{R}^n$ is both closed and bounded. Since $\mathbb{R}^n$ is Hausdorff, every compact subset $K \subseteq \mathbb{R}^n$ is closed by the above proposition. To see that K is also bounded, cover it by open balls $B_N(0)$ centered at the origin with radius N. Since this admits a finite subcover, the set K is contained in a ball with finite radius, which is precisely what it means to be bounded.

Corollary A.55. *A subset of $\mathbb{R}^n$ is compact if and only if it is both closed and bounded.*

The nice thing about compactness is that, quite unlike closedness, it is a property that is preserved by continuous maps.

Corollary A.56. *Let $\phi\colon M \to N$ be a continuous map.*

a) If $K \subseteq M$ is compact, then $\phi(K) \subseteq N$ compact as well.

b) Suppose that M is compact and N is Hausdorff.
Then if $C \subseteq M$ is closed, $\phi(C) \subseteq N$ is closed as well.

c) Suppose that M is compact and N is Hausdorff.
Then every continous bijection $\phi\colon M \to N$ is a homeomorphism.

Proof. For (a), suppose that the open sets U_α cover $\phi(K)$. Since ϕ is continuous, the sets $\phi^{-1}(U_\alpha)$ are open. Since they cover the compact set K, they admit a finite subcover $\phi^{-1}(U_{\alpha_1}), \ldots, \phi^{-1}(U_{\alpha_k})$. Since their images $U_{\alpha_1}, \ldots, U_{\alpha_k}$ constitute a finite cover of $\phi(K)$, the latter is compact.

For (b), note that the closed subset $C \subseteq M$ is compact by Proposition A.53, the continuous image $\phi(C)$ of the compact set C is compact by (a), and the compact subset $\phi(C) \subseteq N$ is closed by Proposition A.54.

For (c), it suffices to prove that the image $\phi(U)$ of every open subset $U \subseteq M$ is open (why?). This follows from (b) as $N - \phi(U) = \phi(M - U)$ is closed. $\square$

If one wants to show that a map $\phi\colon M \to N$ is a homeomorphism, it is often easy to see that it is continuous and bijective, and harder to show that the inverse is continuous. The above corollary allows one to circumvent this.

Problem A.57. Let $\mathbb{T}^2 := \mathbb{R}^2/\mathbb{Z}^2$ be the 2-torus, equipped with the topology defined in §A.3.2. To embed $\mathbb{T}^2$ in $\mathbb{R}^3$, let $0 < r < R$ be the radii of the 'small' and 'big' circle, and let $\phi\colon \mathbb{T}^2 \to \mathbb{R}^3$ be the parameterisation of the torus by two angles x and y,

$$\begin{aligned}\phi(x, y) &= R(\cos(x/2\pi), \sin(x/2\pi), 0)\\ &\quad + r\sin(y/2\pi)(\cos(x/2\pi), \sin(x/2\pi), 0) + r\cos(y/2\pi)(0, 0, 1)\,.\end{aligned}$$

a) Show that $\mathbb{T}^2$ is compact. (Hint: it is the image of a compact subset of $\mathbb{R}^2$ under the continuous projection $\pi\colon \mathbb{R}^2 \to \mathbb{T}^2$.)

b) Show that ϕ is continuous and injective. Conclude that it is a homeomorphism onto its image.

Another interesting application of Corollary A.56 is the following result.

Corollary A.58. *If K is compact, then every continuous map $\phi\colon K \to \mathbb{R}$ assumes a maximal value at some point $x \in K$.*

Proof. Since the image is compact and $\mathbb{R}$ is Hausdorff, $\phi(K) \subseteq \mathbb{R}$ is closed and bounded. Thus the supremum of the set $\phi(K)$ exists, and is a point $\phi(x_{\max})$ in $\phi(K)$. In particular, ϕ assumes a maximal value at the point $x_{\max} \in K$. $\square$

B The inverse function theorem

We prove the Inverse Function Theorem for smooth functions $F\colon U \to \mathbb{R}^n$ on an open neighbourhood $U \subseteq \mathbb{R}^n$ of $p \in \mathbb{R}^n$.

Theorem B.1 (Inverse Function Theorem)**.** *If D_pF is invertible, there exist open neighbourhoods $U_0 \subseteq \mathbb{R}^n$ of $p \in \mathbb{R}^n$ and $V_0 \subseteq \mathbb{R}^n$ of $F(p) \in \mathbb{R}^n$ such that $F|_{U_0}\colon U_0 \to V_0$ is a diffeomorphism.*

To prove the inverse function theorem, we need the following lemma.

Lemma B.2 (Contraction Lemma)**.** *Let $B \subseteq \mathbb{R}^n$ be a nonempty closed set, and let $C\colon B \to B$ be a contraction, meaning that there exists a constant $c < 1$ such that $\|C(x) - C(y)\| \leq c\|x - y\|$ for all $x, y \in B$. Then C has a unique fixed point in B.*

Proof. There is at most one fixed point; if $C(x) = x$ and $C(x') = x'$, then $\|x' - x\| = \|C(x') - C(x)\| < c\|x' - x\|$, so that $x' = x$.

To show that there is at least one fixed point, choose $x_0 \in B$ and let $x_{n+1} := C(x_n)$ for $n \geq 1$. We show that x_n is a Cauchy sequence. Since

$$\|x_{n+1} - x_n\| = \|C(x_n) - C(x_{n-1})\| \leq c\|x_{n+1} - x_n\|,$$

we have $\|x_{n+1} - x_n\| \leq c^n\|x_1 - x_0\|$. For $n > m$, we thus have

$$\begin{aligned}\|x_n - x_m\| &\leq \|x_n - x_{n-1}\| + \ldots + \|x_{m+1} - x_m\| \leq (c^{n-1} + \ldots + c^m)\|x_1 - x_0\| \\ &\leq c^m\big(\textstyle\sum_{k=0}^{\infty} c^k\big)\|x_1 - x_0\| = \dfrac{c^m}{1-c}\|x_1 - x_0\|\,.\end{aligned}$$

It follows that x_n is Cauchy, and hence that $\lim_{n\to\infty} x_n = x$ exists in the closed set B. Since C is continuous (why?), we have $C(x) = \lim_{n\to\infty} C(x_n) = \lim_{n\to\infty} x_{n+1} = x$. ☐

Proof of Theorem B.1. By changing $F(x)$ to $F'(x) = F(x + p) - F(p)$, we may assume that $p = 0$ and $F(p) = 0$. Then, by changing $F(x)$ to $F'(x) = (DF_0)^{-1}F(x)$, we may assume that $DF_0 = \mathbf{1}_n$ is the identity matrix.

First we show that F is *injective* on a closed ball $\overline{B}_\delta(0)$ around the origin. The first order Taylor expansion of F around $x = 0$ is

$$F(x) = x + R^1_0(x).$$

The derivative of the remainder term $R_0(x) = F(x) - x$ is given by $D_xR_0 = D_xF - \mathbf{1}_n$. This expression is continuous in x and zero at $x = 0$, so if we choose $\delta > 0$ sufficiently small, we have $\|D_xR_0\| \leq \frac{1}{2}$ for all $x \in \overline{B}_\delta(0)$. It follows that

$$\|R_0(x') - R_0(x)\| \leq \tfrac{1}{2}\|x' - x\| \tag{139}$$

for all $x, x' \in B_\delta(0)$. To see that F is injective on $\overline{B}_\delta(0)$, note that since $x' - x = (F(x') - F(x)) + (R^1_0(x) - R^1_0(x'))$, we have

$$\|x' - x\| \leq \|F(x') - F(x)\| + \|R^1_0(x) - R^1_0(x')\| \leq \|F(x') - F(x)\| + \tfrac{1}{2}\|x' - x\|,$$

and hence

$$\tfrac{1}{2}\|x' - x\| \leq \|F(x') - F(x)\|. \tag{140}$$

Next, we show that F is *surjective* on an open ball $B_{\delta/2}(0)$. For a given $y \in B_{\delta/2}(0)$, we are looking for an $x \in B_\delta(0)$ such that $F(x) = y$. This is equivalent to $y - F(x) = 0$, and hence to $x - F(x) + y = x$. The good thing about this last equation is that it takes the form $C(x) = x$ for $C(x) := x - F(x) + y$. Since $C(x) = y - R^1_0(x)$ with $\|y\| < \delta/2$ and $\|R^1_0(x)\| \leq \frac{1}{2}\|x\| \leq \delta/2$, we have $C(\overline{B}_\delta(0)) \subseteq B_\delta(0)$. Further, C is a contraction because $\|C(x') - C(x)\| = \|R^1_0(x) - R^1_0(x')\| \leq \frac{1}{2}\|x' - x\|$. By the Contraction Lemma B.2, there is a unique $x \in \overline{B}_\delta(0)$ with $C(x) = x$, and hence $F(x) = y$. Since the image of C is contained in the open ball $B_\delta(0)$, we have $x \in B_\delta(0)$.

We conclude that the restriction of F to $B_{\delta/2}(0)$ is *bijective*, and that F (locally) has an inverse $F^{-1}\colon F(B_{\delta/2}(0)) \to B_{\delta/2}(0)$. To see that this inverse is continuous, substitute $x' = F^{-1}(y')$ and $x = F^{-1}(y)$ in (140) to find

$$\|F^{-1}(y') - F^{-1}(y)\| \leq 2\|y' - y\|. \tag{141}$$

Having shown that $F\colon B_{\delta/2}(0) \to F(B_{\delta/2}(0))$ is a homeomorphism, it remains to prove that F^{-1} is differentiable.

Suppose for a second that we already *knew* that F^{-1} were differentiable. Then the derivative would be easy to calculate: just apply the chain rule to $F^{-1} \circ F(x) = x$. We find that if $y = F(x)$, then $D_y(F^{-1}) \circ D_x F = \mathbf{1}$, so $D_y(F^{-1}) = (D_x F)^{-1}$. So *if* F^{-1} is differentiable in $y \in F(B_{\delta/2}(0))$, then its derivative must be $D_y(F^{-1}) = (D_x F)^{-1}$.

The first thing to prove, then, is that $D_x F$ is invertible for $x \in B_{\delta/2}(0)$. This can be achieved by shrinking δ if necessary. Indeed, since $D_x F$ is invertible if and only if $\det(D_x F) \neq 0$, and since the function $x \mapsto \det(D_x F)$ is continuous with value 1 at $x = 0$, there exists a $\delta > 0$ such that $\det(D_x F) > \frac{1}{2}$ for $\|x\| \leq \delta/2$.

What we have to prove, then, is that F^{-1} is differentiable at y with derivative $(D_x F)^{-1}$. Equivalently, we wish to show that the second order remainder term $\tilde{R}^2_y(y' - y)$ in the Taylor expansion

$$F^{-1}(y') = F^{-1}(y) + (D_x F)^{-1}(y' - y) + \tilde{R}^2_y(y' - y) \tag{142}$$

satifies $\lim_{y' \to y} \|\tilde{R}^2_y(y' - y)\|/\|y' - y\| = 0$. The strategy is to use the similar expansion

$$F(x') = F(x) + (D_x F)(x' - x) + R^2_x(x' - x) \tag{143}$$

for the differentiable function F, where we already *know* that

$$\lim_{x' \to x} \|R^2_x(x' - x)\|/\|x' - x\| = 0.$$

Substituting $x' = F^{-1}(y')$ and $x = F^{-1}(y)$ in (143), we find

$$y' = y + (D_x F)(F^{-1}(y') - F^{-1}(y)) + R^2_x(F^{-1}(y') - F^{-1}(y)).$$

Solving for $F^{-1}(y')$ (using that D_xF is invertible!), this yields

$$F^{-1}(y') = F^{-1}(y) + (D_xF)^{-1}(y'-y) - (D_xF)^{-1}R_x^2(F^{-1}(y') - F^{-1}(y)).$$

Comparing this to (142), we see that the remainder $\tilde{R}_y^2$ in the expansion of F^{-1} can be expressed in terms of the remainder R_x^2 in the expansion of F as

$$\tilde{R}_y^2(y'-y) = -(D_xF)^{-1}R_x^2(F^{-1}(y') - F^{-1}(y)).$$

With $C := \|(D_xF)^{-1}\|$ the matrix norm of $(D_xF)^{-1}$, we have

$$\begin{aligned}\frac{\tilde{R}_y^2(y'-y)}{\|y'-y\|} &\leq \|(D_xF)^{-1}\| \frac{\|R_x^2(F^{-1}(y') - F^{-1}(y))\|}{\|F^{-1}(y') - F^{-1}(y)\|} \frac{\|F^{-1}(y') - F^{-1}(y)\|}{\|y'-y\|} \\ &\leq 2C \frac{\|R_x^2(F^{-1}(y') - F^{-1}(y))\|}{\|F^{-1}(y') - F^{-1}(y)\|}\,, \end{aligned} \tag{144}$$

where the second inequality follows from (141). To show that the limit of (144) for $y' \to y$ is zero, note that since F^{-1} is continuous, $y' \to y$ implies $F^{-1}(y') \to F^{-1}(y)$, so that

$$\lim_{y'\to y} \frac{\|R_x^2(F^{-1}(y') - F^{-1}(y))\|}{\|F^{-1}(y') - F^{-1}(y)\|} = \lim_{x'\to x} \frac{\|R_x^2(x'-x)\|}{\|x'-x\|} = 0\,.$$

It follows that $\lim_{y'\to y} \tilde{R}_y^2(y'-y)/\|y'-y\| = 0$, and we conclude that F^{-1} is differentiable for all $x \in B_{\delta/2}(0)$, with derivative $D_yF^{-1} = (D_xF)^{-1}$. To see that F^{-1} is C^1, note that the derivative $y \mapsto D_yF^{-1}$ is the concatenation of the maps $y \mapsto x := F^{-1}(y)$, $x \mapsto D_xF$, and the matrix inverse $M \mapsto M^{-1}$, all of which are continuous.

We prove by induction that F^{-1} is C^k for every k, and hence that it is smooth. The case $k=1$ is what we proved so far. Suppose by induction that $y \mapsto F^{-1}(y)$ is C^k. Then since the maps $x \mapsto D_xF$ and $M \mapsto M^{-1}$ are C^∞, the concatenation $y \mapsto (D_{F^{-1}(y)}F)^{-1}$ is C^k by the chain rule. It follows that $y \mapsto D_yF^{-1}$ is C^k, and, hence, that F^{-1} is C^{k+1}. $\square$

C Holomorphic functions in several variables

Let D be an open subset of $\mathbb{C}^n$. A function $f\colon D \to \mathbb{C}$ of n complex variables is called *holomorphic* if for every $a \in D$, there exists an open neighbourhood $U \subseteq D$ on which the power series expansion

$$f(z_1, \ldots, z_n) = \sum_{k_1=0}^{\infty} \ldots \sum_{k_n=0}^{\infty} c_{k_1,\ldots,k_n} (z_1 - a_1)^{k_1} \cdots (z_n - a_n)^{k_n} \tag{145}$$

is convergent. A function $F\colon D \to \mathbb{C}^m$ is holomorphic if all of its components $F_j\colon D \to \mathbb{C}$ are holomorphic.

If the power series (145) is convergent at a point $b \in D$ with $r_i := |b_i - a_i| > 0$, then it converges absolutely and uniformly on the open polydisc

$$\Delta(a, r) := \{z \in \mathbb{C}^n\, ;\, |z_i - a_i| < r_i \text{ for } i = 1, \ldots, n\}.$$

In particular, the function f is continuous on D, since it can locally be expressed as a uniform limit of continuous functions. Because the convergence is uniform on $\Delta(a, r)$, the summation in (145) can be performed in any order. From this, it is not hard to see that the function $z_i \mapsto f(z_1, \ldots, z_i, \ldots z_n)$ admits a convergent power series expansion, so that f is holomorphic in each of its arguments.

Lemma C.1 (Osgood's Lemma)**.** *Suppose that $f\colon D \to \mathbb{C}$ is continuous, and that f is holomorphic in each of its variables. Then f is holomorphic on D.*

Proof. For any $a \in D$, we can choose a closed polydisc $\overline{\Delta}(a, r) \subseteq D$ contained in D. Since f is holomorphic in every single variable z_i, we can apply Cauchy's formula to the variables z_1 through z_n successively. For every $z \in \Delta(a, r)$, this yields the repeated integral

$$f(z) = \frac{1}{(2\pi i)^n} \oint_{\gamma_1} \cdots \oint_{\gamma_n} \frac{f(\zeta)}{(\zeta_1 - z_1) \cdots (\zeta_n - z_n)} d\zeta_n \ldots d\zeta_1, \tag{146}$$

where γ_i denotes the curve $\gamma_i = \{\zeta_i \in \mathbb{C}\, ;\, |\zeta_i - a_i| = r_i\}$. Since f is continuous on D, we can use Fubini's Theorem to write this as an n-dimensional integral over the product $\gamma_1 \times \ldots \times \gamma_n \subseteq \mathbb{C}^n$,

$$f(z) = \frac{1}{(2\pi i)^n} \int_{\gamma_1 \times \ldots \times \gamma_n} \frac{f(\zeta)}{(\zeta_1 - z_1) \cdots (\zeta_n - z_n)} d(\zeta_1, \ldots, \zeta_n). \tag{147}$$

Since the geometric series

$$\frac{1}{(\zeta_i - z_i)} = \frac{1}{\zeta_i - a_i} \cdot \frac{1}{1 - \frac{z_i - a_i}{\zeta_i - a_i}} = \sum_{k=0}^{\infty} \frac{(z_i - a_i)^k}{(\zeta_i - a_i)^{k+1}} \tag{148}$$

converges absolutely and uniformly for $\zeta \in \gamma_1 \times \ldots \times \gamma_n$, we can substitute (148) in (147) to obtain a convergent power series expansion (145) with coefficients

$$c_{k_1,\ldots,k_n} := \frac{1}{(2\pi i)^n} \int_{\gamma_1 \times \ldots \times \gamma_n} \frac{f(\zeta)}{(\zeta_1 - a_1)^{k_1+1} \cdots (\zeta_n - a_n)^{k_n+1}} d(\zeta_1, \ldots, \zeta_n).$$

□

Remark C.2. In fact, we have shown that the power series expansion (145) is convergent on any polydisc $\Delta(a, r)$ that is contained in the domain D of f.

By writing $z_i = x_i + iy_i$ as the sum of its real and imaginary parts, we can identify $\mathbb{C}^n = \mathbb{R}^n \oplus i\mathbb{R}^n$ with $\mathbb{R}^{2n}$. Multiplication by i on $\mathbb{R}^{2n}$ is represented by the $2n \times 2n$ matrix

$$J_n = \begin{pmatrix} 0 & \mathbb{I}_n \\ -\mathbb{I}_n & 0 \end{pmatrix},$$

so an $\mathbb{R}$-linear map $A \colon \mathbb{R}^{2n} \to \mathbb{R}^{2m}$ is $\mathbb{C}$-linear if and only if $J_m A = A J_n$. From Osgood's Lemma, we obtain the following useful characterization of holomorphic functions.

Proposition C.3. *A function $F \colon \mathbb{R}^{2n} \supset D \to \mathbb{R}^{2m}$ is holomorphic if and only if it is differentiable, and if the total derivative $D_z F \colon \mathbb{R}^{2n} \to \mathbb{R}^{2m}$ is $\mathbb{C}$-linear for all $z \in D$.*

Proof. By Osgood's Lemma, $F \colon \mathbb{R}^{2n} \supseteq D \to \mathbb{R}^{2m}$ is holomorphic if and only if its components F_j satisfies the Cauchy–Riemann equations in every variable z_i separately. If we split $F_j(z) = u_j(x, y) + iv_j(x, y)$ into a real and imaginary part, the Cauchy–Riemann equations read

$$\frac{\partial u_j}{\partial x_i} = \frac{\partial v_j}{\partial y_i}, \quad \frac{\partial u_j}{\partial y_i} = -\frac{\partial v_j}{\partial x_i}.$$

So F is holomorphic if and only if there exist $m \times n$ matrices A and B such that $D_z F$ takes the form

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix}.$$

This, in turn, is equivalent to $(D_z F) J_n = J_m (D_z F)$. $\square$

In particular, the concatenation $F \circ G$ of two holomorphic functions F and G is holomorphic. Indeed, since $D_{G(z)} F$ and $D_z G$ are $\mathbb{C}$-linear, the same holds for their product $D_z(F \circ G) = D_{G(z)} F \cdot D_z G$.

Just like in the single variable case, holomorphic functions on a connected domain $D \subseteq \mathbb{C}^n$ are completely determined by their values in an arbitrarily small open set.

Proposition C.4. *Let D be a connected, open subset of $\mathbb{C}^n$. If two holomorphic functions $f, g \colon D \to \mathbb{C}$ agree on an open subset of D, then they agree on all of D.*

Proof. Let U be the interior of the closed set $\{z \in D\,;\, f(z) - g(z) = 0\}$. Then U is open in D by definition. It suffices to prove that U is also closed in D. Indeed, $D = U \cup (D \setminus U)$ is then the disjoint union of two open subsets, so Definition A.9 implies that either U or $D \setminus U$ is empty. Since U is nonempty by assumption, $U \setminus D$ must be empty. So $U = D$, and $f(z) = g(z)$ for all $z \in D$.

To prove that U is closed, we show that $a \in U$ for all $a \in \overline{U}$. Let $\Delta(a, r)$ be a polydisc around a that is completely contained in D. Since a is in the closure of U, the set $U \cap \Delta(a, r/2)$ is nonempty. For $b \in U \cap \Delta(a, r/2)$, we have

$a \in \Delta(b, r/2)$, with a polydisc $\Delta(b, r/2)$ that is completely contained in D. By Remark C.2, the power series expansion

$$h(z) = \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} c_{k_1,\dots,k_n}(z_1 - b_1)^{k_1} \cdots (z_n - b_n)^{k_n}$$

for $h(z) := f(z) - g(z)$ is convergent on $\Delta(b, r/2)$. Since the coefficients are given by the partial derivatives

$$c_{k_1,\dots,k_n} = \frac{1}{k_1! \cdots k_n!} \frac{\partial^{k_1+\cdots+k_n}}{\partial z_1 \cdots \partial z_n} h\Big|_{z=b}$$

of h at b, and since h is identically zero on an open neighbourhood of $b \in U$, we have $c_{k_1,\dots,k_n} = 0$. It follows that h is zero on the open neighbourhood $\Delta(b, r/2)$ of a, so in particular $a \in U$. □

Remark C.5. In fact, the proof shows that f and g agree on D if they have the same value and partial derivatives (to all orders) at a single point $a \in D$.

Just like holomorphic functions of a single variable, holomorphic functions of several variables satisfy a maximum modulus principle.

Proposition C.6 (Maximum modulus principle)**.** *Let $D \subseteq \mathbb{C}^n$ be a connected open subset, and let $f\colon D \to \mathbb{C}$ be a holomorphic function on D. If the absolute value $|f(z)|$ has a local maximum on D, then f is constant on D.*

Proof. Suppose that $|f(z)|$ achieves a maximum at $a \in D$, and let $\Delta(a, r) \subseteq D$ be a polydisc in D around a. Note that $g(t) := f((1-t)a + tb))$ is a holomorphic function of a single variable t which has a local maximum at $t = 0$. By the maximum modulus principle for holomorphic functions of a single variable, it is constant on an open neighbourhood of the unit interval $[0, 1] \subseteq \mathbb{C}$. In particular we have $f(b) = g(1) = g(0) = f(a)$ for all $b \in \Delta(a, r)$, so f is constant on $\Delta(a, r)$. By Proposition C.4, this implies that f is constant on D. □

References

[BM94] John C. Baez and Javier P. Muniain, *Gauge fields, knots and gravity*, World scientific, K&E series on Knots and Everything, reprint 2013.

[L03] John M. Lee, *Introduction to smooth manifolds*, Second edition, Graduate texts in Mathematics, Springer, 2013.

[L97] John M. Lee, *Riemannian manifolds: an introduction to curvature*, Graduate texts in Mathematics, Springer, 1997.